

UC Riverside

UC Riverside Electronic Theses and Dissertations

Title

Data Mining Students' Ordinary Handwritten Coursework

Permalink

<https://escholarship.org/uc/item/98p905xs>

Author

Herold, James

Publication Date

2013

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Data Mining Students' Ordinary Handwritten Coursework

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Computer Science

by

James Thomas Herold

June 2013

Dissertation Committee:

Dr. Thomas F. Stahovich, Chairperson
Dr. Christian Shelton
Dr. Victor Zordan
Dr. Eamonn Keogh

Copyright by
James Thomas Herold
2013

The Dissertation of James Thomas Herold is approved:

Committee Chairperson

University of California, Riverside

Acknowledgments

I would like to begin by acknowledging the great mentorship and direction of my advisor Dr. Stahovich, whose guidance has been the primary component of my academic growth. I greatly appreciate all the time, dedication, and wisdom that you have shared with me. I attribute any success I have as a writer and researcher to the great example you have set. Thank you so much for everything and for always doing so with a smile and kindness.

Next I would like to acknowledge my amazing family. Mom, Dad, and Lianne, thank you for your constant love and support. Without you the past five years would have been much more difficult. I often say that I would not be where I am today if it was not for you guys, because its true. I love you very much.

A big thank you to all my friends who have supported me these past years. I appreciate your understanding when I have said, "I can't tonight, I have school stuff."

I want to share my deep appreciation of all the undergraduate students who have helped me. Through your hard work, you have developed some incredible tools and data sets. Alex Zundel, Matt Ung, John Reaver, Rachel Ye, Jean Dibgolongo, Mike Smick, Ray Shum, and Andy Xie thank you so much.

Lastly, I want to thank everyone in the Smart Tools Lab. Working with you all has been both a great pleasure and a fantastic learning experience. I want to thank everyone for all the help you have given and for all the things you have taught me. Big ups to my homies: Weesan Lee, Tyler Bischel, Eric Peterson, Josiah Jordan, Jack Lee, HanLong Lin, Tim VanArsdale, Kevin Rawson, Reyna Garcia, Nicholas Rhodes, and

Justin Gyllen. I will miss you guys.

I dedicate this thesis to the love of my life, my wife. Girls thanks for the smiling support all the years; you have made this whole journey possible for me. I appreciate all your understanding whenever there were late nights and am especially thankful for all the small and large things you have done to support me these past five years. I always feel so blessed whenever I take half a second to think about us. Tracy I love you a bunch and cannot wait to move on to the next phase in our lives.

Matt 22:21

ABSTRACT OF THE DISSERTATION

Data Mining Students' Ordinary Handwritten Coursework

by

James Thomas Herold

Doctor of Philosophy, Graduate Program in Computer Science

University of California, Riverside, June 2013

Dr. Thomas F. Stahovich, Chairperson

Educational Data Mining is a nascent, but rapidly growing field in which data mining techniques are applied to educational data to discover patterns in the ways that students learn. Research in this field has typically been applied to educational data extracted from digital systems such as Learning Content Management Systems and Intelligent Tutoring Systems. The research thus far has been able to identify interesting patterns in the ways students learn when using these systems, but an analysis of students' ordinary problem-solving processes remains unexplored.

In this work we apply data mining and machine learning techniques to a digital data set of students' ordinary, handwritten coursework in the context of a Mechanical Engineering course. This work makes four major contributions. It is the first, to our knowledge, study in which data mining and machine learning techniques have been applied to students' problem-solving processes in their ordinary learning environment, using pen and paper at home or in the classroom. Because the data set is unique, we provide an in-depth description of the large digital collection of students' handwritten

coursework we have collected. Second, we investigate novel, discrete and numerical representations of students' handwritten coursework which characterize different aspects of students' ordinary problem-solving processes. Third, we identify patterns in these representations as well as correlations between them and course performance using various machine learning and data mining techniques. Last and most important, we present insights may be gained from these patterns and correlations. These insights enable instructors of the course, which we have investigated in this work, to improve future offerings.

Contents

List of Figures	xiii
List of Tables	xxxi
1 Introduction	1
2 Related Work	5
2.1 Theories of Learning	5
2.2 Data Mining	8
2.3 Educational Data Mining	10
3 Data and Experimentation	13
3.1 Introduction	13
3.2 Homework Descriptions	14
3.3 2012 Midterm Grading Rubric	24
3.4 Digital Data and Labeling	27
3.4.1 2010 Labeling Scheme	32
3.4.2 2011 Labeling Scheme	32
3.4.3 2012 Labeling Schemes	33
3.5 Descriptive Statistics	40
3.6 Experimental Studies	56
3.6.1 2011 Experimental Studies	56
3.6.2 2012 Experimental Studies	59
3.6.2.1 Homework Three Self-Explanation Prompts	60
3.6.2.2 Homework Four Self-Explanation Prompts	66
3.6.2.3 Homework Five Self-Explanation Prompts	69
3.6.2.4 Homework Eight Self-Explanation Prompts	74
4 Open IE and Self-Explanation Transcripts	75
4.1 Introduction	75
4.2 Open Information Extraction Algorithm	77
4.3 Results and Discussion	79
4.4 Conclusion	82

5	N-Gram Analysis and Problem Number Sequences	83
5.1	Introduction	83
5.2	Analysis	85
5.2.1	Concept Inventory Analysis	86
5.2.2	Grade Analysis	86
5.2.3	Transition Probability Analysis	87
5.2.4	<i>N</i> -gram Analysis	88
5.3	Discussion	92
5.4	Conclusion	93
6	Differential Sequence Mining and Action Sequences	95
6.1	Introduction	95
6.2	Action Sequences	96
6.3	Differential Mining	100
6.4	Performance Prediction	102
6.5	Discussion	103
6.6	Conclusion	106
7	K-Means Clustering and Solution Bitmaps	108
7.1	Introduction	108
7.2	Solution Bitmaps	109
7.3	Clustering	110
7.4	Group Analysis	112
7.5	Discussion	119
7.6	Conclusion	120
8	Linear Regression and Effort-Based, Numerical Features	122
8.1	Introduction	122
8.2	Computing an Estimation of Student Effort	123
8.3	Overall-Effort Features and Performance Models	124
8.4	Per-Problem Features	127
8.5	Discussion	129
8.6	Conclusion	136
9	Information Gain and Effort-Based, Numerical Features	138
9.1	Introduction	138
9.2	Method	139
9.2.1	Features	140
9.2.2	Error groupings	141
9.2.3	Information Gain Prediction	143
9.3	Results	148
9.3.1	Category Error Grouping Information Gain Heat Maps	149
9.3.2	Fine-grained Error Grouping	155
9.3.3	Subcategory Error Grouping	155
9.4	Guided Qualitative Analysis	163
9.4.1	Midterm One Problem One	165

9.4.2	Midterm Two Problem One	169
9.4.3	Midterm Two Problem Two	172
9.4.4	Midterm Two Problem Three	175
9.4.5	Common Qualities of Predictive Homework Problems	177
9.5	Conclusion	178
10	Machine Learning Based Sketch Processing Techniques	180
10.1	Introduction	180
10.2	User Study Data Set	181
10.3	The 1 ^c Recognizer	182
10.3.1	Introduction	182
10.3.2	Related Work	184
10.3.3	Approach	187
10.3.3.1	Resample	187
10.3.3.2	One Dimensional Representation	187
10.3.3.3	Comparing Templates	189
10.3.4	Results	190
10.3.5	Discussion	194
10.3.6	Conclusion	195
10.4	ClassySeg	196
10.4.1	Introduction	196
10.4.2	Related Work	198
10.4.3	Approach	200
10.4.3.1	Candidate Point Selection	200
10.4.3.2	Arc Length	201
10.4.3.3	Curvature	202
10.4.3.4	Straw	203
10.4.3.5	Alpha and Beta	205
10.4.3.6	Region of Support	206
10.4.3.7	Speed	207
10.4.3.8	Arc and Line Fit	208
10.4.3.9	Classifier Training	209
10.4.4	Results	211
10.4.5	Conclusion	214
10.5	Speech-Sketch Alignment	215
10.5.1	Introduction	215
10.5.2	Related Work	220
10.5.3	Background: Distinguishing Gesture and Object Strokes	227
10.5.3.1	Classifier Design	229
10.5.3.2	Sketch Features	230
10.5.3.3	Speech Features	232
10.5.4	Results: Gesture/Object Classification Accuracy	234
10.5.5	Evaluation of Three-Second Technique	234
10.5.6	Manual Alignment	235
10.5.7	Alignment Accuracy of Three-Second Technique	239

10.5.8	Classifier-Based Alignment	242
10.5.8.1	Speech Segmentation	242
10.5.8.2	Segmentation Classifier and Features	244
10.5.8.3	Segmentation Accuracy	247
10.5.9	Stroke-Speech Alignment	249
10.5.9.1	Stroke-Speech Alignment Accuracy	251
10.5.10	Gesture/Object Classification Accuracy	255
10.5.11	Discussion	257
10.5.12	Conclusion	262
11	Conclusions	265
	Bibliography	274

List of Figures

3.1	The Livescribe™ digital pen	15
3.2	A typical Statics problem.	16
3.3	A hypothetical solution to a typical problem.	17
3.4	Typical problem from homework assignment three from 2010	18
3.5	Typical problem from homework assignment four from 2010	18
3.6	Typical problem from homework assignment five from 2010	19
3.7	Typical problem from homework assignment six from 2010	20
3.8	Typical problem from homework assignment seven from 2010	21
3.9	Typical problem from homework assignment eight from 2010	22
3.10	Typical problem from homework assignment nine from 2010	23
3.11	FBD pen strokes	34
3.12	Equation pen strokes	35
3.13	Organization, explanation, and cross-out pen strokes	36
3.14	Sum-of-moments and sum-of-forces pen strokes	37
3.15	Geometry pen strokes	38
3.16	Moment and force direction pen strokes	39
3.17	Problem statement pen strokes	39
3.18	Average total ink for the 2010 data	41
3.19	Average total time spent writing for the 2010 data	41
3.20	Average cross-out ink written for the 2010 data	42
3.21	Average time spent writing cross-outs for the 2010 data	42
3.22	Average duration for the 2010 data	43
3.23	Average equation ink for the 2010 data	43
3.24	Average time writing equations for the 2010 data	44
3.25	Average FBD ink written for the 2010 data	45
3.26	Average time writing FBDs for the 2010 data	45
3.27	Average number of sheets used for the 2010 data	46
3.28	Average time writing each problem 2011 data	47
3.29	Average number of sheets used for the 2011 data	48
3.30	Average amount of ink written on each problem for the 2011 data	48
3.31	Average duration for the 2011 data	49
3.32	Average time to solve each problem for the 2012 data	50

3.33	Average ink written for the 2012 data	51
3.34	Average number of pages used for the 2012 data	51
3.35	Average time spent writing for the 2012 data	52
3.36	Average amount of FBD ink written for the 2012 data	52
3.37	Average time writing equations for the 2012 data	53
3.38	Average amount of equation ink for the 2012 data	53
3.39	Average duration for the 2012 data	54
3.40	Average time crossing out work for the 2012 data	54
3.41	Average cross-out ink for the 2012 data	55
3.42	Self-explanation prompt for homework three problem four	64
3.43	Self-explanation prompt for homework three problem five	65
3.44	Self-explanation prompt for homework five problem three	70
3.45	Self-explanation prompt for homework five problem six	71
3.46	Self-explanation prompt for homework five problem two	73
3.47	Self-explanation prompt for homework five problem five	74
5.1	Pre/posttest scores on the concept inventory 2012	86
5.2	Average score on each homework assignment for the 2012 data	87
5.3	Average transition frequencies for SE and SO groups	88
5.4	Difference in Dice’s coefficients for the SE and SO groups	90
5.5	Difference in mutual information for the SE and SO groups	91
6.1	Histogram of the durations of FBD actions	98
6.2	Histogram of the durations of EQN actions	99
7.1	A student’s handwritten solution and corresponding solution bitmaps.	111
7.2	Typical solution bitmaps from group one of problem two.	113
7.3	Typical solution bitmaps from group two of problem two	114
7.4	Typical solution bitmaps from group three of problem two.	115
7.5	Typical solution bitmaps from group one of problem six.	116
7.6	Typical solution bitmaps from group two of problem six.	117
7.7	Typical solution bitmaps from group three of problem six.	118
7.8	Typical solution bitmaps from groups four, five, and six of problem six.	119
8.1	Typical effort time-series of a single student on a single assignment.	125
8.2	Effort time-series with active and inactive regions identified.	126
8.3	R^2 for linear regression models predicting performance	127
8.4	R^2 of models predicting coursework performance	128
8.5	R^2 of models predicting exam performance	129
8.6	An example of students’ knowledge transfer	135
9.1	Heat map, midterm one problem one, <i>category error grouping</i>	150
9.2	Heat map, midterm two problem one, <i>category error grouping</i>	151
9.3	Heat map, midterm two problem two, <i>category error grouping</i>	152
9.4	Heat map, midterm two problem three, <i>category error grouping</i>	153
9.5	Heat map, midterm one problem one, <i>fine-grained error grouping</i>	156

9.6	Heat map, midterm two problem one, <i>fine-grained error grouping</i>	157
9.7	Heat map, midterm two problem two, <i>fine-grained error grouping</i>	158
9.8	Heat map, midterm two problem three, <i>fine-grained error grouping</i>	159
9.9	Heat map, midterm one problem one, <i>subcategory error groupings</i>	161
9.10	Heat map, midterm two problem, <i>subcategory error groupings</i>	162
9.11	Heat map, midterm two problem two, <i>subcategory error groupings</i>	162
9.12	Heat map, midterm two problem three, <i>subcategory error groupings</i>	163
9.13	Problem description for problem one of midterm one	166
9.14	Problem description for problem two of homework three	167
9.15	Problem description for problem five of homework four	167
9.16	Problem description for problem one of midterm two	168
9.17	Problem description for problem three of homework five	169
9.18	Problem description for problem six of homework five	169
9.19	Problem description for problem four of homework six	170
9.20	Problem description for problem six of homework six	170
9.21	Problem description for problem two of midterm two	173
9.22	Problem description for problem one of homework five	174
9.23	Problem description for problem three of midterm two	176
9.24	Problem description for problem two of homework six	176
10.1	The ten shapes used to evaluate sketch processing techniques	182
10.2	Rotational invariance example	188
10.3	The templatizing process	189
10.4	Average accuracy versus number of training examples	191
10.5	Per participant accuracy results	192
10.6	Average per-shape accuracy results	194
10.7	Example of straw values for a pen stroke	204
10.8	Examples of α and β values of a point	206
10.9	Example region of support features for a point	207
10.10	Hand drawn sketches of simple devices	217
10.11	Pen strokes and corresponding speech	219
10.12	Devices from the device description study	228
10.13	Example of manual segmentation	237
10.14	Gesture pen stroke and associated speech	240
10.15	Accuracy of the three-second technique	241
10.16	The classification of the words in a spoken passage	242
10.17	Matching classifier based segments	248
10.18	Alignment accuracy after the first step of alignment	250
10.19	Classifier-based alignment accuracy	252
10.20	Missing and extra words in aligned speech	253
10.21	Average number of incorrectly aligned words	254
10.22	Classification accuracy vs. alignment technique	256
10.23	Histograms of the thesaurus filter feature	260
10.24	Classification accuracy for strokes with associated speech	261

.1	Histogram of homework two problem one 2010 durations	286
.2	Histogram of homework two problem one 2010 ink written	287
.3	Histogram of homework two problem one 2010 sheet counts	287
.4	Histogram of homework two problem one 2010 time writing	288
.5	Histogram of homework two problem two 2010 durations	288
.6	Histogram of homework two problem two 2010 ink written	288
.7	Histogram of homework two problem two 2010 sheet counts	289
.8	Histogram of homework two problem two 2010 time writing	289
.9	Histogram of homework two problem three 2010 durations	289
.10	Histogram of homework two problem three 2010 ink written	290
.11	Histogram of homework two problem three 2010 sheet counts	290
.12	Histogram of homework two problem three 2010 time writing	291
.13	Histogram of homework two problem four 2010 durations	291
.14	Histogram of homework two problem four 2010 ink written	292
.15	Histogram of homework two problem four 2010 sheet counts	292
.16	Histogram of homework two problem four 2010 time writing	292
.17	Histogram of homework two problem five 2010 durations	293
.18	Histogram of homework two problem five 2010 ink written	293
.19	Histogram of homework two problem five 2010 sheet counts	293
.20	Histogram of homework two problem five 2010 time writing	294
.21	Histogram of homework two problem six 2010 durations	294
.22	Histogram of homework two problem six 2010 ink written	295
.23	Histogram of homework two problem six 2010 sheet counts	295
.24	Histogram of homework two problem six 2010 time writing	296
.25	Histogram of homework two problem seven 2010 durations	296
.26	Histogram of homework two problem seven 2010 ink written	296
.27	Histogram of homework two problem seven 2010 sheet counts	297
.28	Histogram of homework two problem seven 2010 time writing	297
.29	Histogram of homework three problem one 2010 durations	297
.30	Histogram of homework three problem one 2010 ink written	298
.31	Histogram of homework three problem one 2010 sheet counts	298
.32	Histogram of homework three problem one 2010 time writing	299
.33	Histogram of homework three problem two 2010 durations	299
.34	Histogram of homework three problem two 2010 ink written	300
.35	Histogram of homework three problem two 2010 sheet counts	300
.36	Histogram of homework three problem two 2010 time writing	300
.37	Histogram of homework three problem three 2010 durations	301
.38	Histogram of homework three problem three 2010 ink written	301
.39	Histogram of homework three problem three 2010 sheet counts	301
.40	Histogram of homework three problem three 2010 time writing	302
.41	Histogram of homework three problem three 2010 durations	302
.42	Histogram of homework three problem four 2010 ink written	303
.43	Histogram of homework three problem four 2010 sheet counts	303
.44	Histogram of homework three problem four 2010 time writing	304
.45	Histogram of homework three problem five 2010 durations	304

.46	Histogram of homework three problem five 2010 ink written	304
.47	Histogram of homework three problem five 2010 sheet counts	305
.48	Histogram of homework three problem five 2010 time writing	305
.49	Histogram of homework three problem six 2010 durations	305
.50	Histogram of homework three problem six 2010 ink written	306
.51	Histogram of homework three problem six 2010 sheet counts	306
.52	Histogram of homework three problem six 2010 time writing	307
.53	Histogram of homework three problem seven 2010 durations	307
.54	Histogram of homework three problem seven 2010 ink written	308
.55	Histogram of homework three problem seven 2010 sheet counts	308
.56	Histogram of homework three problem seven 2010 time writing	308
.57	Histogram of homework three problem eight 2010 durations	309
.58	Histogram of homework three problem eight 2010 ink written	309
.59	Histogram of homework three problem eight 2010 sheet counts	309
.60	Histogram of homework three problem eight 2010 time writing	310
.61	Histogram of homework four problem one 2010 durations	310
.62	Histogram of homework four problem one 2010 ink written	311
.63	Histogram of homework four problem one 2010 sheet counts	311
.64	Histogram of homework four problem one 2010 time writing	312
.65	Histogram of homework four problem two 2010 durations	312
.66	Histogram of homework four problem two 2010 ink written	312
.67	Histogram of homework four problem two 2010 sheet counts	313
.68	Histogram of homework four problem two 2010 time writing	313
.69	Histogram of homework five problem one 2010 durations	313
.70	Histogram of homework five problem one 2010 ink written	314
.71	Histogram of homework five problem one 2010 sheet counts	314
.72	Histogram of homework five problem one 2010 time writing	315
.73	Histogram of homework five problem two 2010 durations	315
.74	Histogram of homework five problem two 2010 ink written	316
.75	Histogram of homework five problem two 2010 sheet counts	316
.76	Histogram of homework five problem two 2010 time writing	316
.77	Histogram of homework five problem three 2010 durations	317
.78	Histogram of homework five problem three 2010 ink written	317
.79	Histogram of homework five problem three 2010 sheet counts	317
.80	Histogram of homework five problem three 2010 time writing	318
.81	Histogram of homework five problem four 2010 durations	318
.82	Histogram of homework five problem four 2010 ink written	319
.83	Histogram of homework five problem four 2010 sheet counts	319
.84	Histogram of homework five problem four 2010 time writing	320
.85	Histogram of homework five problem five 2010 durations	320
.86	Histogram of homework five problem five 2010 ink written	320
.87	Histogram of homework five problem five 2010 sheet counts	321
.88	Histogram of homework five problem five 2010 time writing	321
.89	Histogram of homework five problem six 2010 durations	321
.90	Histogram of homework five problem six 2010 ink written	322

.91	Histogram of homework five problem six 2010 sheet counts	322
.92	Histogram of homework five problem six 2010 time writing	323
.93	Histogram of homework six problem one 2010 durations	323
.94	Histogram of homework six problem one 2010 ink written	324
.95	Histogram of homework six problem one 2010 sheet counts	324
.96	Histogram of homework six problem one 2010 time writing	324
.97	Histogram of homework six problem two 2010 durations	325
.98	Histogram of homework six problem two 2010 ink written	325
.99	Histogram of homework six problem two 2010 sheet counts	325
.100	Histogram of homework six problem two 2010 time writing	326
.101	Histogram of homework six problem three 2010 durations	326
.102	Histogram of homework six problem three 2010 ink written	327
.103	Histogram of homework six problem three 2010 sheet counts	327
.104	Histogram of homework six problem three 2010 time writing	328
.105	Histogram of homework six problem four 2010 durations	328
.106	Histogram of homework six problem four 2010 ink written	328
.107	Histogram of homework six problem four 2010 sheet counts	329
.108	Histogram of homework six problem four 2010 time writing	329
.109	Histogram of homework six problem five 2010 durations	329
.110	Histogram of homework six problem five 2010 ink written	330
.111	Histogram of homework six problem five 2010 sheet counts	330
.112	Histogram of homework six problem five 2010 time writing	331
.113	Histogram of homework six problem six 2010 durations	331
.114	Histogram of homework six problem six 2010 ink written	332
.115	Histogram of homework six problem six 2010 sheet counts	332
.116	Histogram of homework six problem six 2010 time writing	332
.117	Histogram of homework six problem seven 2010 durations	333
.118	Histogram of homework six problem seven 2010 ink written	333
.119	Histogram of homework six problem seven 2010 sheet counts	333
.120	Histogram of homework six problem seven 2010 time writing	334
.121	Histogram of homework six problem eight 2010 durations	334
.122	Histogram of homework six problem eight 2010 ink written	335
.123	Histogram of homework six problem eight 2010 sheet counts	335
.124	Histogram of homework six problem eight 2010 time writing	336
.125	Histogram of homework seven problem one 2010 durations	336
.126	Histogram of homework seven problem one 2010 ink written	336
.127	Histogram of homework seven problem one 2010 sheet counts	337
.128	Histogram of homework seven problem one 2010 time writing	337
.129	Histogram of homework seven problem two 2010 durations	337
.130	Histogram of homework seven problem two 2010 ink written	338
.131	Histogram of homework seven problem two 2010 sheet counts	338
.132	Histogram of homework seven problem two 2010 time writing	339
.133	Histogram of homework seven problem three 2010 durations	339
.134	Histogram of homework seven problem three 2010 ink written	340
.135	Histogram of homework seven problem three 2010 sheet counts	340

.136	Histogram of homework seven problem three 2010 time writing	340
.137	Histogram of homework seven problem four 2010 durations	341
.138	Histogram of homework seven problem four 2010 ink written	341
.139	Histogram of homework seven problem four 2010 sheet counts	341
.140	Histogram of homework seven problem four 2010 time writing	342
.141	Histogram of homework seven problem five 2010 durations	342
.142	Histogram of homework seven problem five 2010 ink written	343
.143	Histogram of homework seven problem five 2010 sheet counts	343
.144	Histogram of homework seven problem five 2010 time writing	344
.145	Histogram of homework eight problem one 2010 durations	344
.146	Histogram of homework eight problem one 2010 ink written	344
.147	Histogram of homework eight problem one 2010 sheet counts	345
.148	Histogram of homework eight problem one 2010 time writing	345
.149	Histogram of homework eight problem two 2010 durations	345
.150	Histogram of homework eight problem two 2010 ink written	346
.151	Histogram of homework eight problem two 2010 sheet counts	346
.152	Histogram of homework eight problem two 2010 time writing	347
.153	Histogram of homework eight problem three 2010 durations	347
.154	Histogram of homework eight problem three 2010 ink written	348
.155	Histogram of homework eight problem three 2010 sheet counts	348
.156	Histogram of homework eight problem three 2010 time writing	348
.157	Histogram of homework eight problem four 2010 durations	349
.158	Histogram of homework eight problem four 2010 ink written	349
.159	Histogram of homework eight problem four 2010 sheet counts	349
.160	Histogram of homework eight problem four 2010 time writing	350
.161	Histogram of homework eight problem five 2010 durations	350
.162	Histogram of homework eight problem five 2010 ink written	351
.163	Histogram of homework eight problem five 2010 sheet counts	351
.164	Histogram of homework eight problem five 2010 time writing	352
.165	Histogram of homework eight problem six 2010 durations	352
.166	Histogram of homework eight problem six 2010 ink written	352
.167	Histogram of homework eight problem six 2010 sheet counts	353
.168	Histogram of homework eight problem six 2010 time writing	353
.169	Histogram of homework eight problem seven 2010 durations	353
.170	Histogram of homework eight problem seven 2010 ink written	354
.171	Histogram of homework eight problem seven 2010 sheet counts	354
.172	Histogram of homework eight problem seven 2010 time writing	355
.173	Histogram of homework nine problem one 2010 durations	355
.174	Histogram of homework nine problem one 2010 ink written	356
.175	Histogram of homework nine problem one 2010 sheet counts	356
.176	Histogram of homework nine problem one 2010 time writing	356
.177	Histogram of homework nine problem two 2010 durations	357
.178	Histogram of homework nine problem tow 2010 ink written	357
.179	Histogram of homework nine problem two 2010 sheet counts	357
.180	Histogram of homework nine problem two 2010 time writing	358

.181	Histogram of homework nine problem three 2010 durations	358
.182	Histogram of homework nine problem three 2010 ink written	359
.183	Histogram of homework nine problem three 2010 sheet counts	359
.184	Histogram of homework nine problem three 2010 time writing	360
.185	Histogram of homework nine problem four 2010 durations	360
.186	Histogram of homework nine problem four 2010 ink written	360
.187	Histogram of homework nine problem four 2010 sheet counts	361
.188	Histogram of homework nine problem four 2010 time writing	361
.189	Histogram of homework nine problem five 2010 durations	361
.190	Histogram of homework nine problem five 2010 ink written	362
.191	Histogram of homework nine problem five 2010 sheet counts	362
.192	Histogram of homework nine problem five 2010 time writing	363
.193	Histogram of homework one problem one 2011 durations	363
.194	Histogram of homework one problem one 2011 ink written	364
.195	Histogram of homework one problem one 2011 sheet counts	364
.196	Histogram of homework one problem one 2011 time writing	365
.197	Histogram of homework one problem two 2011 durations	365
.198	Histogram of homework one problem two 2011 ink written	365
.199	Histogram of homework one problem two 2011 sheet counts	366
.200	Histogram of homework one problem two 2011 time writing	366
.201	Histogram of homework one problem three 2011 durations	366
.202	Histogram of homework one problem three 2011 ink written	367
.203	Histogram of homework one problem two 2011 sheet counts	367
.204	Histogram of homework one problem three 2011 time writing	368
.205	Histogram of homework one problem four 2011 durations	368
.206	Histogram of homework one problem four 2011 ink written	369
.207	Histogram of homework one problem four 2011 sheet counts	369
.208	Histogram of homework one problem four 2011 time writing	369
.209	Histogram of homework one problem five 2011 durations	370
.210	Histogram of homework one problem five 2011 ink written	370
.211	Histogram of homework one problem five 2011 sheet counts	370
.212	Histogram of homework one problem five 2011 time writing	371
.213	Histogram of homework two problem one 2011 durations	371
.214	Histogram of homework two problem one 2011 ink written	372
.215	Histogram of homework two problem one 2011 sheet counts	372
.216	Histogram of homework two problem one 2011 time writing	373
.217	Histogram of homework two problem two 2011 durations	373
.218	Histogram of homework two problem two 2011 ink written	373
.219	Histogram of homework two problem two 2011 sheet counts	374
.220	Histogram of homework two problem two 2011 time writing	374
.221	Histogram of homework two problem three 2011 durations	374
.222	Histogram of homework two problem three 2011 ink written	375
.223	Histogram of homework two problem three 2011 sheet counts	375
.224	Histogram of homework two problem three 2011 time writing	376
.225	Histogram of homework two problem four 2011 durations	376

.226	Histogram of homework two problem four 2011 ink written	377
.227	Histogram of homework two problem four 2011 sheet counts	377
.228	Histogram of homework two problem four 2011 time writing	377
.229	Histogram of homework two problem five 2011 durations	378
.230	Histogram of homework two problem five 2011 ink written	378
.231	Histogram of homework two problem five 2011 sheet counts	378
.232	Histogram of homework two problem five 2011 time writing	379
.233	Histogram of homework two problem six 2011 durations	379
.234	Histogram of homework two problem six 2011 ink written	380
.235	Histogram of homework two problem six 2011 sheet counts	380
.236	Histogram of homework two problem six 2011 time writing	381
.237	Histogram of homework two problem seven 2011 durations	381
.238	Histogram of homework two problem seven 2011 ink written	381
.239	Histogram of homework two problem seven 2011 sheet counts	382
.240	Histogram of homework two problem seven 2011 time writing	382
.241	Histogram of homework two problem eight 2011 durations	382
.242	Histogram of homework two problem eight 2011 ink written	383
.243	Histogram of homework two problem eight 2011 sheet counts	383
.244	Histogram of homework two problem eight 2011 time writing	384
.245	Histogram of homework three problem one 2011 durations	384
.246	Histogram of homework three problem one 2011 ink written	385
.247	Histogram of homework three problem one 2011 sheet counts	385
.248	Histogram of homework three problem one 2011 time writing	385
.249	Histogram of homework three problem two 2011 durations	386
.250	Histogram of homework three problem two 2011 ink written	386
.251	Histogram of homework three problem two 2011 sheet counts	386
.252	Histogram of homework three problem two 2011 time writing	387
.253	Histogram of homework three problem three 2011 durations	387
.254	Histogram of homework three problem three 2011 ink written	388
.255	Histogram of homework three problem three 2011 sheet counts	388
.256	Histogram of homework three problem three 2011 time writing	389
.257	Histogram of homework three problem four 2011 durations	389
.258	Histogram of homework three problem four 2011 ink written	389
.259	Histogram of homework three problem four 2011 sheet counts	390
.260	Histogram of homework three problem four 2011 time writing	390
.261	Histogram of homework three problem five 2011 durations	390
.262	Histogram of homework three problem five 2011 ink written	391
.263	Histogram of homework three problem five 2011 sheet counts	391
.264	Histogram of homework three problem five 2011 time writing	392
.265	Histogram of homework three problem six 2011 durations	392
.266	Histogram of homework three problem six 2011 ink written	393
.267	Histogram of homework three problem six 2011 sheet counts	393
.268	Histogram of homework three problem six 2011 time writing	393
.269	Histogram of homework three problem seven 2011 durations	394
.270	Histogram of homework three problem seven 2011 ink written	394

.271	Histogram of homework three problem seven 2011 sheet counts	394
.272	Histogram of homework three problem seven 2011 time writing	395
.273	Histogram of homework three problem eight 2011 durations	395
.274	Histogram of homework three problem eight 2011 ink written	396
.275	Histogram of homework three problem eight 2011 sheet counts	396
.276	Histogram of homework three problem eight 2011 time writing	397
.277	Histogram of homework four problem one 2011 durations	397
.278	Histogram of homework four problem one 2011 ink written	397
.279	Histogram of homework four problem one 2011 sheet counts	398
.280	Histogram of homework four problem one 2011 time writing	398
.281	Histogram of homework four problem two 2011 durations	398
.282	Histogram of homework four problem two 2011 ink written	399
.283	Histogram of homework four problem two 2011 sheet counts	399
.284	Histogram of homework four problem two 2011 time writing	400
.285	Histogram of homework four problem three 2011 durations	400
.286	Histogram of homework four problem three 2011 ink written	401
.287	Histogram of homework four problem three 2011 sheet counts	401
.288	Histogram of homework four problem three 2011 time writing	401
.289	Histogram of homework four problem four 2011 durations	402
.290	Histogram of homework four problem four 2011 ink written	402
.291	Histogram of homework four problem four 2011 sheet counts	402
.292	Histogram of homework four problem four 2011 time writing	403
.293	Histogram of homework four problem five 2011 durations	403
.294	Histogram of homework four problem five 2011 ink written	404
.295	Histogram of homework four problem five 2011 sheet counts	404
.296	Histogram of homework four problem five 2011 time writing	405
.297	Histogram of homework four problem six 2011 durations	405
.298	Histogram of homework four problem six 2011 ink written	405
.299	Histogram of homework four problem six 2011 sheet counts	406
.300	Histogram of homework four problem six 2011 time writing	406
.301	Histogram of homework four problem seven 2011 durations	406
.302	Histogram of homework four problem seven 2011 ink written	407
.303	Histogram of homework four problem seven 2011 sheet counts	407
.304	Histogram of homework four problem seven 2011 time writing	408
.305	Histogram of homework five problem one 2011 durations	408
.306	Histogram of homework five problem one 2011 ink written	409
.307	Histogram of homework five problem one 2011 sheet counts	409
.308	Histogram of homework five problem one 2011 time writing	409
.309	Histogram of homework five problem two 2011 durations	410
.310	Histogram of homework five problem two 2011 ink written	410
.311	Histogram of homework five problem two 2011 sheet counts	410
.312	Histogram of homework five problem two 2011 time writing	411
.313	Histogram of homework five problem three 2011 durations	411
.314	Histogram of homework five problem three 2011 ink written	412
.315	Histogram of homework five problem three 2011 sheet counts	412

.316	Histogram of homework five problem three 2011 time writing	413
.317	Histogram of homework five problem four 2011 durations	413
.318	Histogram of homework five problem four 2011 ink written	413
.319	Histogram of homework five problem four 2011 sheet counts	414
.320	Histogram of homework five problem four 2011 time writing	414
.321	Histogram of homework five problem five 2011 durations	414
.322	Histogram of homework five problem five 2011 ink written	415
.323	Histogram of homework five problem five 2011 sheet counts	415
.324	Histogram of homework five problem five 2011 time writing	416
.325	Histogram of homework five problem six 2011 durations	416
.326	Histogram of homework five problem six 2011 ink written	417
.327	Histogram of homework five problem six 2011 sheet counts	417
.328	Histogram of homework five problem six 2011 time writing	417
.329	Histogram of homework five problem seven 2011 durations	418
.330	Histogram of homework five problem seven 2011 ink written	418
.331	Histogram of homework five problem seven 2011 sheet counts	418
.332	Histogram of homework five problem seven 2011 time writing	419
.333	Histogram of homework five problem eight 2011 durations	419
.334	Histogram of homework five problem eight 2011 ink written	420
.335	Histogram of homework five problem eight 2011 sheet counts	420
.336	Histogram of homework five problem eight 2011 time writing	421
.337	Histogram of homework six problem one 2011 durations	421
.338	Histogram of homework six problem one 2011 ink written	421
.339	Histogram of homework six problem one 2011 sheet counts	422
.340	Histogram of homework six problem one 2011 time writing	422
.341	Histogram of homework six problem two 2011 durations	422
.342	Histogram of homework six problem two 2011 ink written	423
.343	Histogram of homework six problem two 2011 sheet counts	423
.344	Histogram of homework six problem two 2011 time writing	424
.345	Histogram of homework six problem three 2011 durations	424
.346	Histogram of homework six problem three 2011 ink written	425
.347	Histogram of homework six problem three 2011 sheet counts	425
.348	Histogram of homework six problem three 2011 time writing	425
.349	Histogram of homework six problem four 2011 durations	426
.350	Histogram of homework six problem four 2011 ink written	426
.351	Histogram of homework six problem four 2011 sheet counts	426
.352	Histogram of homework six problem four 2011 time writing	427
.353	Histogram of homework six problem five 2011 durations	427
.354	Histogram of homework six problem five 2011 ink written	428
.355	Histogram of homework six problem five 2011 sheet counts	428
.356	Histogram of homework six problem five 2011 time writing	429
.357	Histogram of homework six problem six 2011 durations	429
.358	Histogram of homework six problem six 2011 ink written	429
.359	Histogram of homework six problem six 2011 sheet counts	430
.360	Histogram of homework six problem six 2011 time writing	430

.361	Histogram of homework seven problem one 2011 durations	430
.362	Histogram of homework seven problem one 2011 ink written	431
.363	Histogram of homework six problem seven 2011 sheet counts	431
.364	Histogram of homework seven problem one 2011 time writing	432
.365	Histogram of homework seven problem two 2011 durations	432
.366	Histogram of homework seven problem two 2011 ink written	433
.367	Histogram of homework seven problem one 2011 sheet counts	433
.368	Histogram of homework seven problem two 2011 time writing	433
.369	Histogram of homework seven problem three 2011 durations	434
.370	Histogram of homework seven problem three 2011 ink written	434
.371	Histogram of homework seven problem three 2011 sheet counts	434
.372	Histogram of homework seven problem three 2011 time writing	435
.373	Histogram of homework seven problem four 2011 durations	435
.374	Histogram of homework seven problem four 2011 ink written	436
.375	Histogram of homework seven problem four 2011 sheet counts	436
.376	Histogram of homework seven problem four 2011 time writing	437
.377	Histogram of homework seven problem five 2011 durations	437
.378	Histogram of homework seven problem five 2011 ink written	437
.379	Histogram of homework seven problem five 2011 sheet counts	438
.380	Histogram of homework seven problem five 2011 time writing	438
.381	Histogram of homework seven problem six 2011 durations	438
.382	Histogram of homework seven problem six 2011 ink written	439
.383	Histogram of homework seven problem six 2011 sheet counts	439
.384	Histogram of homework seven problem six 2011 time writing	440
.385	Histogram of homework seven problem seven 2011 durations	440
.386	Histogram of homework seven problem seven 2011 ink written	441
.387	Histogram of homework seven problem seven 2011 sheet counts	441
.388	Histogram of homework seven problem seven 2011 time writing	441
.389	Histogram of homework eight problem one 2011 durations	442
.390	Histogram of homework eight problem one 2011 ink written	442
.391	Histogram of homework eight problem one 2011 sheet counts	442
.392	Histogram of homework eight problem one 2011 time writing	443
.393	Histogram of homework eight problem two 2011 durations	443
.394	Histogram of homework eight problem two 2011 ink written	444
.395	Histogram of homework eight problem two 2011 sheet counts	444
.396	Histogram of homework eight problem two 2011 time writing	445
.397	Histogram of homework eight problem three 2011 durations	445
.398	Histogram of homework eight problem three 2011 ink written	445
.399	Histogram of homework eight problem three 2011 sheet counts	446
.400	Histogram of homework eight problem three 2011 time writing	446
.401	Histogram of homework eight problem four 2011 durations	446
.402	Histogram of homework eight problem four 2011 ink written	447
.403	Histogram of homework eight problem four 2011 sheet counts	447
.404	Histogram of homework eight problem four 2011 time writing	448
.405	Histogram of homework eight problem five 2011 durations	448

.406	Histogram of homework eight problem five 2011 ink written	449
.407	Histogram of homework eight problem five 2011 sheet counts	449
.408	Histogram of homework eight problem five 2011 time writing	449
.409	Histogram of homework one problem one 2012 durations	450
.410	Histogram of homework one problem one 2012 ink written	451
.411	Histogram of homework one problem one 2012 sheet counts	451
.412	Histogram of homework one problem one 2012 time writing	452
.413	Histogram of homework one problem two 2012 durations	452
.414	Histogram of homework one problem two 2012 ink written	452
.415	Histogram of homework one problem two 2012 sheet counts	453
.416	Histogram of homework one problem two 2012 time writing	453
.417	Histogram of homework one problem three 2012 durations	453
.418	Histogram of homework one problem three 2012 ink written	454
.419	Histogram of homework one problem three 2012 sheet counts	454
.420	Histogram of homework one problem three 2012 time writing	455
.421	Histogram of homework one problem four 2012 durations	455
.422	Histogram of homework one problem four 2012 ink written	456
.423	Histogram of homework one problem four 2012 sheet counts	456
.424	Histogram of homework one problem four 2012 time writing	456
.425	Histogram of homework one problem five 2012 durations	457
.426	Histogram of homework one problem five 2012 ink written	457
.427	Histogram of homework one problem five 2012 sheet counts	457
.428	Histogram of homework one problem five 2012 time writing	458
.429	Histogram of homework one problem six 2012 durations	458
.430	Histogram of homework one problem six 2012 ink written	459
.431	Histogram of homework one problem six 2012 sheet counts	459
.432	Histogram of homework one problem six 2012 time writing	460
.433	Histogram of homework one problem seven 2012 durations	460
.434	Histogram of homework one problem seven 2012 ink written	460
.435	Histogram of homework one problem seven 2012 sheet counts	461
.436	Histogram of homework one problem seven 2012 time writing	461
.437	Histogram of homework one problem eight 2012 durations	461
.438	Histogram of homework one problem eight 2012 ink written	462
.439	Histogram of homework one problem eight 2012 sheet counts	462
.440	Histogram of homework one problem eight 2012 time writing	463
.441	Histogram of homework two problem one 2012 durations	463
.442	Histogram of homework two problem one 2012 ink written	464
.443	Histogram of homework two problem one 2012 sheet counts	464
.444	Histogram of homework two problem one 2012 time writing	464
.445	Histogram of homework two problem two 2012 durations	465
.446	Histogram of homework two problem two 2012 ink written	465
.447	Histogram of homework two problem two 2012 sheet counts	465
.448	Histogram of homework two problem two 2012 time writing	466
.449	Histogram of homework two problem three 2012 durations	466
.450	Histogram of homework two problem three 2012 ink written	467

.451	Histogram of homework two problem three 2012 sheet counts	467
.452	Histogram of homework two problem three 2012 time writing	468
.453	Histogram of homework two problem four 2012 durations	468
.454	Histogram of homework two problem four 2012 ink written	468
.455	Histogram of homework two problem four 2012 sheet counts	469
.456	Histogram of homework two problem four 2012 time writing	469
.457	Histogram of homework two problem five 2012 durations	469
.458	Histogram of homework two problem five 2012 ink written	470
.459	Histogram of homework two problem five 2012 sheet counts	470
.460	Histogram of homework two problem five 2012 time writing	471
.461	Histogram of homework two problem six 2012 durations	471
.462	Histogram of homework two problem six 2012 ink written	472
.463	Histogram of homework two problem six 2012 sheet counts	472
.464	Histogram of homework two problem six 2012 time writing	472
.465	Histogram of homework two problem seven 2012 durations	473
.466	Histogram of homework two problem seven 2012 ink written	473
.467	Histogram of homework two problem seven 2012 sheet counts	473
.468	Histogram of homework two problem seven 2012 time writing	474
.469	Histogram of homework two problem eight 2012 durations	474
.470	Histogram of homework two problem eight 2012 ink written	475
.471	Histogram of homework two problem eight 2012 sheet counts	475
.472	Histogram of homework two problem eight 2012 time writing	476
.473	Histogram of homework three problem one 2012 durations	476
.474	Histogram of homework three problem one 2012 ink written	476
.475	Histogram of homework three problem one 2012 sheet counts	477
.476	Histogram of homework three problem one 2012 time writing	477
.477	Histogram of homework three problem two 2012 durations	477
.478	Histogram of homework three problem two 2012 ink written	478
.479	Histogram of homework three problem two 2012 sheet counts	478
.480	Histogram of homework three problem two 2012 time writing	479
.481	Histogram of homework three problem three 2012 durations	479
.482	Histogram of homework three problem three 2012 ink written	480
.483	Histogram of homework three problem three 2012 sheet counts	480
.484	Histogram of homework three problem three 2012 time writing	480
.485	Histogram of homework three problem four 2012 durations	481
.486	Histogram of homework three problem four 2012 ink written	481
.487	Histogram of homework three problem four 2012 sheet counts	481
.488	Histogram of homework three problem four 2012 time writing	482
.489	Histogram of homework three problem five 2012 durations	482
.490	Histogram of homework three problem five 2012 ink written	483
.491	Histogram of homework three problem five 2012 sheet counts	483
.492	Histogram of homework three problem five 012 time writing	484
.493	Histogram of homework three problem six 2012 durations	484
.494	Histogram of homework three problem six 2012 ink written	484
.495	Histogram of homework three problem six 2012 sheet counts	485

.496	Histogram of homework three problem six 2012 time writing	485
.497	Histogram of homework three problem seven 2012 durations	485
.498	Histogram of homework three problem seven 2012 ink written	486
.499	Histogram of homework three problem seven 2012 sheet counts	486
.500	Histogram of homework three problem seven 2012 time writing	487
.501	Histogram of homework three problem eight 2012 durations	487
.502	Histogram of homework three problem eight 2012 ink written	488
.503	Histogram of homework three problem eight 2012 sheet counts	488
.504	Histogram of homework three problem eight 2012 time writing	488
.505	Histogram of homework four problem one 2012 durations	489
.506	Histogram of homework four problem one 2012 ink written	489
.507	Histogram of homework four problem one 2012 sheet counts	489
.508	Histogram of homework four problem one 2012 time writing	490
.509	Histogram of homework four problem two 2012 durations	490
.510	Histogram of homework four problem two 2012 ink written	491
.511	Histogram of homework four problem two 2012 sheet counts	491
.512	Histogram of homework four problem two 2012 time writing	492
.513	Histogram of homework four problem three 2012 durations	492
.514	Histogram of homework four problem three 2012 ink written	492
.515	Histogram of homework four problem three 2012 sheet counts	493
.516	Histogram of homework four problem thre 2012 time writing	493
.517	Histogram of homework four problem four 2012 durations	493
.518	Histogram of homework four problem four 2012 ink written	494
.519	Histogram of homework four problem four 2012 sheet counts	494
.520	Histogram of homework four problem four 2012 time writing	495
.521	Histogram of homework four problem five 2012 durations	495
.522	Histogram of homework four problem five 2012 ink written	496
.523	Histogram of homework four problem five 2012 sheet counts	496
.524	Histogram of homework four problem five 2012 time writing	496
.525	Histogram of homework four problem six 2012 durations	497
.526	Histogram of homework four problem six 2012 ink written	497
.527	Histogram of homework four problem six 2012 sheet counts	497
.528	Histogram of homework four problem six 2012 time writing	498
.529	Histogram of homework four problem seven 2012 durations	498
.530	Histogram of homework four problem seven 2012 ink written	499
.531	Histogram of homework four problem seven 2012 sheet counts	499
.532	Histogram of homework four problem seven 2012 time writing	500
.533	Histogram of homework five problem one 2012 durations	500
.534	Histogram of homework five problem one 2012 ink written	500
.535	Histogram of homework five problem one 2012 sheet counts	501
.536	Histogram of homework five problem one 2012 time writing	501
.537	Histogram of homework five problem two 2012 durations	501
.538	Histogram of homework five problem two 2012 ink written	502
.539	Histogram of homework five problem two 2012 sheet counts	502
.540	Histogram of homework five problem two 2012 time writing	503

.541	Histogram of homework five problem three 2012 durations	503
.542	Histogram of homework five problem three 2012 ink written	504
.543	Histogram of homework five problem three 2012 sheet counts	504
.544	Histogram of homework five problem three 2012 time writing	504
.545	Histogram of homework five problem four 2012 durations	505
.546	Histogram of homework five problem four 2012 ink written	505
.547	Histogram of homework five problem four 2012 sheet counts	505
.548	Histogram of homework five problem four 2012 time writing	506
.549	Histogram of homework five problem five 2012 durations	506
.550	Histogram of homework five problem five 2012 ink written	507
.551	Histogram of homework five problem five 2012 sheet counts	507
.552	Histogram of homework five problem five 2012 time writing	508
.553	Histogram of homework five problem six 2012 durations	508
.554	Histogram of homework five problem six 2012 ink written	508
.555	Histogram of homework five problem six 2012 sheet counts	509
.556	Histogram of homework five problem six 2012 time writing	509
.557	Histogram of homework five problem seven 2012 durations	509
.558	Histogram of homework five problem seven 2012 ink written	510
.559	Histogram of homework five problem seven 2012 sheet counts	510
.560	Histogram of homework five problem seven 2012 time writing	511
.561	Histogram of homework five problem eight 2012 durations	511
.562	Histogram of homework five problem eight 2012 ink written	512
.563	Histogram of homework five problem eight 2012 sheet counts	512
.564	Histogram of homework five problem eight 2012 time writing	512
.565	Histogram of homework six problem one 2012 durations	513
.566	Histogram of homework six problem one 2012 ink written	513
.567	Histogram of homework six problem one 2012 sheet counts	513
.568	Histogram of homework six problem one 2012 time writing	514
.569	Histogram of homework six problem two 2012 durations	514
.570	Histogram of homework six problem two 2012 ink written	515
.571	Histogram of homework six problem two 2012 sheet counts	515
.572	Histogram of homework six problem two 2012 time writing	516
.573	Histogram of homework six problem three 2012 durations	516
.574	Histogram of homework six problem three 2012 ink written	516
.575	Histogram of homework six problem three 2012 sheet counts	517
.576	Histogram of homework six problem three 2012 time writing	517
.577	Histogram of homework six problem four 2012 durations	517
.578	Histogram of homework six problem four 2012 ink written	518
.579	Histogram of homework six problem four 2012 sheet counts	518
.580	Histogram of homework six problem four 2012 time writing	519
.581	Histogram of homework six problem five 2012 durations	519
.582	Histogram of homework six problem five 2012 ink written	520
.583	Histogram of homework six problem five 2012 sheet counts	520
.584	Histogram of homework six problem five 2012 time writing	520
.585	Histogram of homework six problem six 2012 durations	521

.586	Histogram of homework six problem six 2012 ink written	521
.587	Histogram of homework six problem six 2012 sheet counts	521
.588	Histogram of homework six problem six 2012 time writing	522
.589	Histogram of homework seven problem one 2012 durations	522
.590	Histogram of homework seven problem one 2012 ink written	523
.591	Histogram of homework seven problem one 2012 sheet counts	523
.592	Histogram of homework seven problem one 2012 time writing	524
.593	Histogram of homework seven problem two 2012 durations	524
.594	Histogram of homework seven problem two 2012 ink written	524
.595	Histogram of homework seven problem two 2012 sheet counts	525
.596	Histogram of homework seven problem two 2012 time writing	525
.597	Histogram of homework seven problem three 2012 durations	525
.598	Histogram of homework seven problem three 2012 ink written	526
.599	Histogram of homework seven problem three 2012 sheet counts	526
.600	Histogram of homework seven problem three 2012 time writing	527
.601	Histogram of homework seven problem four 2012 durations	527
.602	Histogram of homework seven problem four 2012 ink written	528
.603	Histogram of homework seven problem four 2012 sheet counts	528
.604	Histogram of homework seven problem four 2012 time writing	528
.605	Histogram of homework seven problem five 2012 durations	529
.606	Histogram of homework seven problem five 2012 ink written	529
.607	Histogram of homework seven problem five 2012 sheet counts	529
.608	Histogram of homework seven problem five 2012 time writing	530
.609	Histogram of homework seven problem six 2012 durations	530
.610	Histogram of homework seven problem six 2012 ink written	531
.611	Histogram of homework seven problem six 2012 sheet counts	531
.612	Histogram of homework seven problem six 2012 time writing	532
.613	Histogram of homework seven problem seven 2012 durations	532
.614	Histogram of homework seven problem seven 2012 ink written	532
.615	Histogram of homework seven problem seven 2012 sheet counts	533
.616	Histogram of homework seven problem seven 2012 time writing	533
.617	Histogram of homework eight problem one 2012 durations	533
.618	Histogram of homework eight problem one 2012 ink written	534
.619	Histogram of homework eight problem one 2012 sheet counts	534
.620	Histogram of homework eight problem one 2012 time writing	535
.621	Histogram of homework eight problem two 2012 durations	535
.622	Histogram of homework eight problem two 2012 ink written	536
.623	Histogram of homework eight problem two 2012 sheet counts	536
.624	Histogram of homework eight problem two 2012 time writing	536
.625	Histogram of homework eight problem three 2012 durations	537
.626	Histogram of homework eight problem three 2012 ink written	537
.627	Histogram of homework eight problem three 2012 sheet counts	537
.628	Histogram of homework eight problem three 2012 time writing	538
.629	Histogram of homework eight problem four 2012 durations	538
.630	Histogram of homework eight problem four 2012 ink written	539

.631	Histogram of homework eight problem four 2012 sheet counts	539
.632	Histogram of homework eight problem four 2012 time writing	541
.633	Histogram of homework eight problem five 2012 durations	541
.634	Histogram of homework eight problem five 2012 ink written	543
.635	Histogram of homework eight problem five 2012 sheet counts	543

List of Tables

1.1	Overview of analyses and representations for each chapter	4
3.1	Students who completed assignments with a digital pen 2010.	24
3.2	Students who completed assignments with a digital pen 2011	25
3.3	Students who completed assignments with a digital pen 2012.	26
3.4	Midterm one problem one rubric 2012	28
3.5	Midterm two problem one rubric 2012	29
3.6	Midterm two problem two rubric 2012	30
3.7	Midterm two problem three rubric 2012	31
3.8	An overview of the 2012 scaffolded prompts	61
4.1	Open IE target classes based on expert self-explanations	78
4.2	Accuracy of concept recognition for prompt four of homework three . .	80
4.3	Accuracy of concept recognition for prompt four of homework three . .	81
4.4	Accuracy of concept recognition for prompt three of homework eight . .	81
6.1	Descriptive statistics for the lengths of action sequences	100
6.2	Features selected using the CFS algorithm	103
6.3	Important coefficients for model predicting student performance	105
8.1	Features selection for models predicting homework performance	126
8.2	Descriptives for clusters for homework assignment three	131
8.3	Descriptives for clusters for homework assignment four.	132
8.4	Descriptives for clusters for homework assignment five.	132
8.5	Descriptives for clusters for homework assignment six.	133
8.6	Descriptives for clusters for homework assignment seven.	133
8.7	Descriptives for clusters for homework assignment eight.	134
9.1	<i>Category error groupings</i> for midterm one problem one	142
9.2	<i>Category error groupings</i> for midterm two problem one	142
9.3	<i>Category error groupings</i> for midterm two problem two	142
9.4	<i>Category error groupings</i> for midterm two problem three	143
9.5	<i>Subcategory error groupings</i> for midterm one problem one	144

9.6	<i>Subcategory error groupings</i> for midterm two problem one	144
9.7	<i>Subcategory error groupings</i> for midterm two problem two	145
9.8	<i>Subcategory error grouping</i> for midterm two problem three	146
9.9	10 most predictive features for the <i>category error grouping</i>	154
9.10	10 least predictive feature for the <i>category error grouping</i>	154
9.11	10 most predictive features for the <i>fine-grained error groupings</i>	160
9.12	10 least predictive features for the <i>fine-grained error groupings</i>	160
9.13	10 most predictive features for the <i>subcategory error groupings</i>	164
9.14	10 least predictive features for the <i>subcategory error groupings</i>	164
10.1	Comparison of computation times	190
10.2	Per-shape confusion matrix of 1 ^c	192
10.3	Per-shape confusion matrix of \$1	193
10.4	Average accuracy of each gesture recognition technique	193
10.5	Segmentation accuracy for various segmentation techniques	213
10.6	Features used for gesture/object classification	231
10.7	Properties of manually aligned speech	238
10.8	Features for speech segmenter	246
10.9	Features for segment-stroke alignment	251

Chapter 1

Introduction

In this work we apply data mining and machine learning techniques to digital records of students' ordinary, handwritten coursework. This thesis makes four important contributions:

1. The data set investigated in this work is the first of its kind. We present, in great detail, the large digital data set of students' handwritten coursework which we collected over the past three years.
2. We investigate novel, discrete and numerical representations of students' handwritten coursework which characterize different aspects of students' ordinary problem-solving processes.
3. We identify patterns in these representations as well as correlations between them and performance in the course by applying a breadth of machine learning and data mining techniques.
4. Most importantly, we present interesting insights that may be gained from these

patterns and correlations. These insights allow instructors of the course, which we investigate in this work, to both rapidly assess students' learning as well as improve future course offerings.

We begin with a motivating example of how data mining has been used to identify interesting patterns in a data set of ant behavior. Mersch et al. [90] mined data collected from ants to identify interesting social organization patterns. The authors physically applied a tiny, unique marker to the backs of ants which, using a video camera, allowed the authors to individually track the location and orientation of each ant within the confines of a nest chamber. Using the location and orientation data the authors were able to both track the ants' trajectories each day as well as infer when two ants interacted with one another. The authors represented the data using a graph structure and applied a network analysis technique that allowed them to identify three distinct social groups: nurses, cleaners, and foragers. This is quite interesting, as it shows that by digitally instrumenting (in this case, by placing markers on) ants' ordinary behavior, the authors were able to discover new ways in which ants interact with one another. The question then becomes, if this type of knowledge discovery can be accomplished for ants, why not do so for college students?

While applying markers to students' backs may not be feasible, instrumenting their ordinary learning behavior is a tractable task that, as we show in this work, leads to interesting knowledge discovery about students' problem-solving habits.

In fact, Educational Data Mining is a nascent, but rapidly growing field in which data mining techniques are applied to educational data to discover patterns in the ways that students learn. Current research has typically focused on educational

data extracted from either Learning Content Management Systems or Intelligent Tutoring Systems. Both of these environments are digital and do not always directly represent a students' natural learning environment. These digital environments are convenient though, as some form of digital instrument is required to digitally record the students' learning processes. A primary goal of the research presented in this thesis is to instead use a minimally intrusive digital device that records students' work as they solve problems the way they would if left to their own devices. We use LiveScribe™ digital pens for this purpose. These pens digitize students' handwriting, allowing them to solve problems as they ordinarily would with pen and paper and on their own schedule.

In this thesis we explore a breadth of analyses that have hitherto not been possible. The goal is to investigate both novel representations of students' handwritten coursework as well as analysis techniques to identify the types of correlations and patterns that can be identified in this unique data base.

We begin, in chapter two, by placing our work in the context of related work. This research comprises an intersection of quite a few research areas: pedagogical research, general purpose data mining and machine learning, and educational data mining.

Next, in chapter three, we provide an in-depth description of the data we have collected. Additionally, several experiments were conducted during the data collection process each year. We describe each of these experiments and their purposes.

Lastly, in chapters four through nine, we present each of the novel data mining and machine learning techniques and data representations we have developed. Each chapter comprises a different combination of data representation and analysis technique. Table 1.1 summarizes each chapter and the data representation and analysis technique

Chapter	Data	Representation	Analysis	Publication
4	2011 Homework	Text	Info. Extraction	[55]
5	2011 Homework	Prob. No. Sequence	N-Gram	[56]
6	2012 Homework	Action Sequence	Pattern Mining	[51]
7	2012 Exams	Solution Bitmap	Clustering	[58]
8	2012 Homework	Effort-based features	Linear Regression	[57]
9	2012 Homework	Effort-based features	Decision Stump	n/a

Table 1.1: An overview of the research presented in each chapter. Each row corresponds to a chapter and shows for that chapter: the year and type of data used (Data column); the way the data was represented (Rep. column); the data mining or machine learning technique used to analyze that representation (Analysis column); and the publication in which this research has appeared if one exists (Publication column).

used, as well as which portion of the data that was used.

In chapter ten, we take a detour to describe novel, machine learning based sketch processing techniques. These low-level techniques serve as a preprocessing step in several of the educational data mining techniques presented in this thesis, and additionally may be used other sketch understanding tasks.

Lastly, we discuss the review the insights gained by applying each of our Educational Data Mining techniques and compare them to one another in chapter eleven.

Chapter 2

Related Work

2.1 Theories of Learning

Chi et al. [24] argue that “the metacognitive component of training is important in that it allows students to understand and take control of their learning process.” Metacognition is the ability to be aware of one’s own learning process and it serves as a major foundation for the research we have performed on self-explanation described in Chapters 4 and 5.

Chi et al. [24] made comparisons between two groups of students: “poor” and “good” performing students. These students were asked to generate self-explanation after studying worked out example problems. The results of this study demonstrated that students who perform poorly are typically unable to generate sufficient self-explanation of the worked out example problems.

Steif et al. [116] present and evaluate a strategy for teaching statics concepts which focuses on student’s conceptual knowledge. During instruction, students are given

example free body diagrams and asked whether they are correct. Students are then shown a video explaining what errors are present in the diagram. Additionally, students in an experimental group are asked questions eliciting an explanation from the student pertaining to the relationships between the diagram and the forces which act upon it. This work showed a significantly lower error rate amongst students who generated self-explanation. Additionally though, the analysis of the content of self-explanations which we present in Chapter 4. Our method enables automatic analysis of the content of self-explanation, which may enable the creation of intelligent tutoring systems that probe a student's understanding if that student's self-explanation is lacking.

Numerous studies have demonstrated the positive impact self-explanation has on student performance. Bielaczyc et al. [9] studied the impact of different self-explanation strategies on a student's ability to learn LISP programming. In their experiment, students were given instruction via an intelligent tutoring system. Some students were also trained to ask themselves a series of questions regarding their own understanding of worked out examples they viewed with the tutorial. The experiment revealed a significant difference between the learning gains from the pre- to posttest between students that did and did not generate self-explanation. This differs from the way in which we presented students with self-explanation, as in our study, students generate self-explanation throughout their entire problem-solving process.

Chi et al. [23] stated that, "generating explanations to oneself facilitates the integration of new knowledge." To verify this statement, the authors conducted a study in which eighth grade students were asked to provide self-explanation as they read passages from a text on the circulatory system. This demonstrates that students who

generated self-explanation performed significantly better than those who did not. This study differs from ours in that students explained passages they read, whereas in our study, students explained their own solution processes.

Weerasinghe and Mitrovic[127] investigated the impact that self-explanation, paired with the use of an intelligent tutor, has on student performance in a database design course. In the study, students in the experimental group were prompted for self-explanation by the tutoring system whenever the student made a mistake. This protocol was used as the authors claim that prompting students to explain most of their problem steps would be “too burdensome,” although no evidence for this is provided. Because there was no statistical analysis, the results were inconclusive.

Hall and Vance [46] investigated the impact that self-explanation has on student performance as well as self-efficacy in a statistics course. Students in the experimental group collaboratively solved problems in teams of three, providing self-explanations of the reasoning behind their answers to one another. Students in a control group solved the same problems individually. This study showed that students who generated collaborative self-explanation perform significantly better at solving problems than students who did not.

These studies demonstrate that self-explanation can have a strong positive impact on students’ learning. In the current research, we are interested in using our unique database of students’ handwritten coursework to identify not only that self-explanation can lead to improved performance, but we seek to discover changes in behavior caused by generating self-explanation that may provide insight as to why self-explanation can lead to improved performance

Mayer [87] examine differences between retention and transfer. The former is the application of knowledge from one problem to an identical problem, while the latter is the application of that knowledge to a different problem. Mayer argues that metaskill, the ability to control and monitor the cognitive processes, is an essential part of transfer. Metaskill strategies may be taught just as any other skill such as arithmetic via strategy instruction. For example, students who are taught basic reading skills as well strategies for summarizing their own reading, perform better on transfer questions[16]. In later chapters, we use the results of various data mining and machine learning techniques to better understand the types of knowledge transfer which students make from homework problems to exam problems. In this sense, the idea of transfer provides a foundation for interpreting our results.

2.2 Data Mining

Information extraction (IE) is the process by which target relations are extracted from machine readable documents, such as text transcripts. This is distinguishable from attempting to understand the entire content of such documents. There is a long history of research in IE techniques [63]. Older techniques have typically relied a great deal on domain dependent attributes and were usually rule-based[91] or applied machine learning techniques[28]. While these systems achieved high accuracy, their domain dependent nature required a great deal of manual effort in order to adapt them to new domains. More recently, researchers have focused on automatic IE techniques intended for use with the world wide web. These techniques are more general and exten-

sible than prior methods and are thus called open IE techniques. We apply an open IE technique in Chapter 4 to automatically identify the content type present in transcripts of students' self-explanations.

Sequential pattern mining [2] is a technique used to identify significant patterns in sequences of discrete items, e.g., consumer transaction records [2] or DNA transcripts [8]. These techniques have typically been used to mine patterns from a single database of sequences. In Educational Data Mining, it is often the case that researchers seek to find patterns that best distinguish students who do and do not perform well in the course. Thus there is a need for novel pattern mining techniques aimed at differentiating between two databases of sequences.

More recently, Ye and Keogh [139] developed a novel technique which identifies patterns which best separate two time-series databases. This technique identifies frequently occurring patterns within each database, as traditional pattern mining techniques have, but furthermore evaluates each pattern by using it to separate sequences from the two databases. If a sequence contains the pattern, that sequence is identified as being part of the same database that the pattern came from. The pattern which provides the greatest information gain is kept as the "shapelet" that best separates the two databases. We apply a similar sequential pattern mining technique to sequences of the actions student took while solving homework problems in Chapter 6.

2.3 Educational Data Mining

Data-driven educational research has traditionally been limited by the time-consuming process of monitoring students' learning. For example, substantial research has been performed which investigates the correlation between performance and the amount of time and effort spent on homework assignments [5, 14, 31, 103, 112]. Manually watching each student solve each homework assignment would require an intractable amount of time and, additionally, may skew the results of the study. Instead, each of these researchers relied on students or their parents to self-report the amount of time spent on each homework assignment.

Cooper et al. [26] compared the results of each of these studies and found an average correlation of $r = 0.14$ with a range from -0.25 to 0.65 . Cooper et al. summarize this inconsistency in findings when they state that, "to date, the role of research in forming homework policies and practices has been minimal. This is because the influences on homework are complex, and no simple, general finding applicable to all students is possible." This underlies the impact that Educational Data Mining can have on the educational research community. By instrumenting students' natural problem-solving processes, we are able to capture a precise measurement of the actions students perform when solving their homework assignments.

More recently, researchers have applied data mining techniques to Intelligent Tutoring System (ITS) and Course Management System (CMS) data. For example, Romero et al. [105] applied data mining techniques to data collected with the Moodle CMS. This system allows students to both view and submit various assignments,

e.g., homework and exams, and records detailed logs of students' interactions. These interaction logs were mined for rare association rules, that is, patterns which appear infrequently in the data. The resulting rules were then manually inspected to identify fringe behaviors exhibited by students.

Similarly Mostow et al. [92] applied data mining techniques to interaction logs taken from Project LISTEN's Reading Tutor, an ITS. This system tutors young students as they learn to read by listening to them read stories aloud and providing feedback. The authors developed a system which automatically identified meaningful features from these logs which were then used to train classifiers to predict students' future behavior with the system.

Similarly, Kinnebrew and Biswas [75] have developed a novel differential pattern mining technique used to identify patterns that differentiate between the interactions of different groups of students with the Betty's Brain ITS. This technique begins by using SPAM [7] to identify patterns that occur in a significant number of sequences in either database. A *t*-test for each pattern is then performed to determine if there is a significant difference in the frequency of that pattern in each sequence of each of the two databases. This algorithm can identify patterns that occur significantly frequently in one database and not the other.

The work of Oviatt et al. [97] suggests that natural work environments are critical to student performance. Their examination of computer interfaces for completing geometry problems suggests that, "as the interfaces departed more from familiar work practice..., students would experience greater cognitive load such that performance would deteriorate in speed, attentional focus, meta-cognitive control, correctness

of problem solutions, and memory.” Thus, our goal is to apply Educational Data Mining techniques to data collected in natural work environments.

To that end, recent research has focused on mining ordinary, handwritten coursework data. For example, Van Arsdale and Stahovich [124] demonstrated that a correlation exists between the temporal and spatial organization of students’ handwritten problem solutions and the correctness of the work. The organization of exam solutions was characterized by a set of quantitative features, which were then used to predict performance on those problems. On average these features accounted for 40.0% of the variance in students’ performance on exam problems.

Our work continues this trend of instrumenting students’ learning behaviors so that those processes may be mined, but we employ a minimally invasive device which allows us to capture students’ ordinary learning.

Chapter 3

Data and Experimentation

3.1 Introduction

Since 2010 we have digitized and stored students' ordinary, handwritten course-work collected from an undergraduate Mechanical Engineering course. This course is offered at the University of California, Riverside as ME 10 Statics. The course description reads, "Covers equilibrium of coplanar force systems; analysis of frames and trusses; noncoplanar force systems; friction; and distributed loads." While ME 02 is the first course to cover core Statics concepts, it does so briefly, with approximately three lectures on the topic. ME 10, on the other hand provides students with their first in-depth coverage of Statics concepts and provides a crucial foundation for several other courses in Mechanical Engineering, such as, Dynamics, Strength and Materials, Machine Design, Senior Design, and even Fluids.

Students who took this course during the winter quarters of 2010, 2011, 2012,

and 2013¹ received Livescribe™ digital pens with which they completed their coursework. These pens serve as traditional pens, allowing students to solve problems by writing ink on paper, but additionally digitize the writing, producing a time-stamped, digital record of the students' work. Figure 3.1 a digital pen and the special paper it is used with. This technology requires students to write on special Anoto™ paper. Each page in a notebook of this paper contains a unique dot-matrix pattern. An integrated camera reads tiny regions of this pattern enabling the digital pen to identify both its location on the page as well as the page (in a notebook) on which it is writing.

Each year, students were given a pen and notebook of Anoto™ paper at the beginning of the quarter starting in week three. Students were asked to complete all coursework with the pens, including quizzes, midterms, homework assignments, and final exams. Some years, students were grouped into different experiments, with students from each group, for example, receiving variations in homework assignments throughout the year.

In the following sections the details of the data collected each year are presented, as well as an in-depth explanation of each of the experiments.

3.2 Homework Descriptions

Students completed seven of their nine homework assignments in the 2010 year with their digital pens. Below is a brief explanation of the types of problems students were asked to solve and an example problem for each assignment. We finish by explaining

¹The data collected from the 2013 winter quarter is still being processed, and thus is not discussed in this work.



Figure 3.1: The Livescribe™ digital pen. This pen allows users to write ink on paper, but additionally digitizes the writing, producing a time-stamped record of the ink. This technology is enabled by the special Anoto™ paper on which the user must write. This paper comprises a tiny dot-matrix pattern which a tiny camera, located near the tip of the digital pen, uses to determine where on the page the pen is writing.

how the assignments from the 2011 and 2012 years are similar and different from those given in 2010.

We begin by showing a sample problem that characterizes the general problem-solving process required by students in this course. Figure 3.2 shows a typical problem from this course. Problems always include a picture of a two or three dimensional system in a state of equilibrium under the action of forces. Students are shown forces which act

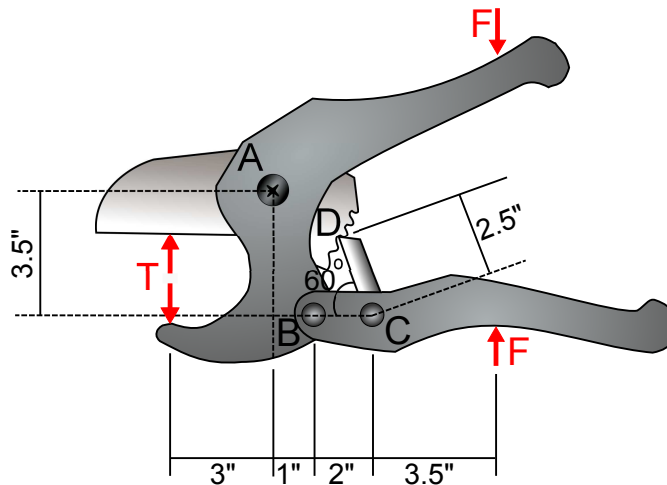


Figure 3.2: A typical Statics problem. The problem statement reads, “The device shown is used for cutting PVC pipe. If a force, $F = 15lb$, is applied to each handle as shown, determine the cutting force T . Also, determine the magnitude and the direction of the force that the pivot at A applies to the blade.”

upon that system and are then required to solve for the unknown forces.

Students typically begin solving a problem by drawing a free body diagram (FBD) representing the boundary of the system and the forces which act upon it. The FBD is then used as a guide for constructing force and moment equilibrium equations. Most pen strokes in a solution correspond to either a free body diagram, an equation, or a cross-out. Because the Livescribe pens use ink, students cannot erase errors and must instead cross them out. Figure 3.3 shows a hypothetical² solution to the Statics problem in Figure 3.2.

Students were typically given a week to complete each assignment which comprised five to eight problems.

²To protect the identity of the students who generated this data we do not show any of their original handwritten work here. Instead we present a redrawing that is similar to the students’ work. This is true of all handwritten examples shown in the remainder of the paper.

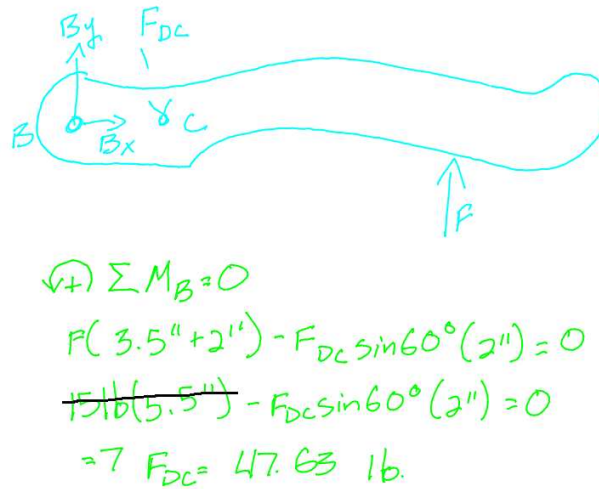


Figure 3.3: A hypothetical solution to a typical Statics problem. The color of each pen stroke identifies the type of solution element: cyan = FBD, green = equation, and cross-out = black.

Homework assignment three provided students with their first opportunity to practice core Statics concepts. They were required to solve for unknown forces acting on simple, single-body systems in equilibrium. A typical problem from this assignment is shown in Figure 3.4 in which students are asked to solve for the maximum load that can be supported by a bracket.

Homework assignment four was similar to assignment three, in that students were again required to solve for unknown forces acting on a simple system. In this case however the systems were three dimensional. A typical problem from this assignment is shown in Figure 3.5 in which students are asked to solve for the loads acting on a pipe.

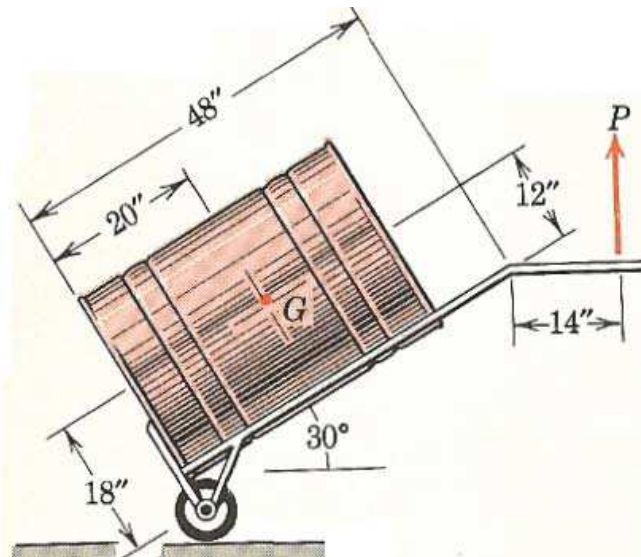


Figure 3.4: Typical problem from homework assignment three from 2010 taken from [89]. The problem description reads, “The oil drum weighs 620 lb when full and has a mass center at G. Calculate the vertical force P required to maintain equilibrium of the drum and dolly in the position shown. The weight of the dolly may be neglected compared with that of the drum.”

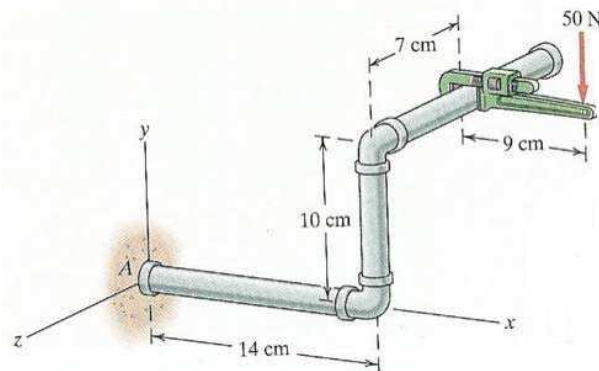


Figure 3.5: Typical problem from homework assignment four from 2010. The problem description reads, “A 50 N force is applied to the pipe wrench attached to the pipe system shown. If the pipe is in equilibrium, determine the loads acting on the pipe at support A.”

Homework assignment five required students to solve more complicated three dimensional single-body problems and additionally required students to model tension forces as three dimensional vectors. A typical problem from this assignment is shown in Figure 3.6 in which students are asked to compute the tension in a wire holding a door open.

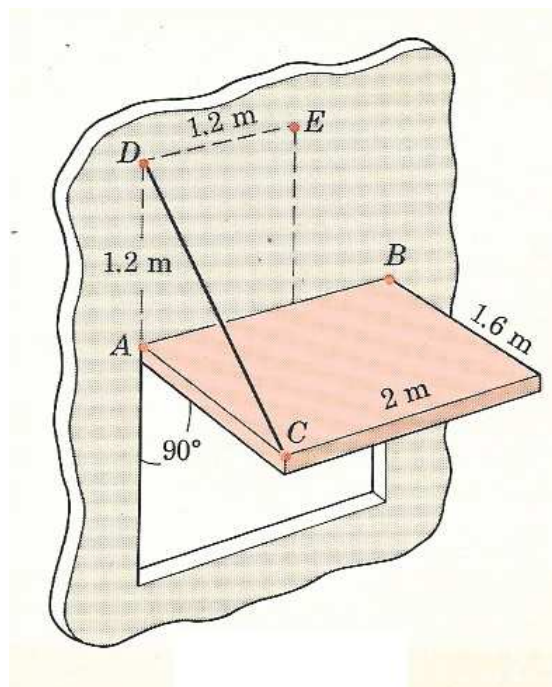


Figure 3.6: Typical problem from homework assignment five from 2010 taken from [89]. The problem description reads, “The uniform ventilator door has a mass of 200 kg and is hinged at the corners A and B of its upper edge. The door is held open in a horizontal position by the wire from C to point D in the vertical wall. Compute the tension in the wire and the forces normal to the hinge axis supported by the hinge pins at A and B.”

Homework assignment six required students to solve multi-body frame and machine systems. These systems often included two-force members. A typical problem from this assignment is shown in Figure 3.7 in which students are asked to determine

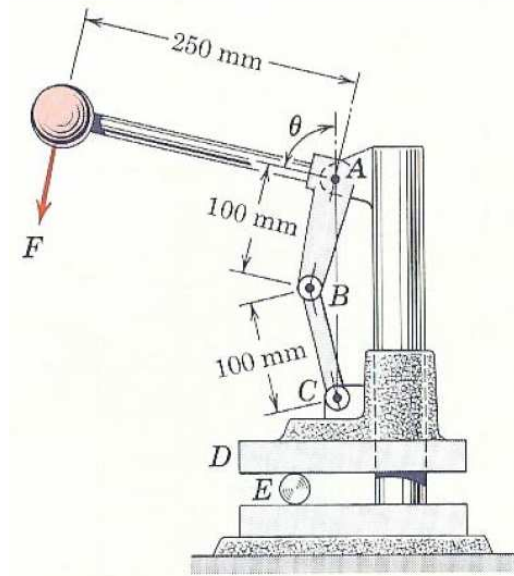


Figure 3.7: Typical problem from homework assignment six from 2010 taken from [89]. The problem description reads, “The upper jaw D of the toggle press slides with negligible frictional resistance along the fixed vertical column. Calculate the compressive force R exerted on the cylinder E and the force supported by the pin at A if a force $F = 200 \text{ N}$ is applied to the handle at an angle of $\theta = 75^\circ$ ”

the torque exerted by a motor.

Homework assignment seven required students to solve truss problems and identify whether two-force members were in tension or compression. A typical problem from this assignment is shown in Figure 3.8 in which students were asked determine the force in all members of a truss.

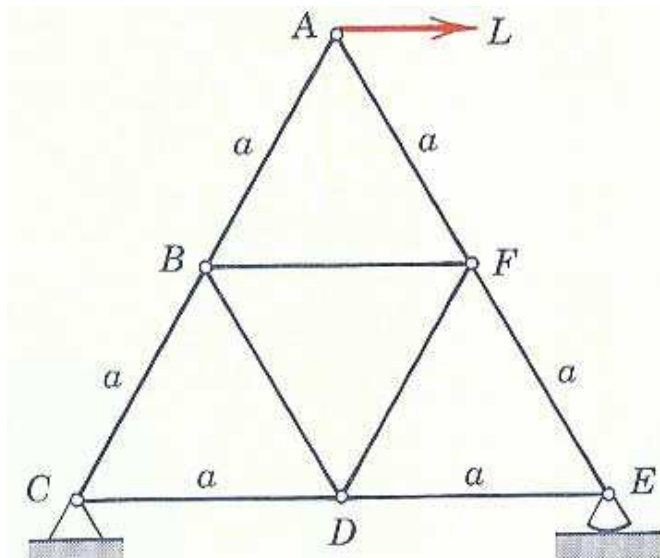


Figure 3.8: Typical problem from homework assignment seven from 2010 taken from [89]. The problem description reads, “The equiangular truss is loaded and supported as shown. Use joint method to determine the force in all members in terms of the horizontal load L . State if the members are in tension or compression.”

Homework assignment eight required students to identify the centroids of two dimensional areas and three dimensional volumes. A typical problem from this assignment is shown in Figure 3.9.

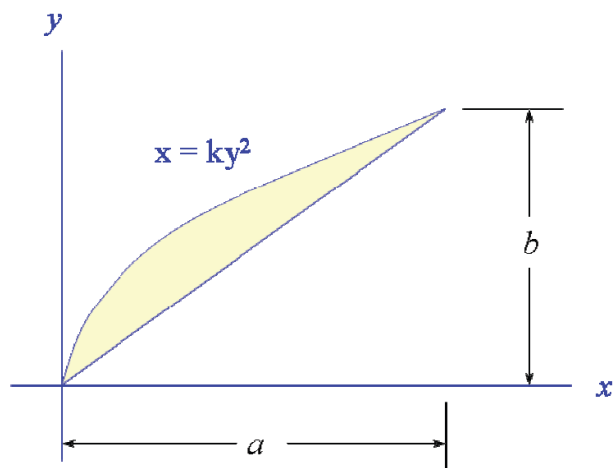


Figure 3.9: Typical problem from homework assignment eight from 2010 taken from [89]. The problem description reads, “Calculate the coordinates of the centroid of the area shown.”

Homework assignment nine required students to solve systems involving friction. A typical problem from this assignment is shown in Figure 3.10 in which students are asked to find the minimum coefficient of static friction between a block and the incline upon which it rests.

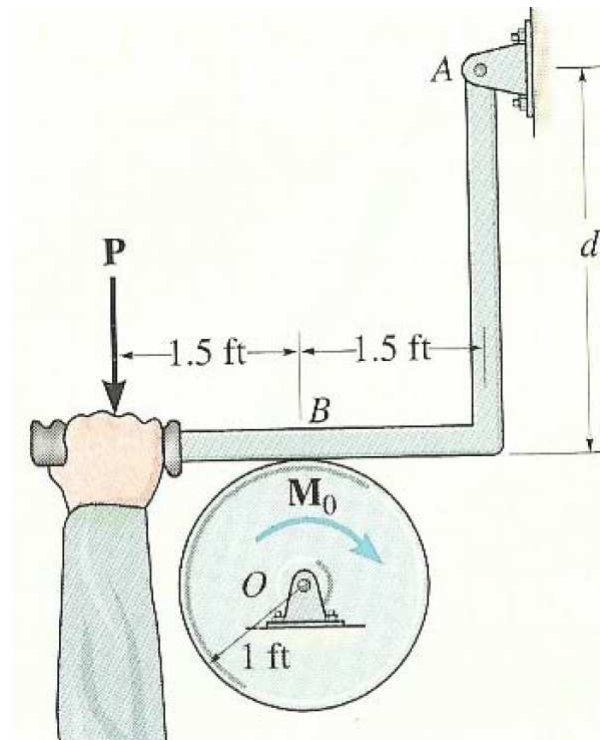


Figure 3.10: Typical problem from homework assignment nine from 2010 taken from [61]. The problem description reads, “The brake is designed to be self-locking, that is, it will not rotate when no load P is applied to it when the disk is subjected to a clockwise couple moment M_o . Determine the distance d of the lever that will allow this to happen. The coefficient of static friction at B is 0.5.”

Nine homework assignments were assigned in 2010 and eight were assigned in both 2011 and 2012. In all three years, homework assignments one through three covered the same topics using similar problems. Homework assignment four from 2011 and 2012 covered the topics and problems presented in assignments four and five from 2010. As

Assignment	No. of Students
Homework 3	82
Homework 4	74
Homework 5	76
Homework 6	79
Homework 7	74
Homework 8	69
Homework 9	66
Quiz 1	59
Quiz 2	72
Quiz 3	77
Quiz 4	73
Quiz 5	79
Quiz 6	78
Quiz 7	76
Midterm 1	84
Midterm 2	83
Final	85

Table 3.1: The number of students who completed each assignment with their digital pen from the 2010 course offering. For example, no students completed Homework assignment one with their digital pen while 77 students did complete Quiz three. Students were not instructed to complete assignments one or two with their digital pen.

a result, homework assignments six, seven, eight, and nine from 2010 corresponded to homework assignments five, six, seven and eight from 2011 and 2012 respectively.

3.3 2012 Midterm Grading Rubric

During the 2012 course, problem one from midterm one and all problems from midterm two were each graded using a rigorous, fine-grained rubric. Each rubric comprised an exhaustive list of errors that could be made while solving a given midterm problem. Each error is binary with one indicating that a student made a particular

Assignment	G_a	G_b	G_c	G_d	Total
Homework 3	34	33	34	15	116
Homework 4	34	31	32	17	114
Homework 5	26	31	33	15	120
Homework 6	27	30	32	18	107
Homework 7	22	27	30	17	96
Homework 8	24	30	31	13	98
Quiz 2	23	28	22	11	84
Quiz 3	27	27	28	12	94
Quiz 4	25	30	28	14	97
Quiz 5	25	29	30	16	100
Quiz 6	20	28	27	12	87
Quiz 7	19	26	27	12	84
Midterm 1	36	33	36	17	122
Midterm 2	28	31	34	18	111
Final	25	31	33	16	105

Table 3.2: The number of students in each group who completed each assignment with their digital pen from the 2011 course offering. For example, 24 students from G_a , 30 students from G_b , 31 students from G_c , and 13 students from G_d , completed homework assignment eight with their digital pen. The group numbers, G_a through G_d , correspond to the experiment groupings described in Section 3.6. Students were not instructed to complete assignments one or two or quiz one with their digital pen.

Assignment	G_1	G_2	G_3	G_4	G_5	G_6	Total
Homework 3	23	22	21	17	19	19	121
Homework 4	23	23	20	19	18	17	120
Homework 5	20	23	19	18	18	14	112
Homework 6	23	23	21	18	20	18	123
Homework 7	22	21	20	19	17	18	117
Homework 8	22	20	16	16	17	11	102
Quiz 3	17	18	15	10	12	9	81
Quiz 4	18	19	15	11	16	10	89
Quiz 5	17	18	17	13	18	12	95
Quiz 6	19	17	18	13	17	13	97
Quiz 7	19	18	15	12	16	14	94
Midterm 1	22	22	19	16	19	15	113
Midterm 2	21	20	20	15	18	12	106
Final	22	23	20	16	19	15	115

Table 3.3: The number of students in each group who completed each assignment with their digital pen from the 2012 course offering. For example, 22 students from G_1 , 20 students from G_2 , 16 students from G_3 , and 16 students from G_4 , 17 students from G_5 , and 11 students from G_6 completed homework assignment eight with their digital pen. The group numbers, G_1 through G_6 , correspond to the experiment groupings described in Section 3.6. Students were not instructed to complete assignments one or two or quizzes one or two with their digital pen.

error and a zero indicating that the student did not. Each error is also associated with a penalty value which indicates the number of points a student would lose if he or she made that particular error. Grading these midterm problems was accomplished by checking for the presence of each error and summing their penalty values. These errors provide a rich source of data for the analyses presented in Chapter 9. The full list of errors and their descriptions are shown in Table 3.4, Table 3.5, Table 3.6, and Table 3.7 for midterm one problem one, midterm two problem one, midterm two problem two, and midterm two problem three respectively. Each error in the tables is given a brief description, along with a category and subcategory. The category identifies which solution element the error corresponds to, e.g., FBD or equation. The subcategory distinguishes between different types of the same solution element, e.g., there is more than one type of equation in a typical solution, namely the sum of forces in the x- and y-direction equations and the sum of the moments equation.

3.4 Digital Data and Labeling

In this section, we describe the underlying representation of the sketch data as well as describe the semantic labeling that was either automatically or manually performed.

Each sketch corresponds to a single page of work from a student. Each sketch, $K = \{s_1, \dots, s_m\}$, comprises a series of pen strokes. Each pen stroke, $s_i = \{p_1, \dots, p_n\}$, comprises a series of points. Each point $p_j = \{x, y, t\}$ is a triple where x and y are two-

No.	Category	Subcategory	Description
1	FBD	FBD	Missing
2	FBD	FBD	Diagram does not qualify as an FBD
3	FBD	Body	Incorrect body selected
4	FBD	Pivot	Missing
5	FBD	Pivot	Incorrect
6	FBD	Tension	Missing
7	FBD	Tension	Incorrect
8	FBD	Boom Weight	Missing
9	FBD	Boom Weight	Incorrect
10	FBD	Bulldozer Weight	Missing
11	FBD	Bulldozer Weight	Incorrect
12	Angle	Tension	Angle for tension at C incorrectly calculated
13	Angle	Tension	Angle for tension at C not calculated
14	Equation	Moments about A	Sum of moments about A missing
15	Equation	Moments about A	Horizontal distance A
16	Equation	Moments about A	Tx term missing
17	Equation	Moments about A	Tx term incorrect
18	Equation	Moments about A	Ty term missing
19	Equation	Moments about A	Ty term incorrect
20	Equation	Moments about A	Weight of boom term missing
21	Equation	Moments about A	Weight of boom term incorrect
22	Equation	Moments about A	Weight of bull dozer term missing
23	Equation	Moments about A	Weight of bull dozer term incorrect
24	Equation	Moments about A	Included extraneous force
25	Equation	Sum of X-Forces	Sum of forces in the x-direction equation missing
26	Equation	Sum of X-Forces	Dx term missing
27	Equation	Sum of X-Forces	Dx term incorrect
28	Equation	Sum of X-Forces	Tx term missing
29	Equation	Sum of X-Forces	Tx term incorrect
30	Equation	Sum of X-Forces	Included extraneous force
31	Equation	Sum of X-Forces	Sum of forces in the y-direction equation missing
32	Equation	Sum of Y-Forces	Dy term missing
33	Equation	Sum of Y-Forces	Dy term incorrect
34	Equation	Sum of Y-Forces	Ty term missing
35	Equation	Sum of Y-Forces	Ty term incorrect
36	Equation	Sum of Y-Forces	Weight of boom term missing
37	Equation	Sum of Y-Forces	Weight of boom term incorrect
38	Equation	Sum of Y-Forces	Weight of bull dozer term missing
39	Equation	Sum of Y-Forces	Weight of bull dozer term incorrect
40	Equation	Sum of Y-Forces	Included extraneous force
41	Magnitude	Magnitude of D	No calculation of the magnitude at D
42	Magnitude	Magnitude of D	Incorrect calculation of the magnitude at D
43	Algebra	Algebra	Minor algebra error
44	Algebra	Algebra	Major algebra error

Table 3.4: List of the errors used to grade problems from the first problem of the first midterm from the 2012 course.

Category	Subcategory	Description	
1	1 st FBD	FBD	Missing
2	1 st FBD	FBD	Diagram does not qualify as a FBD
3	1 st FBD	Body	Incorrect body selected
4	1 st FBD	Reaction	Missing
5	1 st FBD	Reaction	Incorrect
6	1 st FBD	Reaction	Extra
7	1 st FBD	Reaction	Included internal forces
8	2 nd FBD	FBD	Missing
9	2 nd FBD	FBD	Diagram does not qualify as a FBD
10	2 nd FBD	Body	Incorrect body selected
11	2 nd FBD	Reaction	Missing
12	2 nd FBD	Reaction	Incorrect
13	2 nd FBD	Reaction	Extra
14	2 nd FBD	Reaction	Included internal forces
15	Equation	Sum of Moments	Sum of moments equation missing
16	Equation	Sum of Moments	Moment arm incorrect
17	Equation	Sum of Moments	Term missing
18	Equation	Sum of Moments	Extra term
19	Equation	Sum of Moments	Force component incorrect
20	Equation	Sum of Moments	Direction incorrect
21	Equation	Sum of Moments	Other error
22	Equation	Sum of X-Forces	Sum of forces in the x-direction equation missing
23	Equation	Sum of X-Forces	Term missing
24	Equation	Sum of X-Forces	Term incorrect
25	Equation	Sum of X-Forces	Extra term
26	Equation	Sum of Y-Forces	Sum of forces in the y-direction equation missing
27	Equation	Sum of Y-Forces	Term missing
28	Equation	Sum of Y-Forces	Term incorrect
29	Equation	Sum of Y-Forces	Extra term
30	Algebra	Algebra	Minor algebra error
31	Answer	Algebra	Did not solve Equation
32	Answer	Algebra	Incorrectly solved Equation
33	Answer	2-force member	Incorrectly identified tension or compression

Table 3.5: List of the errors used to grade problems from the first problem of the second midterm from the 2012 course. The No. column provides a unique number for each error. The Category and Subcategory provide a simple grouping of errors by the problem solving component to which they refer. Lastly, the Description column presents a simple explanation of what the error entailed.

No.	Category	Subcategory	Description
1	FBD	FBD	Missing
2	FBD	FBD	Not qualify as an FBD
3	FBD	Body	Incorrect body selected
4	FBD	Force	Applied force missing
5	FBD	Force	Applied force incorrect
6	FBD	Force	Normal at B missing
7	FBD	Force	Normal at B incorrect
8	FBD	Force	Normal at A missing
9	FBD	Force	Normal at A incorrect
10	FBD	Moment	Moment at A missing
11	FBD	Moment	Moment at A incorrect
12	FBD	Moment	Applied moment missing
13	FBD	Moment	Applied incorrect
14	Angle	Point A	Angle at A incorrect
15	Angle	Point B	Angle at B Incorrect
16	Eqn.	Sum Mom.	Sum of moments missing
17	Eqn.	Sum Mom.	Normal x- term missing
18	Eqn.	Sum Mom.	Normal x- term incorrect
19	Eqn.	Sum Mom.	Normal y- term missing
20	Eqn.	Sum Mom.	Normal y- term incorrect
21	Eqn.	Sum Mom.	Applied moment term missing
22	Eqn.	Sum Mom.	Applied moment term incorrect
23	Eqn.	Sum Mom.	Reaction moment at A term missing
24	Eqn.	Sum Mom.	Reaction moment at A term incorrect
25	Eqn.	Sum Mom.	Included extraneous term
26	Eqn.	Sum X-For.	X-forces eqn. missing
27	Eqn.	Sum X-For.	NA_x term missing
28	Eqn.	Sum X-For.	NA_x term incorrect
29	Eqn.	Sum X-For.	NB_x term missing
30	Eqn.	Sum X-For.	NB_x term incorrect
31	Eqn.	Sum X-For.	Included extraneous force
32	Eqn.	Sum X-For.	Y-forces eqn. missing
33	Eqn.	Sum Y-For.	NA_y term missing
34	Eqn.	Sum Y-For.	NA_y term incorrect
35	Eqn.	Sum Y-For.	NB_y term missing
36	Eqn.	Sum Y-For.	NB_y term incorrect
37	Eqn.	Sum Y-For.	Applied force term missing
38	Eqn.	Sum Y-For.	Applied force incorrect
39	Eqn. Sol.	Sim. Eqn.	No calculation to simultaneously solve equation
40	Eqn. Sol.	Sim. Eqn.	Error in calculation to simultaneously solve equation
41	Algebra	Algebra	Minor algebra error
42	Algebra	Algebra	Major algebra error

Table 3.6: List of the errors used to grade problems from the second problem of the second midterm from the 2012 course.

No.	Category	Subcategory	Description
1	1 st FBD	FBD	Missing
2	1 st FBD	FBD	Diagram does not qualify as a FBD
3	1 st FBD	Body	Incorrect body selected
4	1 st FBD	Reaction	Missing
5	1 st FBD	Reaction	Incorrect
6	1 st FBD	Reaction	Extra
7	1 st FBD	Reaction	Included internal forces
8	2 nd FBD	FBD	Missing
9	2 nd FBD	FBD	Diagram does not qualify as a FBD
10	2 nd FBD	Body	Incorrect body selected
11	2 nd FBD	Reaction	Missing
12	2 nd FBD	Reaction	Incorrect
13	2 nd FBD	Reaction	Extra
14	2 nd FBD	Reaction	Included internal forces
15	3 rd FBD	FBD	Missing
16	3 rd FBD	FBD	Diagram does not qualify as a FBD
17	3 rd FBD	Body	Incorrect body selected
18	3 rd FBD	Reaction	Missing
19	3 rd FBD	Reaction	Incorrect
20	3 rd FBD	Reaction	Extra
21	3 rd FBD	Reaction	Included internal forces
22	Eqn.	Sum Mom.	Sum of moments equation missing
23	Eqn.	Sum Mom.	Moment arm incorrect
24	Eqn.	Sum Mom.	Term missing
25	Eqn.	Sum Mom.	Extra term
26	Eqn.	Sum Mom.	Force component incorrect
27	Eqn.	Sum Mom.	Direction incorrect
28	Eqn.	Sum Mom.	Other error
29	Eqn.	Sum X-For.	Sum of forces in the x-direction equation missing
30	Eqn.	Sum X-For.	Term missing
31	Eqn.	Sum X-For.	Term incorrect
32	Eqn.	Sum X-For.	Extra term
33	Eqn.	Sum Y-For.	Sum of forces in the y-direction equation missing
34	Eqn.	Sum Y-For.	Term missing
35	Eqn.	Sum Y-For.	Term incorrect
36	Eqn.	Sum Y-For.	Extra term
37	Algebra	Algebra	Minor algebra error
38	Answer	Algebra	Did not solve equation
39	Answer	Algebra	Incorrectly solved equation
40	Answer	2FM	Incorrectly identified tension or compression
41	Angle	Angle	the angle of forces at ED and EC incorrect

Table 3.7: List of the errors used to grade problems from the third problem of the second midterm from the 2012 course.

dimensional Cartesian coordinates, and t is the time-stamp of that point. All points within a pen stroke, and all pen strokes within a sketch, are ordered by increasing time-stamp. The time-stamp of the first point in a pen stroke signifies the start time of that pen stroke and the last point is used to signify its end time.

3.4.1 2010 Labeling Scheme

This data has been automatically labeled using the technique developed by Lung et al. [48]. This produced a sequence of labels for each sketch, $L = \{l_1, \dots, l_m\} | l \in \{FBD, EQN, CRO\}$. Each label, l_i , identifies stroke, s_i , by its semantic content: free body diagram (FBD), equation (EQN), or cross-out (CRO).

3.4.2 2011 Labeling Scheme

None of the homework assignments from this course have been labeled, but all quizzes, midterms, and the final exam were manually labeled. This produced a sequence of labels. L , for each sketch:

$$L = \{l_1, \dots, l_m\} | l \in \{FBD, EQN, EQN_m, EQN_f, EXP, ORG, \\ CRO, PID, SUM_m, SUM_f, GEO_f, GEO_e, PRB, PST, DIR_m, DIR_f\}$$

Each label, l_i , identifies stroke, s_i , by its semantic content: free body diagram (*FBD*); general equations not including force and moment equations (*EQN*); moment equations (*EQN_m*); force equations (*EQN_f*); unprompted written explanation of a student's problem-solving process (*EXP*); organizational writing, e.g., boxing answers (*ORG*); cross-out pen strokes (*CRO*); personally identifying information, such

as a student's name (*PID*)³; sum of moment equation declaration, typically a student will begin solving a moment equation by labeling it as $\sum M = 0$ (*SUM_m*); sum of forces equation declaration, typically a student will begin solving a force equation by labeling it as $\sum F_x = 0$ or $\sum F_y = 0$ (*SUM_f*); geometric figures used to assist with trigonometric calculations (*GEO_f*); equations accompanying geometric figures (*GEO_e*); problem number labels (*PRB*); rewritten copy of the problem statement text (*PST*); moment direction symbol which indicates the direction of positive moments (*DIR_m*); force direction symbol which indicates the positive direction of the forces (*DIR_f*). Typical examples of each label are shown in Figure 3.11 through Figure 3.17; differences in pen stroke color for each diagram is used to distinguish pen strokes of different labels.

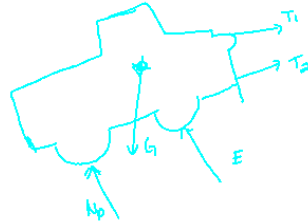
3.4.3 2012 Labeling Schemes

All data from this course were manually labeled; quizzes, midterms, and the final exam were labeled with a different scheme than the homework assignments.

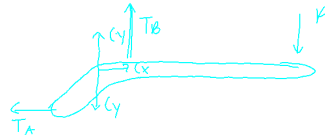
For the homework assignments, this produced a sequence of labels for each sketch, $L = \{l_1, \dots, l_m\} | l \in \{FBD, EQN, CRO\}$. Each label, l_i , identifies stroke, s_i , by its semantic content: free body diagram (FBD), equation (EQN), or cross-out (CRO).

For the quizzes, midterms and final exam, this produced a sequence of labels for each sketch, $L = \{l_1, \dots, l_m\} | l \in \{FBD, EQN, EXP, ORG, CRO, PID, GEO, PRB, PST\}$.

³All personally identifying information has been removed from all data (not just the 2011 data set) to protect the anonymity of the students who generated this data



(a) Example of a free body diagram pen stroke, labeled as *FBD* in all data sets.



(b) Example of a free body diagram pen stroke, labeled as *FBD* in all data sets.



(c) Example of a free body diagram pen stroke, labeled as *FBD* in all data sets.

Figure 3.11: Examples of pen strokes labeled as *FBD*.

$$T_B = T_A e^{.3\left(\frac{3}{2}\right)}$$

(a) General equation example, labeled as *EQN* for all data sets.

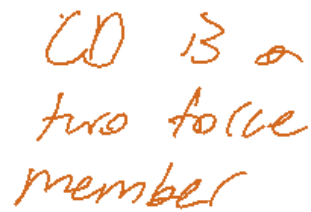
$$-By(-.75in) - F(1.75) = 0$$

(b) Moment equation example, labeled as: *EQN* for the 2010 data set; *EQN_m* for the 2011 data set; *EQN* for the 2012 homework and exam data sets.

$$-(y + T_B)P = 0$$

(c) Force equation example, labeled as: *EQN* for the 2010 data set; *EQN_f* for the 2011 data set; *EQN* for the 2012 homework and exam data sets.

Figure 3.12: Examples different pen strokes labeled as equation.



CD is a
two force
member

(a) Example of explanation pen strokes, labeled as: *EQN* for the 2010 data set; *EXP* for the 2011 and 2012 exam data set; *EQN* for the 2012 homework data set.



Next Page →

(b) Organization pen stroke example, labeled as: either *FBD* or *EQN* for the 2010 and 2012 homework data set, depending upon whether or not the pen stroke was used to organize FBD or equation work respectively; *ORG* for the 2011 and 2012 exam data set.



~~Eg = 2 kW~~

(c) Cross-out example, labeled as *CRO* in all data sets.

Figure 3.13: Examples pen strokes labeled as organization, explanation and cross-out.

$$\sum M_B = 0$$

(a) Example of sum of moment label pen strokes, labeled as: *EQN* in the 2010 and both 2012 data sets; *SUM_m* in the 2011 data set.

$$\sum F_y = 0$$

(b) Example of sum of force label pen strokes, labeled as: *EQN* in the 2010 and both 2012 datasets; *SUM_f* in the 2011 data set.

Figure 3.14: Examples of the pen strokes labeled as sum-of-moments and sum-of-forces.



(a) Example of geometry figure pen strokes, labeled as: *FBD* in the 2010 and 2012 homework data set; *GEO_f* in the 2011 data set; *GEO* in the 2012 data set.

$$\theta = \tan^{-1}\left(\frac{1.2}{2.1}\right) = 29.7^\circ$$

(b) Example of geometry equation pen strokes, labeled as: *FBD* in the 2010 and 2012 homework data set; *GEO_e* in the 2011 data set; *GEO* in the 2012 data set.

Figure 3.15: Examples of pen strokes labeled as geometry figure and and geometry text.



(a) Example of moment direction pen strokes, labeled as: DIR_m in the 2011 data set; EQN in the 2010 and both 2012 data sets.



(b) Example of force direction pen strokes, labeled as: DIR_f in the 2011 data set; EQN in the 2010 and both 2012 data sets.

Figure 3.16: Examples of pen strokes labeled as moment and force direction.

A vertical force of 60 N is applied to one end of a crowbar to pull a nail from a floor. With the crowbar skew in the figure, find the force applied to the floor at A. How does the answer change if the 60 N force is applied at the same location \perp to the crowbar?

Figure 3.17: Example of pen strokes labeled as problem statement, labeled as: PST in the 2011 and 2012 exam data sets; EQN in the 2010 and 2012 homework data sets.

Each label, l_i , identifies stroke, s_i , by its semantic content: free body diagram (*FBD*); all equations pen strokes (*EQN*); unprompted written explanation of a student's problem-solving process (*EXP*); organizational writing, e.g., boxing answers (*ORG*); cross-out pen strokes (*CRO*); personally identifying information, such as a student's name (*PID*)⁴; geometric figures and equations used to assist with trigonometric calculations (*GEO*); problem number labels (*PRB*); rewritten copy of the problem statement text (*PST*).

3.5 Descriptive Statistics

As this data set is the first of its kind, we present in this section, a rigorous set of descriptive statistics for simple numerical features. These features are simply the amount of time spent and ink written on each homework problem. We compute the mean and standard deviation for these features and present them in Figure 3.18 through Figure 3.41. For example, Figure 3.18 presents the average and standard deviation of the total ink written on every problem of every homework assignment. We present an extensive collection of the histograms for each feature in the Appendix. This analysis provides instructors with a very precise window into how students complete their homework assignments in this course.

⁴All personally identifying information has since been removed from all data (not just the 2011 data set) to protect the anonymity of the students who generated this data

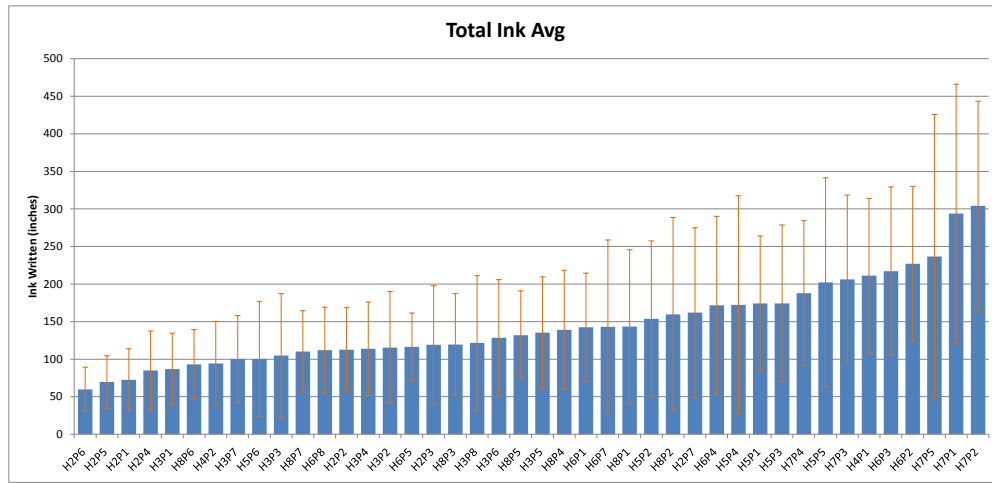


Figure 3.18: Average total ink written for each assignment from the 2010 data. Each bar shows the average amount of ink written on a single homework problem. The label given to each bar indicates the homework number and problem number, for example, H2P6 corresponds to homework two problem six.

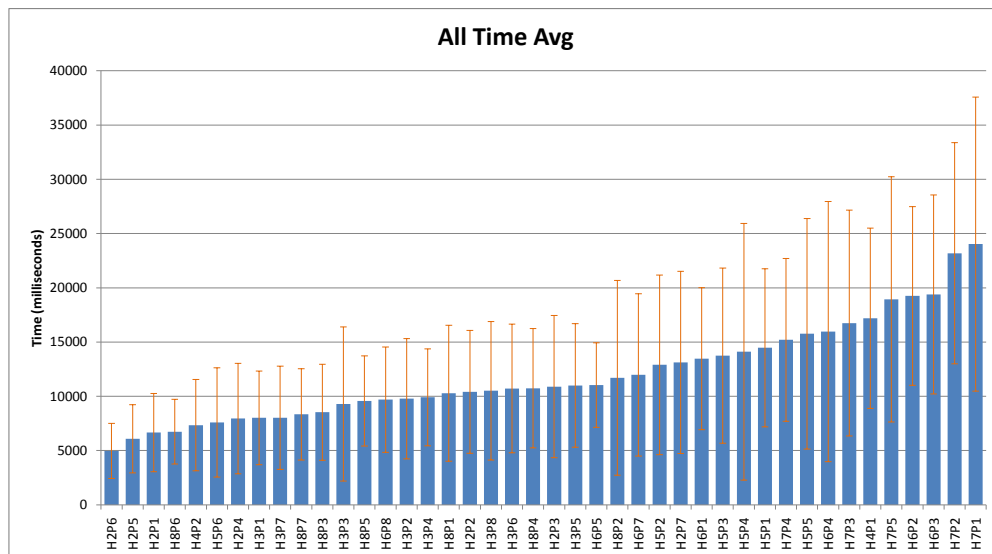


Figure 3.19: Average total time spent writing for each assignment from the 2010 data. Each bar shows the average amount of time spent writing on a single homework problem. The label given to each bar indicates the homework number and problem number, for example, H2P6 corresponds to homework two problem six.

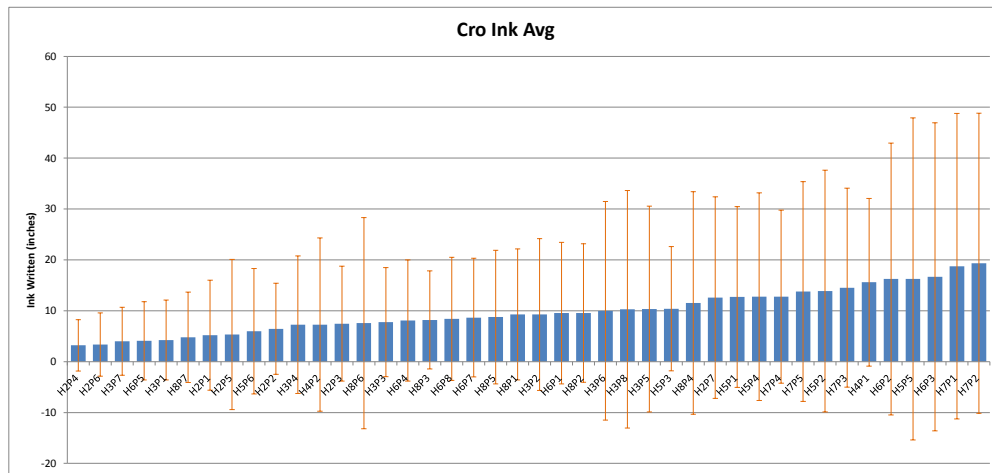


Figure 3.20: Average cross-out ink written for each assignment from the 2010 data. Each bar shows the average amount of cross-out ink written on a single homework problem. The label given to each bar indicates the homework number and problem number, for example, H2P6 corresponds to homework two problem six.

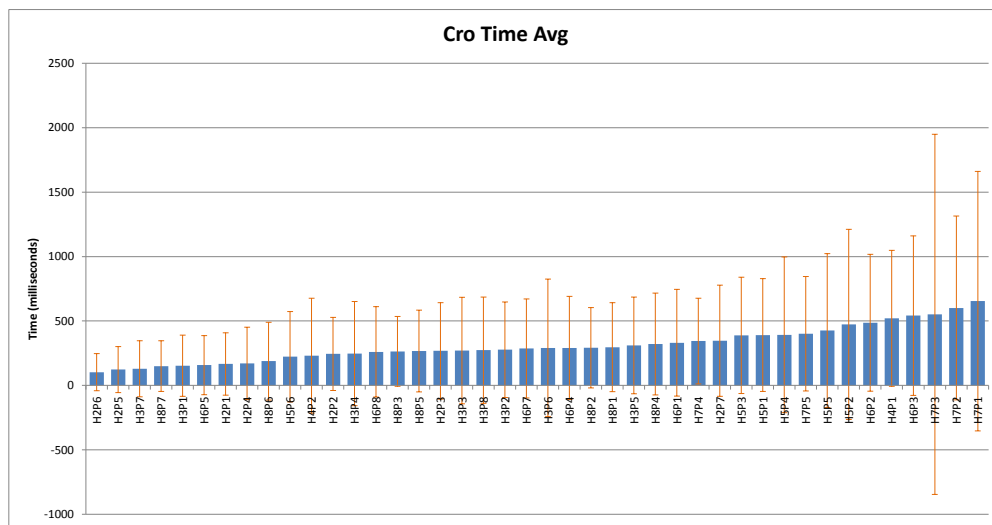


Figure 3.21: Average time spent writing cross-outs for each assignment from the 2010 data. Each bar shows the average time spent writing cross-outs on a single homework problem. The label given to each bar indicates the homework number and problem number, for example, H2P6 corresponds to homework two problem six.

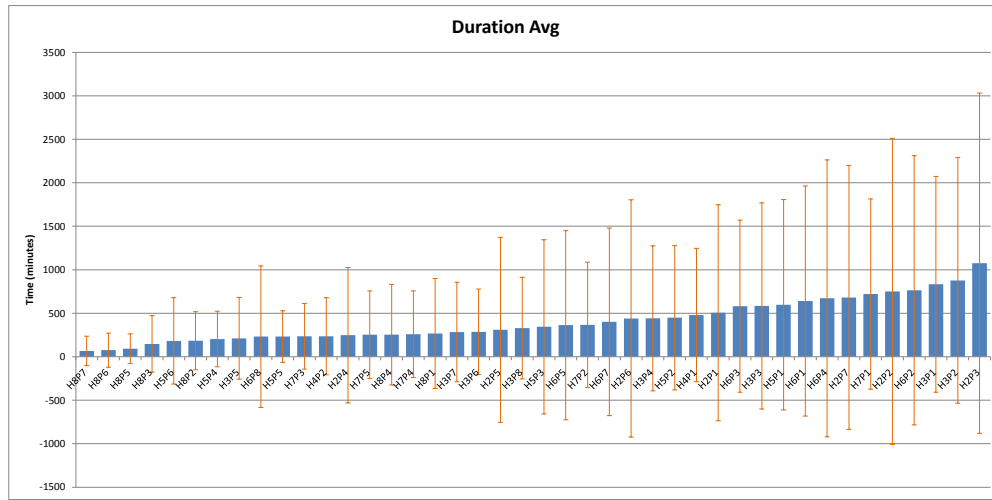


Figure 3.22: Average duration (time from the first to last stroke in an assignment) of each assignment from the 2010 data. Each bar shows the average duration (time from the first pen stroke to the last of a homework solution) of a single homework problem solution. The label given to each bar indicates the homework number and problem number, for example, H2P6 corresponds to homework two problem six.

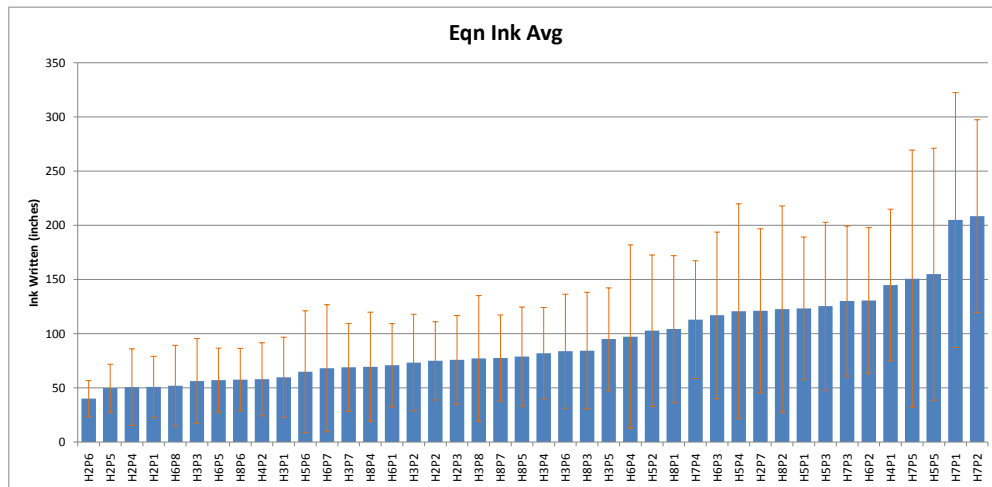


Figure 3.23: Average equation ink written on each assignment from the 2010 data. Each bar shows the average amount of equation ink written on a single homework problem. The label given to each bar indicates the homework number and problem number, for example, H2P6 corresponds to homework two problem six.

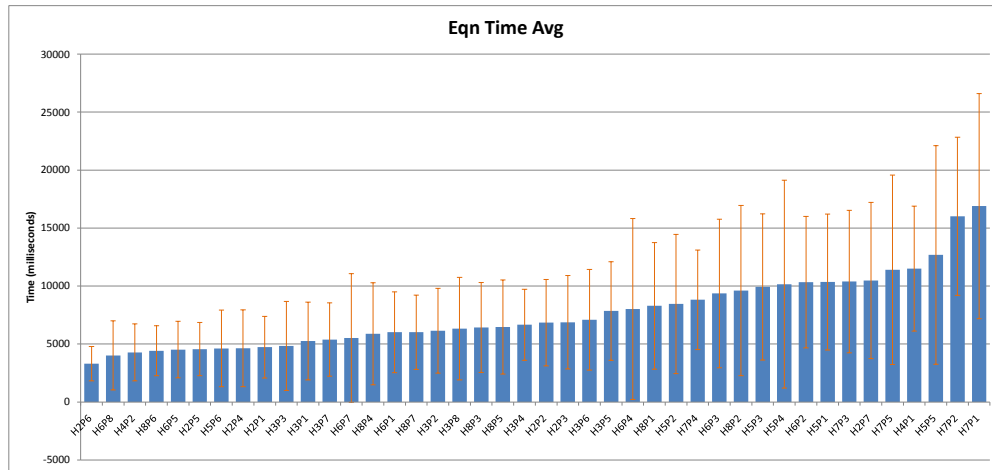


Figure 3.24: Average time spent writing equations on each assignment from the 2010 data. Each bar shows the average time spent writing equations on a single homework problem. The label given to each bar indicates the homework number and problem number, for example, H2P6 corresponds to homework two problem six.

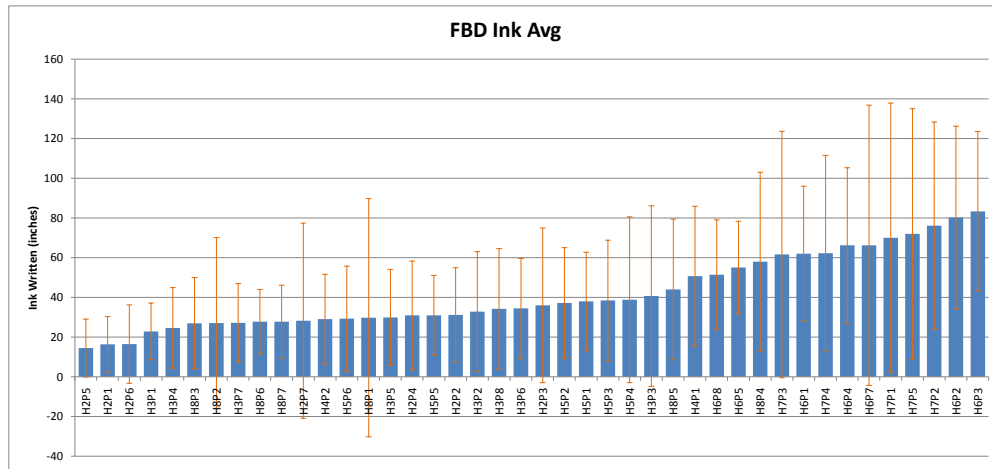


Figure 3.25: Average FBD ink written for each assignment from the 2010 data. Each bar shows the average amount of FBD ink written on a single homework problem. The label given to each bar indicates the homework number and problem number, for example, H2P6 corresponds to homework two problem six.

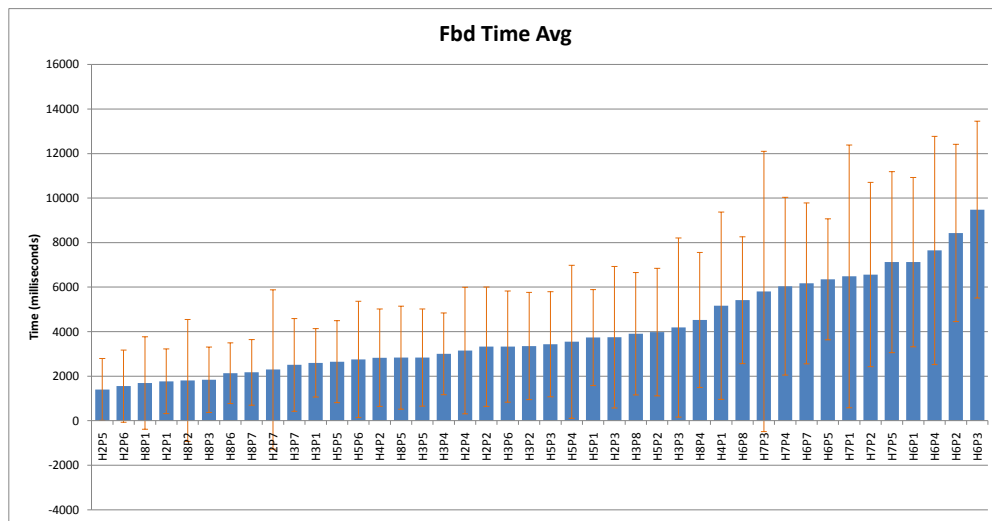


Figure 3.26: Average time spent writing FBDs for each assignment from the 2010 data. Each bar shows the average time spent writing FBDs on a single homework problem. The label given to each bar indicates the homework number and problem number, for example, H2P6 corresponds to homework two problem six.

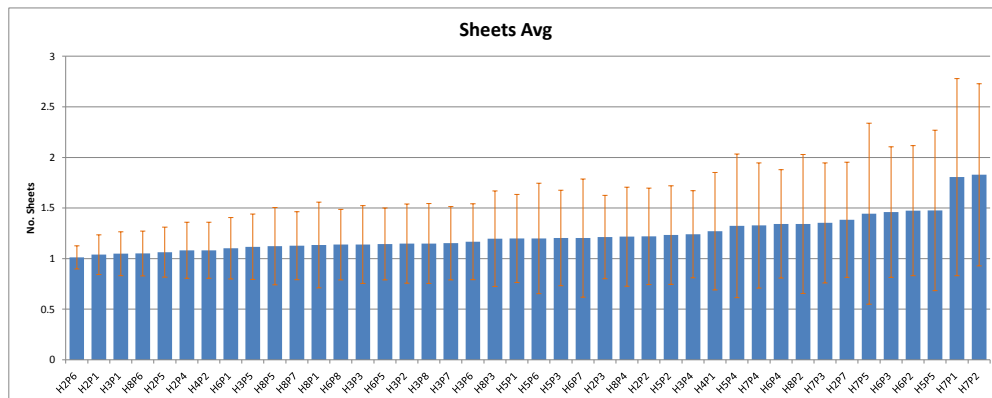


Figure 3.27: Average number of sheets taken to solve each problem of each assignment from the 2010 data. Each bar shows the average number of sheets written on for a single homework problem. The label given to each bar indicates the homework number and problem number, for example, H2P6 corresponds to homework two problem six.

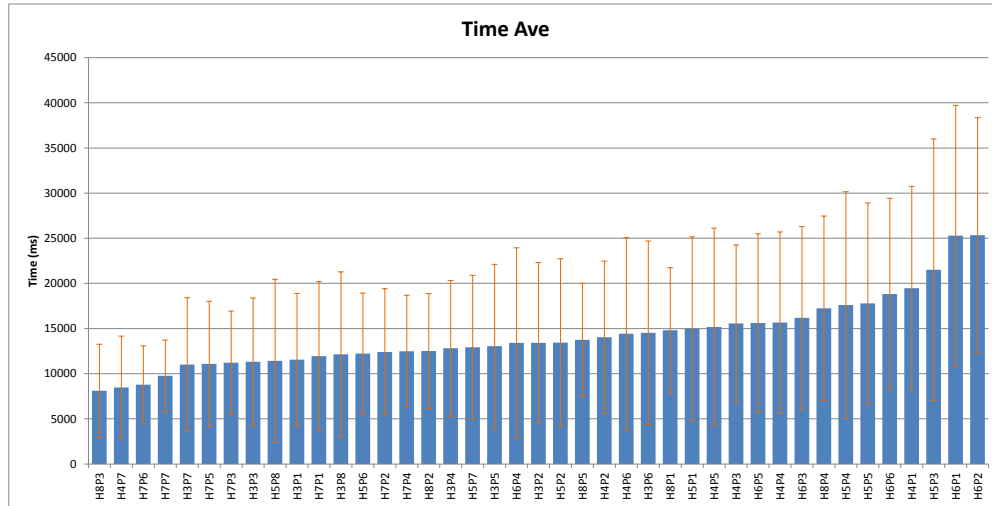


Figure 3.28: Average time spent writing each problem of each assignment from the 2011 data. Each bar shows the average amount of ink written on a single homework problem. The label given to each bar indicates the homework number and problem number, for example, H2P6 corresponds to homework two problem six.

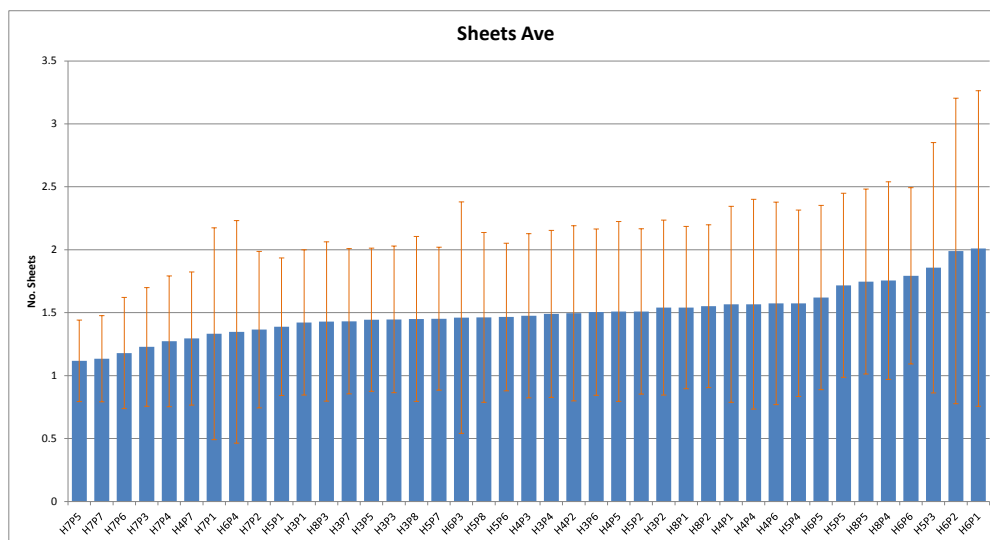


Figure 3.29: Average number of sheets taken to solve each problem of each assignment from the 2011 data. Each bar shows the average number of sheets used in solving a single homework problem. The label given to each bar indicates the homework number and problem number, for example, H2P6 corresponds to homework two problem six.

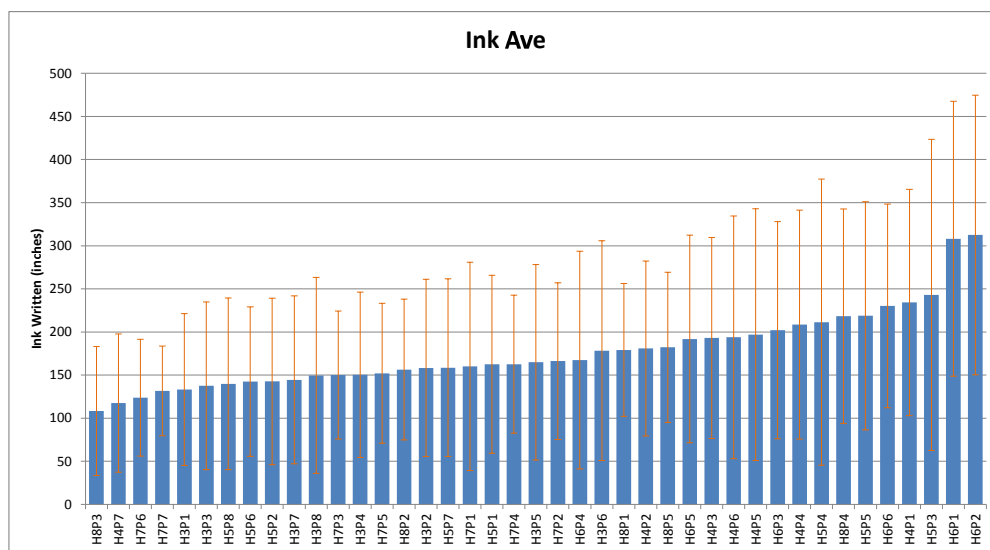


Figure 3.30: Average amount of ink written on each problem of each assignment from the 2011 data. Each bar shows the average amount of ink written on a single homework problem. The label given to each bar indicates the homework number and problem number, for example, H2P6 corresponds to homework two problem six.

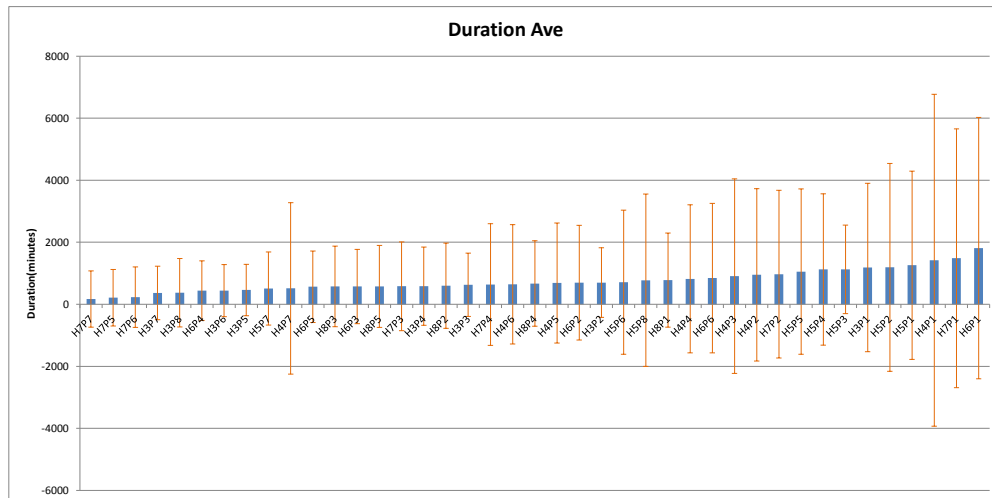


Figure 3.31: Average duration (time from the first to last stroke) of each problem from each assignment of the 2011 data. Each bar shows the average solution duration (time from first pen stroke to last) for a single homework problem. The label given to each bar indicates the homework number and problem number, for example, H2P6 corresponds to homework two problem six.

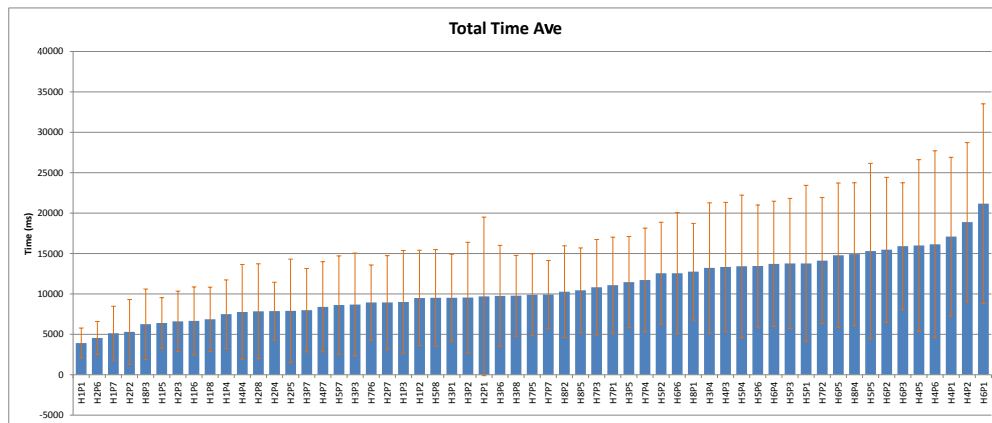


Figure 3.32: Average time taken to solve each problem of each assignment from the 2012 data. Each bar shows the average amount of time spent writing on a single homework problem. The label given to each bar indicates the homework number and problem number, for example, H2P6 corresponds to homework two problem six.

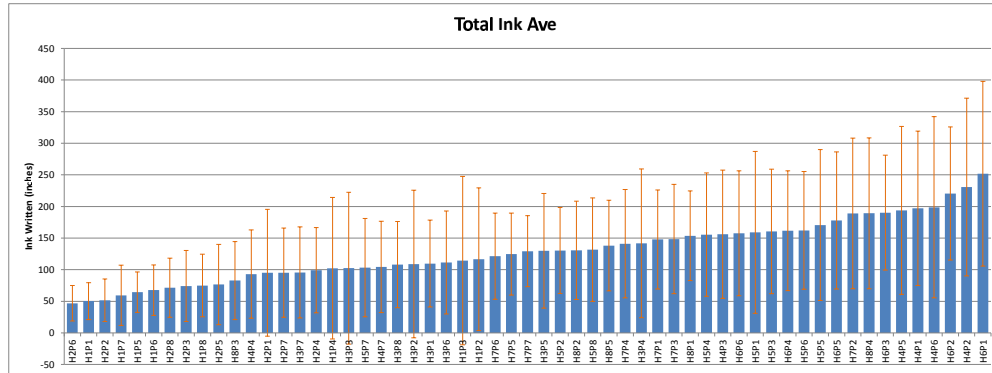


Figure 3.33: Average total amount of ink written to solve each problem of each assignment from the 2012 data. Each bar shows the average amount of ink written on a single homework problem. The label given to each bar indicates the homework number and problem number, for example, H2P6 corresponds to homework two problem six.

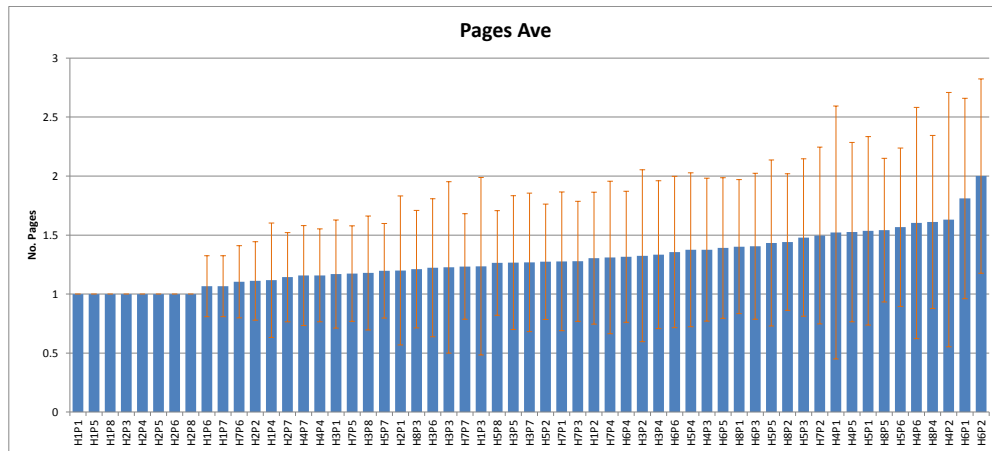


Figure 3.34: Average number of pages taken to solve each problem of each assignment from the 2012 data. Each bar shows to the average number of pages of a solution to a single homework problem. The label given to each bar indicates the homework number and problem number, for example, H2P6 corresponds to homework two problem six.

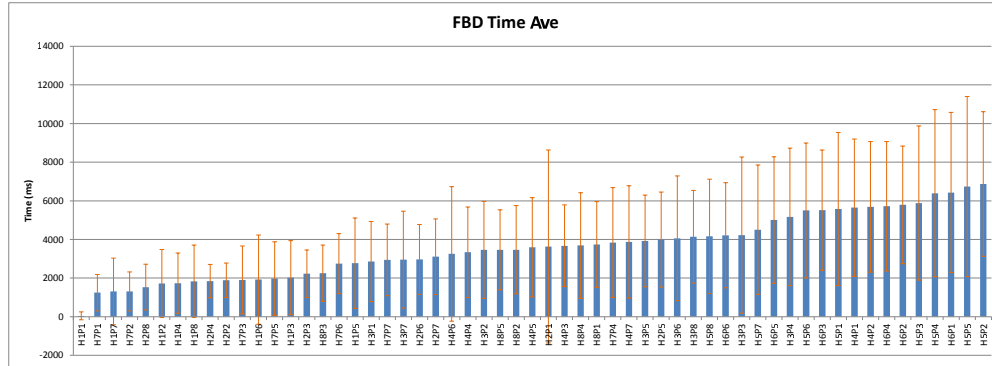


Figure 3.35: Average amount of time spent writing FBDs for each problem of each assignment from the 2012 data. Each bar shows the average amount of time spent writing FBDs on a single homework problem. The label given to each bar indicates the homework number and problem number, for example, H2P6 corresponds to homework two problem six.

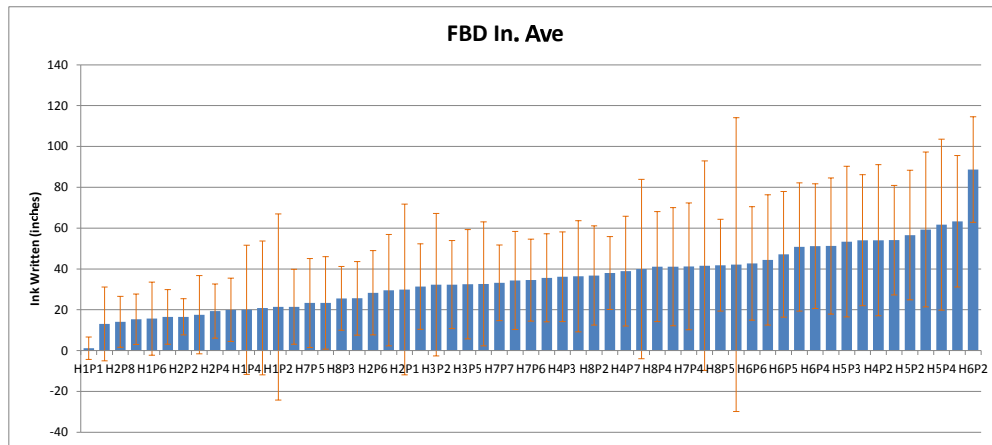


Figure 3.36: Average amount of ink written FBDs for each problem of each assignment from the 2012 data. Each bar shows the average amount of FBD ink written on a single homework problem. The label given to each bar indicates the homework number and problem number, for example, H2P6 corresponds to homework two problem six.

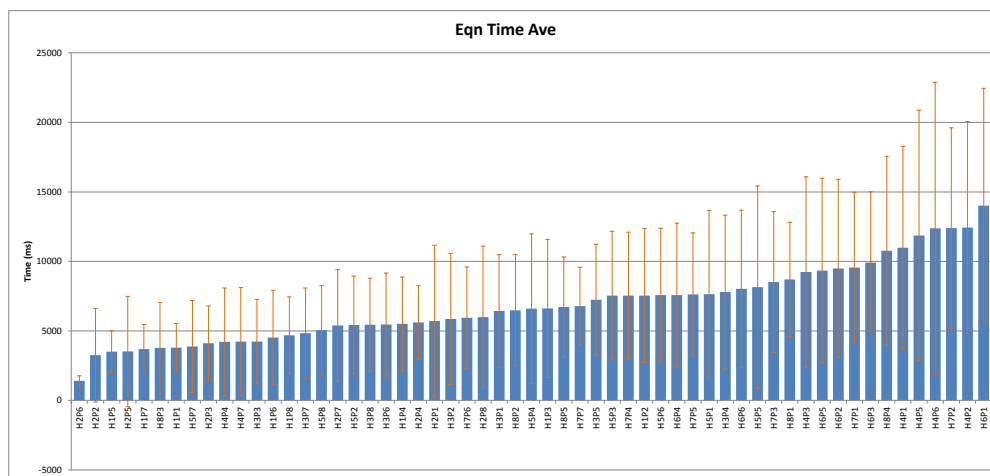


Figure 3.37: Average amount of time spent writing equations for each problem of each assignment from the 2012 data. Each bar shows the amount of time spent writing equations on a single homework problem. The label given to each bar indicates the homework number and problem number, for example, H2P6 corresponds to homework two problem six.

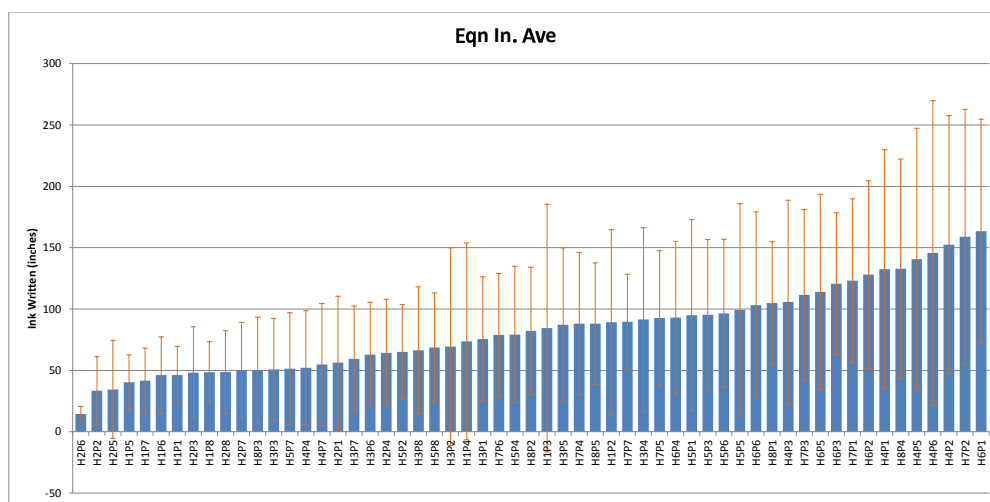


Figure 3.38: Average amount of equation ink written for each problem of each assignment from the 2012 data. Each bar shows the average amount of equation ink written on a single homework problem. The label given to each bar indicates the homework number and problem number, for example, H2P6 corresponds to homework two problem six.

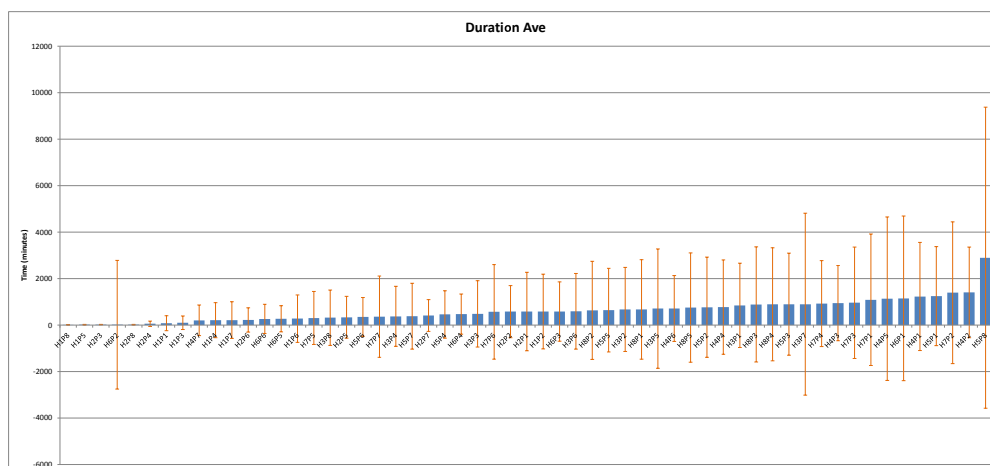


Figure 3.39: Average duration (time from the first to last stroke) of each problem of each assignment from the 2012 data. Each bar shows the average duration of a solution to a single homework problem. The label given to each bar indicates the homework number and problem number, for example, H2P6 corresponds to homework two problem six.

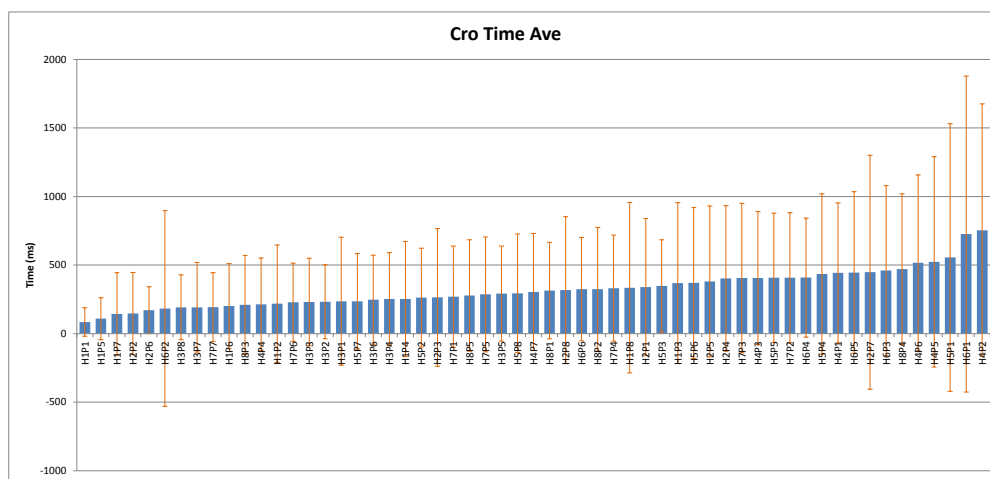


Figure 3.40: Average time spent crossing out work for each problem of each assignment from the 2012 data. Each bar shows the average amount of time spent writing cross-outs on a single homework problem. The label given to each bar indicates the homework number and problem number, for example, H2P6 corresponds to homework two problem six.

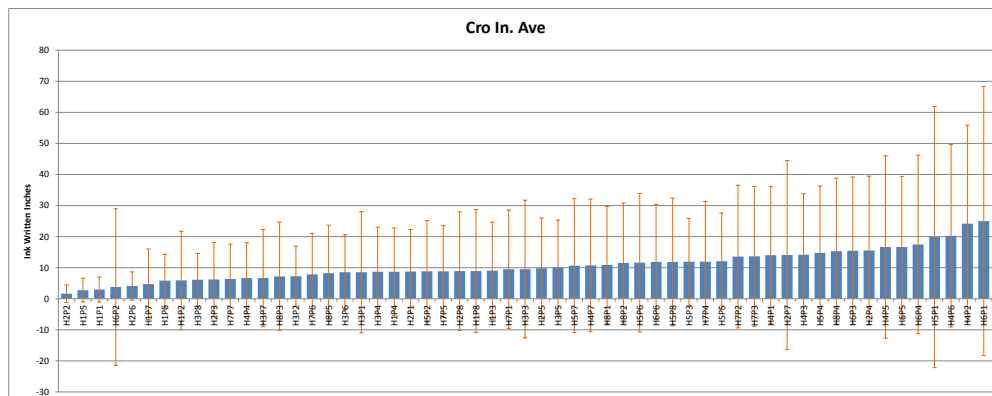


Figure 3.41: Average amount of cross-out ink written for each problem of each assignment from the 2012 data. Each bar shows the average amount of cross-out ink written on a single homework problem. The label given to each bar indicates the homework number and problem number, for example, H2P6 corresponds to homework two problem six.

3.6 Experimental Studies

The students enrolled in the 2011 and 2012 course offerings were split into different groups, each of which was given a different experimental treatment. There were no experimental treatments given to students enrolled in the 2010 course, and thus no groupings either. These treatments were typically supplemental learning tasks, such as spending time with a tutoring system. In this section, we describe each of the groups from each course offering and the experimental treatments they received.

3.6.1 2011 Experimental Studies

The students from this course were organized into four experimental groups according to their discussion section. While all students enrolled in this course attended the same lecture led by the course professor, there were four mandatory discussion sections in which students received supplemental material from a teaching assistant. Each discussion section constituted an experimental group. Thus there were four experimental groups for this course offering, identified as: G_a , G_b , G_c , and G_d . Students in G_a were the control group, and received no special instructional materials during the course. G_a comprised 36 students. Students in G_b were given supplemental instruction with Newton's Pen II [79], an intelligent, pen-based Statics tutoring system. Group G_b comprised 37 students. The final two groups, G_c and G_d , were given self-explanation prompts along with homework assignments three, four, five, six, and eight. These prompts elicited from students a handwritten explanation of the reasoning behind their problem-solving steps on each homework problem. Students were told to respond to these prompts for each

problem after completing that problem. G_c comprised 37 students and G_d comprised 20.

To ground later analyses, we asked three experts to complete homework assignments three and eight as well as generate self-explanations. These experts comprised one graduate and two undergraduate mechanical engineering students, the latter two of whom had solved these exact homework problems two years prior. We manually transcribed the student and expert self-explanations. Spelling errors were corrected, but grammatical errors were left as is. The experts' self-explanation transcripts exemplify the types of responses we expect from students who possess an expert-stance on statics concepts.

We elicited self-explanations only for those homework assignments that dealt primarily with equilibrium. We excluded, for example, the first two assignments which covered prerequisite topics such as vectors and moments. Students from G_c and G_d generated self-explanation for homework assignments three, four, five, six, and eight.

The prompts for homework assignment four were:

1. Why did you select the system that you used for your free-body diagram?
2. Could you have selected some other system and still solved the problem?
3. How did you model each of the reaction forces? For example, did you consider the reaction to be a pivot, roller, contact with friction, etc?
4. When computing moments for the moment equilibrium equation(s), why did you choose the particular point that you used to compute moments about? For example, if you computed moments about point A, why did you pick A and not some

other point?

5. Could you have simplified the analysis by picking some other point to take moments about?
6. Why did you choose to solve the equilibrium equations in the order that you did? For example, if you solved the x-equilibrium equation first, the moment equilibrium equation about the z-axis second, and so on, why did you choose this particular order?

The prompts for homework assignment five were as follows:

1. Why did you decompose the system into the particular set of free body diagrams that you used?
2. Which bodies are two-force members?
3. Why did you analyze the free body diagrams in the order that you did? For example, if you analyzed the equilibrium conditions for bar A first, and bar B second, why did you choose this order?
4. When computing moments for a moment equilibrium equation, why did you choose the particular point that you used to compute moments about? For example, if you computed moments about point A, why did you pick A and not some other point?

The prompts for homework assignment six were only required for problems five and six and were as follows:

1. Why did you decompose the system into the particular set of free body diagrams that you used?
2. Which bodies are two-force members?
3. Why did you analyze the free body diagrams in the order that you did? For example, if you analyzed the equilibrium conditions for bar A first, and bar B second, why did you choose this order?
4. When computing moments for a moment equilibrium equation, why did you choose the particular point that you used to compute moments about? For example, if you computed moments about point A, why did you pick A and not some other point?

The prompts for homework assignment eight were as follows:

1. To begin your solution, you must make several assumptions. Which surfaces, if any, did you assume were on the verge of slip? Why did you think these surfaces were on the verge of slip?
2. If slip did occur, which direction would each slipping member move?
3. For problem 3 only: Why is the device self-locking?

3.6.2 2012 Experimental Studies

In 2012, we made several extensions to the self-explanation study conducted in the 2011 course offering.

First, there were six groups instead of four. These six groups were assigned at random, instead of based upon section number, and are identified as G_1 , G_2 , G_3 , G_4 , G_5 , and G_6 . G_2 received no self-explanation prompts but instead received several treatments with the Newton's Pen II[79] tutorial system. G_3 , G_4 , G_5 , G_6 all received self-explanation prompts on homework assignments three, four, five, and eight, though a different set of prompts was assigned to each group for each assignment. The differences between the prompts given to each group comprise the sequencing and degree to which those prompts were scaffolded. Scaffolding is a pedagogical approach in which a student is guided but not told the correct solution. Finally, G_1 was the "delayed self-explanation" group, meaning that they received no self-explanation prompts on assignments three and four, but received the same prompts as did G_3 for assignments five and eight.

3.6.2.1 Homework Three Self-Explanation Prompts

Below are the self-explanation prompts that G_3 received for homework three. None of the prompts are scaffolded.

1. Why did you select the system that you used for your free-body diagram? Could you have selected some other system and still solved the problem?
2. How did you model each of the reaction forces?
3. When computing moments for the equilibrium, why did you choose the particular point that you used to compute moments about? For example, if you computed moments about A, why did you pick A and not some other point? Could you have simplified the analysis by picking some other point to take moments about?

Hw - Prob.	G_1	G_2	G_3	G_4	G_5	G_6
3 - 1	N	N	U	U	S	S
3 - 2	N	N	U	S	U	S
3 - 3	N	N	U	U	U	S
3 - 4	N	N	U	U	S	S
3 - 5	N	N	U	S	U	S
3 - 6	N	N	U	U	U	S
3 - 7	N	N	U	U	U	U
3 - 8	N	N	U	U	U	U
4 - 1	N	N	U	U	U	U
4 - 2	N	N	U	U	S	S
4 - 3	N	N	U	S	U	S
4 - 4	N	N	U	U	U	U
4 - 5	N	N	U	U	S	S
4 - 6	N	N	U	S	U	S
4 - 4	N	N	U	U	U	U
4 - 5	N	N	U	U	U	U
4 - 6	N	N	U	U	U	U
4 - 7	N	N	U	U	U	U
4 - 8	N	N	U	U	U	U
5 - 1	N	U	U	U	U	U
5 - 2	N	U	U	U	S	S
5 - 3	N	U	U	S	U	S
5 - 4	N	U	U	U	U	U
5 - 5	N	U	U	U	S	S
5 - 6	N	U	U	S	U	S
5 - 7	N	U	U	U	U	U
5 - 8	N	U	U	U	U	U
8 - 1	N	U	U	U	U	U
8 - 2	N	U	U	U	U	U
8 - 3	N	U	U	U	U	U
8 - 4	N	U	U	U	U	U
8 - 5	N	U	U	U	U	U

Table 3.8: This table presents an overview of which students received scaffolded self-explanation prompts on particular homework assignment problems. Each row corresponds to a single homework assignment problem. Each column of that row indicates whether a particular group received a scaffolded self-explanation prompt (S) or a prompt that was unscaffolded (U), or received no self-explanation prompt (N)

4. Why did you choose to solve the equilibrium equations in the order that you did?
For example, if you solved the x-equilibrium equation, then the y-equilibrium equation, and finally the moment equilibrium equation, why did you chose this order?

Below are the self-explanation prompts for problem one of homework three, which were given to G_5 and G_6 . The second prompt is scaffolded.

1. Why did you select the system that you used for your free-body diagram? Could you have selected some other system and still solved the problem?
2. Reaction forces can be modeled as contact with smooth or rough surfaces, roller supports, freely sliding guides, pin connections, fixed supports, and connections to flexible elements. How did you model the reactions forces at A and B? Why?
3. When computing moments for the equilibrium equation, why did you choose the particular point that you used to compute moments about? For example, if you computed moments about A, why did you pick A and not some other point? Could you have simplified the analysis by picking some other point to take moments about?
4. Why did you choose to solve the equilibrium equations in the order that you did?
For example, if you solved the x-equilibrium equation, then the y-equilibrium equation, and finally the moment equilibrium equation, why did you choose this order?

Below are the self-explanation prompts for problem two of homework three which were given to G_4 and G_6 . The second prompt is scaffolded.

1. Why did you select the system that you used for your free-body diagram? Could you have selected some other system and still solved the problem?
2. Reaction forces can be modeled as contact with smooth or rough surfaces, roller supports, freely sliding guides, pin connections, fixed supports, and connections to flexible elements. How did you model the reaction forces at A and at the nail? Why?
3. When computing moments for the equilibrium, why did you choose the particular point that you used to compute moments about? For example, if you computed moments about A, why did you pick A and not some other point? Could you have simplified the analysis by picking some other point to take moments about?
4. Why did you choose to solve the equilibrium equations in the order that you did? For example, if you solved the x-equilibrium equation, then the y-equilibrium equation, and finally the moment equilibrium equation, why did you chose this order?

Below are the self-explanation prompts for problem four of homework three which were given to G_5 and G_6 . The first and second prompts are scaffolded.

1. Below are three examples of systems that you could select for a free-body diagram. Each system is highlighted in green. Which is the best? Why? What is lacking or incorrect about the remaining two choices?

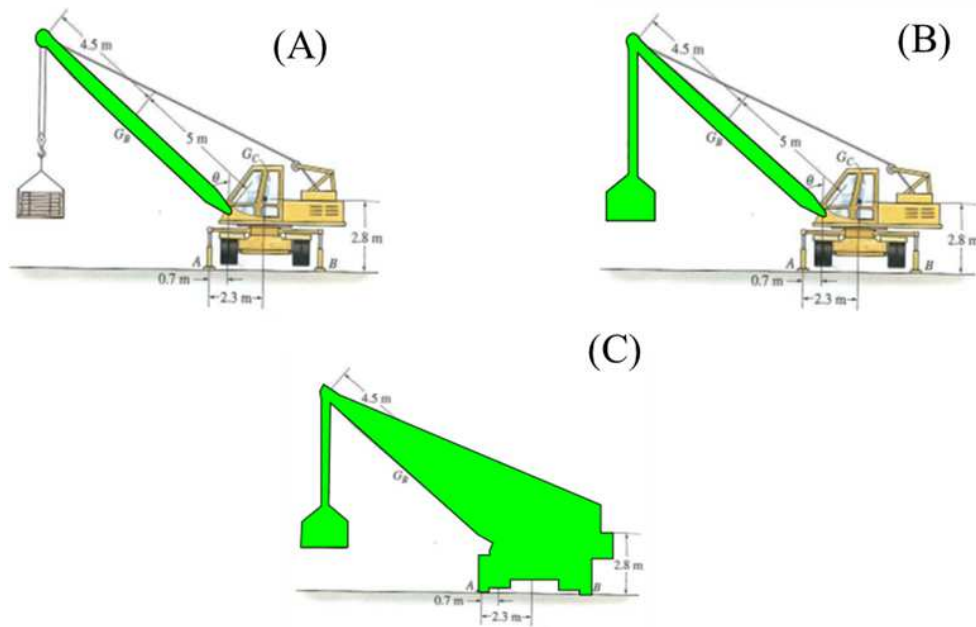


Figure 3.42: Self-explanation prompt for homework three problem four. Students were shown different potential FBDs to consider in their self-explanation.

2. Reaction forces can be modeled as contact with smooth or rough surfaces, roller supports, freely sliding guides, pin connections, fixed supports, and connections to flexible elements. How did you model the reactions forces at A and B? Why?
3. When computing moments for the equilibrium equation, why did you choose the particular point that you used to compute moments about? For example, if you computed moments about A, why did you pick A and not some other point? Could you have simplified the analysis by picking some other point to take moments about?
4. Why did you choose to solve the equilibrium equations in the order that you did? For example, if you solved the x-equilibrium equation, then the y-equilibrium equation, and finally the moment equilibrium equation, why did you choose this

order?

Below are the self-explanation prompts for problem five of homework three which were given to G_5 and G_6 . The first prompt is scaffolded.

1. Below are three examples of systems that you could select for a free-body diagram. Each system is highlighted in green. Which is the best? Why? What is lacking or incorrect about the remaining two choices?

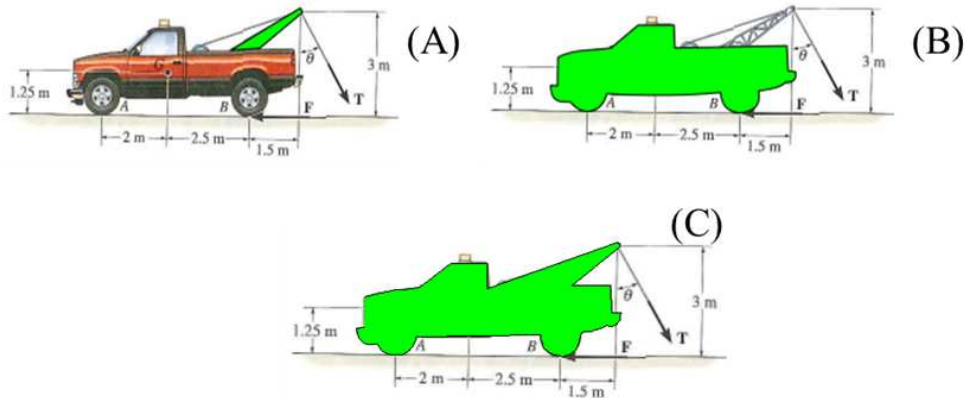


Figure 3.43: Self-explanation prompt for homework three problem five. Students were shown different potential FBDs to consider in their self-explanation.

2. How did you model each of the reaction forces?
3. When computing moments for the equilibrium, why did you choose the particular point that you used to compute moments about? For example, if you computed moments about A, why did you pick A and not some other point? Could you have simplified the analysis by picking some other point to take moments about?
4. Why did you choose to solve the equilibrium equations in the order that you did? For example, if you solved the x-equilibrium equation, then the y-equilibrium

equation, and finally the moment equilibrium equation, why did you chose this order?

3.6.2.2 Homework Four Self-Explanation Prompts

G_3 received the following prompts for each problem on homework four:

1. Why did you select the system that you used for your free-body diagram? Could you have selected some other system and still solved the problem?
2. How did you model each of the reaction forces?
3. When computing moments for the equilibrium equation, why did you choose the particular points or axes that you used to compute moments about? Could you have simplified the analysis by picking some other point or axis to take moments about?
4. If you solved this problem with the scalar approach, why did you choose to solve the equilibrium equations in the order that you did? For example, if you first solved the x-equilibrium equation and then solved the moment equilibrium equation about the z-axis, why did you choose this order?

Below are the self-explanation prompts for problem two of homework four which were given to G_5 and G_6 . The second and third prompts are scaffolded.

1. Why did you select the system that you used for your free-body diagram? Could you have selected some other system and still solved the problem?

2. Reaction forces can be modeled as contact with smooth or rough surfaces, ball-and-socket joints, fixed connections, thrust or non-thrust bearings, or flexible elements.
How did you model the reaction forces at A, B, and C? Why?
3. If a moment is computed about point A, which unknowns will remain? If a moment is computed about the z-axis, which unknowns will remain? Which points or axes did you chose to take moments about for your equilibrium equations? Why?
4. If you solved this problem with the scalar approach, why did you choose to solve the equilibrium equations in the order that you did? For example, if you first solved the x-equilibrium equation and then solved the moment equilibrium equation about the z-axis, why did you choose this order?

Below are the self-explanation prompts for problem three of homework four which were given to G_4 and G_6 . The second and third prompts are scaffolded.

1. Why did you select the system that you used for your free-body diagram? Could you have selected some other system and still solved the problem?
2. Reaction forces can be modeled as contact with smooth or rough surfaces, ball-and-socket joints, fixed connections, thrust or non-thrust bearings, or flexible elements.
How did you model the reaction forces at A, C, and D? Why?
3. If a moment is computed about point A, which unknowns will remain? If a moment is computed about the x-axis, which unknowns will remain? Which points or axes did you chose to take moments about for your equilibrium equation? Why?
4. If you solved this problem with the scalar approach, why did you choose to solve the

equilibrium equations in the order that you did? For example, if you first solved the x-equilibrium equation and then solved the moment equilibrium equation about the z-axis, why did you choose this order?

Below are the self-explanation prompts for problem five of homework four which were given to G_5 and G_6 . The second and third prompts are scaffolded.

1. Why did you select the system that you used for your free-body diagram? Could you have selected some other system and still solved the problem?
2. Reaction forces can be modeled as contact with smooth or rough surfaces, ball-and-socket joints, fixed connections, thrust or non-thrust bearings, or flexible elements. How did you model the reaction forces at A, B, and C? Why?
3. If a moment is computed about C, which unknowns will remain? If a moment is computed about the z-axis, which unknowns will remain? Which points or axes did you chose to take moments about for your equilibrium equations? Why?
4. If you solved this problem with the scalar approach, why did you choose to solve the equilibrium equations in the order that you did? For example, if you first solved the x-equilibrium equation and then solved the moment equilibrium equation about the z-axis, why did you choose this order?

Below are the self-explanation prompts for problem six of homework four which were given to G_4 and G_6 . The second and third prompts are scaffolded.

1. Why did you select the system that you used for your free-body diagram? Could you have selected some other system and still solved the problem?

2. Reaction forces can be modeled as contact with smooth or rough surfaces, ball-and-socket joints, fixed connections, thrust or non-thrust bearings, or flexible elements.
How did you model the reaction forces at O and A? Why?
3. If a moment is computed about point A, which unknowns will remain? If a moment is computed about point O, which unknowns will remain? What points or axes did you chose to take moments about for your equilibrium equations? Why?
4. Why did you choose to solve the equilibrium equations in the order that you did? For example, if you solved the x-equilibrium equation, then the y-equilibrium equation, and finally the moment equilibrium equation, why did you choose this order?

3.6.2.3 Homework Five Self-Explanation Prompts

Below are the self-explanation prompts given to G_3 and G_2 for each problem on homework five. None of the prompts are scaffolded.

1. Why did you decompose the system into the particular set of free body diagrams that you used?
2. Which bodies are two-force members?
3. Why did you analyze the free body diagrams in the order that you did?
4. When computing moments for your first moment equilibrium equation, why did you choose the particular point that you used to compute moments about?

Below are the self-explanation prompts for problem three of homework five which were given to G_4 and G_6 . The second and third prompts are scaffolded.

1. Why did you decompose the system into the particular set of free body diagrams that you used?
2. Below are the components that comprise the pipe cutter. How many forces act on Component A? How many forces act on Component B? How many forces act on Component C? How many forces act on Component D? In your solution, which of these components, if any, did you model as a two-force member?

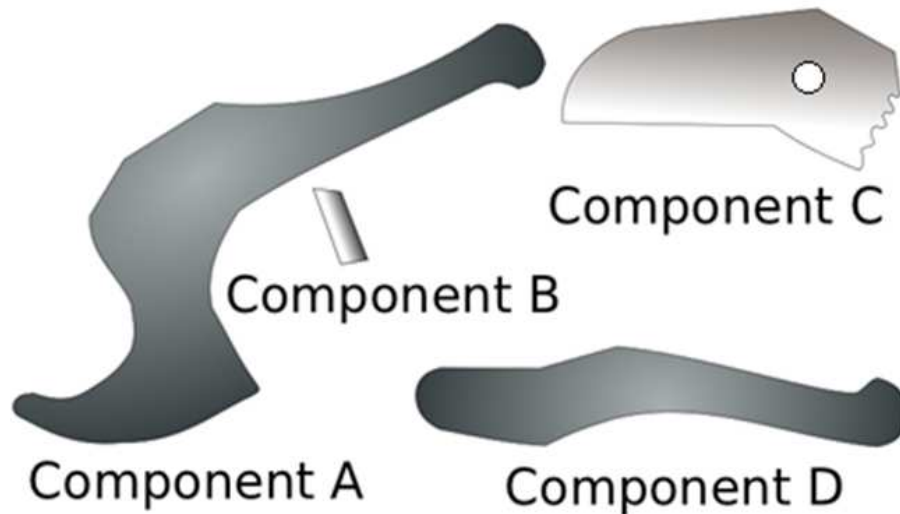


Figure 3.44: Self-explanation prompt for homework five problem three. Students were shown the individual components comprising a device and were asked to indicate which were two-force members.

3. Why did you analyze the free body diagrams in the order that you did?
4. When computing moments for your first moment equilibrium equation, why did you choose the particular point that you used to compute moments about?

Below are the self-explanation prompts for problem six of homework five which were given to G_4 and G_6 . The second and third prompts are scaffolded.

1. Why did you decompose the system into the particular set of free body diagrams that you used?
2. Below are the components that comprise the device in problem 6. How many forces act on Component A? How many forces act on Component B? How many forces act on Component C? In your solution, which of these components, if any, did you model as a two-force member?

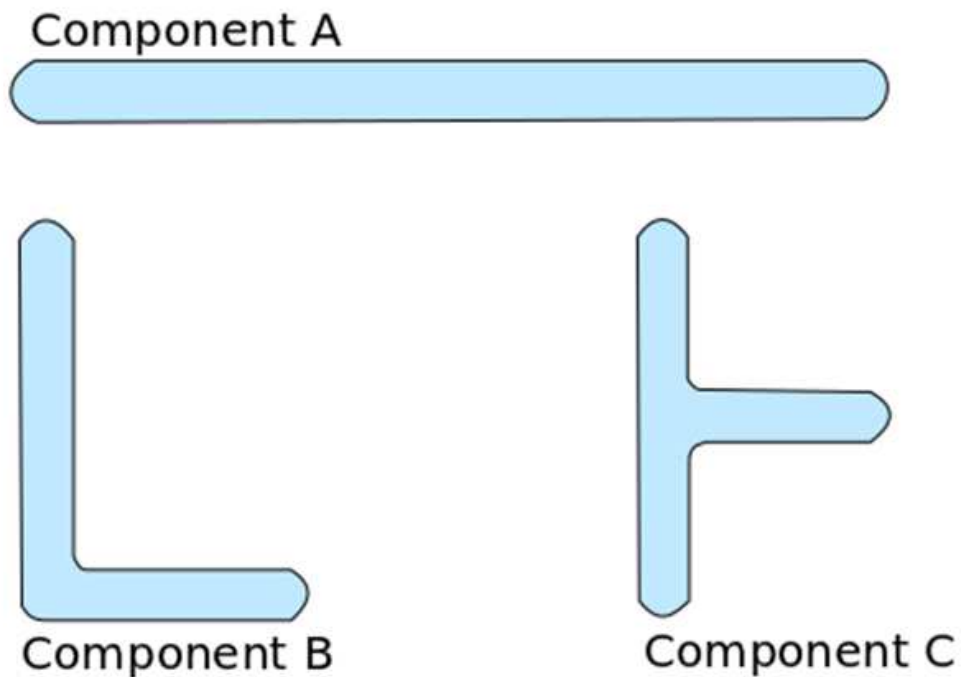


Figure 3.45: Self-explanation prompt for homework five problem six. Students were shown the individual components comprising a device and were asked to indicate which were two-force members.

3. Why did you analyze the free body diagrams in the order that you did?

4. When computing moments for your first moment equilibrium equation, why did you choose the particular point that you used to compute moments about?

Self-explanation prompts for problem two given to G_5 and G_6 : Below are the self-explanation prompts for problem two of homework five which were given to G_5 and G_6 . The second and third prompts are scaffolded.

1. Why did you decompose the system into the particular set of free body diagrams that you used?
2. Below are the components that comprise the vise. How many forces act on Component A? How many forces act on Component B? How many forces act on Component C? How many forces act on Component D? In your solution, which of these components, if any, did you model as a two-force member?
3. Why did you analyze the free body diagrams in the order that you did?
4. When computing moments for your first moment equilibrium equation, why did you choose the particular point that you used to compute moments about?

Self-explanation prompts for problem 5 given to G_5 and G_6 . None of the prompts are scaffolded.

1. Why did you decompose the system into the particular set of free body diagrams that you used?
2. Below are the components that comprise the log hoist. How many forces act on Component A? How many forces act on Component B? How many forces act on

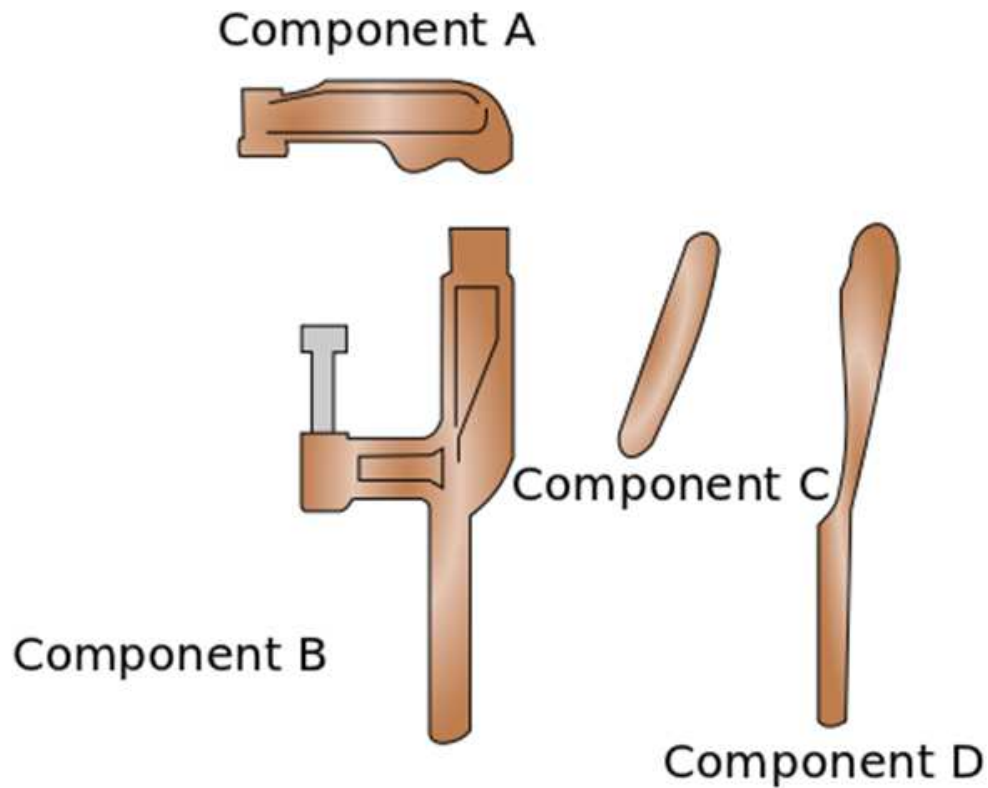


Figure 3.46: Self-explanation prompt for homework five problem two. Students were shown the components comprising a device and were asked to indicate which were two-force members.

Component C? How many forces act on Component D? How many forces act on Component E? In your solution, which of these components, if any, did you model as a two-force member?

3. Why did you analyze the free body diagrams in the order that you did?
4. When computing moments for your first moment equilibrium equation, why did you choose the particular point that you used to compute moments about?

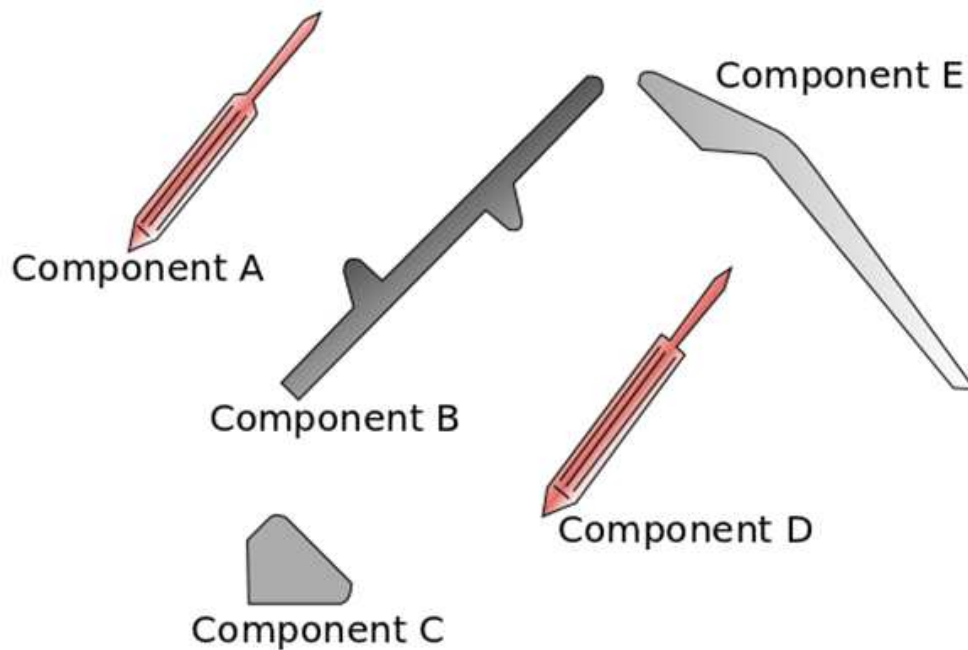


Figure 3.47: Self-explanation prompt for homework five problem five. Students were shown the components comprising a device and were asked to indicate which were two-force members.

3.6.2.4 Homework Eight Self-Explanation Prompts

G_2 , G_3 , G_4 , G_5 , and G_6 were all given the following prompts for all problems on homework assignment eight. None of these prompts were scaffolded.

1. To begin your solution, you must make several assumptions. Which surfaces, if any, did you assume were on the verge of slip? Why did you think these surfaces were on the verge of slip?
2. If slip did occur, which direction would each slipping member move?
3. For problem 3 only: Why is the device self-locking?

Chapter 4

Open IE and Self-Explanation

Transcripts

4.1 Introduction

Research has demonstrated that self-explanation hones student's metacognitive skills and increases student performance. We have found, however, that not all self-explanation is substantive. Our goal, in this chapter, is to develop computational techniques capable of determining if a student's explanation is relevant or not. This will enable, for example, a system to automatically process students' handwritten self-explanations and present to the instructor those that are and are not relevant, providing the instructor with an opportunity to rapidly address deficiencies in students' understanding. Additionally, this technique may enable an interactive tutoring system to prompt students to continue their explanations when necessary. This is a tractable task as self-explanations typically contain only a small number of possible concepts. The

language used to express concepts in each self-explanation can vary greatly, but our task is only to identify the existence of the concepts, not to perform general machine interpretation. In this chapter, we present work on the automatic understanding of students' handwritten self-explanation of their solutions to homework problems.

In this chapter¹, we examine data extracted from the 2011 course offering. In particular, we investigate the self-explanations of students from G_c as described in Section 3.6.1.

To provide a benchmark for the self-explanations, we asked three experts to solve some of the same problems and generate their own self-explanations. We manually analyzed these and identified the concepts used. We found that the experts used only a small set of concepts in their explanation of any particular problem-solving step. We would expect that a student with an expert-stance would utilize the same set of concepts in their explanations.

We employ an information extraction technique to automatically identify whether a student's self-explanation responses contain the same concepts used by the experts. For example, this technique can determine if a student assumed that bodies in a friction problem were on the verge of slip, a concept that experts often included in their self-explanations.

In our experiments, this technique has proven to be quite reliable, achieving an accuracy of up to 97% on a particular explanation. This level of accuracy can be attributed to the regular nature of the students' self-explanation. Furthermore this high-level of accuracy suggests that it may be feasible to develop automated systems to

¹The work presented in this chapter has been published and appears in [55]

elicit meaningful self-explanations from students.

4.2 Open Information Extraction Algorithm

For our analysis, we implemented the open IE algorithm developed by Soderland et al. [113]. This technique learns a set of rules which maps self-explanations to content types. These rules comprise constraints on the existence of words in the self-explanations and the locations of those words. If the correct set of words exists in the correct locations, the rule assumes that a particular concept has been expressed.

This technique begins by using the TextRunner software package [35] to extract all noun phrases present in each self-explanation sentence. Noun phrases take the form of a tuple, $(arg_1, pred, arg_2)$, where arg_1 is the subject, $pred$ is the predicate, and arg_2 is the object.

A variety of words can be used to express the same concepts. For example, “on the verge of slip” and “impending slip” have the same meaning. To accommodate these sorts of variations, Soderland’s algorithm relies on lists of synonymous words. More precisely, it requires the identification of word classes, and the enumeration of the words within those classes. Table 4.1 lists the word classes we use in our analysis. For example, the *variable* class contains the various words that are frequently used to describe the unknown forces to be computed in a statics problem. These words include “variable,” “force,” “unknown,” and “component.” Note that identifying the existence of a concept in a self-explanation is more complex than simply identifying the existence of specific words. The relationships between those words is essential.

HWK-Prompt	Class Name	Words in Class
3-1,3-4	Eliminate	cancel eliminate, rid ignore, took out, avoid
3-1	Contains	contains, touches
3-1	Need	need, require, necessary
3-1, 3-4	Variable	variable, force, unknown, component
3-1	Only	only
3-4	Only Unknown	only unknown, only one
3-4	Direct	direct, one step
3-4	Solve	solve, give, gave
8-3	Assumption	assume, occur, think, assumption
8-3	Slip	slip
8-3	FBD Component	block, point, arm, crate brake, surface, member, box
8-3	Negative	wasn't, isn't, didn't, not
8-3	Verge	verge, impend, about

Table 4.1: The word classes and the words they contain for the self-explanation prompts for the problems in homework assignments three and eight.

A rule learning process is used to learn these relationships. The rules attempt to infer the commonalities between different expressions of the same concept. Initially, the technique creates an overly-specified rule for each tuple. The rule, in effect, assumes that for another tuple to have the same meaning, it must have the same words in the same order. More precisely, the rule contains a constraint for every word class and preposition found in both that tuple and the sentence that contains it. The constraints govern both the existence and locations of those words. This technique recognizes five possible locations for word classes and prepositions: arg_1 , $pred$, arg_2 , the portion of the sentence preceding the tuple, and the portion proceeding. Each overly-specific rule will likely match only a few other tuples in the training data, if any. To find a more accurate rule, the technique repeatedly relaxes constraints so that the rule has higher precision in identifying the concept. Here, precision is defined as:

$$\text{precision} = \frac{\text{true positives}}{\text{true positives} + \text{false positives}} \quad (4.1)$$

where “true positives” are tuples that were correctly identified, and “false positives” are tuples that were incorrectly classified as this concept.

A beam search is used to find the most precise version of a rule. This search begins by dropping constraints from the overly-specified rule, one at a time, and computing the precision of each resulting, relaxed rule over the training set. The k most precise rules are kept, where k is called the beam width. The process repeats for each of these k relaxed rules. The process ultimately terminates when an empty rule is reached. In our implementation, we use a beam width of 10.

4.3 Results and Discussion

We performed leave-one-out cross-validation to train and test this technique. In each fold of cross-validation, the data from one subject (either an expert or a student) is selected for testing, and the data from the other subjects is used for training. In this way, the data used to train and test the system are never the same as each other.

Table 4.2 shows the accuracy results for identifying concepts for prompt one of homework three. For example, the technique correctly identified 30 self-explanations that expressed the *needed-forces* concept, and incorrectly identified seven other self-explanations that also expressed this concept. Thus, the technique achieved 81.1% accuracy at identifying this concept. Overall, the technique achieved 75.9% accuracy at identifying concepts used in the expert’s self-explanations. Similarly, the technique

Concept	Correct	Incorrect	Accuracy
Needed Forces	30	7	81.1%
Only One	9	4	69.2%
Least Forces	2	1	66.7%
Alternative Difficult	0	1	0.0%
All Concepts	41	13	75.9%
None	52	22	70.2%

Table 4.2: Accuracy of concept recognition in self-explanations for prompt four of homework three. The “All Concepts” row of the table contains the overall accuracy for identifying self-explanations that contain a concept used in the expert’s self-explanations. The “None” row is the accuracy for identifying self-explanations that contained none of the concepts used in the expert’s self-explanations.

achieved 70.2% accuracy at identifying self-explanations that contained none of the concepts used in the expert’s self-explanations.

Table 4.3 contains the accuracy results for identifying concepts for prompt four of homework three. Overall, the technique achieved 87.7% accuracy at identifying concepts used in the expert’s self-explanations. Similarly, the technique achieved 68.6% accuracy at identifying self-explanations that contained none of the concepts used in the expert’s self-explanations. Finally, Table 4.4 contains the accuracy results for identifying concepts for prompt three of homework eight. Overall, the technique achieved 84.2% accuracy at identifying concepts used in the expert’s self-explanations. Similarly, the technique achieved 97.3% accuracy at identifying self-explanations that contained none of the concepts used in the expert’s self-explanations.

This technique does perform better with more training data. For example, in Table 4.2, there were numerous examples of the *needed-forces* concept, and only one for the *alternative-difficult* concept. The program performed accurately on the former and

Concept	Correct	Incorrect	Accuracy
Directly Solves	4	2	66.7%
Only Unknown	6	2	75.0%
Eliminate Forces	40	1	97.6%
All Concepts	50	7	87.7%
None	70	32	68.6%

Table 4.3: Accuracy of concept recognition in self-explanations for prompt four of homework three. The “All Concepts” row of the table contains the overall accuracy for identifying self-explanations that contain a concept used in the expert’s self-explanations. The “None” row is the accuracy for identifying self-explanations that contained none of the concepts used in the expert’s self-explanations.

Concept	Correct	Incorrect	Accuracy
Slip	11	7	61.1%
No-slip	6	6	50.0%
Verge	63	2	96.9%
All Concepts	80	15	84.2%
None	36	1	97.3%

Table 4.4: Accuracy of concept recognition in self-explanations for prompt three of homework eight. The “All Concepts” row of the table contains the overall accuracy for identifying self-explanations that contain a concept used in the expert’s self-explanations. The “None” row is the accuracy for identifying self-explanations that contained none of the concepts used in the expert’s self-explanations.

poorly on the latter.

This technique currently works with manual transcriptions. In order for this technique to work in a completely automated system the manual transcriptions will need to be replaced with automatically recognized handwriting made possible by techniques such as the image-based recognizer[72] or the dollar recognizer [129]. Future work will need to account for the errors that may be introduced by such processes.

4.4 Conclusion

In this chapter, we have presented a technique which is able to accurately identify whether a student's self-explanation contains the same concepts used in self-explanations generated by experts. The technique correctly identified the existence of such concepts with an accuracy that ranged from 75.9% to 87.7%. Similarly, the technique correctly identified the lack of such concepts with an accuracy that ranged 68.6% to 97.3%.

This work has several applications. For instance in an intelligent system that engages students to generate meaningful self-explanations of their work, thus honing their metacognitive skills and increasing their mastery of the subject. In the future, these techniques may be implemented within an interactive tutoring system, enabling it to determine if a student has provided meaningful self-explanation. Such a system may then prompt the student to continue his explanations when necessary.

Chapter 5

N-Gram Analysis and Problem Number Sequences

5.1 Introduction

Self-explanation is the process by which a student provides, in words, a summary of their own understanding. In related work, students have been asked to generate self-explanations of the steps of a worked-out example or the rationale behind the students' own solutions to a problem. These self-explanations serve a metacognitive purpose, allowing students to evaluate and monitor their own understanding of concepts and enabling them to guide their own learning process. We demonstrate in this chapter, as numerous other studies have demonstrated, that self-explanation positively impacts student performance. Additionally, we demonstrate the positive impact self-explanation has on a student's ordinary solution process.

In this chapter¹, we compare the course performance and homework problem-solving behaviors of G_a and G_c . Because students in G_c generated self-explanation, we refer to them, in this chapter, as the SE (self-explanation) group. Similarly, because students in G_a did only generated homework solutions, and did not generate self-explanation, we refer to them as the SO (solution only) group. In addition to course performance, we present performance results on Steif's statics concept inventory [115] for the two groups.

A comparison of homework performance for the two groups demonstrates the expected result that students who generate self-explanations perform significantly better on their homework assignments than those who did not. Similarly, a comparison of the performance on the statics concept inventory showed that the experimental group had significantly greater learning gains for the fundamental statics concepts than did the control group.

While improvements in learning gains are an important result, the unique nature of our data set – time-stamped pen strokes – enables a much richer analysis of differences between the groups. In particular, it enables us to examine the process by which each student completes the problems in an assignment. In this chapter, we examine the order in which students solve the problems in each assignment.

As mentioned in the previous chapter, to ground our analysis, we asked three experts to use digital pens to complete some of the same homework assignments our experimental and control group students completed. We then used statistical analysis techniques to compare the work from the control and experimental groups to that of the

¹The work presented in this chapter has been published and appears in [56]

experts. This unique form of educational informatics is enabled by our novel database of student work. This analysis revealed that students who generated self-explanations solved problems more like the experts than did the students in the control group.

5.2 Analysis

Our first two analyses consider student performance on both homework grades and Steif's concept inventory [115]. We performed repeated measures analysis of variance (ANOVA) on both the homework and concept inventory performance. A number of the students had incomplete records. Some failed to complete all of the homework assignments and some failed to complete both the pre- and posttest concept inventory. As a remedy, we performed a missing values analysis to estimate missing homework and concept inventory scores. This technique estimates missing values by regressing on known values. The ANOVA results presented in both the Concept Inventory Analysis and Grade Analysis sections were done using these estimated missing values.

Our second two analyses consider the order in which students completed the problems in an assignment. We found that the experts always solved the problems in sequential order. Consequently, our analysis examines the extent to which the students solve problems out of order, an indication of novice behavior. To perform this analysis, we employ two sequence analysis methods commonly used in such disciplines as natural language processing and bioinformatics[18, 117, 107].

5.2.1 Concept Inventory Analysis

The average number of questions students correctly answered on the concept inventory is shown in Figure 5.1. ANOVA revealed that the difference in the pre- to posttest learning gains between the two groups is significant ($p = 0.011$).

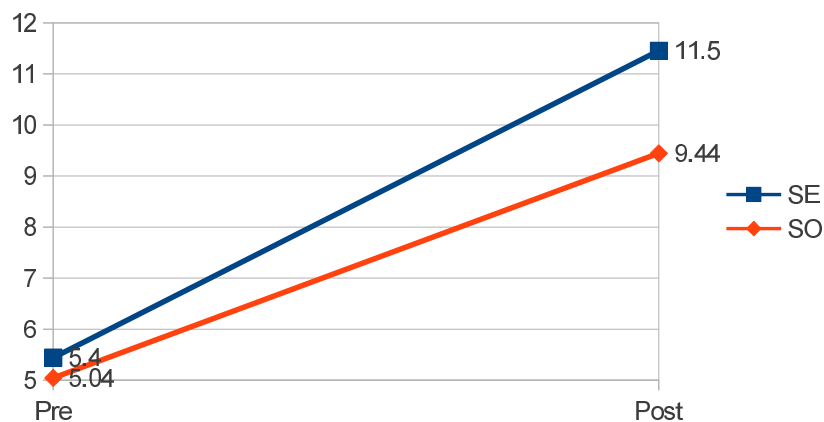


Figure 5.1: Pre/posttest scores for the SE (blue) and SO (orange) groups on the concept inventory for the 2012 course offering.

5.2.2 Grade Analysis

Next, we compared the homework performance of the two groups. The average scores for each homework assignment are shown in Figure 5.2. It is important to note that in this course, the grade on a homework assignment was determined by the performance on one problem – the other problems were not graded. ANOVA revealed that the differences between the homework grades of the two groups is significant. More specifically, there is a significant difference ($p < 0.01$) in the slopes of the linear best fit

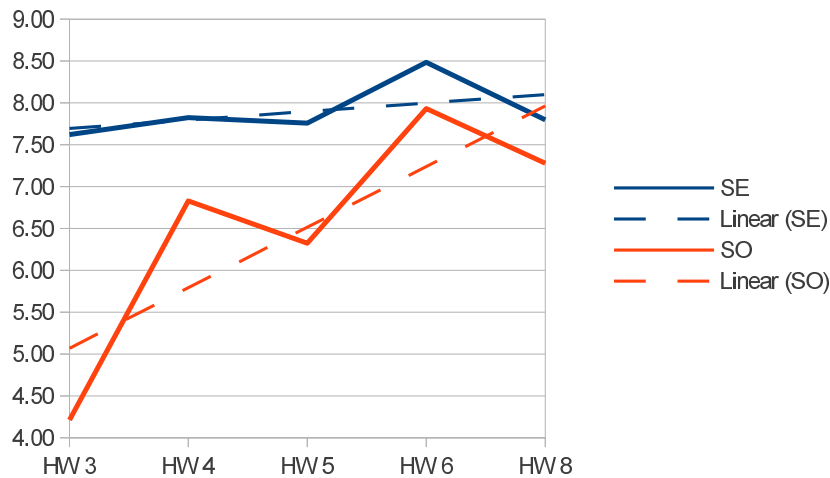


Figure 5.2: Average score for each homework assignment of both the SE (blue) and SO (orange) groups. The dashed lines represent the linear best fit of each. ANOVA revealed a significant difference in the slope of these two lines.

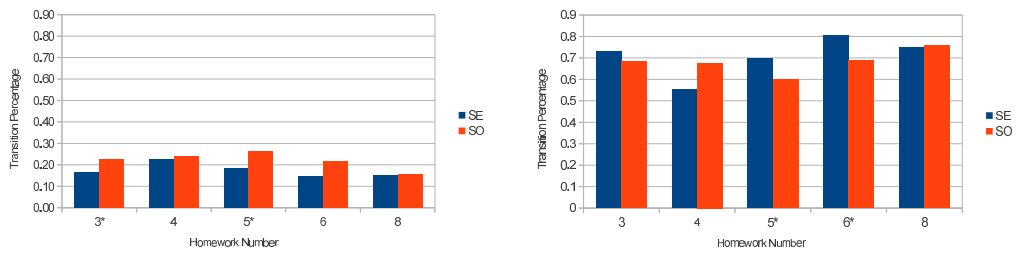
lines for the homework scores as shown in Figure 5.2.

5.2.3 Transition Probability Analysis

To examine the order in which students solve the problems in an assignment, we represent their work as a sequence of problem numbers. For example, if a student begins with problem one, moves on to problem two, and then returns to problem one before working on problem three, the sequence would be “(1, 2, 1, 3)”. In this example, there are two out-of-order problem transitions: 2-1 and 1-3. In our analysis, we consider three types of transitions: *in-order* – a transition to the immediately next problem, such as from 3 to 4; *skip* – a transition to a future problem, such as from 3 to 5; and *backtrack* – a transition to any earlier problem, such as from 3 to 2.

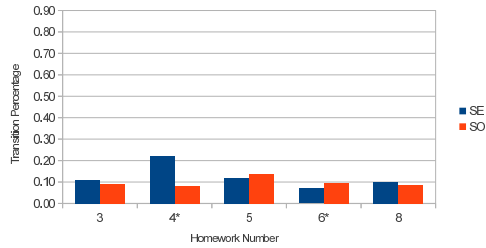
We compute the occurrences of each of these kinds of transitions for each

student and normalize by the total number of transitions, yielding a transition frequency. Figure 5.3 shows the average transition frequencies on each homework assignment for both the SE and SO students. We used a t-test to determine if the differences between transition frequencies for the two groups are significant. The problems for which the differences are significant are indicated in the figure.



(a) Backtrack

(b) In-Order



(c) Skip

Figure 5.3: Average transition frequencies for the SE and SO groups for each homework assignment. An asterisk (*) next to the homework number indicates that the difference between the two groups is significant ($p < 0.1$) as determined by a t-test.

5.2.4 N-gram Analysis

Whereas the previous analysis considered the frequency of out-of-order transitions, here we consider the frequency of transitions between particular problems, using

an N -gram analysis. An N -gram is a subsequence of length N taken from a student’s problem number sequence. For example, a two-gram or bigram is a subsequence of length two and a three-gram or trigram is a subsequence of length three. Consider, for instance, the student problem number sequence (1, 2, 3, 2, 4) which contains 4 bigrams: (1, 2), (2, 3), (3, 2), and (2, 4), and 3 trigrams: (1, 2, 3), (2, 3, 2), and (3, 2, 4).

Our analysis focuses on the likelihood that the elements in the N -gram occur together. First we consider Dice’s coefficient, which is defined only for bigrams. Consider two sets of bigrams: the set of bigrams in which a particular problem number, p_1 , is the first element of each bigram and another set of bigrams in which some other problem number, p_2 , is the second element of each bigram. Dice’s coefficient provides a measure of “similarity” for these two sets, computed as:

$$S = \frac{2|X \cup Y|}{|X| + |Y|} \quad (5.1)$$

Here, $|X|$ is the number of times some problem, p_1 , appears as the first element in a bigram, $|Y|$ is the number of times a second problem number, p_2 , appears as the second element in a bigram, and $|X \cup Y|$ is the number of times the two problems appear in the same bigram, (p_1, p_2) . S is a number between 0 and 1; the closer it is to unity, the greater the similarity between the two sets, or in other words, the more likely it is that the two problem numbers appear together.

We created two corpora, one containing every problem sequence from the SE group and one with all sequences from the SO group. We computed Dice’s coefficient for each bigram in these corpora separately. The differences between the Dice’s coefficients of the SE and SO groups are shown in Figure 5.4. Here, we compare the Dice coefficients

only for bigrams whose problem numbers were out of order. A negative value indicates that the problem numbers in a bigram appear more frequently together in the SO corpus than in the SE corpus. Note that we do not present results for bigrams that occur fewer than five times as Dice’s coefficients for such cases would be unreliable.

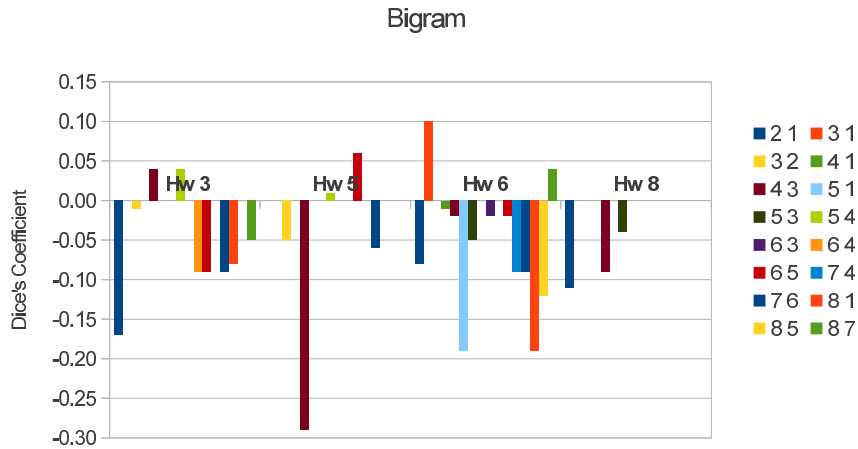


Figure 5.4: The difference in Dice’s coefficients (SE - SO) for bigrams occurring more than five times. Negative values indicate bigrams that are more similar in the SO than in the SE corpus.

Because Dice’s coefficient is limited to bigrams, we consider mutual information when examining trigrams. This measures the statistical dependence of elements in a trigram. Mutual information can be thought of as the difference between the marginal entropy of a random variable, and the conditional entropy of that variable given a second:

$$I(X; Y) = H(X) - H(X|Y) \tag{5.2}$$

$H(X)$ represents the uncertainty that some problem number, p_1 , appears as the first element in any bigram and $H(X|Y)$ represents the uncertainty that, p_1 is the

first element of a bigram given the occurrence of some particular second element of that bigram, say p_2 . In this case, $I(X;Y)$ gives an indication of the amount of information gained from knowing that p_2 follows p_1 . This calculation naturally generalizes to three variables, in which it is a measure of the amount of information gained about p_1 knowing that p_2 and some other problem, p_3 , occur after it. $I(X;Y)$ is also a number between 0 and 1 such that the closer it is to unity, the more those problems are dependent on each other.

We compute mutual information separately for the SE and SO corpora. The differences in mutual information between the two corpora for each trigram are shown in Figure 5.5. Trigrams which appear fewer than five times do not appear in these results. A negative value in Figure 5.5 indicates that a sequence of problem numbers appears more frequently in the SO corpus than in the SE corpus.

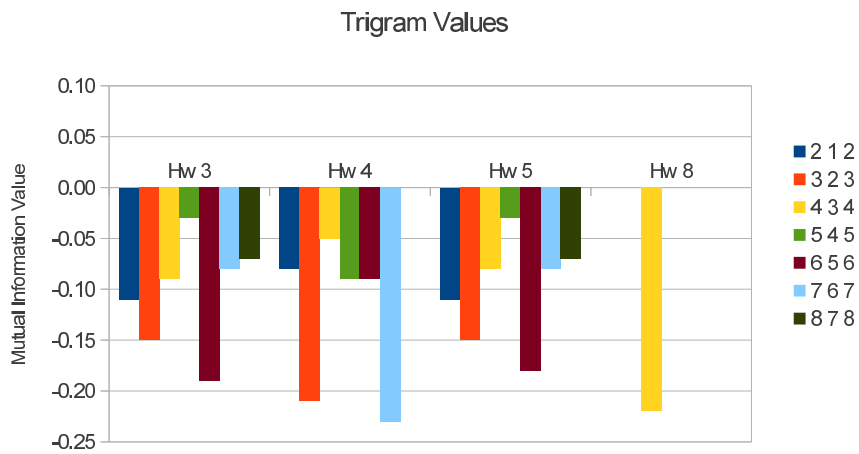


Figure 5.5: Difference in mutual information values between problem sequence trigrams of the SE and SO groups.

5.3 Discussion

Figure 5.1 shows that students in the SE and SO groups begin with equivalent conceptual understanding of statics concepts. The SE and SO groups on average got 5.4 and 5.04 questions right on the inventory pretest, respectively. At the end of the quarter, students from the SE group on average correctly answered 11.5 questions, while the SO group correctly answered on average 9.44. The difference in learning gains between the two groups is significant as revealed by ANOVA ($p = 0.011$). This evidence suggests that self-explanation may lead to greater conceptual understanding in this statics course. This corroborates intuition that self-explanation sharpens metaskills, leading to greater performance on transfer problems.

Figure 5.2 corroborates the well known story that self-explanation positively impacts performance. However, there is more to this story. Figure 5.2 shows that self-explanation can lead to a large boost in performance, but as time goes on, that difference dwindles to an insignificant level. The difference in the grades between the two groups is significant ($p < 0.1$) in the first three homework assignments, but not in the last two. This suggests that there may be a ceiling on student performance and that self-explanation, and the metaskills that it fosters, lead students to reach that ceiling quicker. This is an important benefit of self-explanation, especially in the context of a fast paced quarter system.

Our analysis of problem number sequences suggests that self-explanation can more quickly lead students to an expert-stance in the way that they solve problems. Both the transition, bigram, and trigram analyses show that students in the SE group typically

solve problems out of order less frequently than the SO students. As indicated in Figure 5.3, whenever there is a significant difference between the transition frequencies of the two groups, SE students transition out of order less frequently. Similarly, Figure 5.4 shows that specific out-of-order bigrams appear more frequently for SO students than for SE student. Figure 5.5 presents perhaps the clearest distinction between the two groups; SO students always had more out-of-order trigrams than the SE students.

One possible explanation for this behavior may be related to the fact that the homework assignments contained groups of three consecutive problems that differed only superficially. It is possible that the SO students gained insights on subsequent problems, enabling them to revisit earlier problems in the set to correct their work. The SE students, on the other hand, may have better understood their work on the first attempt of a problem, making revisits to prior work unnecessary.

5.4 Conclusion

In this chapter we have compared, through a variety of analyses, the differences in performance and solution processes of students who did and did not generate self-explanation with their homework solutions.

In our first analysis, we compared the homework grades and pre- and posttest scores on the statics concept inventory. We found that students who generated self-explanation performed better on both the homework and the concept inventory posttest. This is an expected result, as prior work has demonstrated similar effects.

We used sequential analysis techniques commonly used in bioinformatics and

natural language processing to discover meaningful patterns in the order in which students solved homework problems. In doing so, we have demonstrated that self-explanation not only leads to greater performance gains, it also leads students to solve homework problems more like an expert.

Chapter 6

Differential Sequence Mining and Action Sequences

6.1 Introduction

In this chapter¹, we investigate a novel representation of a student's handwritten assignment which characterizes the sequence of actions the student took to solve each problem. This representation comprises an alphabet of canonical actions that a student may perform when solving a homework assignment. Each action is characterized by its duration, problem number, and semantic content. This representation allows us to apply traditional data mining techniques to our database of students' handwritten homework solutions.

Our analyses focus on two separate groups of students from the 2012 course offering, those who scored in the top third of the class on exams, and those who scored

¹The work presented in this chapter has been published and appears in [51]

in the bottom third. We applied a differential data mining technique to the sequences of each of these groups and identified behaviors that are more frequently exhibited by one group than the other.

These patterns serve as the basis for a number of numeric features used to train a linear regression model to predict students' performance in the course. This model achieves an R^2 of 0.34. More importantly, the underlying parameters of this model provide valuable insights as to which of the patterns most correlate with performance. Using these patterns, we are able to gain insight into the high-level, cognitive behaviors exhibited by the students.

6.2 Action Sequences

In this section, we describe how each sketch may be transformed into an *action sequence*, comprising discrete actions, that is suitable for differential pattern mining. Each action is an element of a predefined alphabet of canonical actions. Each element in the alphabet represents an uninterrupted period of problem-solving performed by a student as he or she solves a homework assignment. We seek to characterize the duration, semantic content, and homework problem number for each action.

We begin by segmenting the pen strokes of each sketch by semantic type. To do so, we simply identify each index, i , in L such that $l_i \neq l_{i+1}$, and segment the series of pen strokes at each identified index. Each resulting segment contains a sequence of actions corresponding to the same semantic type. The resulting segments do not yet satisfy the above definition of an action, as they do not necessarily contain uninterrupted work.

Thus, we further segment the sketch at each index, j , such that the difference between the start time of s_j and the end time of s_{j-1} is greater than a specified threshold. In this chapter, we use a threshold of five minutes, which was determined *a priori*; five minutes is a sufficiently large gap to be considered an intended break in the problem-solving process.

Each segment is then labeled with an element from the alphabet of canonical actions. If the segment comprises cross-out pen strokes, then it is given the cross-out label, C , regardless of its length or problem number. The remaining groups are labeled with a triple, $\{P, T, D\}$, where P represents the problem number, T represents the semantic type, and D , represents the duration of the action. $P \in \{1, \dots, 8\}$ as there are never more than eight problems on a given homework assignment. $T \in \{F, E\}$ where F represents a FBD action and E represents an equation action. Lastly, $D \in \{S, M, L\}$, where S , M , and L indicate an action of small, medium, or large duration respectively. Take for example, the label $\langle 1-E-S \rangle$. This indicates a small action on the equations from problem one of an assignment.

The cut-off points for each duration category were determined by studying the distribution of lengths of all the FBD and equation actions. Figures 6.1 and 6.2 show a histogram for the duration of FBD and equation actions respectively. We partition each distribution into three segments such that the area under the curve for each segment is equal. The resulting thresholds are 11.26 and 80.1 seconds for FBDs and 29.59 and 147.82 seconds for equations. There are 49 unique labels in the canonical action alphabet, comprising the 48 possible combinations for a given triple and the additional

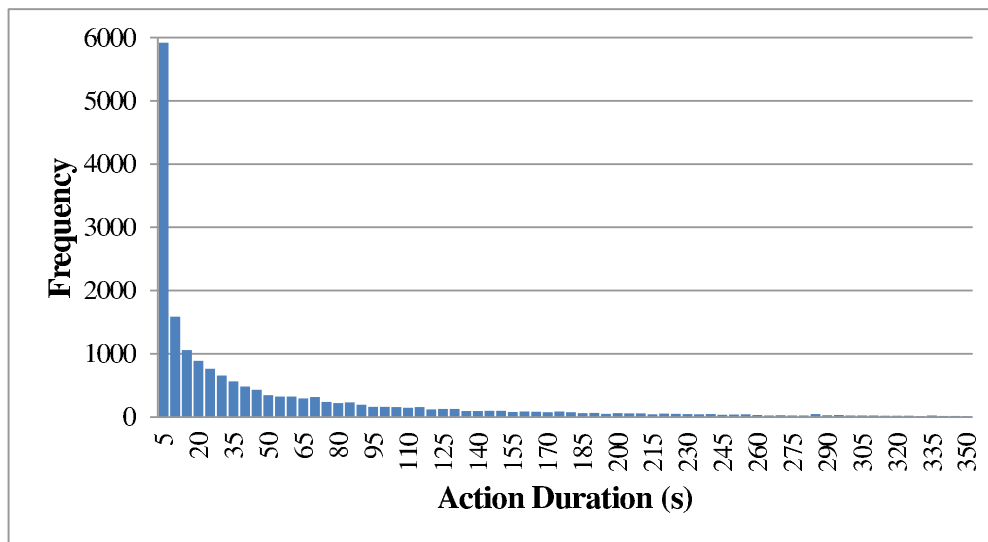


Figure 6.1: A histogram of the durations of FBD actions across all homework assignments. For example, the first (leftmost) bar indicates that approximately 6,000 FBD actions were between zero and five seconds long.

cross-out label.

We seek to assign the action sequences of a student to a *performance group* based on that student’s performance. In particular, we group a student’s action sequence for an assignment by that student’s performance on the most relevant exam, which we defined as the one that occurred most recently after that assignment was due.

Students completed homework assignments three and four prior to the first midterm exam. Students completed homework assignments five and six after the first midterm exam and before the second. Students completed homework assignment eight after the second midterm exam and before the final exam. Each midterm exam only comprised problems similar to those encountered on the homework assignments leading up to it. Thus the first midterm exam required that students solve problems similar to those found on homework assignment three and four and the second midterm exam

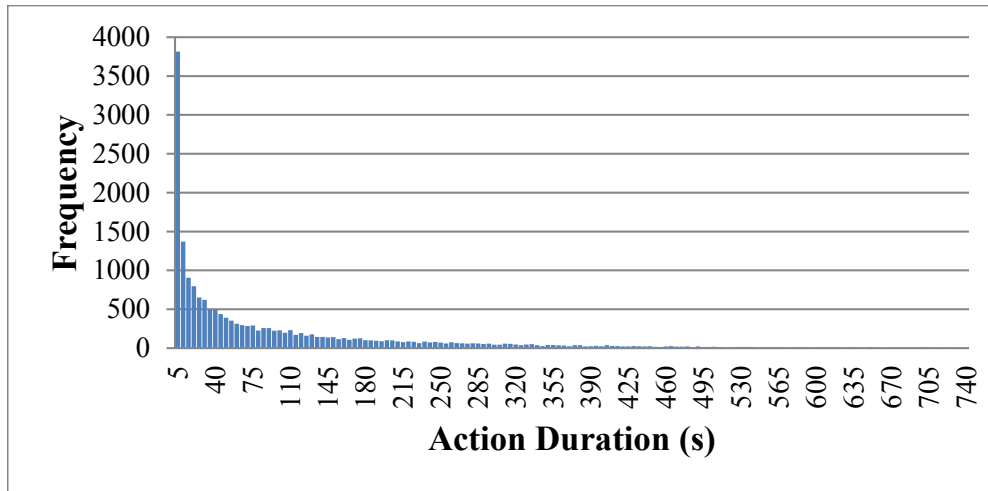


Figure 6.2: A histogram of the durations of EQN actions across all homework assignments. For example, the first (leftmost) bar indicates that approximately 3,500 EQN actions were between zero and five seconds long.

required students to solve problems similar to those found on homework assignments five and six. The final exam comprised problems similar to all those encountered on all homework assignments.

Using this schedule of exams and homework assignments, we assign each action sequence to a group based on performance. An action sequence is assigned to the top-performing group if the student who performed those actions scored in the top third on the relevant exam. Similarly, an action sequence is assigned to the bottom-performing group if the student scored in the the bottom third of the class. The differential mining technique employed in this chapter requires exactly two databases as input, thus the remaining middle-performing students are excluded from our analysis to help accentuate the differences in problem-solving behaviors of top- and bottom-performing students.

Descriptive statistics of the lengths of the action sequences for the two performance group for each assignment are shown in Table 6.1. It is interesting to note

Group	Average	Median	Std. Dev.	<i>t</i> -test
HW3 Bot.	89.52	88	43.30	0.21
HW3 Top	75.38	54.5	56.12	–
HW4 Bot.	130.25	128	63.97	0.00
HW4 Top	83.28	78	46.67	–
HW5 Bot.	127.88	119.5	70.89	0.01
HW5 Top	87.14	72	54.42	–
HW6 Bot.	144.73	140	52.55	0.171
HW6 Top	126.52	122	66.05	–
HW8 Bot.	82.45	72	73.94	0.17
HW8 Top	62.28	53.5	39.64	–

Table 6.1: Average, median, and standard deviation of the sequences for each grouping of sequences on each assignment. The fourth column contains the *p*-value of a *t*-test comparing the bottom-performing and top-performing groups on each assignment.

that the average action sequences of the bottom-performing group are always longer than those of the top-performing group, and in two cases this difference is significant ($p < 0.01$).

6.3 Differential Mining

To identify patterns that distinguish good performance from poor performance we employ the differential pattern mining technique developed by Kinnebrew and Biswas [75]. This algorithm identifies patterns that are differentially frequent with respect to two databases of sequences, called the left and right databases.

This algorithm uses two metrics to measure the frequency of a pattern, *s*-frequency and *i*-frequency. *s*-frequency is defined as the number of sequences in a database that contains a specific pattern. *i*-frequency is defined as the number of times a pattern appears within a single sequence. Take for example, a database of ten sequences in which the first seven sequences contain one instance of a particular pattern and the

last three sequences contain two instances of that same pattern. This pattern would then have a s -frequency of 10. This pattern would have an i -frequency of one in the first pattern and an i -frequency of two in the last pattern.

This algorithm begins by finding all patterns that meet a specified s -frequency threshold in the left and right database separately. Each such pattern is called an s -frequent pattern. A modified implementation of the SPAM algorithm [7] is used to identify the initial set of s -frequent patterns constrained by the a maximum gap between subsequent elements within a pattern. We use a maximum gap constraint of two in our study.

The i -frequency of each s -frequent pattern is computed for each sequence in each database. A separate t -test is computed for each s -frequent pattern to determine if the i -frequency values computed using the left database are significantly different from those computed for the right database. If the resulting p -value of the t -test is below a certain threshold, called the p -value threshold, it is considered to be *differentially frequent*. This algorithm identifies four types of differentially frequent patterns: those that are s -frequent in both sets but whose average i -frequency is higher in the left database; those that are s -frequent in both sets but whose average i -frequency is higher in the right database; those that are only s -frequent in the left database; and those that are only s -frequent in the right database. In this study, we consider only the sequences from the last two cases as they are the most most useful for distinguishing between good- and poor-performing students.

In our implementation, we use the set of sequences from the bottom-performing group as the left database and those from the top-performing group as the right database.

We use a s -frequency threshold of 0.6, meaning that a pattern must appear in at least 60% of the sequences in a database in order to be considered s -frequent. We use a p -value threshold of 0.1.

6.4 Performance Prediction

The differential pattern mining technique identified 98 patterns in total: 6 that were s -frequent in the top-performing group but not in the bottom-performing group, and 92 that were s -frequent in the bottom-performing group but not in the top-performing group.

Our goal is to use these 98 patterns to construct a model to distinguish between good- and poor-performing students. We represent each student with 98 binary features. Each feature indicates whether a particular differential pattern from a particular assignment is contained within a student’s action sequence for that assignment. To avoid computing a model that over-fits the data, we used the Correlation-based Feature Selection (CFS) algorithm with 10-fold cross-validation to identify the subset of the 98 features with the most predictive power. Those features that were selected in more than six of the ten folds by the CFS algorithm were included in the final feature subset. Table 6.2 shows the 20 features that were ultimately selected in this way.

We then used these 20 features to construct a linear regression model which predicts students’ overall performance in the course. While more robust, non-linear classifiers could have been used, e.g., AdaBoost [27] or Support Vector Machines [39], we use a linear regression model because of the ease of interpretation; the coefficients

HW No.	Perf. Group	Sequence
3	Top	1-E-M 1-F-S
3	Top	1-F-M 1-E-M
3	Bot	2-F-M 2-E-S
3	Bot	C 5-E-S
3	Bot	5-E-M 5-F-S
3	Bot	5-E-S 5-F-M
3	Bot	C 4-E-L
4	Bot	C 1-E-M
4	Bot	1-E-L C
4	Bot	C 5-E-L
4	Bot	1-E-M 1-E-S
4	Bot	C C
5	Bot	4-E-S 4-E-S
6	Bot	1-F-M 1-E-S 1-E-M
6	Bot	1-E-S 1-E-M 1-F-M
6	Bot	1-F-M 1-F-S
6	Bot	1-F-S 1-F-M
8	Bot	5-F-M 5-F-S
8	Bot	5-F-S 5-E-M
8	Bot	5-F-M 5-E-S

Table 6.2: Features selected using the CFS algorithm. Each feature corresponds to a pattern identified by the differential pattern mining algorithm. Each line shows the homework number and group (top or bottom) from which the pattern was identified. The final column shows the pattern that was used to compute the feature.

that comprise the model give insight into the predictive power of the features used to train it. We used the linear regression package available in the WEKA machine learning software suite [44] to train the model. Our predictive model achieves an R^2 of 0.343 and includes seven features with non-zero coefficients. Table 6.3 lists these seven features.

6.5 Discussion

We manually inspected each of the 98 patterns identified by the differential pattern mining algorithm and categorized the different types of cognitive processes they

demonstrate. We identified seven distinct categories. *Difficulty* is the category in which students seem to encounter difficulties with a particular problem, evidenced by either repeated cross-outs or repeated attempts at the same component of the same problem. For example, the pattern $\langle C, 1-E-S, C \rangle$ describes a scenario in which the student crossed out work, worked on equations for problem one for a short time, and then again crossed out work.

Three categories describe patterns in which actions are repeated: *Repeated Equation*, *Repeated FBD*, and *Repeated Cross-out*. For instance, $\langle 2-E-S 2-E-S \rangle$ is an example of a *Repeated Equation* action. Such sequences may be an indication that a student is taking a break in the middle of a particular activity to think more carefully before continuing with that activity.

Two categories describe patterns suggesting that a student may be revising either a FBD (*FBD Revision*) or an equation (*Equation revision*). These patterns comprise a cross-out followed by either the FBD or equation they are most likely revising. Also, when a student moves from working on an equation back to a FBD, this is likely an indication that the FBD is being revised; students typically attempt to complete their FBD before moving on to equations.

Lastly, is the *Normal* category. This is the category for all patterns in which a FBD is followed by an equation of the same problem number. A differential pattern belonging to the *Normal* category is particularly informative when one group exhibits significantly more normal sequences – it is an indication that the other group is solving their homework assignment out-of-order more often.

Sequence	HW	Weight	Category
1-F-S 1-F-M	3	48.8	Repeated FBD
C 5-E-L	3	51.0	EQN Revision
C 5-E-S	4	51.2	EQN Revision
1-E-M 1-F-S	4	55.5	FBD Revision
C 1-E-M	6	62.7	EQN Revision
1-F-M 1-E-S 1-E-M	6	63.1	Difficulty
5-F-M 5-F-S	8	73.1	Repeated FBD

Table 6.3: Non-zero feature coefficients for the linear regression model trained to predict student performance.

The non-zero weighted features of the linear regression model (Table 6.3) help identify the patterns which are most predictive of students' grades, and thus provide insight into the behaviors which best correlate with students' performance. In Table 6.3, Patterns 1, 4, 5, and 6 are all similar in that they comprise actions pertaining to the first problem on a homework assignment, and suggest that a student may be having difficulty or is frequently revising his or her work. This is an indication that when students encounter difficulty on the first problem, which is typically the easiest problem of the homework assignment, that they may continue to encounter those difficulties throughout the quarter.

Patterns 2, 3, and 7 in Table 6.3 are all similar in that they pertain to problems that are very similar to problems that appear on either a later midterm, the final exam, or both. (These problems differ only superficially from exam problems. For example, the geometry may be rotated.) These patterns all describe situations in which the student is revising his or her equations or FBDs. The features suggest that students who frequently revise problems which are similar to an exam problem are likely to have difficulty with those problems later on during an exam.

It would be difficult to use the linear regression model to predict performance for students of a future section in Statics. To do so would require that the instruction, assignments, and exams, be identical. This is not likely to be the case, as some of the homework problems are modified each year to prevent copying solutions from the previous offering.

Instead, the patterns and correlations discovered in this chapter may be used to guide future offerings of this course. For example, if a student's work contains patterns which indicate difficulty, similar to those found in this study, on the first problem of an assignment or on a problem which is similar to one that will appear in a future exam, the instructor can provide targeted materials for that student to address that difficulty. Furthermore, the results here indicate which problems have a strong bearing on students' performance. For example, students who seemed to have difficulty constructing a FBD on problem five of homework eight typically did not perform well in the course. This indicates to the instructors of future offerings this course, that more time should be spent in class reviewing how the FBD for this problem should be constructed.

6.6 Conclusion

In this chapter, we have presented a novel representation of students' handwritten work on an assignment which characterizes the sequence of actions the student took to solve that problem. This representation comprises an alphabet of 49 canonical actions that a student may make when solving his or her homework assignment. Each action is characterized by its duration, problem number, and semantic content. This

representation allows us for the first time, to apply traditional data mining techniques to sequences of students' handwritten problem solutions.

We assigned these sequences into top- and bottom-performing groups according to performance on each sequence's most relevant exam. The most relevant exam for a sequence from a particular homework assignment is the exam which occurs most recently after that homework assignment was due. Sequences from students who performed in the top third of the class on that assignment comprise the top-performing group and sequences from students who performed in the bottom third comprise the bottom-performing group. We applied a differential data mining technique to the sequences from the students in each of these groups and identified patterns that are more frequently exhibited by one group than the other.

These patterns serve as the basis for features used to train a linear regression model to predict students' performance in the course. This model achieves an R^2 of 0.34. Furthermore, the underlying parameters of this model provide valuable insights as to which of the patterns best correlate with performance. From these best-correlating patterns, we have manually identified high-level cognitive behaviors exhibited by the students. These behaviors provide insight as to when students may be experiencing difficulty in the course. These techniques may be applied in future sections of this course to identify when students are having difficulty in class, enabling the instructor to rapidly address those difficulties.

Chapter 7

K-Means Clustering and Solution

Bitmaps

7.1 Introduction

It has been shown in prior work that both the temporal and spatial organization of a student's solution to a homework or exam problem correlates with his or her performance on that solution. This result supports the intuition that the way in which a student organizes his or her work provides a view into the cognitive processes by which that student solved that problem.

In this chapter¹, we seek to develop taxonomy for the organization exhibited by students. The organizational categories we identify serve as a basis for examining the cognitive processes employed by students as they solve homework problems. While we could manually inspect student work to identify typical organizational patterns, such

¹The work presented in this chapter has been published and appears in [58]

an approach is prohibitively time-consuming. Also, the results may be subject to coder-bias as they would rely on a particular inspectors judgment. Instead, we employ a data-driven approach to automatically discover patterns latent in the organization of students from the 2012 course offering.

To capture the spatial organization exhibited by the students, we represent each page of a solution as a low-resolution bitmap. This process removes small variations in the students solutions, capturing a general representation of the layout of the ink on a page. We may then compute the distance between two bitmaps using the Hausdorff distance, a popular bitmap distance metric. Having a distance metric for bitmaps allows us to cluster them using the K-Means[85] clustering algorithm. This algorithm identifies groupings of bitmaps that are more similar with one another than with the bitmaps of other groups. Each of these groups represents a distinct spatial organization type. We then manually examine the pages which comprise each grouping, and describe the high-level organizational habits that are present. From these habits, we gain insights into the cognitive processes employed by students as they solve homework problems.

7.2 Solution Bitmaps

Our approach begins by converting each handwritten solution into two binary images: one containing ink for all FBDs and one for all equations. Figure 7.1 4 shows an example of a students solution and the resulting two FBD and equation bitmaps. To construct a solution bitmap, first a minimum bounding box is constructed around the entire solution. This box is divided into a 10 x 10 bitmap. Each pixel in the bitmap is

then marked with a value of either 1 or 0, indicating whether any ink exists in that pixel or not respectively. This produces a coarse-grained representation of the students work as it is effectively a down-sampling of the original sketch. This representation naturally removes minute variances between students solutions and captures the general spatial organization present in each solution.

7.3 Clustering

We have performed eight separate clustering processes, one for each problem on the final exam. In each process, all pages of solutions corresponding to the same final exam problem number are used as input to the K-Means clustering algorithm.

The clustering algorithm employs the Hausdorff distance to measure the similarity between two sets. The Hausdorff distance between two bitmaps is defined as:

$$H(A, B) = \max(h(A, B), h(B, A)) \quad (7.1)$$

where:

$$h(A, B) = \arg \max_{a \in A} (\arg \min_{b \in B} \text{distance}(a, b)) \quad (7.2)$$

is called the directed Hausdorff distance. Note that $h(A, B) \neq h(B, A)$. Here, distance is the Manhattan distance between two bitmap pixels a and b. Intuitively, the Hausdorff distance identifies the distance, d, such that each pixel in A is at most d from some pixel in B, providing a general measure of similarity between bitmaps A and B. The K-Means clustering algorithm is used to optimally group n objects into k clusters,

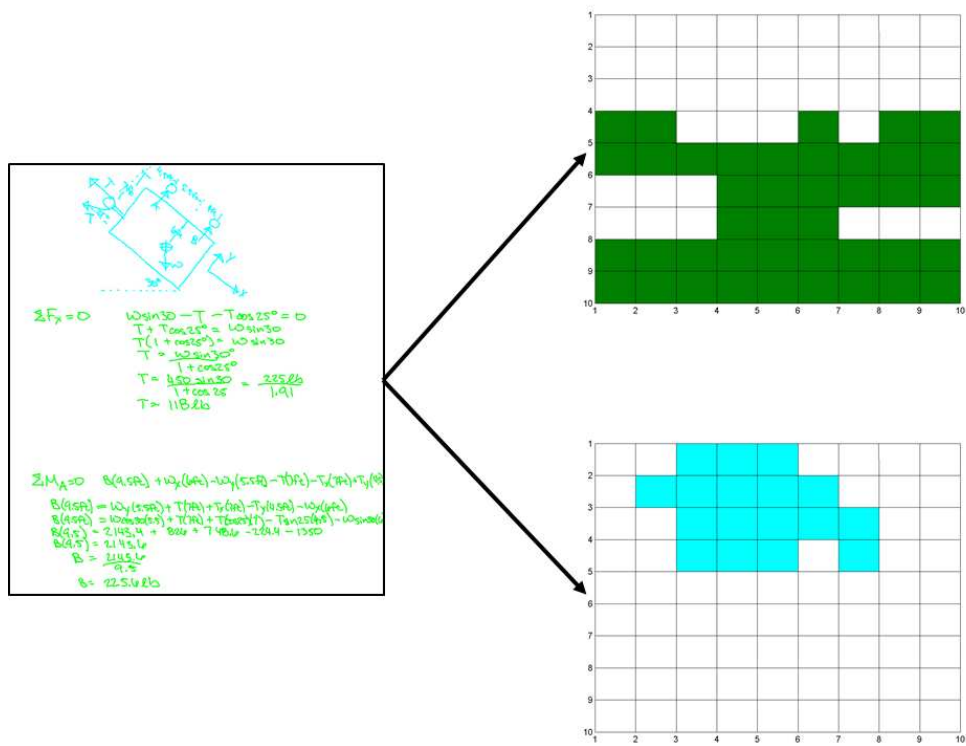


Figure 7.1: Example of a student's handwritten solution (left) and resulting equation (green) and FBD (cyan) bitmap.

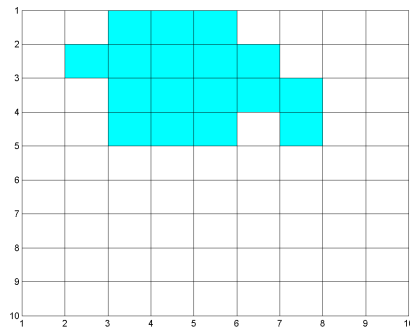
such that each object within a cluster is closer to the mean of all objects in that cluster than the mean of objects in any other cluster. Here, the objects being clustered are the solution bitmaps, and the distance between bitmaps is defined using the Hasudorff distance. The number of clusters, k , must be selected a priori. We use a value of $k = 9$ for each run of the clustering algorithm. This value proves to be sufficiently high for the data we have collected. We use the K-Means implementation of the WEKA[44] data mining software suite.

7.4 Group Analysis

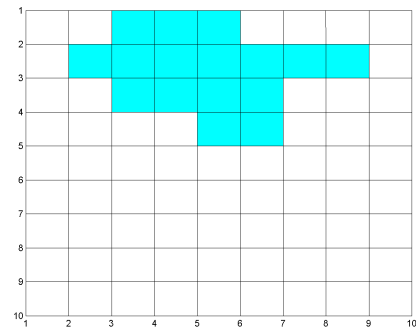
Presenting the clustering results for each grouping for each problem would prove to be intractably large. Instead we present in this section a manual analysis of the results of clustering the FBD solution bitmaps from final exam problem two and six and discuss the high-level behaviors exhibited by the typical solutions of each group. The results for these problems are sufficient to characterize the types of conclusions that can be drawn using our unsupervised analysis. While we used a cluster value, k , of nine, it is not always the case that the algorithm will identify nine meaningful groups. In some cases, there were less than nine groupings present in the data and thus some groups are empty, that is, they contain no solution bitmaps.

For problem two, the K-Means algorithm identified three non-empty groups. The first group comprised 24 sketches. The FBDs of the sketches within the first group typically were small in comparison to FBDs of other groups. Furthermore, these FBDs typically depicted a single element comprising an outline of the entire system shown

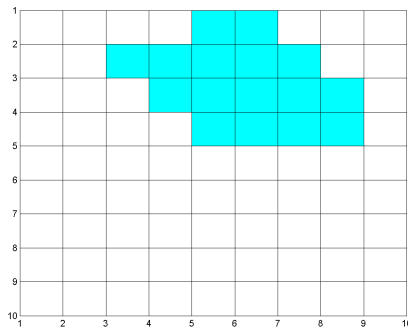
in the problem description image. Three typical solution bitmaps from this group are shown in Figure 7.2. This grouping characterizes solutions by students who are having difficulty solving the problem and are unable to divide the system into components which allow the unknown forces to be solved.



(a)



(b)

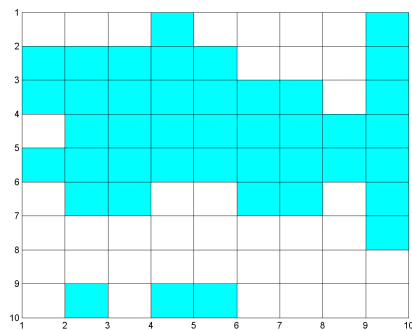


(c)

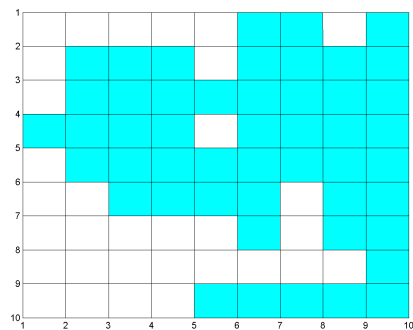
Figure 7.2: Typical solution bitmaps from group one of problem two of the final exam from the 2012 course offering.

The second group comprised 92 sketches. The FBDs of the sketches within this group typically spanned the entire page and were often either a single large FBD outlining the entire problem description image or comprised several small components

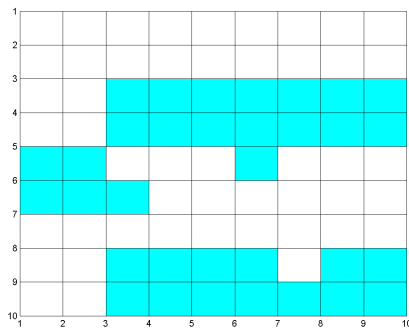
scattered across the entire page. This group is characterized by solutions written by students who perhaps were struggling to identify the FBD that would best lead to the correct solution, and spent an entire page testing different FBDs until they came across an acceptable one. Three typical bitmaps from this group are shown in Figure 7.3.



(a)



(b)



(c)

Figure 7.3: Typical solution bitmaps from group two of problem two of the final exam from the 2012 course offering.

Lastly, the third group comprised 24 sketches. Sketches from this group typically contained the same FBD component redrawn at least once. This group is characterized by students who made mistakes in their first FBD and had to later start over,

perhaps indicating a lack of understanding. Three typical bitmaps from this group are shown in Figure 7.4.

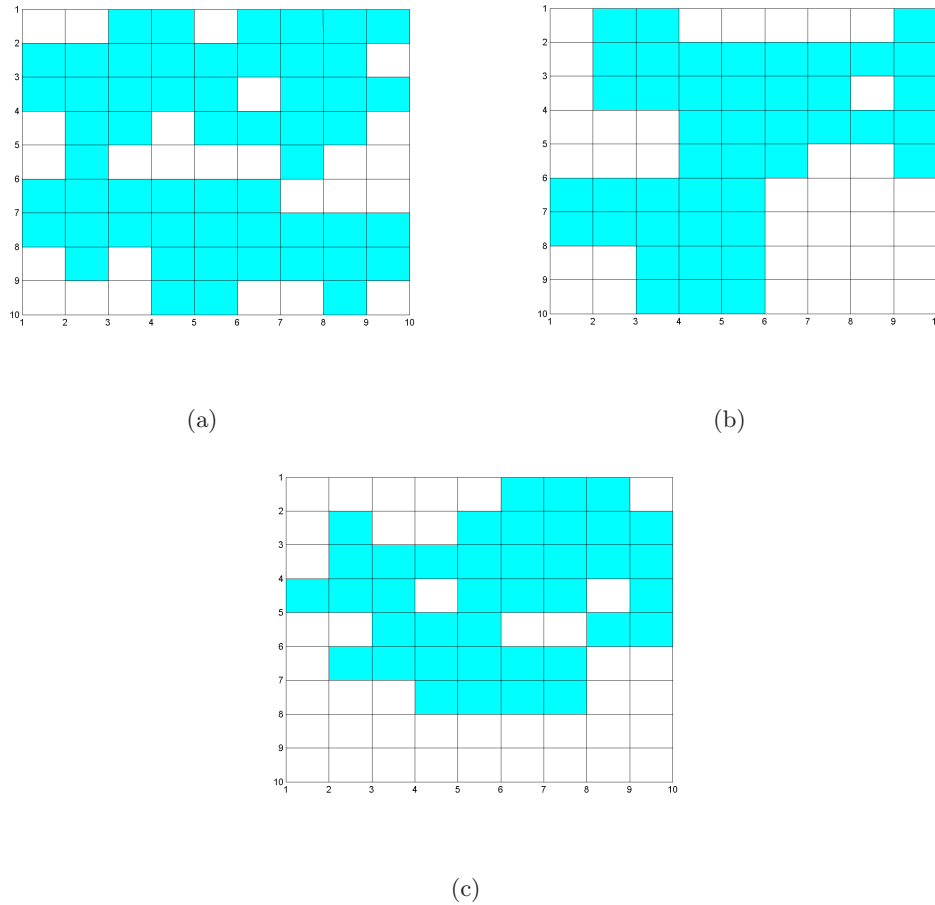


Figure 7.4: Typical solution bitmaps from group three of problem two of the final exam from the 2012 course offering.

For problem six, the K-Means clustering algorithm identified five non-empty groups. The first group comprised 18 sketches, all of which came from the second page of a solution to a problem. Sketches in this group typically contained a single, small FBD in the upper left corner of the page, representing the two-force-member present in that problem. This grouping is characterized by solutions of students who saved solving the

two-force-member as the final step. Three typical bitmaps from this group are shown in Figure 7.5.

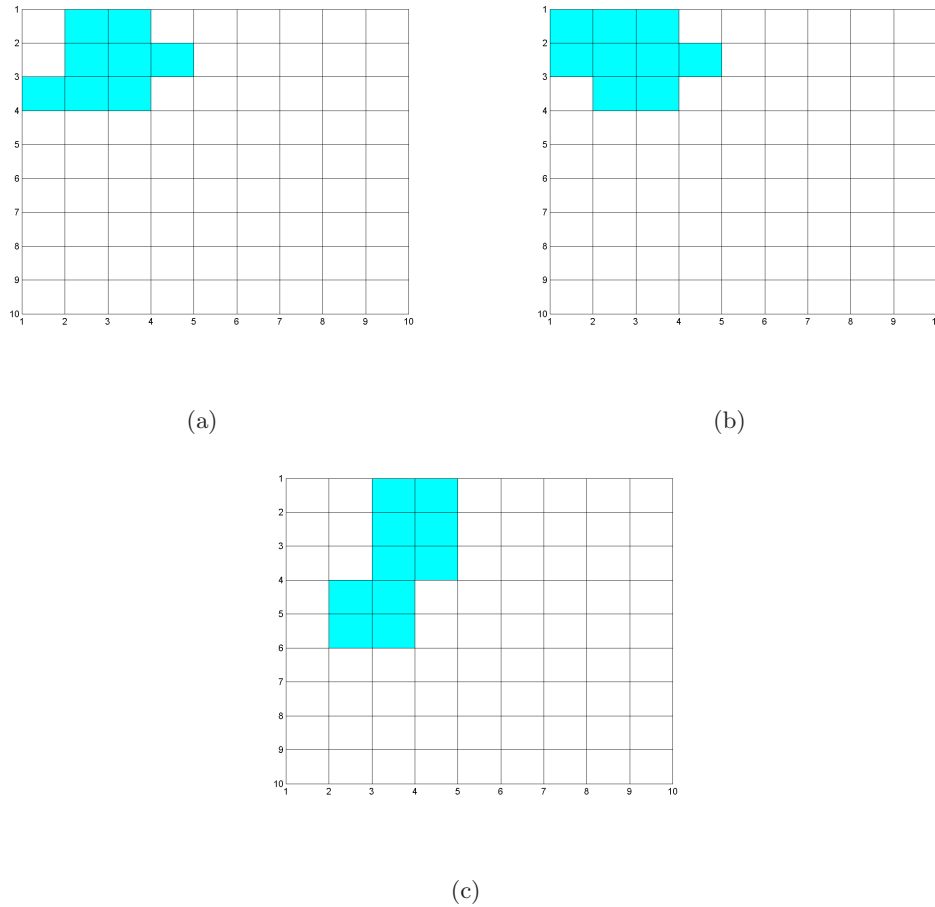
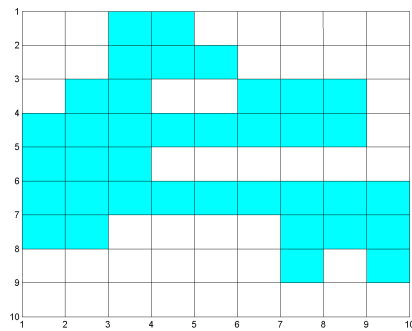
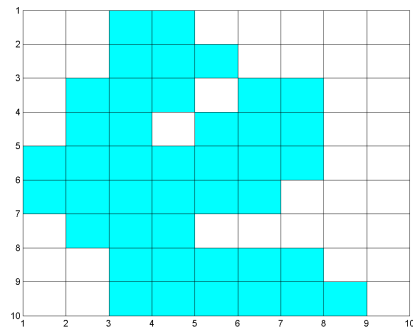


Figure 7.5: Typical solution bitmaps from group one of problem six of the final exam for the 2012 course offering.

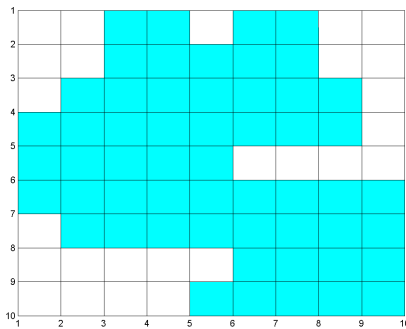
The second group comprised 42 sketches. These sketches typically contained a single, large FBD which was simply an outline of the entire system shown in the problem description image. This group is characterized by a lack of understanding by students, as they were unable to identify proper boundaries which exposed the forces required to



(a)



(b)



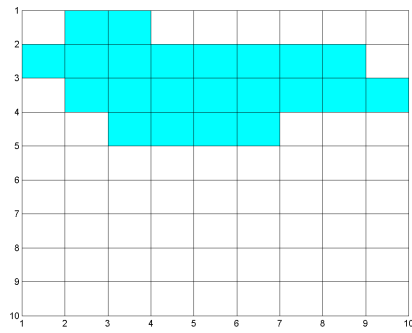
(c)

Figure 7.6: Typical solution bitmaps from group two of problem six of the final exam for the 2012 course offering.

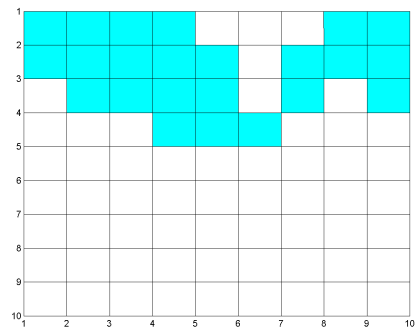
solve for the unknowns. Three typical bitmaps from this group are shown in Figure 7.6.

The third group comprised 19 sketches. These sketches typically contained just two FBDs that horizontally spanned the top of the page. This group is characterized by students who completely finished their FBDs prior to beginning work on the equations. Three typical images of solutions from this group are shown in Figure 7.7.

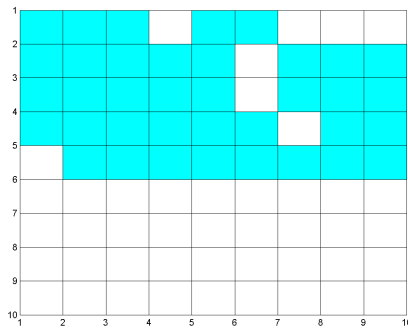
Groups four, five, and six are very similar and comprise 39, 47, and 8 sketches respectively. Each of these groups typically contains three to four FBDs spread out



(a)



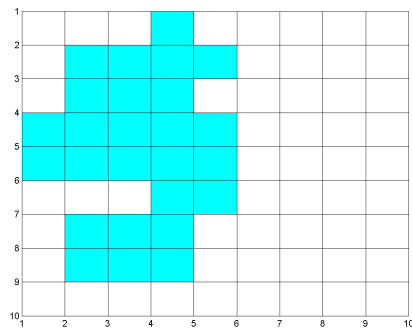
(b)



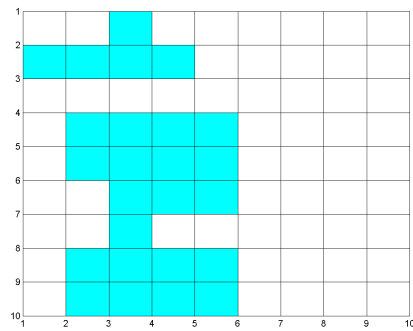
(c)

Figure 7.7: Typical solution bitmaps from group three of problem six of the final exam for the 2012 course offering.

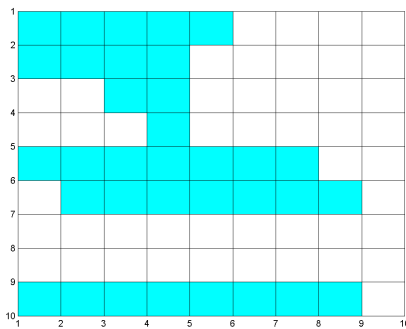
vertically along the left margin of the page. These groups are characterized by students who drew a single component of the FBD, solved equilibrium equations for that component, then moved on to the next component, and so forth. Three typical bitmaps from this group are shown in Figure 7.8.



(a)



(b)



(c)

Figure 7.8: Typical solution bitmaps from groups four, five, and six of problem six of the final exam for the 2012 course offering.

7.5 Discussion

It is important to note that the results presented in the previous section concern typical templates from each grouping and are not a generalization that necessarily applies to every template in a single grouping. We strove to manually identify common themes present in many of the templates but which may not hold true for all templates in that group.

Furthermore, the clustering results do not directly reflect the quality or content

of the FBDs or equations written by the students. For example, two FBDs of similar shape, size, and location, drawn by two different students, would be grouped together by this algorithm, even if those FBDs corresponded to solutions of a student who did and did not perform well.

Instead, these clusters identify different organizational behaviors exhibited by students. Using these clusters, we identified the common, high-level organizational patterns exhibited by the students via a manual inspection of the actual solutions. While it has been shown in previous work that students organizational habits are indicative of their performance, the goal of this work is not to automatically identify students performance given the organization of their FBD and equation writing, but instead to develop a taxonomy of the types of behaviors exhibited by students as they solve problems. By better understanding the typical behaviors students employ when solving Statics problems, instructors may gain insights into the cognitive processes employed by students. This work paves the way for future work to analyze the performance of students who exhibit particular organizational patterns to see if there is a significant correlation between the types of organization employed and performance.

7.6 Conclusion

In this work, we have taken first steps towards developing a taxonomy for the spatial organization employed by students. We applied educational informatics techniques to automatically identify groupings of students who organized their solutions in a similar way.

We represent each page of a solution as a low-resolution bitmap. We compute distances between bitmaps using the Hausdorff distance and cluster the bitmaps by that distance using the K-Means clustering algorithm. This algorithm identifies groupings of bitmaps that are similar with one another and distinct from bitmaps of other groups. Each of these groups represents a distinct spatial organization type. Having examined the pages which comprise each grouping, we were able to describe the higher-level organizational habits exhibited by the students. These habits have important pedagogical implications for instructors, and provide insights into the process by which students solve their problems.

Chapter 8

Linear Regression and Effort-Based, Numerical Features

8.1 Introduction

Homework serves a number of purposes. It provides students with an opportunity to practice methods they have learned in the classroom, familiarizes them with new material before it is covered in lecture, and helps them synthesize concepts and apply them in new ways. Despite its widespread use, there is contention as to whether homework leads to better course performance. Numerous studies have examined the existence of correlations between a student's effort on homework and performance in a course, yet the results of these studies are mixed.

Variations in the nature of these studies may partially account for these inconsistencies. For example, these studies vary in the grade-level of the students, the type of homework assigned, and the subject matter. Additionally, bias and inconsistencies

in the measurement of homework effort may also confound the results. Most previous work relies on the students themselves, or their parents, to report the amount of time spent on homework.

In this chapter¹, we present more precise and objective measures of homework effort enabled by the pen data extracted from the 2012 course offering. We compute numerical features from these records which estimate the effort students expended on each homework assignment. We use these features to predict students performance via a number of measures, including homework, quiz, and exam scores. We show that these effort-based features explain up to 39.9% of the variance in the students performance. These results have several pedagogical implications. The correlations we identify provide insights into the types of transfer students make from homework to exam problems. The results can also be used to evaluate the effectiveness of homework assignments, allowing instructors to improve homework assignments for future course offerings. Lastly, our analyses can be used to identify patterns of homework effort exhibited by students who perform well in the course.

8.2 Computing an Estimation of Student Effort

Here we describe the two types of novel quantitative features we use to estimate students effort on homework. The overall-effort features are coarse-grained, and characterize the total effort a student spent on a particular assignment. The per-problem features are fine-grained, and characterize the amount of effort spent on each individual problem.

¹The work presented in this chapter has been published and appears in [57]

8.3 Overall-Effort Features and Performance Models

The overall-effort features characterize the distribution of effort a student expends on his or her homework assignment. For example, some students may begin an assignment early and put substantial work into it each day, resulting in several homework “episodes.” Conversely, other students may put off the homework until shortly before it is due, resulting in a single, large homework episode.

To compute the overall-effort features, we first create a time-series representing the effort a student exerted on an assignment. The series begins with the first pen stroke written and ends with the last. This time span is divided into five-minute intervals. Each interval is characterized by the amount of ink written, which is defined as the distance the pen tip travels on the paper during that interval. In this way, effort is characterized by the amount of writing rather than simply the amount of time elapsed.

Figure 8.1 shows a typical effort time-series. Effort time-series are typically flat and punctuated with a few, large episodes of activity. We compute four features from each time-series. The first feature is the total amount of ink written, which characterizes the total effort spent on that assignment. The remaining three features characterize the distribution of this effort. To compute these features, we first identify active episodes, in which a student is writing, and inactive episodes in which no writing occurs. Each contiguous sequence of non-zero intervals (i.e., intervals containing writing) forms an active episode. To prevent small breaks in writing from splitting an episode, active episodes may contain subsequences of up to two zero-valued intervals. Thus a break of ten minutes or less does not break a problem-solving episode. All remaining contiguous

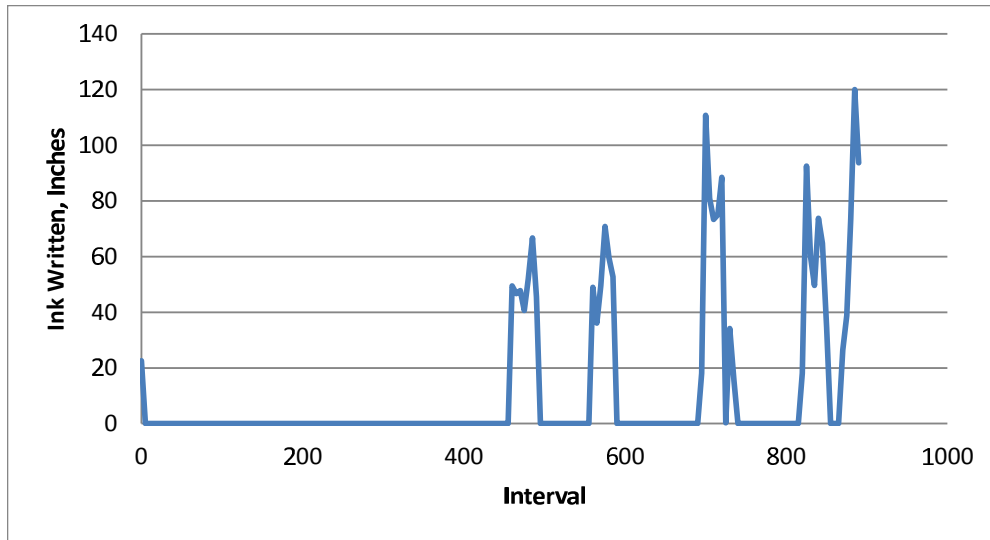


Figure 8.1: Typical effort time-series of a single student on a single assignment. The abscissa denotes the index of the five-minute intervals, not the actual time stamp of the interval. For example, interval one comprises the first five minutes of the students problem solution. The ordinate denotes the total ink written during a given five-minute interval.

sequences of zero-valued intervals are identified as inactive episodes. The effort time-series is characterized by the number of active episodes, the average length of the active episodes, and the average length of the inactive episodes. Figure 8.1 shows the active and inactive episodes for the effort time-series from Figure 8.1.

We used the four overall-effort features to construct models relating students effort on a particular assignment to performance on that assignment. We computed these models using the linear regression package in the WEKA Data Mining Software suite[44]. Figure 8.3 presents the coefficient of determination for the models constructed for each homework assignment. WEKA's linear regression package employs a greedy feature selection algorithm. Features are removed from the model until there is no improvement in the error estimation, as determined by the Akaike information criterion[4]. The

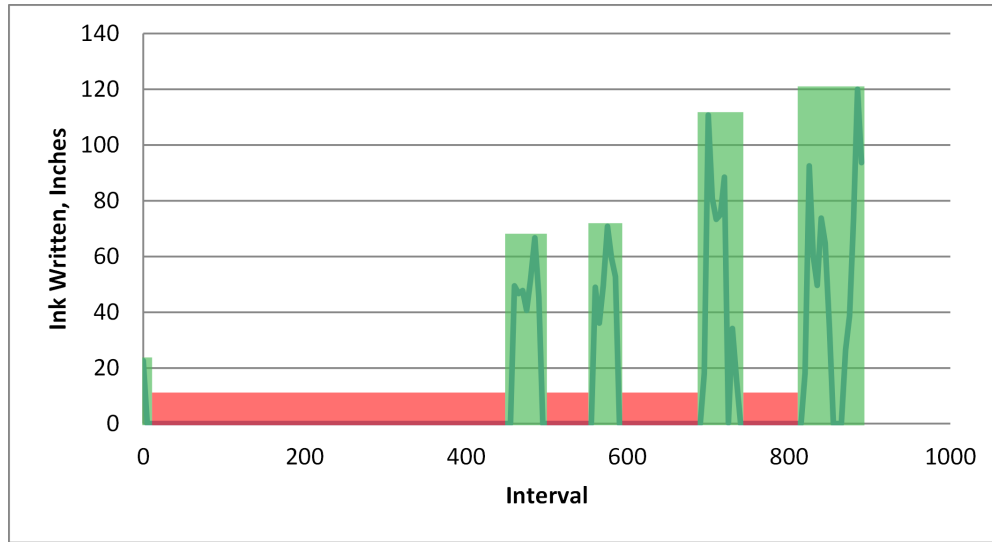


Figure 8.2: Effort time-series previously shown in Figure Figure 8.1 with active (green regions) and inactive (red regions) episodes identified.

HW No.	Avg. Act. Length	Avg. Inact. Length	No. Act. Intervals	Ink (in.)
3			Selected	
4			Selected	
5		Selected		Selected
6	Selected			Selected
7		Selected		Selected
8		Selected	Selected	

Table 8.1: Overall-effort features selected for linear regression models for homework performance.

features selected for each homework model are presented in Table 8.1.

We also used WEKAs implementation of the expectation maximization (EM) clustering[30] algorithm to group students by similarities in both the effort they exerted on a homework assignment and they performance they achieved on it. The clusters identified for each assignment are listed in Table 8.2 through Table 8.7. Each cluster is characterized by the average and standard deviation of the homework grade (the

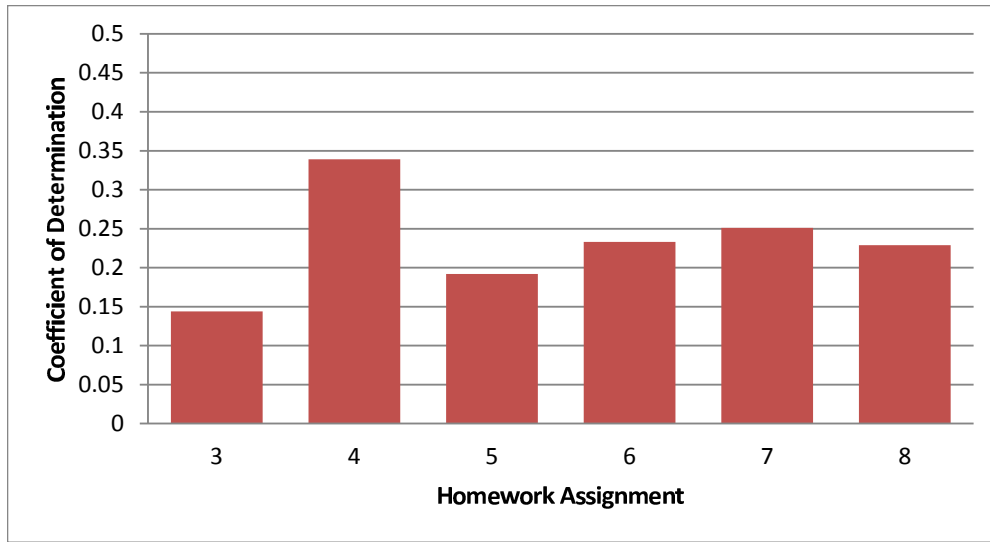


Figure 8.3: Coefficients of determination for linear regression models relating overall-effort features to homework performance. The average coefficient of determination is 0.231.

maximum grade is 10.0) and the overall-effort features of the data points contained in that cluster.

8.4 Per-Problem Features

We use the overall-effort features to examine the relationship between the total effort on an assignment and performance on that assignment. Here we examine how effort on individual homework problems relates to performance on subsequent homework assignments, exams (midterms and final), and quizzes. We estimate the effort on a single homework problem as the total time during which the pen is in contact with the paper. The time between pen strokes is not included in this value. We once again employ the linear regression package available in WEKA to compute regression models. The resulting coefficients of determination are shown in Figure 8.4. In the figure, the

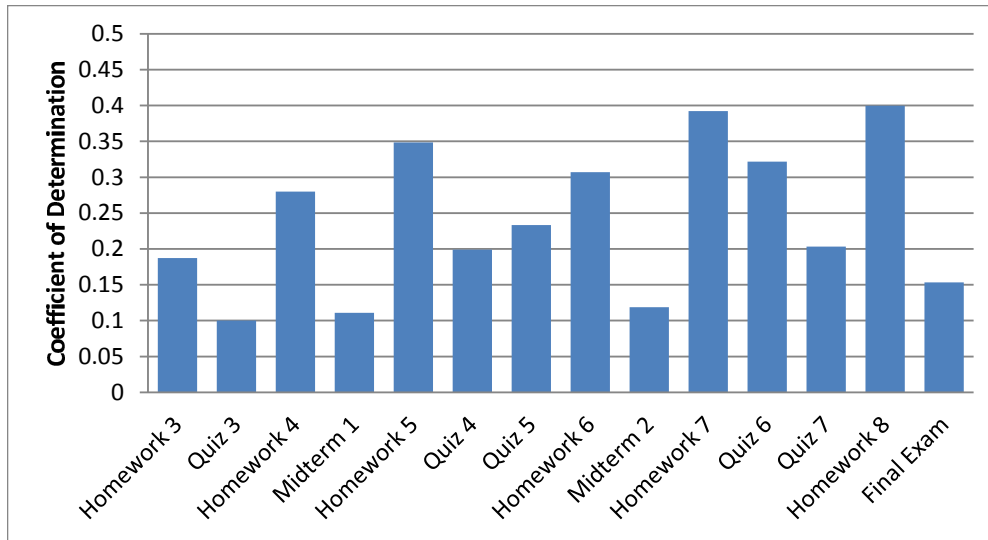


Figure 8.4: Coefficients of determination of models using per-problem features to predict performance on coursework. The average coefficient of determination is 0.239. The coursework is listed in the order completed.

coursework is listed in the order it was completed. For example, quiz six was completed just after homework seven was due. Furthermore, a model using the effort on each of the problems from homework assignments three through seven predicts performance on quiz six with a coefficient of determination of 0.32.

In a similar way, we used the per-problem effort features to predict performance on individual problems on the midterm and the final exams. The results are shown in Figure 8.5. The features selected by WEKA's greedy feature selection algorithm provide insights about learning transfer. For example, the features selected for predicting performance on the first problem of the first midterm were effort on homework assignment three, problem four; effort on homework three, problem five; and effort on homework four, problem five. (Because of the large number of features we consider, space constraints prevent inclusion of the complete feature selection results correspond-

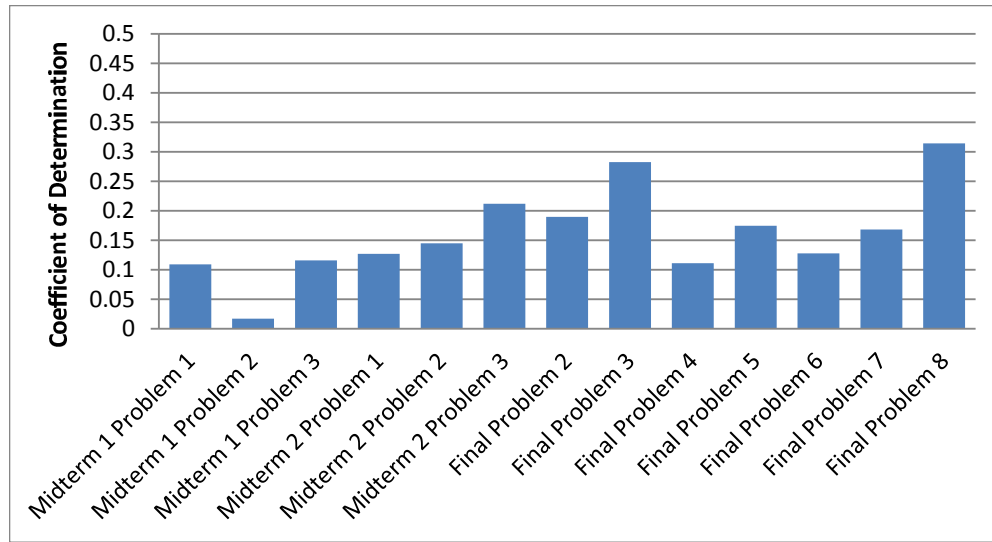


Figure 8.5: Coefficients of determination of models using per-problem features to predict performance on individual exam problems. The average coefficient of determination is 0.161. The first problem of the final exam concerned professional ethics question and thus was excluded from the analysis.

ing to Figure 8.4 and Figure 8.5.)

8.5 Discussion

It is important to note that our effort features capture only a portion of the effort expended by students on studying. Other elements of studying, such as the amount of time spent reading the textbook or working on scratch paper, are not captured by the digital pens we use. However, we believe that the amount of time spent problem solving on homework provides a useful measure of a students effort in a course.

The results of the linear regression analysis of the overall-effort features indicate that students effort does account for a considerable portion of the variance in performance. For example, in the best case, the effort-based features accounted for 33.9%

of the variance in performance on homework. This correlation is considerably stronger than that found in previous studies. Interestingly, the feature selection results indicate that the number of active intervals and the total ink written are the most important features. Each is selected for the models for at least half of the homework assignments. This suggests that the more often a student sits down to work on an assignment and the more writing he or she does, the more likely it is that the student will do well on that assignment.

Each of the linear regression models showed a positive correlation between performance and each effort feature, indicating that the more time a student spent on a particular problem the better he or she performed. This may demonstrate that students who spent more time on their homework were better prepared for the exam problems. It is important to note that this may not always be the case. It is entirely possible that a particular effort feature could negatively correlate with performance. Such a case may indicate that a student spent a large amount of time on a problem as because he or she had difficulty understanding it and as a result that student did not perform well on related exam problems.

The clustering results reveal a similar story to the linear regression analysis. The clusters with the highest average grade are typically those which also have the highest average number of active intervals. The clustering results serve as an easy-to-read summary of typical solution behaviors exhibited by students on a particular assignment. Instructors can use these sorts of results to quickly determine which groups of students are spending a sufficient amount of time on the homework. More importantly, this analysis reveals just how much time is needed to do well on an assignment. This

Cluster	1	2	3
Avg. Grade	5.35	6.2	7.23
Std. Grade	2.92	3.3	3.08
Avg. Active Length	393.86	14.8	100.16
Std. Active Length	412.1	9.83	45.42
Avg. Inactive Length	779.13	80.72	76.69
Std. Inactive Length	273	82.56	41.81
Avg. No. Active Intervals	2.5	2.88	2.03
Std. No. Active Intervals	0.5	1.77	0.18
Avg. Ink (Inches)	905.9	1045.4	1216.0
Std. Ink (Inches)	247.4	692.7	1168.4

Table 8.2: The average and standard deviation for each of the EM clusters for homework assignment three.

will enable an instructor, for example, to identify when a large number of students are performing poorly on a problem despite spending a great deal of effort on it, a strong indication that a widespread misconception or difficulty exists in the class.

The results of the per-problem linear regression analysis reveal that the amount of effort spent by students on individual homework problems can account for up to 39.9% of the variance of students performance on subsequent homework assignments. Furthermore, when the per-problem features are used to predict performance on individual exam problems, they can account for up to 31.4% of the variance in that grade. This is an interesting result as these features do not consider the semantic content of the students solutions.

The per-problem features selected by the final model of each linear regression are an indication of the importance of individual homework problems. (In the present

Cluster	1	2	3
Avg. Grade	3.53	3.75	6.87
Std. Grade	3.57	3.36	1.48
Avg. Active Length	654.97	56.78	14.86
Std. Active Length	115.82	64.93	9.97
Avg. Inactive Length	636.38	83.58	174.18
Std. Inactive Length	119.52	86.68	115.59
Avg. No. Active Intervals	2	2.06	4.5
Std. No. Active Intervals	1.81	0.95	1.61
Avg. Ink (Inches)	804.8	1001.1	166.9
Std. Ink (Inches)	358.6	652.1	758.5

Table 8.3: The average and standard deviation for each of the EM clusters for homework assignment four.

Cluster	1	2	3	4
Avg. Grade	5.04	5.56	2.45	1.61
Std. Grade	1.7	2	3.11	2.16
Avg. Active Length	109.08	12.23	873.58	23.12
Std. Active Length	84.55	8.3	303.81	36.64
Avg. Inactive Length	93.79	164.16	855.98	276.44
Std. Inactive Length	71.28	118.78	308.34	217.46
Avg. No. Active Intervals	1.96	4.81	2	2.93
Std. No. Active Intervals	0.62	1.69	1.77	1.03
Avg. Ink (Inches)	1088.9	176.5	675.4	925.1
Std. Ink (Inches)	680.2	756.6	401.6	539.6

Table 8.4: The average and standard deviation for each of the EM clusters for homework assignment five.

Cluster	1	2	3	4
Avg. Grade	5.57	8.02	3.89	8.23
Std. Grade	2.97	1.68	4.04	1.54
Avg. Active Length	973.42	10.93	97.08	33.38
Std. Active Length	408.61	5.16	86.93	22.12
Avg. Inactive Length	1149.78	168.8	109.29	27.02
Std. Inactive Length	152.99	139.52	87.22	17.19
Avg. No. Active Intervals	2.14	4.74	2.07	1.75
Std. No. Active Intervals	0.34	1.61	0.79	0.76
Avg. Ink (Inches)	1188.2	660.3	928.1	1290.1
Std. Ink (Inches)	297.1	686.4	602.6	6124.7

Table 8.5: The average and standard deviation for each of the EM clusters for homework assignment six.

Cluster	1	2
Avg. Grade	6.38	7.36
Std. Grade	3.47	2.94
Avg. Active Length	10.02	686.56
Std. Active Length	6.26	980.46
Avg. Inactive Length	357.31	658.08
Std. Inactive Length	416.62	983.1
Avg. No. Active Intervals	3.68	2
Std. No. Active Intervals	2.3	2.17
Avg. Ink (Inches)	1237.1	1246.3
Std. Ink (Inches)	667.1	555.9

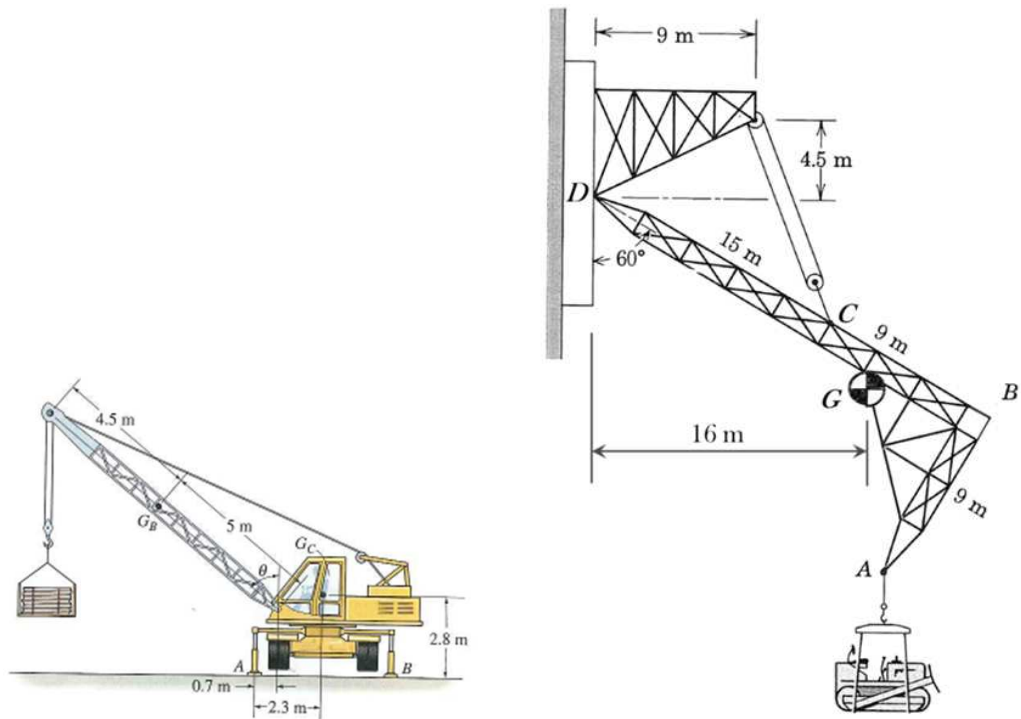
Table 8.6: The average and standard deviation for each of the EM clusters for homework assignment seven.

Cluster	1	2	3	4
Avg. Grade	7.61	8.09	0.78	8.81
Std. Grade	1.37	2.37	0.41	1.78
Avg. Active Length	650.91	24.07	73.21	26.36
Std. Active Length	343.74	32.02	140.13	31.26
Avg. Inactive Length	628.91	18.98	100.74	183.19
Std. Inactive Length	340.41	24.58	165.08	159.16
Avg. No. Active Intervals	2	1.24	1.43	3.01
Std. No. Active Intervals	1.07	0.43	0.64	1.05
Avg. Ink (Inches)	769.6	610	486.5	1119
Std. Ink (Inches)	343.8	202.8	293.3	342.8

Table 8.7: The average and standard deviation for each of the EM clusters for homework assignment eight.

data, all of these features were positively correlated with performance.) The selected features reveal which homework problems lead to success in learning particular concepts in the course. More specifically, they reveal the transfer taking place from particular homework problems to particular exam problems. Consider, for example, the first problem of the first midterm, shown in Figure 8.6. The amount of effort exerted on homework three, problem four was one of the three features selected to predict the performance on this midterm problem. Interesting, the midterm problem can be considered a rotated version of the homework problem. This clearly shows students transferring knowledge from the homework problem to solve the midterm problem.

These results indicate practical changes instructors can make to homework assignments. Namely, this suggests that exam problems which comprise simple extensions to homework problems can be used to identify students ability to transfer knowledge. Instructors should examine students performance on the homework problems that will be similar to upcoming exam problems. If students are spending insufficient effort on



(a) Homework three, problem four

(b) Midterm one, problem one

Figure 8.6: An example of the knowledge transfer students made. Problem (b) is effectively a rotated version of problem (a).

these problems, they should be further examined in class. This analysis can serve as an invaluable tool to the instructor of a course. Using it, the instructor may review which features (i.e., homework problems) are selected in the per-problem models, identify the types of transfer students made, and use that knowledge to shape exam and homework problems in future course offerings. A manual analysis of the transfer revealed by our present data will be an important element of future work on this project.

Overall, the linear regression analysis results for both the overall-effort and per-problem models provide correlations that are much stronger than those found in prior work. As mentioned earlier, those studies typically relied on either the students or their

parents to report the amount of time spent working on each homework assignment. The Livescribe digital pens provide a more reliable measure of the amount of time students spend on their homework assignments which may account for the higher coefficients of determination we obtain.

8.6 Conclusion

In this chapter, we have presented novel, data-driven methods for analyzing students homework habits. Namely, we have computed a number of numerical features which characterized the effort that students exerted when solving their homework assignments. We have applied machine learning techniques, namely a linear regression classifier, in order to use these features to both predict and cluster students' performance in this course.

There were two major types of features computed: overall-effort features and per-problem features. Four overall-effort features were computed which characterized both the amount and distribution of effort exerted on a single assignment: the total amount of ink written in an assignment; the number of active problem-solving episodes; the average length of the active problem-solving episodes; and the average length of the inactive episodes. These features were used to predict the performance on the homework assignment from which they were computed. These overall-effort features explain up to 33.9% of the variance in students performance on a particular assignment. Additionally, these features and the homework assignment grade were used as input to the EM clustering algorithm. This algorithm identified groups of students who both

displayed similar effort behaviors and assignment performance. These groupings may be used as behavior-performance profiles that are valuable feedback for an instructor.

The per-problem features comprised the amount of time spent writing the solution to each problem of a homework assignment. These features were used to predict performance on subsequent homework assignments, quizzes, and exams. These features accounted for up to 39.9% of the variance on a particular item. Additionally, the per-problem features were used to predict performance on individual exam problems. They accounted for up to 31.4% of the variance of the performance on individual problems. More importantly, the features selected by these linear regression models provide important insights for the instructor, indicating which of the homework problems lead to good performance on the exam questions. By analyzing which homework problems most account for the variance on a particular exam problem, instructors may identify the types of transfer students make from homework to exam problems. This information is an invaluable source of feedback for the instructor as well as a guide to how homework and exam problems should be designed for future course offerings.

Chapter 9

Information Gain and Effort-Based, Numerical Features

9.1 Introduction

In this chapter we investigate the ability of the effort students expend on their homework assignments to predict errors those students make on the midterm exams. Based on the work presented in the previous chapter, we sought to more directly measure the predictive ability of effort on each homework assignment. Whereas in the last chapter, we consider models which combined features from different homework problems to predict performance, here we consider the predictive power of each feature individually. This provides a clearer understanding of the value of each feature as opposed to considering its value when used in combination with other features. Furthermore, this more clearly allows us to identify which homework assignments were most effective in preparing students for the exams.

As mentioned in Section 3.3, these exams were graded according to a precise, error-based rubric. The goal of this work is to identify how well the amount of effort expended on a homework assignment, estimated as ink written on that assignment, predicts whether or not a student will make each of the errors which comprise the rubric.

We compute features for every homework problem, which characterizes the effort spent on that problem in terms of the amount of ink written. We then use each feature individually to predict whether or not each student will make each particular error on each midterm problem. This prediction is made using a simple decision stump, which bases its decisions on the amount of information gained.

Given that there are several features computed for each homework problem and that there are several binary errors to be predicted, we compile the large number of results from using every feature to predict every error by creating easy-to-read heat maps. These heat maps present all the resulting information gain values in a sorted fashion that allows the reader to quickly identify the most consistently predictive features. By doing so, we have developed a novel method by which instructors may quickly identify the homework problems which best lead to good performance on midterm exams. This provides a powerful, formative assessment tool.

9.2 Method

In this section we begin by describing the features computed for each problem of each homework assignment to characterize the amount of effort a student expended

on that problem. Next, we describe the different ways that the target midterm errors were grouped. Lastly, we describe how we use the information gained by each feature to predict if a student will make each error.

9.2.1 Features

We characterized effort on each homework assignment by computing three features: Ink_{FBD} , Ink_{EQN} , and Ink_{E2F} . Ink_{FBD} is the amount of ink, measured in inches, spent writing FBDs on a single problem of a particular homework assignment. Ink_{EQN} is the amount of ink, again measured in inches, spent writing equations on a single problem of a particular homework assignment. Lastly, Ink_{E2F} is the ratio of the two, computed as $Ink_{EQN}/(Ink_{FBD} + Ink_{EQN})$.

We use these features to predict if students will make particular errors on midterm problems. We use only features computed from the two homework assignments completed most recently before a midterm when predicting errors on that midterm. In this course, the content for midterm problems was based on the two previous homework assignments, making them the most relevant. Thus, when predicting errors for midterm one, features from homework assignments three and four were used. Since homework assignment three had eight problems and homework assignment four had seven problems, there were 45 total features $((8 + 7) \times 3)$ considered when predicting errors on midterm one. Similarly, when predicting errors for midterm two, features from homework assignments five and six were used. Since homework assignment five had eight problems and homework assignment six had six problems, there were 42 total features $((8 + 6) \times 3)$ considered when predicting errors on midterm two.

9.2.2 Error groupings

Because the number of errors described in Section 3.3 is large compared to the low number of students who made each individual error, we have grouped the errors. We have investigated a number of different groupings by applying different consolidation schemes. The idea for each scheme is to assign similar errors to the same group. Each grouping is still binary and indicates whether a student made any of the errors in that group (1) or made none of the errors in that group (0). The prediction task is then to determine whether a student made any of the errors in each group or not. We describe below the three different consolidation schemes by which we have grouped the midterm errors.

The first scheme did not merge any errors at all and retained the fine-grained errors exactly as they were described in Table 3.4 through Table 3.7. This is called the *fine-grained error grouping* scheme.

The next scheme grouped errors as broadly as possible, this is called the *category error grouping* scheme. In this grouping scheme, errors were grouped by category (column one of Table 3.4 through Table 3.7). For example, using this scheme, midterm one problem one was assigned five different error types, one for each error category in Table 3.4: FBD Error, Angle Error, Equation Error, Magnitude Error, and Algebra Error. The errors that have been placed into each group using this scheme are shown in Table 9.1 through Table 9.4 for each of the midterm-problem pairs.

The next scheme groups errors at a level of granularity that is between the prior

Error Group	Errors in Group (Table 3.4)
FBD Error	1 - 11
Angle Error	12 - 13
Equation Error	14 - 40
Magnitude Error	41 - 42
Algebra Error	43 - 44

Table 9.1: The error groups for midterm one problem one using the *category error grouping* scheme. The Error Group column shows the name of each of the five error groups, and the Errors in Group column shows the numbers of the errors placed in each group. These numbers refer to errors in Table 3.4.

Error Group	Errors in Group (Table 3.5)
1 st FBD Error	1 - 7
2 nd FBD Error	8 - 14
Equation Error	15 - 29
Algebra Error	30
Answer Error	31 - 33

Table 9.2: The error groups for midterm two problem one using the *category error grouping* scheme. The Error Group column shows the name of each of the five error groups, and the Errors in Group column shows the numbers of the errors placed in each group. These numbers refer to errors in Table 3.5.

Error Group	Errors in Group (Table 3.6)
1 st FBD Error	1 - 7
2 nd FBD Error	8 - 14
3 rd FBD Error	15 - 21
Equation Error	22 - 36
Algebra Error	37
Answer Error	38 - 41

Table 9.3: The error groups for midterm two problem two using the *category error grouping* scheme. The Error Group column shows the name of each of the six error groups, and the Errors in Group column shows the numbers of the errors placed in each group. These numbers refer to errors in Table 3.6.

Error Group	Errors in Group (Table 3.7)
1 st FBD Error	1 - 7
2 nd FBD Error	8 - 14
3 rd FBD Error	15 - 21
Equation Error	22 - 36
Algebra Error	37
Answer Error	38 - 41

Table 9.4: The error groups for midterm two problem three using the *category error grouping* scheme. The Error Group column shows the name of each of the six error groups, and the Errors in Group column shows the numbers of the errors placed in each group. These numbers refer to errors in Table 3.7.

two schemes. This scheme is called the *subcategory error grouping* scheme. Using this scheme, errors were grouped according to both their category and subcategory (column two of Table 3.4 through Table 3.7). Additionally though, any error that corresponded to a missing element was given its own error group. The errors that have been placed into each group using this scheme are shown in Table 9.5 through Table 9.8 for each of the midterm-problem pairs.

9.2.3 Information Gain Prediction

We use each feature individually to separate those students who did make a particular error from those who did not. To do this given a single feature, F , and error E , we create a data set, D_{FE} of all the students. Each student, i , is represented in D_{FE} as the pair (f_i, e_i) , where f_i is that students' feature, F , value, and e_i is a binary label identifying whether the student made error E . All students in D_{FE} are sorted by increasing F value. A decision stump[67] is then created which best partitions the

Error Group	Errors in Group (Table 3.4)
FBD Not Qualify	1
Incorrect Body Selected	2
Reaction Error	3 - 9
Angle Error	11 - 12
Sum of Moments Missing	13
Sum of Moments Error	14 - 23
X-Force Eqn. Missing	24
X-Force Eqn. Error	25 - 30
Y-Force Eqn. Missing	31
Y-Force Eqn. Error	32 - 39
Magnitude Error	40 - 41
Algebra Error	42 - 43

Table 9.5: The error groups for midterm one problem one using the *subcategory error grouping* scheme. The Error Group column shows the name of each of the twelve error groups, and the Errors in Group column shows the numbers of the errors placed in each group. These numbers refer to errors in Table 3.4.

Error Group	Errors in Group (Table 3.5)
1 st FBD Not Qualify	2
Incorrect 1 st Body Selected	3
1 st FBD Reaction Error	4 - 7
2 nd FBD Not Qualify	9
Incorrect 2 nd Body Selected	10
2 nd FBD Error	11 - 14
Moment Eqn. Missing	15
Moment Eqn. Error	16 - 21
X-Force Eqn. Missing	22
X-Force Eqn. Error	24 - 25
Y-Force Eqn. Missing	26
Y-Force Eqn. Error	27 - 29
Algebra Error	30
Answer Error	31 - 33

Table 9.6: The error groups for midterm two problem one using the *subcategory error grouping* scheme. The Error Group column shows the name of each of the 14 error groups, and the Errors in Group column shows the numbers of the errors placed in each group. These numbers refer to errors in Table 3.5.

Error Group	Errors in Group (Table 3.6)
1 st FBD Missing	1
1 st FBD Not Qualify	2
Incorrect 1 st Body Selected	3
1 st FBD Reaction Error	4 - 7
2 nd FBD Missing	8
2 nd FBD Not Qualify	9
Incorrect 2 nd Body Selected	10
2 nd FBD Error	11 - 14
3 rd FBD Missing	15
3 rd FBD Not Qualify	16
Incorrect 3 rd Body Selected	17
3 rd FBD Error	19 - 21
Moment Eqn. Missing	22
Moment Eqn. Error	23 - 28
X-Force Eqn. Missing	29
X-Force Eqn. Error	30 - 32
Y-Force Eqn. Missing	33
Y-Force Eqn. Error	34 - 36
Algebra Error	37
Answer Error	38 - 41

Table 9.7: The error groups for midterm two problem two using the *subcategory error grouping* scheme. The Error Group column shows the name of each of the 20 error groups, and the Errors in Group column shows the numbers of the errors placed in each group. These numbers refer to errors in Table 3.6.

Error Group	Errors in Group (Table 3.7)
1 st FBD Missing	1
1 st FBD Not Qualify	2
Incorrect 1 st Body Selected	3
1 st FBD Reaction Error	4 - 7
2 nd FBD Missing	8
2 nd FBD Not Qualify	9
Incorrect 2 nd Body Selected	10
2 nd FBD Error	11 - 14
3 rd FBD Missing	15
3 rd FBD Not Qualify	16
Incorrect 3 rd Body Selected	17
3 rd FBD Error	19 - 21
Moment Eqn. Missing	22
Moment Eqn. Error	23 - 28
X-Force Eqn. Missing	29
X-Force Eqn. Error	30 - 32
Y-Force Eqn. Missing	33
Y-Force Eqn. Error	34 - 36
Algebra Error	37
Answer Error	38 - 41

Table 9.8: The error groups for midterm two problem three using the *subcategory error grouping* scheme. The Error Group column shows the name of each of the 20 error groups, and the Errors in Group column shows the numbers of the errors placed in each group. These numbers refer to errors in Table 3.7.

data into two regions, left and right. The pivot used to separate the data is that which provides the best Information Gain. This measure compares the entropy of the entire data set to the entropy of the two partitions.

The entropy of data set, D_{FE} , provides a measure of the uncertainty about whether or not the students in D made E :

$$E(D_{FE}) = -(p(a)\log(p(a)) + p(b)\log(p(b))) \quad (9.1)$$

Where $p(a)$ is the percentage of students in D_{FE} that made error E and $p(b)$ is the percentage that did not. Entropy is a real-valued number $\in [0, 1]$ (when using a logarithm of base 2) where “0” signifies that there is no uncertainty, i.e., all students in the data set either did or did not commit the error, and “1” signifies complete uncertainty, i.e., there are as many students in the data set that did make an error as did not.

The information gain for a given split of D_{FE} , which results in data sets D_{left} and D_{right} is:

$$I(D_{FE}, D_{left}, D_{right}) = E(D_{FE}) - p(D_{left})E(D_{left}) + p(D_{right})E(D_{right}) \quad (9.2)$$

where $p(D_{left}) = |D_{left}|/|D_{FE}|$ and $p(D_{right}) = |D_{right}|/|D_{FE}|$. The information gain then shows the amount of uncertainty that was reduced by splitting the data set into the two groups, or said conversely, the information gained by doing so. For each data set, the optimal pivot point is found by exhaustively computing the information gained by splitting the data set at each student.

9.3 Results

We have created a $m \times n$ matrix for each (midterm, problem, error grouping) triple. Each matrix shows the information gained using each of the m features to predict each of the n errors on a particular midterm problem given a particular error grouping scheme. For example, consider the *category error grouping* scheme on midterm one problem one. In this case, 45 effort-based features were used to predict five different error groups, resulting in a 45×5 matrix in which every element (i, j) is the information gained by using feature f_i to split students according to error e_j . We sum the information gain values of each row and then sort the rows of the matrix by increasing information gain sum. Similarly, we sum the information gain values of each column and then sort the columns of the matrix by increasing information gain sum. By sorting the matrix in this way, the more consistently predictive features appear towards the bottom of the matrix and the more consistently predictable errors appear towards the right of the matrix. To illustrate this pattern, we have created colored heat-maps for each matrix. These heat-maps are colored with a three-color gradient, which maps the lowest information gain value to red, highest information gain value to green, and intermediate information gain value to yellow.

We present the heat map created for each (midterm problem, consolidation scheme) pair in the following sections.

9.3.1 Category Error Grouping Information Gain Heat Maps

In this section we show the information gain heat maps created for midterm one problem one and midterm two problems one through three using the *category error grouping* consolidation scheme. These heat maps are shown in Figure 9.1 through Figure 9.4.

Additionally, we highlight the first and last ten features (rows) of each heat map. The first ten rows in the matrix are those features with the smallest sum of information gain values. These features can be considered the 10 least consistently predictive features and will be important for the discussion that follows. Similarly, the last ten rows of the matrix are those features with the largest sum of information gain values. These features can be considered the 10 most consistently predictive features and will be important for the discussion that follows. The 10 most and the 10 least predictive features from each heat map are shown in Table 9.9 and Table 9.10 respectively. These tables also count the number of occurrences of each feature type found in the top-10 list for each midterm problem. This gives an indication of how often the Ink_{EQN} , Ink_{FBD} , and Ink_{E2F} features are the most predictive of errors on a particular midterm problem.

ErrorName	Eqn Error	Ang Err	Mag Err	FBD Err	Alg Err
Hw3-Pr1-EQN	0.028	0.019	0.030	0.011	0.016
Hw3-Pr8-FBD	0.026	0.015	0.016	0.030	0.019
Hw3-Pr1-FBD	0.033	0.017	0.024	0.018	0.018
Hw3-Pr8-EQN	0.037	0.025	0.024	0.019	0.014
Hw3-Pr1-E2F	0.019	0.033	0.036	0.020	0.013
Hw3-Pr5-FBD	0.029	0.028	0.024	0.013	0.031
Hw3-Pr5-E2F	0.014	0.022	0.030	0.020	0.041
Hw3-Pr8-E2F	0.037	0.041	0.024	0.013	0.014
Hw4-Pr7-FBD	0.028	0.019	0.039	0.016	0.032
Hw4-Pr1-FBD	0.018	0.027	0.032	0.036	0.023
Hw4-Pr3-FBD	0.027	0.025	0.019	0.039	0.030
Hw4-Pr6-FBD	0.016	0.052	0.032	0.015	0.025
Hw4-Pr1-E2F	0.028	0.022	0.024	0.039	0.027
Hw4-Pr1-EQN	0.028	0.018	0.032	0.039	0.025
Hw4-Pr5-FBD	0.014	0.038	0.036	0.030	0.027
Hw3-Pr7-EQN	0.029	0.015	0.031	0.014	0.059
Hw3-Pr6-E2F	0.032	0.015	0.020	0.030	0.055
Hw3-Pr6-EQN	0.028	0.015	0.020	0.035	0.055
Hw4-Pr3-EQN	0.029	0.025	0.017	0.053	0.029
Hw3-Pr5-EQN	0.019	0.032	0.030	0.029	0.045
Hw3-Pr4-E2F	0.013	0.025	0.045	0.024	0.050
Hw3-Pr7-FBD	0.032	0.015	0.042	0.018	0.052
Hw4-Pr7-E2F	0.052	0.025	0.025	0.020	0.038
Hw3-Pr3-FBD	0.032	0.022	0.028	0.027	0.052
Hw4-Pr2-E2F	0.025	0.022	0.035	0.065	0.013
Hw3-Pr6-FBD	0.028	0.015	0.030	0.036	0.052
Hw4-Pr4-FBD	0.013	0.025	0.039	0.034	0.052
Hw4-Pr4-EQN	0.013	0.017	0.045	0.027	0.063
Hw4-Pr6-EQN	0.016	0.052	0.024	0.051	0.031
Hw3-Pr3-EQN	0.037	0.028	0.028	0.030	0.052
Hw3-Pr7-E2F	0.029	0.031	0.031	0.026	0.059
Hw4-Pr6-E2F	0.016	0.052	0.024	0.058	0.027
Hw3-Pr4-FBD	0.028	0.028	0.048	0.030	0.043
Hw3-Pr2-FBD	0.016	0.032	0.051	0.036	0.043
Hw4-Pr2-FBD	0.040	0.026	0.030	0.066	0.017
Hw3-Pr3-E2F	0.028	0.031	0.026	0.043	0.052
Hw4-Pr7-EQN	0.052	0.013	0.067	0.020	0.038
Hw4-Pr2-EQN	0.040	0.043	0.009	0.074	0.027
Hw4-Pr3-E2F	0.029	0.049	0.021	0.057	0.038
Hw4-Pr5-EQN	0.032	0.070	0.021	0.053	0.021
Hw3-Pr4-EQN	0.013	0.033	0.067	0.038	0.050
Hw4-Pr5-E2F	0.025	0.072	0.058	0.019	0.035
Hw3-Pr2-E2F	0.028	0.031	0.050	0.050	0.055
Hw3-Pr2-EQN	0.026	0.035	0.050	0.053	0.055
Hw4-Pr4-E2F	0.018	0.070	0.045	0.037	0.063

Figure 9.1: Information gain heat map for midterm one problem one using the *category error grouping* scheme. Each element (i, j) is the information gain value from using feature i to predict error j . This heat-map is colored with a three-color gradient, which maps the lowest information gain value to red, highest information gain value to green, and intermediate information gain value to yellow. Each row is labeled by homework number, problem number and feature type, e.g., Hw3-Pr2-EQN corresponds to the amount of equation ink written on homework three problem two.

ErrorName	Alg Err	Ans Err	Fbd2 Err	Fbd1 Err	Eqn Err
Hw5-Pr8-EQN	0.010	0.018	0.015	0.018	0.017
Hw5-Pr8-E2F	0.019	0.020	0.011	0.020	0.017
Hw5-Pr8-FBD	0.023	0.024	0.017	0.029	0.017
Hw6-Pr5-FBD	0.012	0.020	0.017	0.034	0.037
Hw6-Pr3-FBD	0.014	0.044	0.021	0.026	0.017
Hw6-Pr2-FBD	0.015	0.040	0.011	0.031	0.026
Hw6-Pr2-EQN	0.017	0.020	0.024	0.040	0.023
Hw6-Pr1-EQN	0.039	0.018	0.009	0.051	0.015
Hw5-Pr1-E2F	0.018	0.020	0.024	0.015	0.056
Hw5-Pr5-FBD	0.025	0.027	0.020	0.019	0.043
Hw5-Pr4-E2F	0.012	0.022	0.021	0.038	0.042
Hw5-Pr2-EQN	0.028	0.014	0.035	0.020	0.039
Hw6-Pr3-E2F	0.033	0.044	0.026	0.012	0.027
Hw5-Pr2-FBD	0.020	0.012	0.035	0.053	0.030
Hw6-Pr2-E2F	0.012	0.020	0.020	0.021	0.078
Hw5-Pr4-FBD	0.012	0.043	0.017	0.032	0.048
Hw6-Pr4-FBD	0.021	0.029	0.021	0.061	0.022
Hw6-Pr3-EQN	0.020	0.044	0.032	0.029	0.031
Hw6-Pr1-FBD	0.027	0.036	0.007	0.060	0.029
Hw5-Pr7-E2F	0.019	0.025	0.014	0.049	0.053
Hw5-Pr7-EQN	0.010	0.018	0.025	0.055	0.053
Hw6-Pr5-E2F	0.018	0.019	0.048	0.024	0.055
Hw5-Pr7-FBD	0.022	0.040	0.025	0.034	0.051
Hw5-Pr5-EQN	0.025	0.027	0.047	0.029	0.043
Hw5-Pr6-E2F	0.019	0.011	0.025	0.031	0.088
Hw5-Pr2-E2F	0.037	0.041	0.035	0.023	0.039
Hw5-Pr4-EQN	0.014	0.024	0.051	0.031	0.057
Hw6-Pr6-FBD	0.019	0.033	0.038	0.047	0.049
Hw5-Pr3-FBD	0.030	0.019	0.059	0.019	0.063
Hw5-Pr1-FBD	0.030	0.009	0.034	0.051	0.066
Hw6-Pr5-EQN	0.018	0.027	0.059	0.041	0.067
Hw6-Pr1-E2F	0.064	0.055	0.037	0.045	0.016
Hw5-Pr1-EQN	0.018	0.012	0.069	0.057	0.063
Hw5-Pr3-E2F	0.040	0.027	0.047	0.015	0.090
Hw5-Pr5-E2F	0.039	0.027	0.047	0.019	0.089
Hw5-Pr3-EQN	0.047	0.020	0.048	0.027	0.080
Hw6-Pr4-EQN	0.037	0.020	0.020	0.061	0.089
Hw6-Pr6-EQN	0.022	0.029	0.064	0.033	0.080
Hw5-Pr6-EQN	0.019	0.013	0.025	0.079	0.106
Hw6-Pr4-E2F	0.022	0.025	0.053	0.061	0.090
Hw6-Pr6-E2F	0.017	0.015	0.089	0.033	0.096
Hw5-Pr6-FBD	0.022	0.057	0.031	0.068	0.089

Figure 9.2: Information gain heat map for midterm two problem one using the *category error grouping scheme*. Each element (i, j) is the information gain value from using feature i to predict error j . This heat-map is colored with a three-color gradient, which maps the lowest information gain value to red, highest information gain value to green, and intermediate information gain value to yellow. Each row is labeled by homework number, problem number and feature type, e.g., Hw3-Pr2-EQN corresponds to the amount of equation ink written on homework three problem two.

ErrorName	Fbd2 Err	FBD3 Err	Alg Err	Ans Err	Fbd1 Err	Eqn Err
Hw6-Pr5-E2F	0.019	0.014	0.022	0.024	0.018	0.020
Hw6-Pr6-FBD	0.021	0.020	0.040	0.016	0.013	0.017
Hw6-Pr5-EQN	0.018	0.020	0.020	0.039	0.011	0.020
Hw6-Pr2-FBD	0.018	0.014	0.017	0.032	0.040	0.019
Hw6-Pr4-FBD	0.020	0.024	0.027	0.032	0.021	0.017
Hw5-Pr5-FBD	0.022	0.014	0.017	0.023	0.025	0.042
Hw5-Pr5-E2F	0.016	0.023	0.028	0.019	0.027	0.032
Hw6-Pr3-EQN	0.016	0.020	0.015	0.049	0.021	0.024
Hw6-Pr4-E2F	0.020	0.025	0.030	0.036	0.021	0.017
Hw5-Pr5-EQN	0.020	0.023	0.026	0.016	0.027	0.038
Hw5-Pr4-FBD	0.032	0.014	0.018	0.031	0.029	0.026
Hw5-Pr4-E2F	0.032	0.020	0.016	0.031	0.030	0.027
Hw5-Pr8-E2F	0.014	0.035	0.032	0.024	0.029	0.023
Hw6-Pr3-E2F	0.016	0.033	0.018	0.049	0.021	0.020
Hw5-Pr2-EQN	0.032	0.014	0.020	0.023	0.055	0.014
Hw5-Pr8-EQN	0.020	0.030	0.032	0.026	0.017	0.033
Hw5-Pr2-E2F	0.032	0.014	0.028	0.023	0.053	0.011
Hw5-Pr7-FBD	0.013	0.010	0.040	0.018	0.035	0.048
Hw5-Pr8-FBD	0.032	0.028	0.032	0.032	0.018	0.024
Hw6-Pr2-EQN	0.016	0.014	0.017	0.032	0.058	0.031
Hw6-Pr1-E2F	0.031	0.021	0.021	0.040	0.023	0.033
Hw6-Pr2-E2F	0.021	0.026	0.022	0.032	0.045	0.025
Hw5-Pr4-EQN	0.033	0.016	0.016	0.031	0.052	0.027
Hw6-Pr3-FBD	0.016	0.020	0.040	0.060	0.021	0.017
Hw5-Pr6-EQN	0.016	0.028	0.042	0.023	0.039	0.029
Hw5-Pr3-FBD	0.026	0.040	0.016	0.031	0.023	0.042
Hw5-Pr7-E2F	0.016	0.016	0.030	0.031	0.068	0.017
Hw6-Pr1-EQN	0.022	0.030	0.017	0.040	0.023	0.048
Hw6-Pr5-FBD	0.059	0.014	0.024	0.032	0.016	0.041
Hw5-Pr3-E2F	0.033	0.040	0.022	0.031	0.026	0.037
Hw5-Pr6-E2F	0.039	0.022	0.042	0.019	0.039	0.029
Hw5-Pr3-EQN	0.028	0.040	0.016	0.031	0.018	0.059
Hw6-Pr6-EQN	0.026	0.022	0.047	0.041	0.040	0.020
Hw6-Pr6-E2F	0.026	0.042	0.031	0.060	0.018	0.021
Hw5-Pr2-FBD	0.032	0.042	0.020	0.023	0.045	0.041
Hw6-Pr1-FBD	0.032	0.029	0.016	0.063	0.018	0.045
Hw6-Pr4-EQN	0.020	0.024	0.022	0.070	0.026	0.042
Hw5-Pr7-EQN	0.048	0.028	0.030	0.017	0.068	0.024
Hw5-Pr6-FBD	0.016	0.029	0.042	0.017	0.023	0.093
Hw5-Pr1-E2F	0.025	0.044	0.047	0.016	0.054	0.041
Hw5-Pr1-EQN	0.016	0.043	0.047	0.016	0.054	0.082
Hw5-Pr1-FBD	0.032	0.046	0.047	0.016	0.063	0.134

Figure 9.3: Information gain results for midterm two problem two using the *category error grouping* scheme. Each element (i, j) is the information gain value from using feature i to predict error j . This heat-map is colored with a three-color gradient, which maps the lowest information gain value to red, highest information gain value to green, and intermediate information gain value to yellow. Each row is labeled by homework number, problem number and feature type, e.g., Hw3-Pr2-EQN corresponds to the amount of equation ink written on homework three problem two.

ErrorName	Alg Err	FBD3 Err	Ans Err	EQN Err	FBD1 Err	FBD2 Err
Hw5-Pr6-E2F	0.008	0.011	0.015	0.021	0.015	0.014
Hw5-Pr3-E2F	0.015	0.009	0.025	0.027	0.019	0.016
Hw5-Pr4-FBD	0.009	0.017	0.025	0.036	0.026	0.012
Hw5-Pr7-E2F	0.012	0.016	0.040	0.040	0.012	0.015
Hw6-Pr3-E2F	0.012	0.010	0.031	0.015	0.018	0.052
Hw6-Pr5-EQN	0.011	0.027	0.015	0.013	0.039	0.036
Hw6-Pr3-EQN	0.010	0.024	0.031	0.011	0.017	0.052
Hw5-Pr2-E2F	0.018	0.025	0.038	0.013	0.019	0.035
Hw6-Pr3-FBD	0.012	0.016	0.031	0.014	0.023	0.052
Hw5-Pr6-FBD	0.010	0.033	0.037	0.043	0.020	0.010
Hw5-Pr1-E2F	0.026	0.041	0.020	0.031	0.020	0.016
Hw6-Pr5-FBD	0.013	0.025	0.015	0.025	0.028	0.049
Hw6-Pr5-E2F	0.017	0.032	0.017	0.013	0.042	0.036
Hw5-Pr2-FBD	0.015	0.025	0.040	0.027	0.020	0.035
Hw5-Pr8EQN	0.006	0.025	0.047	0.042	0.014	0.029
Hw5-Pr3-FBD	0.018	0.013	0.052	0.032	0.026	0.024
Hw5-Pr3-EQN	0.018	0.024	0.047	0.032	0.031	0.016
Hw5-Pr5-FBD	0.012	0.024	0.063	0.048	0.020	0.007
Hw5-Pr7-EQN	0.011	0.038	0.040	0.052	0.020	0.016
Hw5-Pr6-EQN	0.022	0.067	0.022	0.023	0.027	0.016
Hw6-Pr1-FBD	0.025	0.016	0.037	0.033	0.024	0.044
Hw5-Pr4-EQN	0.023	0.016	0.020	0.055	0.022	0.045
Hw5-Pr2-EQN	0.015	0.025	0.039	0.024	0.045	0.035
Hw5-Pr8FBD	0.014	0.024	0.052	0.040	0.024	0.029
Hw6-Pr6-E2F	0.018	0.020	0.013	0.017	0.063	0.055
Hw5-Pr4-E2F	0.011	0.076	0.019	0.044	0.027	0.032
Hw5-Pr5-EQN	0.011	0.041	0.022	0.076	0.030	0.034
Hw5-Pr1-FBD	0.012	0.041	0.036	0.054	0.060	0.016
Hw5-Pr7-FBD	0.008	0.038	0.020	0.072	0.052	0.030
Hw5-Pr1-EQN	0.012	0.041	0.023	0.057	0.064	0.025
Hw6-Pr1-EQN	0.023	0.028	0.031	0.031	0.056	0.057
Hw5-Pr5-E2F	0.009	0.039	0.023	0.042	0.046	0.067
Hw6-Pr6-FBD	0.028	0.067	0.032	0.017	0.030	0.054
Hw5-Pr8E2F	0.013	0.047	0.047	0.040	0.048	0.036
Hw6-Pr2-FBD	0.057	0.015	0.031	0.031	0.017	0.100
Hw6-Pr1-E2F	0.020	0.050	0.031	0.035	0.066	0.050
Hw6-Pr4-FBD	0.030	0.023	0.047	0.043	0.080	0.048
Hw6-Pr2-EQN	0.046	0.028	0.031	0.035	0.037	0.100
Hw6-Pr2-E2F	0.022	0.038	0.040	0.046	0.050	0.100
Hw6-Pr6-EQN	0.032	0.091	0.037	0.032	0.061	0.072
Hw6-Pr4-E2F	0.012	0.029	0.047	0.098	0.104	0.068
Hw6-Pr4-EQN	0.013	0.059	0.047	0.071	0.104	0.079

Figure 9.4: Information gain heat map for midterm two problem three using the category error grouping scheme. Each element (i, j) is the information gain value from using feature i to predict error j . This heat-map is colored with a three-color gradient, which maps the lowest information gain value to red, highest information gain value to green, and intermediate information gain value to yellow. Each row is labeled by homework number, problem number and feature type, e.g., Hw3-Pr2-EQN corresponds to the amount of equation ink written on homework three problem two.

	M1P1	M2P1	M2P2	M2P3
	Hw3-Pr3- <i>Ink</i> _{E2F}	Hw5-Pr1- <i>Ink</i> _{EQN}	Hw6-Pr6- <i>Ink</i> _{EQN}	Hw6-Pr6- <i>Ink</i> _{FBD}
	Hw4-Pr7- <i>Ink</i> _{EQN}	Hw5-Pr3- <i>Ink</i> _{E2F}	Hw6-Pr6- <i>Ink</i> _{E2F}	Hw5-Pr8- <i>Ink</i> _{E2F}
	Hw4-Pr2- <i>Ink</i> _{EQN}	Hw5-Pr5- <i>Ink</i> _{E2F}	Hw5-Pr2- <i>Ink</i> _{FBD}	Hw6-Pr2- <i>Ink</i> _{FBD}
	Hw4-Pr3- <i>Ink</i> _{E2F}	Hw5-Pr3- <i>Ink</i> _{EQN}	Hw6-Pr1- <i>Ink</i> _{FBD}	Hw6-Pr1- <i>Ink</i> _{E2F}
	Hw4-Pr5- <i>Ink</i> _{EQN}	Hw6-Pr4- <i>Ink</i> _{EQN}	Hw6-Pr4- <i>Ink</i> _{EQN}	Hw6-Pr4- <i>Ink</i> _{FBD}
	Hw3-Pr4- <i>Ink</i> _{EQN}	Hw6-Pr6- <i>Ink</i> _{EQN}	Hw5-Pr7- <i>Ink</i> _{EQN}	Hw6-Pr2- <i>Ink</i> _{EQN}
	Hw4-Pr5- <i>Ink</i> _{E2F}	Hw5-Pr6- <i>Ink</i> _{EQN}	Hw5-Pr6- <i>Ink</i> _{FBD}	Hw6-Pr2- <i>Ink</i> _{E2F}
	Hw3-Pr2- <i>Ink</i> _{E2F}	Hw6-Pr4- <i>Ink</i> _{E2F}	Hw5-Pr1- <i>Ink</i> _{E2F}	Hw6-Pr6- <i>Ink</i> _{EQN}
	Hw3-Pr2- <i>Ink</i> _{EQN}	Hw6-Pr6- <i>Ink</i> _{E2F}	Hw5-Pr1- <i>Ink</i> _{EQN}	Hw6-Pr4- <i>Ink</i> _{E2F}
	Hw4-Pr4- <i>Ink</i> _{E2F}	Hw5-Pr6- <i>Ink</i> _{FBD}	Hw5-Pr1- <i>Ink</i> _{FBD}	Hw6-Pr4- <i>Ink</i> _{EQN}
No. <i>Ink</i> _{FBD}	0	1	2	3
No. <i>Ink</i> _{EQN}	5	5	4	3
No. <i>Ink</i> _{E2F}	5	4	2	4

Table 9.9: The 10 most predictive features for each midterm and problem, for the *category error grouping* scheme. The header row denotes the midterm and problem number, e.g., M1P1 corresponds to midterm one problem one. The last three rows count the number of features for each midterm problem that correspond to free body diagrams, equations, and the ratio of equation to free body diagram time respectively. Each feature is represented by three pieces of information: the homework number, problem number, and type, e.g., Hw3-Pr3-EQN corresponds to the amount of equation ink written for homework three problem three.

	M1P1	M2P1	M2P2	M2P3
	Hw3-Pr1- <i>Ink</i> _{EQN}	Hw5-Pr8- <i>Ink</i> _{EQN}	Hw6-Pr5- <i>Ink</i> _{E2F}	Hw5-Pr6- <i>Ink</i> _{E2F}
	Hw3-Pr8- <i>Ink</i> _{FBD}	Hw5-Pr8- <i>Ink</i> _{E2F}	Hw6-Pr6- <i>Ink</i> _{FBD}	Hw5-Pr3- <i>Ink</i> _{E2F}
	Hw3-Pr1- <i>Ink</i> _{FBD}	Hw5-Pr8- <i>Ink</i> _{FBD}	Hw6-Pr5- <i>Ink</i> _{EQN}	Hw5-Pr4- <i>Ink</i> _{FBD}
	Hw3-Pr8- <i>Ink</i> _{EQN}	Hw6-Pr5- <i>Ink</i> _{FBD}	Hw6-Pr2- <i>Ink</i> _{FBD}	Hw5-Pr7- <i>Ink</i> _{E2F}
	Hw3-Pr1- <i>Ink</i> _{E2F}	Hw6-Pr3- <i>Ink</i> _{FBD}	Hw6-Pr4- <i>Ink</i> _{FBD}	Hw6-Pr3- <i>Ink</i> _{E2F}
	Hw3-Pr5- <i>Ink</i> _{FBD}	Hw6-Pr2- <i>Ink</i> _{FBD}	Hw5-Pr5- <i>Ink</i> _{FBD}	Hw6-Pr5- <i>Ink</i> _{EQN}
	Hw3-Pr5- <i>Ink</i> _{E2F}	Hw6-Pr2- <i>Ink</i> _{EQN}	Hw5-Pr5- <i>Ink</i> _{E2F}	Hw6-Pr3- <i>Ink</i> _{EQN}
	Hw3-Pr8- <i>Ink</i> _{E2F}	Hw6-Pr1- <i>Ink</i> _{EQN}	Hw6-Pr3- <i>Ink</i> _{EQN}	Hw5-Pr2- <i>Ink</i> _{E2F}
	Hw4-Pr7- <i>Ink</i> _{FBD}	Hw5-Pr1- <i>Ink</i> _{E2F}	Hw6-Pr4- <i>Ink</i> _{E2F}	Hw6-Pr3- <i>Ink</i> _{FBD}
	Hw4-Pr1- <i>Ink</i> _{FBD}	Hw5-Pr5- <i>Ink</i> _{FBD}	Hw5-Pr5- <i>Ink</i> _{EQN}	Hw5-Pr6- <i>Ink</i> _{FBD}
No. FBD	5	5	4	3
No. EQN	2	3	3	2
No. E2F	3	2	3	5

Table 9.10: The 10 least predictive features for each midterm and problem, for the *category error grouping* categories. The header row denotes the midterm and problem number, e.g., M1P1 corresponds to midterm one problem one. The last three rows count the number of features for each midterm problem that correspond to free body diagrams, equations, and the ratio of equation to free body diagram time respectively. Each feature is represented by three pieces of information: the homework number, problem number, and type, e.g., Hw3-Pr3-EQN corresponds to the amount of equation ink written for homework three problem three.

9.3.2 Fine-grained Error Grouping

In this section we show the information gain heat maps created for midterm one problem one and midterm two problems one through three using the *fine-grained error grouping* consolidation scheme. These heat maps are shown in Figure 9.5 through Figure 9.8.

Additionally, we highlight the first and last ten features (rows) of each heat map. The first ten rows in the matrix are those features with the smallest sum of information gain values. These features can be considered the 10 least consistently predictive features and will be important for the discussion that follows. Similarly, the last ten rows of the matrix are those features with the largest sum of information gain values. These features can be considered the 10 most consistently predictive features and will be important for the discussion that follows. The top and 10 least predictive features from each heat map are shown in Table 9.9 and Table 9.10 respectively. These tables also count the number of each feature type found in the top-10 list for each midterm problem. This gives an indication of how often the Ink_{EQN} , Ink_{FBD} , and Ink_{E2F} features are the most predictive for errors on a particular midterm problem.

9.3.3 Subcategory Error Grouping

In this section we show the information gain heat maps created for midterm one problem one and midterm two problems one through three using the *subcategory*



Figure 9.5: Information gain heat maps for midterm one problem one predicting the *fine-grained error grouping* scheme. Each element (i, j) is the information gain value from using feature i to predict error j . This heat-map is colored with a three-color gradient, which maps the lowest information gain value to red, highest information gain value to green, and intermediate information gain value to yellow. Each row is labeled by homework number, problem number and feature type, e.g., Hw3-Pr2-EQN corresponds to the amount of equation ink written on homework three problem two.

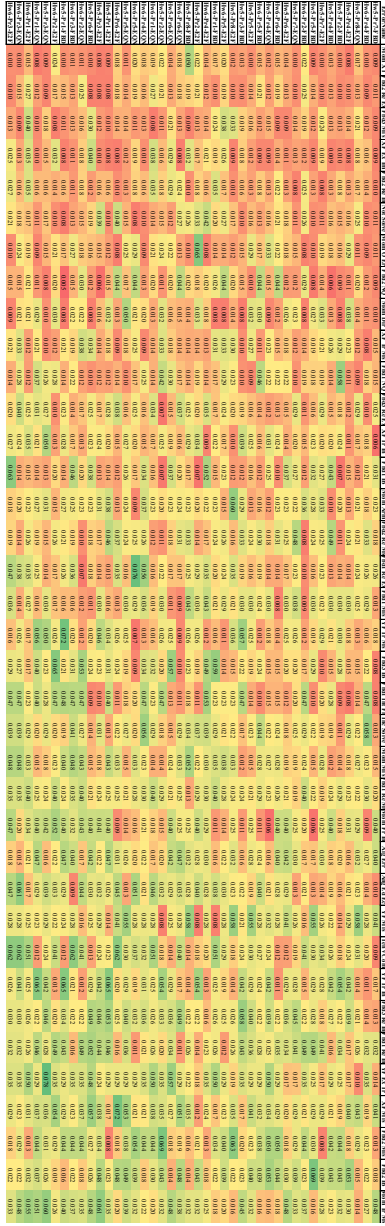


Figure 9.7: Information gain heat maps for midterm two problem two predicting the *fine-grained error grouping* categories. Each element (i, j) is the information gain value from using feature i to predict error j . This heat-map is colored with a three-color gradient, which maps the lowest information gain value to red, highest information gain value to green, and intermediate information gain value to yellow. Each row is labeled by homework number, problem number and feature type, e.g., Hw3-Pr2-EQN corresponds to the amount of equation ink written on homework three problem two.



Figure 9.8: Information gain heat maps for midterm two problem three predicting the *fine-grained error grouping* categories. Each element (i, j) is the information gain value from using feature i to predict error j . This heat-map is colored with a three-color gradient, which maps the lowest information gain value to red, highest information gain value to green, and intermediate information gain value to yellow. Each row is labeled by homework number, problem number and feature type, e.g., Hw3-Pr2-EQN corresponds to the amount of equation ink written on homework three problem two.

	M1P1	M2P1	M2P2	M2P3
	Hw3-Pr4- <i>Ink_{E2F}</i>	Hw5-Pr1- <i>Ink_{E2F}</i>	Hw5-Pr1- <i>Ink_{EQN}</i>	Hw5-Pr3- <i>Ink_{FBD}</i>
	Hw4-Pr1- <i>Ink_{E2F}</i>	Hw5-Pr2- <i>Ink_{E2F}</i>	Hw5-Pr1- <i>Ink_{FBD}</i>	Hw5-Pr8- <i>Ink_{E2F}</i>
	Hw4-Pr2- <i>Ink_{EQN}</i>	Hw5-Pr2- <i>Ink_{EQN}</i>	Hw5-Pr3- <i>Ink_{EQN}</i>	Hw5-Pr8- <i>Ink_{FBD}</i>
	Hw4-Pr3- <i>Ink_{E2F}</i>	Hw5-Pr3- <i>Ink_{E2F}</i>	Hw5-Pr6- <i>Ink_{FBD}</i>	Hw6-Pr1- <i>Ink_{EQN}</i>
	Hw4-Pr3- <i>Ink_{EQN}</i>	Hw5-Pr3- <i>Ink_{EQN}</i>	Hw6-Pr1- <i>Ink_{E2F}</i>	Hw6-Pr4- <i>Ink_{E2F}</i>
	Hw4-Pr5- <i>Ink_{E2F}</i>	Hw5-Pr6- <i>Ink_{EQN}</i>	Hw6-Pr2- <i>Ink_{E2F}</i>	Hw6-Pr4- <i>Ink_{EQN}</i>
	Hw4-Pr5- <i>Ink_{EQN}</i>	Hw5-Pr6- <i>Ink_{FBD}</i>	Hw6-Pr2- <i>Ink_{EQN}</i>	Hw6-Pr4- <i>Ink_{FBD}</i>
	Hw4-Pr6- <i>Ink_{E2F}</i>	Hw5-Pr7- <i>Ink_{FBD}</i>	Hw6-Pr3- <i>Ink_{E2F}</i>	Hw6-Pr6- <i>Ink_{E2F}</i>
	Hw4-Pr6- <i>Ink_{EQN}</i>	Hw6-Pr1- <i>Ink_{E2F}</i>	Hw6-Pr3- <i>Ink_{EQN}</i>	Hw6-Pr6- <i>Ink_{EQN}</i>
	Hw4-Pr6- <i>Ink_{FBD}</i>	Hw6-Pr4- <i>Ink_{E2F}</i>	Hw6-Pr3- <i>Ink_{FBD}</i>	Hw6-Pr6- <i>Ink_{FBD}</i>
No. FBD	1	2	3	4
No. EQN	4	3	4	3
No. E2F	5	5	3	3

Table 9.11: The 10 most predictive features for each midterm and problem, for the *fine-grained error groupings*. The header row denotes the midterm and problem number, e.g., M1P1 corresponds to midterm one problem one. The last three rows count the number of features for each midterm problem that correspond to free body diagrams, equations, and the ratio of equation to free body diagram time respectively. Each feature is represented by three pieces of information: the homework number, problem number, and type, e.g., Hw3-Pr3-EQN corresponds to the amount of equation ink written for homework three problem three.

	M1P1	M2P1	M2P2	M2P3
	Hw3-Pr1- <i>Ink_{E2F}</i>	Hw5-Pr4- <i>Ink_{FBD}</i>	Hw5-Pr2- <i>Ink_{FBD}</i>	Hw5-Pr1- <i>Ink_{E2F}</i>
	Hw3-Pr1- <i>Ink_{EQN}</i>	Hw5-Pr8- <i>Ink_{E2F}</i>	Hw5-Pr4- <i>Ink_{E2F}</i>	Hw5-Pr2- <i>Ink_{FBD}</i>
	Hw3-Pr1- <i>Ink_{FBD}</i>	Hw5-Pr8- <i>Ink_{EQN}</i>	Hw5-Pr4- <i>Ink_{FBD}</i>	Hw5-Pr4- <i>Ink_{EQN}</i>
	Hw3-Pr5- <i>Ink_{FBD}</i>	Hw5-Pr8- <i>Ink_{FBD}</i>	Hw5-Pr5- <i>Ink_{FBD}</i>	Hw5-Pr4- <i>Ink_{FBD}</i>
	Hw3-Pr6- <i>Ink_{E2F}</i>	Hw6-Pr2- <i>Ink_{E2F}</i>	Hw5-Pr7- <i>Ink_{E2F}</i>	Hw5-Pr6- <i>Ink_{E2F}</i>
	Hw3-Pr6- <i>Ink_{EQN}</i>	Hw6-Pr2- <i>Ink_{FBD}</i>	Hw5-Pr7- <i>Ink_{EQN}</i>	Hw5-Pr6- <i>Ink_{FBD}</i>
	Hw3-Pr6- <i>Ink_{FBD}</i>	Hw6-Pr3- <i>Ink_{E2F}</i>	Hw5-Pr7- <i>Ink_{FBD}</i>	Hw5-Pr7- <i>Ink_{FBD}</i>
	Hw3-Pr7- <i>Ink_{E2F}</i>	Hw6-Pr5- <i>Ink_{E2F}</i>	Hw5-Pr8- <i>Ink_{EQN}</i>	Hw6-Pr1- <i>Ink_{E2F}</i>
	Hw3-Pr8- <i>Ink_{FBD}</i>	Hw6-Pr5- <i>Ink_{EQN}</i>	Hw6-Pr2- <i>Ink_{FBD}</i>	Hw6-Pr3- <i>Ink_{E2F}</i>
	Hw4-Pr2- <i>Ink_{FBD}</i>	Hw6-Pr5- <i>Ink_{FBD}</i>	Hw6-Pr4- <i>Ink_{FBD}</i>	Hw6-Pr3- <i>Ink_{FBD}</i>
No. FBD	5	4	6	5
No. EQN	2	2	2	1
No. E2F	3	4	2	4

Table 9.12: The 10 least predictive features for each midterm and problem, for the *fine-grained error groupings*. The header row denotes the midterm and problem number, e.g., M1P1 corresponds to midterm one problem one. The last three rows count the number of features for each midterm problem that correspond to free body diagrams, equations, and the ratio of equation to free body diagram time respectively. Each feature is represented by three pieces of information: the homework number, problem number, and type, e.g., Hw3-Pr3-EQN corresponds to the amount of equation ink written for homework three problem three.

Error/Share	YE-MIS	XE-MIS	REA-ERR	IME-BUD	MAE-ERR	RMD-NQ	NUM-MIS	SUM-ERR	MAE-ERR	YE-ERR	AE-ERR	XE-ERR
Hw3-Pr1-EQN	0.014	0.026	0.077	0.029	0.019	0.011	0.019	0.030	0.030	0.017	0.016	0.028
Hw4-Pr2-E2E	0.012	0.019	0.024	0.035	0.022	0.024	0.021	0.023	0.035	0.025	0.013	0.012
Hw4-Pr1-EQN	0.010	0.008	0.046	0.033	0.018	0.042	0.015	0.015	0.032	0.027	0.025	0.017
Hw3-Pr1-E2E	0.023	0.026	0.015	0.029	0.033	0.024	0.013	0.023	0.036	0.030	0.013	0.028
Hw3-Pr1-BBD	0.009	0.026	0.015	0.029	0.017	0.013	0.021	0.030	0.024	0.029	0.018	0.067
Hw3-Pr4-BBD	0.024	0.063	0.015	0.020	0.015	0.015	0.027	0.043	0.016	0.017	0.019	0.032
Hw3-Pr7-E2E	0.012	0.029	0.017	0.014	0.031	0.028	0.021	0.023	0.031	0.021	0.059	0.023
Hw3-Pr2-BBD	0.021	0.026	0.015	0.014	0.026	0.033	0.038	0.035	0.030	0.026	0.017	0.018
Hw3-Pr5-E2E	0.012	0.030	0.024	0.016	0.022	0.021	0.022	0.035	0.030	0.031	0.041	0.039
Hw3-Pr6-E2E	0.030	0.034	0.023	0.017	0.015	0.018	0.021	0.047	0.020	0.027	0.055	0.017
Hw3-Pr7-EQN	0.025	0.031	0.047	0.016	0.015	0.011	0.016	0.023	0.031	0.020	0.059	0.032
Hw3-Pr8-EQN	0.030	0.038	0.035	0.033	0.025	0.025	0.018	0.030	0.024	0.027	0.014	0.029
Hw4-Pr4-EQN	0.021	0.008	0.017	0.023	0.017	0.035	0.013	0.011	0.045	0.025	0.063	0.051
Hw4-Pr2-E2E	0.014	0.016	0.021	0.023	0.025	0.013	0.064	0.045	0.025	0.018	0.038	0.030
Hw4-Pr1-E2E	0.015	0.023	0.046	0.033	0.022	0.018	0.024	0.045	0.024	0.026	0.027	0.029
Hw3-Pr5-BBD	0.014	0.028	0.032	0.027	0.028	0.010	0.024	0.035	0.024	0.031	0.031	0.032
Hw3-Pr8-E2E	0.030	0.040	0.006	0.033	0.041	0.029	0.034	0.039	0.024	0.036	0.014	0.012
Hw3-Pr7-BBD	0.023	0.023	0.031	0.019	0.015	0.032	0.021	0.036	0.042	0.024	0.052	0.029
Hw3-Pr5-EQN	0.012	0.031	0.029	0.016	0.032	0.018	0.026	0.035	0.030	0.031	0.045	0.040
Hw4-Pr7-EQN	0.047	0.022	0.019	0.018	0.013	0.018	0.019	0.033	0.067	0.018	0.038	0.039
Hw4-Pr4-BBD	0.043	0.019	0.044	0.020	0.025	0.043	0.023	0.015	0.017	0.039	0.029	0.036
Hw3-Pr4-EQN	0.027	0.039	0.018	0.026	0.019	0.048	0.013	0.026	0.030	0.029	0.052	0.034
Hw3-Pr6-BBD	0.029	0.024	0.012	0.030	0.020	0.024	0.038	0.052	0.039	0.028	0.032	0.032
Hw4-Pr2-BBD	0.043	0.032	0.037	0.020	0.025	0.043	0.024	0.043	0.019	0.029	0.030	0.032
Hw3-Pr2-E2E	0.010	0.021	0.025	0.025	0.031	0.034	0.030	0.037	0.050	0.031	0.055	0.021
Hw3-Pr6-EQN	0.029	0.034	0.054	0.016	0.015	0.031	0.022	0.039	0.020	0.027	0.055	0.031
Hw3-Pr3-E2E	0.027	0.034	0.024	0.029	0.031	0.013	0.045	0.035	0.026	0.031	0.052	0.035
Hw3-Pr4-BBD	0.016	0.016	0.032	0.030	0.028	0.018	0.044	0.047	0.048	0.044	0.043	0.030
Hw4-Pr1-BBD	0.033	0.022	0.030	0.042	0.027	0.016	0.037	0.029	0.032	0.032	0.023	0.020
Hw4-Pr3-E2E	0.035	0.022	0.044	0.020	0.049	0.043	0.023	0.030	0.021	0.039	0.038	0.032
Hw3-Pr3-BBD	0.012	0.034	0.024	0.016	0.022	0.038	0.054	0.035	0.028	0.031	0.052	0.050
Hw4-Pr4-BBD	0.023	0.012	0.030	0.020	0.025	0.043	0.013	0.029	0.039	0.025	0.052	0.032
Hw4-Pr2-EQN	0.027	0.056	0.024	0.030	0.043	0.056	0.077	0.029	0.009	0.025	0.027	0.097
Hw4-Pr6-BBD	0.010	0.012	0.027	0.046	0.052	0.027	0.029	0.025	0.032	0.074	0.025	0.056
Hw4-Pr4-E2E	0.036	0.016	0.017	0.023	0.070	0.035	0.021	0.017	0.045	0.023	0.063	0.051
Hw3-Pr3-EQN	0.040	0.039	0.026	0.029	0.028	0.032	0.043	0.035	0.028	0.031	0.052	0.035
Hw3-Pr4-E2E	0.023	0.032	0.024	0.028	0.025	0.038	0.040	0.047	0.045	0.031	0.050	0.035
Hw3-Pr2-EQN	0.030	0.024	0.015	0.023	0.035	0.055	0.030	0.037	0.050	0.031	0.055	0.036
Hw3-Pr2-BBD	0.024	0.011	0.021	0.051	0.032	0.065	0.024	0.026	0.051	0.037	0.043	0.055
Hw4-Pr5-BBD	0.040	0.015	0.037	0.064	0.038	0.024	0.034	0.030	0.036	0.038	0.027	0.058
Hw4-Pr5-EQN	0.031	0.018	0.037	0.069	0.070	0.037	0.032	0.023	0.021	0.042	0.021	0.064
Hw4-Pr6-E2E	0.013	0.014	0.046	0.036	0.052	0.050	0.047	0.026	0.024	0.070	0.027	0.066
Hw4-Pr6-EQN	0.015	0.024	0.018	0.036	0.052	0.058	0.047	0.026	0.024	0.070	0.031	0.079
Hw3-Pr4-EQN	0.025	0.028	0.026	0.017	0.033	0.047	0.093	0.049	0.067	0.031	0.050	0.066
Hw4-Pr5-E2E	0.055	0.029	0.037	0.069	0.072	0.047	0.040	0.035	0.058	0.021	0.035	0.065

Figure 9.9: Information gain heat map for midterm one problem one predicting the *subcategory error grouping* categories. Each element (i, j) is the information gain value from using feature i to predict error j . This heat-map is colored with a three-color gradient, which maps the lowest information gain value to red, highest information gain value to green, and intermediate information gain value to yellow. Each row is labeled by homework number, problem number and feature type, e.g., Hw3-Pr2-EQN corresponds to the amount of equation ink written on homework three problem two.

error grouping consolidation scheme. These heat maps are shown in Figure 9.9 through Figure 9.12.

Additionally, we highlight the first and last ten features (rows) of each heat map. The first ten rows in the matrix are those features with the smallest sum of information gain values. These features can be considered the 10 least consistently

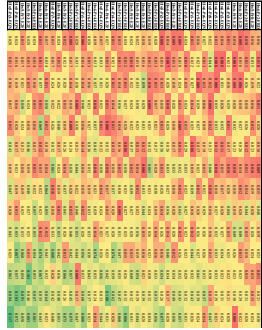


Figure 9.10: Information gain heat map for midterm two problem one predicting the *subcategory error grouping* categories. Each element (i, j) is the information gain value from using feature i to predict error j . This heat-map is colored with a three-color gradient, which maps the lowest information gain value to red, highest information gain value to green, and intermediate information gain value to yellow. Each row is labeled by homework number, problem number and feature type, e.g., Hw3-Pr2-EQN corresponds to the amount of equation ink written on homework three problem two.



Figure 9.11: Information gain heat map for midterm two problem two predicting the *subcategory error grouping* categories. Each element (i, j) is the information gain value from using feature i to predict error j . This heat-map is colored with a three-color gradient, which maps the lowest information gain value to red, highest information gain value to green, and intermediate information gain value to yellow. Each row is labeled by homework number, problem number and feature type, e.g., Hw3-Pr2-EQN corresponds to the amount of equation ink written on homework three problem two.

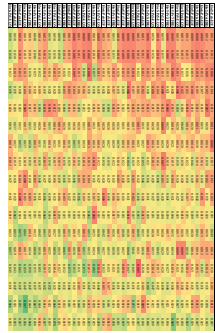


Figure 9.12: Information gain heat map for midterm two problem three predicting the *subcategory error grouping* categories. Each element (i, j) is the information gain value from using feature i to predict error j . This heat-map is colored with a three-color gradient, which maps the lowest information gain value to red, highest information gain value to green, and intermediate information gain value to yellow. Each row is labeled by homework number, problem number and feature type, e.g., Hw3-Pr2-EQN corresponds to the amount of equation ink written on homework three problem two.

predictive features and will be important for the discussion that follows. Similarly, the last ten rows of the matrix are those features with the largest sum of information gain values. These features can be considered the 10 most consistently predictive features and will be important for the discussion that follows. The top and 10 least predictive features from each heat map are shown in Table 9.13 and Table 9.14 respectively. These tables also count the number of each feature type found in the 10 most predictive list for each midterm problem. This gives an indication of how often the Ink_{EQN} , Ink_{FBD} , and Ink_{E2F} features are the most predictive for errors on a particular midterm problem.

9.4 Guided Qualitative Analysis

Using the results from category error grouping seen in the previous section, we are able to provide an analysis of the qualities of the homework problems that were

	M1P1	M2P1	M2P2	M2P3
	Hw3-Pr2- <i>Ink_{EQN}</i>	Hw5-Pr2- <i>Ink_{E2F}</i>	Hw5-Pr1- <i>Ink_{EQN}</i>	Hw5-Pr1- <i>Ink_{EQN}</i>
	Hw3-Pr2- <i>Ink_{FBD}</i>	Hw5-Pr3- <i>Ink_{E2F}</i>	Hw5-Pr1- <i>Ink_{FBD}</i>	Hw5-Pr3- <i>Ink_{FBD}</i>
	Hw3-Pr3- <i>Ink_{EQN}</i>	Hw5-Pr3- <i>Ink_{EQN}</i>	Hw5-Pr3- <i>Ink_{E2F}</i>	Hw5-Pr8- <i>Ink_{E2F}</i>
	Hw3-Pr4- <i>Ink_{E2F}</i>	Hw5-Pr6- <i>Ink_{E2F}</i>	Hw6-Pr2- <i>Ink_{E2F}</i>	Hw5-Pr8- <i>Ink_{FBD}</i>
	Hw3-Pr4- <i>Ink_{EQN}</i>	Hw5-Pr6- <i>Ink_{EQN}</i>	Hw6-Pr2- <i>Ink_{EQN}</i>	Hw6-Pr1- <i>Ink_{EQN}</i>
	Hw4-Pr5- <i>Ink_{E2F}</i>	Hw5-Pr6- <i>Ink_{FBD}</i>	Hw6-Pr3- <i>Ink_{E2F}</i>	Hw6-Pr1- <i>Ink_{FBD}</i>
	Hw4-Pr5- <i>Ink_{EQN}</i>	Hw5-Pr7- <i>Ink_{E2F}</i>	Hw6-Pr3- <i>Ink_{FBD}</i>	Hw6-Pr4- <i>Ink_{E2F}</i>
	Hw4-Pr5- <i>Ink_{FBD}</i>	Hw5-Pr7- <i>Ink_{EQN}</i>	Hw6-Pr4- <i>Ink_{EQN}</i>	Hw6-Pr4- <i>Ink_{EQN}</i>
	Hw4-Pr6- <i>Ink_{E2F}</i>	Hw6-Pr1- <i>Ink_{E2F}</i>	Hw6-Pr5- <i>Ink_{E2F}</i>	Hw6-Pr4- <i>Ink_{FBD}</i>
	Hw4-Pr6- <i>Ink_{EQN}</i>	Hw6-Pr4- <i>Ink_{E2F}</i>	Hw6-Pr6- <i>Ink_{E2F}</i>	Hw6-Pr6- <i>Ink_{EQN}</i>
No. FBD	2	1	2	4
No. EQN	5	3	3	4
No. E2F	3	6	5	2

Table 9.13: The 10 most predictive features for each midterm and problem, for the *subcategory error grouping* scheme. The header row denotes the midterm and problem number, e.g., M1P1 corresponds to midterm one problem one. The last three rows sum the number of features for each midterm problem that correspond to free body diagrams, equations, and the ratio of equation to free body diagram time respectively. Each feature is represented by three pieces of information: the homework number, problem number, and type, e.g., Hw3-Pr3-EQN corresponds to the amount of equation ink written for homework three problem three.

	M1P1	M2P1	M2P2	M2P3
	Hw3-Pr1- <i>Ink_{E2F}</i>	Hw5-Pr8- <i>Ink_{E2F}</i>	Hw5-Pr2- <i>Ink_{E2F}</i>	Hw5-Pr1- <i>Ink_{E2F}</i>
	Hw3-Pr1- <i>Ink_{EQN}</i>	Hw5-Pr8- <i>Ink_{EQN}</i>	Hw5-Pr2- <i>Ink_{FBD}</i>	Hw5-Pr2- <i>Ink_{FBD}</i>
	Hw3-Pr1- <i>Ink_{FBD}</i>	Hw6-Pr2- <i>Ink_{E2F}</i>	Hw5-Pr3- <i>Ink_{FBD}</i>	Hw5-Pr3- <i>Ink_{E2F}</i>
	Hw3-Pr5- <i>Ink_{E2F}</i>	Hw6-Pr2- <i>Ink_{EQN}</i>	Hw5-Pr4- <i>Ink_{E2F}</i>	Hw5-Pr4- <i>Ink_{EQN}</i>
	Hw3-Pr6- <i>Ink_{E2F}</i>	Hw6-Pr2- <i>Ink_{FBD}</i>	Hw5-Pr4- <i>Ink_{FBD}</i>	Hw5-Pr4- <i>Ink_{FBD}</i>
	Hw3-Pr7- <i>Ink_{E2F}</i>	Hw6-Pr3- <i>Ink_{EQN}</i>	Hw5-Pr6- <i>Ink_{E2F}</i>	Hw5-Pr6- <i>Ink_{E2F}</i>
	Hw3-Pr8- <i>Ink_{FBD}</i>	Hw6-Pr3- <i>Ink_{FBD}</i>	Hw5-Pr7- <i>Ink_{FBD}</i>	Hw5-Pr6- <i>Ink_{FBD}</i>
	Hw4-Pr1- <i>Ink_{EQN}</i>	Hw6-Pr5- <i>Ink_{E2F}</i>	Hw5-Pr8- <i>Ink_{EQN}</i>	Hw5-Pr7- <i>Ink_{FBD}</i>
	Hw4-Pr2- <i>Ink_{E2F}</i>	Hw6-Pr5- <i>Ink_{EQN}</i>	Hw6-Pr1- <i>Ink_{EQN}</i>	Hw6-Pr3- <i>Ink_{E2F}</i>
	Hw4-Pr2- <i>Ink_{FBD}</i>	Hw6-Pr6- <i>Ink_{EQN}</i>	Hw6-Pr4- <i>Ink_{FBD}</i>	Hw6-Pr5- <i>Ink_{EQN}</i>
No. FBD	3	2	5	4
No. EQN	2	5	2	2
No. E2F	5	3	3	4

Table 9.14: The 10 least predictive features for each midterm and problem, for the *subcategory error grouping*. The header row denotes the midterm and problem number, e.g., M1P1 corresponds to midterm one problem one. The last three rows sum the number of features for each midterm problem that correspond to free body diagrams, equations, and the ratio of equation to free body diagram time respectively. Each feature is represented by three pieces of information: the homework number, problem number, and type, e.g., Hw3-Pr3-EQN corresponds to the amount of equation ink written for homework three problem three.

found to most consistently predict whether or not students would make particular errors.

9.4.1 Midterm One Problem One

Looking at the midterm one, problem one (M1P1) column of Table 9.9 we can see that features from certain homework problems appear more than once. This indicates that effort spent on those problems may better predict performance on this midterm problem. To further examine this, we show the problem description for problem one of midterm one in Figure 9.13.

Next we examined each problem-homework pairs that appeared in the M1P1 column more than once, namely, homework three problem two and homework four problem five, whose problem descriptions are shown in Figure 9.14 and Figure 9.15 respectively.

The problems which appeared more than once in the 10 most predictive list for midterm one problem one which were: Homework three problem two, and homework four problem five. The features computed from these problems that appeared in the 10 most predictive list were both Ink_{E2F} and Ink_{EQN} . This indicates that the amount of equation ink written and the ratio of equation ink to FBD ink were the most predictive features for errors made on midterm one problem one.

Homework three problem two is nearly identical to a rotated version of the midterm problem, as they are both effectively a single body lever. Thus, it makes sense that the amount of time spent on this problem would be the best predictor of whether or not students made errors on the midterm problem.

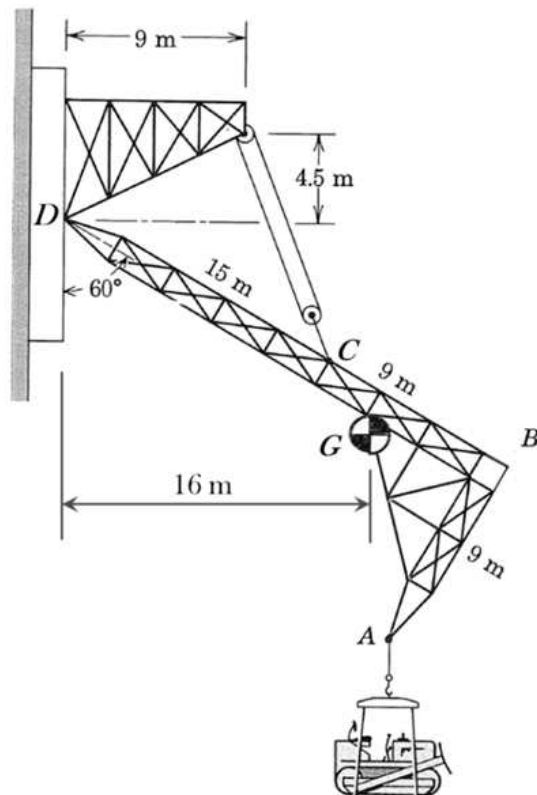


Figure 9.13: Problem description for problem one of midterm one which reads, "The crane is hoisting a 4000 kg bulldozer. The mass center of the 2000 kg boom is located at G. The system is in equilibrium in the configuration shown. In your analysis, neglect the width of the boom. a) Draw a large, clearly-labeled free body diagram. b) Determine the tension T in the cable where it attaches at C. c) Determine the magnitude of the force applied to the boom at its hinge D."

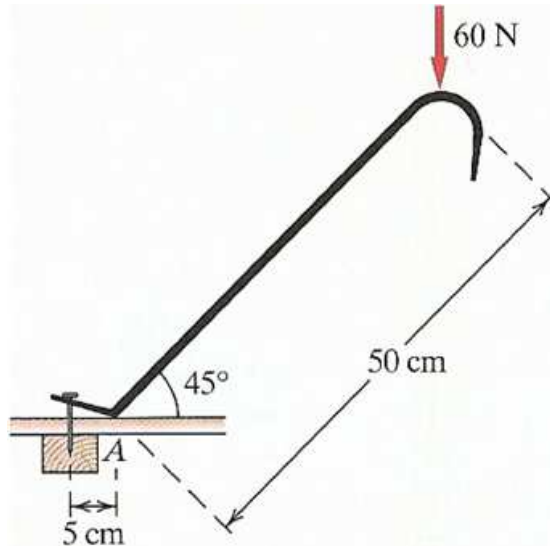


Figure 9.14: Problem description for problem two of homework three which reads, "A vertical force of 60 N is applied to one end of a crowbar to pull a nail from a floor. With the crowbar show in the figure, find the force applied to the floor at A. How does the answer change if the 60 N force is applied at the same location perpendicular to the crowbar?"

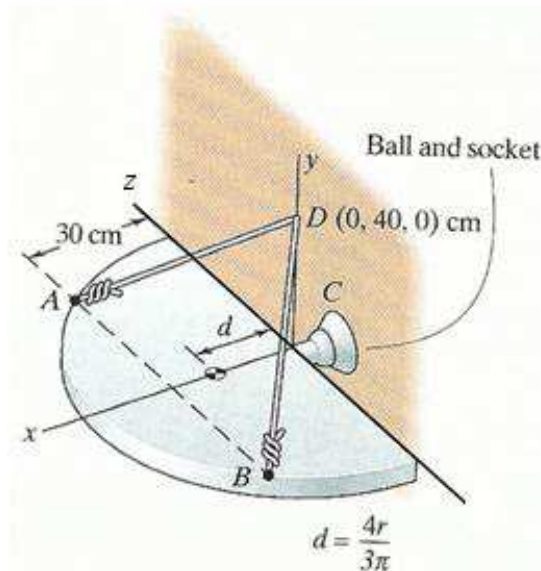


Figure 9.15: Problem description for problem five of homework four which reads, "Consider the semi-circular plate shown. The plate has a 40 cm radius and weighs 100 N, with center of gravity at $(d, 0, 0)$. Determine the tension in the cables and at A and B and the loads acting on the plate at C, which is a ball-and-socket connection. Note that $d = (4r)/(3\pi)$."

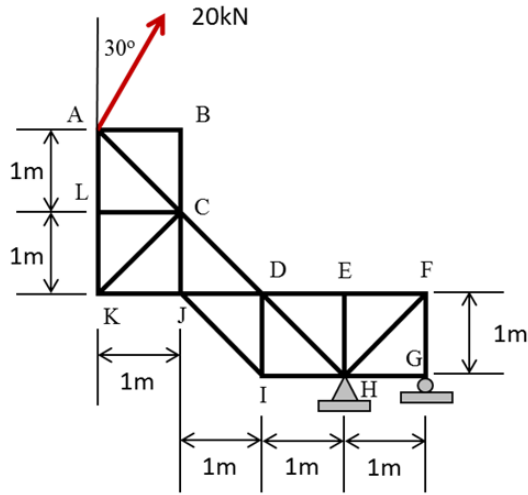


Figure 9.16: Problem description for problem one of midterm two which reads, "Determine the forces in members CD, JI, and DI and indicate if they are in tension or compression."

Homework four problem five has very little in common with the midterm problem on the surface. Upon further inspection though, it turns out that there is an important geometry calculation that must be done in order to successfully solve homework four problem five. This becomes more interesting, when, we notice that homework four problem five is most predictive of whether or not a student will make a geometry error on the midterm problem. This indicates that this problem may lead students to be more proficient with their geometry calculations.

Furthermore, all features from both homework three problem one and homework three problem eight appeared in the list of 10 least predictive features. This indicates that these homework assignments may be ineffective and should be modified in the subsequent quarter.

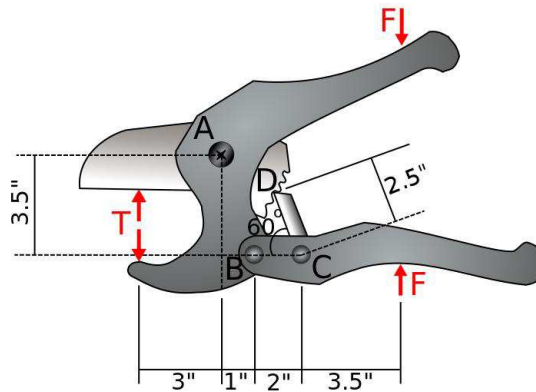


Figure 9.17: Problem description for problem three of homework five which reads, "The device shown is used for cutting PVC pipe. If a force, $F = 15 \text{ lb}$, is applied to each handle as shown, determine the cutting force T . Also, determine the magnitude and the direction of the force that the pivot at A applies to the blade."

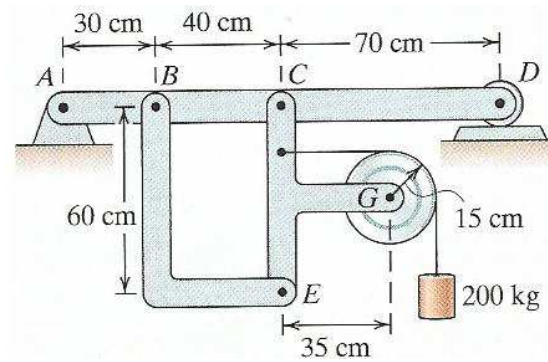


Figure 9.18: Problem description for problem six of homework five which reads, "Determine the forces acting on each of the members, and on the frame at A and D ."

9.4.2 Midterm Two Problem One

Next we examined each problem-homework pair that appeared in the M2P1 column more than once, namely: homework five problem three, homework five problem six, homework six problem four, homework six problem six, shown in Figure 9.17, Figure 9.18, Figure 9.19, and Figure 9.20 respectively.

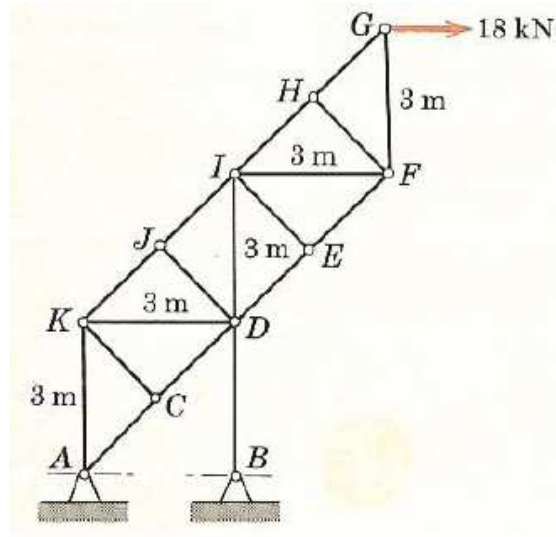


Figure 9.19: Problem description for problem four of homework six which reads, "Use the method of sections to calculate the forces in members DI, DE and EI for the loaded truss shown. State if the members are in tension or compression."

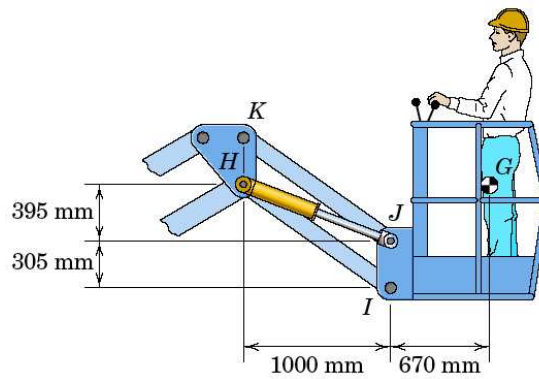


Figure 9.20: Problem description for problem six of homework six which reads, "The figure shows a portion of a "cherry-picker," a machine used to lift workers to elevated locations. The worker and the bucket have a combined mass of 200 kg with mass center at G. Determine the forces in KJ, HJ, and HI and indicate if they are in compression or tension."

There were four problems which showed up more than once in the 10 most predictive list for midterm two problem one. Homework five problem three appeared with Ink_{E2F} and Ink_{EQN} features. Homework five problem six appeared with Ink_{EQN} and Ink_{FBD} features. Homework six problem four appeared with Ink_{E2F} and Ink_{EQN} features. Homework six problem six appeared with Ink_{E2F} and Ink_{EQN} features.

Homework five problem three has little in common with midterm two problem one. The major overlap between the two problems is that homework five problem three has a two force member that students must properly model in order to correctly solve. The body in midterm two problem one is comprised entirely of two force members. Furthermore, Homework five problem three is one of the first opportunities students have to solve a problem with a two force member in it. This may indicate that students who spent more time to learn to solve problems with two force members as soon as they were available, are better equipped later on to solve similar problems.

Similarly to homework five problem three, homework five problem six has little in common with midterm two problem one. The most apparent similarity between these problems is the types of joints present in each: both have exactly one pin (element A in homework five problem six and element H in midterm two problem one) and one roller joint (element D in homework five problem six and element G in midterm two problem one).

In contrast with the previous two problems, homework six problem four has a great deal in common with midterm two problem one; in fact, the solution path for each is nearly identical.

Lastly, homework six problem six shares a great deal in common with midterm

two problem one even though the former is a machine problem and the latter is a truss problem.

Interestingly, all the problems identified as the most predictive of success on this midterm problem share one important trait in common that may explain why they were the most predictive. Midterm two problem one is best solved using the “method of sections” which requires students to decompose the truss into sections by “cutting” through elements. Solving a truss problem in this way requires that a student have a stronger mastery of concepts than the alternative solution approach, called “method of joints.” These homework problems are all similar in that they require students to carefully identify their solution path before attempting to solve the problem, knowing the order in which components and equations will be solved ahead of time. This foresight is a crucial component of the “method of sections” approach to solving; this may indicate that students who spend more time on problems which require them to better plan ahead in their solution will perform better on this midterm problem.

Furthermore, all features from homework five, problem eight appeared in the list of 10 least predictive features. This indicates that this homework assignment problem may be ineffective and should be modified or replaced in the subsequent quarter.

9.4.3 Midterm Two Problem Two

Next we examined each problem-homework pair that appeared in the M2P2 column more than once, namely, homework five problem one and homework six problem six, whose problem descriptions are shown in Figure 9.22 and Figure 9.20 respectively.

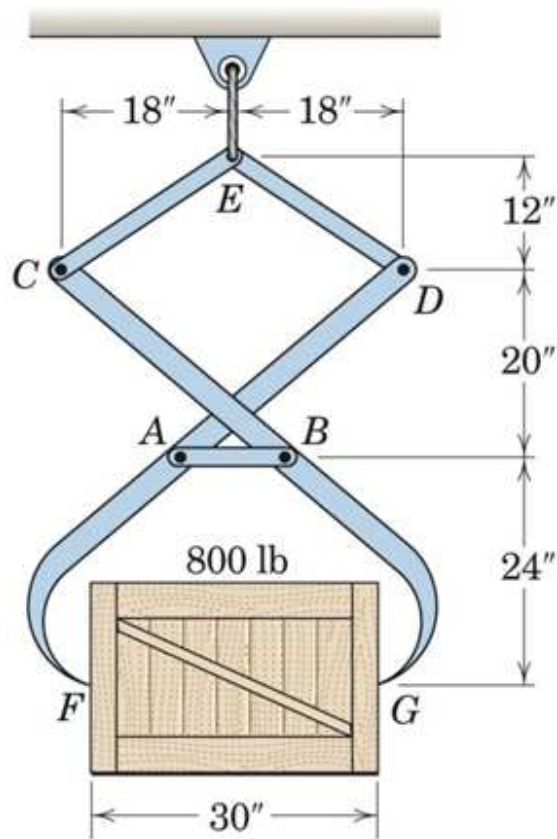


Figure 9.21: Problem description for problem two of midterm two which reads, "The 800-lb crate is held in equilibrium by the lifting device. (a) Determine the forces in members CE and ED and indicate if they are in tension or compression. (b) Determine the force in member AB and indicate if it is in tension or compression. "

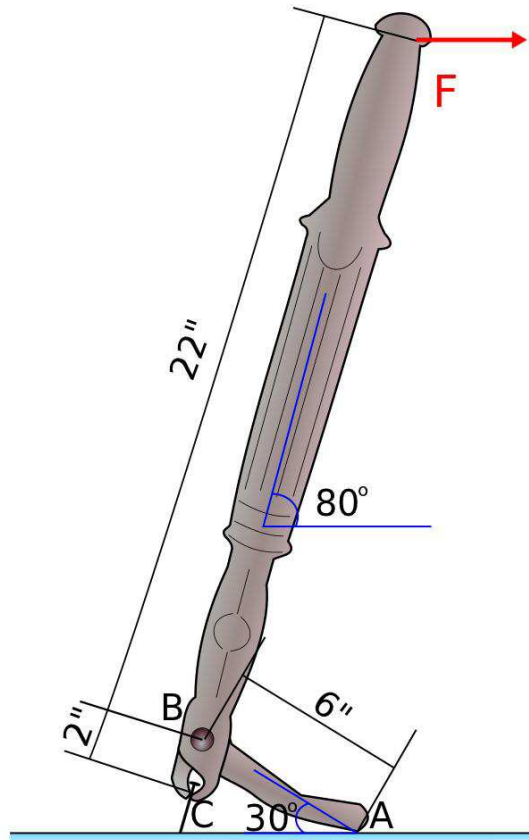


Figure 9.22: Problem description for problem one of homework five which reads, "The device shown is used to remove nails. When a force is applied to the handle, the device grips the nail and extracts it from the board. In this case, a horizontal force $F = 10$ lb is required to extract the nail. (a) Determine the axial force applied to the nail by the puller. (b) Determine the clamping force that each jaw of the puller applies to the nail. (c) Determine the magnitude and direction of the reaction force at A."

There were two problems whose features appeared more than once in the 10 most predictive feature list for midterm two problem two. Homework five problem one appeared with Ink_{EQN} and Ink_{FBD} features. Homework six problem six appeared with Ink_{E2F} and Ink_{EQN} features.

Homework five problem one two important elements in common with midterm two problem two. First, both problems requires students to correctly model two force members in order to solve the problem. Secondly, and more interestingly, homework five problem one requires students to solve for the clamping force of the jars on the nail at C; midterm two problem two similarly requires students to solve for the clamping force of the jaws on the box. This may indicate that students who spend time on problems with two force members and clamping forces may do better on this midterm problem.

Homework six problem six shares a great deal in common with midterm two problem two; even though the figures are quite different looking, the solution path for both is quite similar.

Furthermore, all features from homework five problem five appeared in the list of 10 least predictive features. This indicates that this homework assignment problem may be ineffective and should be modified or replaced in the subsequent course offering.

9.4.4 Midterm Two Problem Three

Next we examined each problem-homework pair that appeared in the M2P3 column more than once, namely, homework six problem two, homework six problem four, and homework six problem six, shown in Figure 9.24, Figure 9.19 and Figure 9.20

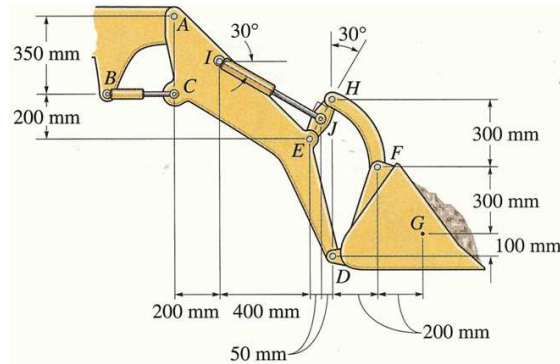


Figure 9.23: Problem description for problem three of midterm two which reads, "The tractor shovel carries a 600-kg load of soil, having a center of gravity at G. $EJ = 100\text{mm}$. $JH = 100\text{mm}$. (a) Identify all of the bodies that are two-force members. List them by the points on the bodies, e.g., DFG. (b) Determine the force in hydraulic cylinder BC and indicate if it is in tension or compression. (c) Determine the magnitude of the force acting on the shovel at point F. (d) Determine the force in hydraulic cylinder IJ and indicate if it is in tension or compression."

respectively.

There were three problems whose features appeared more than once in the list of the 10 most predictive features for midterm two problem three. Homework six problem two appeared with Ink_{E2F} , Ink_{EQN} , and Ink_{FBD} features. Homework six problem four appeared with Ink_{E2F} , Ink_{EQN} , and Ink_{FBD} features. Homework six

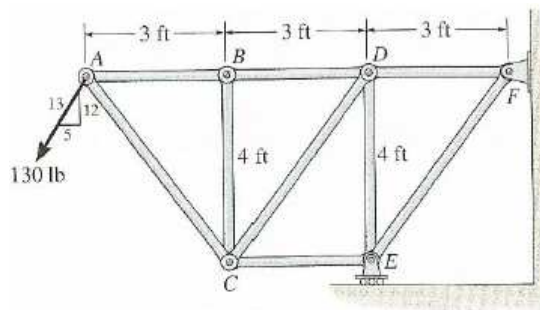


Figure 9.24: Problem description for problem two of homework six which reads, "Use the method of joints to determine the force in each member of the truss. State if the members are in tension or compression."

problem six appeared with Ink_{EQN} and Ink_{FBD} features.

In all cases, these homework problems are most effectively solved using the method of sections, as is the case with midterm two problem three. This may indicate that students who spend more time on homework assignments which lend themselves to being solved via the method of sections are less likely to make errors on this midterm problem.

9.4.5 Common Qualities of Predictive Homework Problems

In comparing each of the homework problems identified as being most predictive on each midterm problem, we find two interesting patterns worth discussing.

First, homework six problem six appears as the most predictive problem for every one of the midterm two problems. This is a clear indication that this problem is effective in teaching students important statics concepts. This problem had the most in common with each of the midterm problems, especially since the solution paths of homework six problem six and each of the midterm problems were very similar.

Secondly, in comparing each of the prompts of the identified homework problems, nearly all of them require students to solve for most if not all of the elements present in the figure. This is a contrast to most homework assignment problems which require students to only solve for a single element in the problem. This, is an exciting discovery as it suggests that students in this class should be presented with problems that require them to identify all unknown forces in a problem, instead of having them only solve for a single unknown.

9.5 Conclusion

In this work, we have presented an automated technique for identifying which homework problems assigned to students in a Statics course were most predictive of students performance on midterm exam problems. Using semantically labeled pen stroke data from students homework solutions, we computed the amount of ink written on each problem corresponding to either free body diagrams and equations. We used these amounts as simple features which were used to predict which errors students would make on each problem of their midterms. We did this by using each feature to compute the amount of information gained about a particular error using the information gain algorithm. We then created a information gain heat map which, for each midterm problem, contained the amount of information gained by each feature for each error. We sorted this matrix by the sum of the row and column values, resulting in a matrix in which the most predictive features appeared at the bottom of the matrix and the least predictive appeared at the top. Using these matrices, we identified the homework assignment problems which were most predictive of performance on a midterm problem as the homework assignments whose features appeared more than twice in the top feature list for that midterm problem. This top feature list guided a qualitative analysis, in which the most predictive homework problems were manually compared to the midterm problem so as to understand why each homework problem was most predictive of errors on that exam problem.

This analysis technique has several implications both within the Statics course for which it was developed as well as for the Educational Data Mining community.

Within the context of the Statics course, we have identified the homework problems which are most and least predictive of performance on midterm problems. By comparing these homework problems, we are able to identify the types of transfer students likely made when solving the midterm exam. This knowledge may be used to develop new, similar homework problems that will lead to the same types of transfer and hopefully be even more predictive of students performance next year.

The results of this study present an easy-to-read digest form of the amount of effort students spent on each assignment and resulting performance. This technique will be valuable in future years in which the digital pens are distributed to students to provide the instructor with formative assessments (after each midterm and each quiz) that can be used to guide both the current course and help develop the next one.

More importantly, this is a generalizable technique that is of interest to the entire educational data mining community. Any instructor capable of providing his or her students with digital pens with which those students may complete their homework assignments, may apply this technique to better understand the relationship between the homework assignments students complete and their performance. More importantly, the results of this technique are applicable not only to that instructor but also to any instructor in that subject matter.

Chapter 10

Machine Learning Based Sketch Processing Techniques

10.1 Introduction

In this chapter, we present techniques used to process low-level representations of handwritten pen strokes. In particular we present three techniques: one to segment pen strokes, one to recognize single-stroke gestures, and one to align each pen stroke to the speech with which it is associated. These techniques are all important first steps in larger systems which accept users handwritten sketching as input. We have found, as we will show in the subsequent related work sections, that prior approaches typically rely on heuristic-based approaches. It is the goal of our work, presented in this chapter, to show the effectiveness of general and extensible machine learning based approaches to replace these prior approaches. First we present the 1^c recognizer, an easy-to-implement gesture recognition technique which is enabled by a simple nearest neighbor search. Next

we present ClassySeg, which employs a statistical classifier to detect intended points of segmentation in hand written pen strokes. Lastly we present a machine learning based stroke-speech alignment technique. This technique is intended for multimodal (speech and sketch) interfaces, and identifies aligns speech with the pen strokes to which it refers.

10.2 User Study Data Set

We used the pen stroke data from the study described in [59] to evaluate the performance of 1^c and benchmark it against \$1. In that study, an HP TC4400 Tablet PC with a digitizer resolution of 1024x768 pixels was used for data collection. The participants were asked to draw each of the ten symbols in Figure 10.1 18 times. Thus, there were 180 strokes drawn by each participant, yielding 2,520 strokes in total.

The participants were informed that the purpose of the study was to collect data to evaluate the performance of an algorithm. Participants were instructed to “draw naturally with ordinary accuracy,” and to not attempt to “trick or break the system.” Before beginning the exercise, each participant was given a few minutes to practice drawing on the Tablet PC.

Each point from the data set is a triple containing an x-coordinate, y-coordinate, and time value. To facilitate training and testing of algorithms, the true segment points on each pen stroke were manually labeled using an approach analogous to that in [132]. Specifically, the pen stroke data was initially segmented using the segmentation technique described in [114]. Then segment points were manually added, deleted, and moved as necessary to achieve the correct segmentation. Additionally, each gesture was manu-

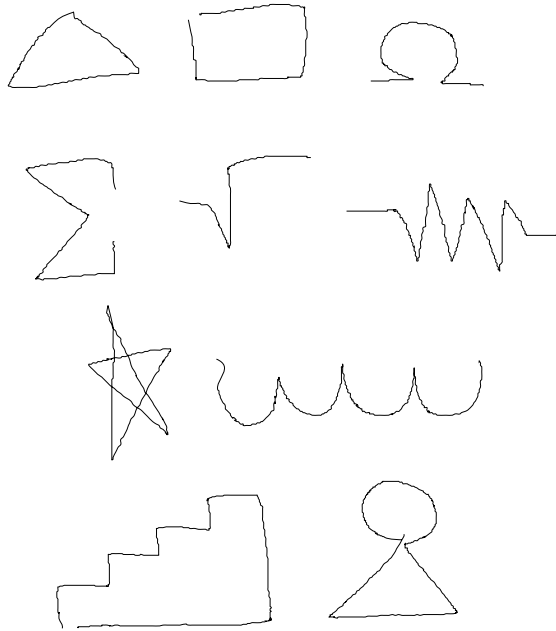


Figure 10.1: The ten shapes used to evaluate our sketch processing techniques: triangle, rectangle, omega, sigma, square root, resistor, star, inductor, stoop, not.

ally labeled according to its shape.

10.3 The 1¢ Recognizer

10.3.1 Introduction

Gesture recognition is a primary enabling technology in pen-based computer interfaces and sketch understanding¹. Gesture recognition is the process by which a user's handwritten pen stroke is recognized as being one of a number of predefined gesture types.

As we show in the related work section, numerous handwritten gesture recognition techniques have been developed. Arguably, the most popular of these is the Dollar

¹This work originally appeared in [60]

Recognizer (\$1). This aptly named technique’s popularity can be attributed to its ease of implementation, efficiency, and high performance.

In this work, we improve on \$1 by introducing the One Cent Recognizer (1[¢]), a technique that requires less code, is more efficient, and performs at least as well as \$1. This technique is enabled by a novel transformation of raw pen stroke data into a one-dimensional feature vector. A simple time series classification technique can then be used to find a feature vector’s closest match in a training corpus. This novel transformation is intrinsically rotation invariant, eliminating the need for rotate-and-check searches typically employed in previous approaches, resulting in a substantial speed advantage for the 1[¢] technique.

We perform two rigorous train-and-test cross-validation schemes to evaluate the efficiency and accuracy of 1[¢] and \$1. The first is a user-dependent scheme, which demonstrates each recognizer’s performance when the testing and training data are both generated by the same participant. The second is a user-independent scheme, which demonstrates each recognizer’s performance when the testing data comes from a participant who generated none of the training data. These tests show that 1[¢] is always at least as accurate as \$1, and is always two orders of magnitude faster.

In the following section we discuss related work in handwritten gesture recognition as well as time series classification techniques that serve as inspiration for 1[¢]. In Section 10.3.3 we provide the algorithmic details of our technique. The results of the user-dependent and user-independent evaluations are presented in Section 10.3.4. We discuss the results and present conclusions in Section 10.3.5 and Section 10.3.6. Lastly, in the Appendix, we present concise pseudocode for 1[¢] that can be used to quickly im-

plement this technique. Additionally, the appendix contains the URL to an open-source implementation of 1^c.

10.3.2 Related Work

Wobbrock et al. [129] developed \$1, which is so called because of its ease of implementation; the pseudocode for \$1 contains only 72 functioning lines. This handwritten gesture recognition technique comprises two major steps: transformation and recognition. In the first step, a raw stroke is transformed into a template. This transformation begins by resampling a stroke to a fixed number of points such that the distances between all successive pairs of points are equal. Next, the resampled stroke is rotated so that its indicative angle lies at 0°. Lastly, the points are scaled to a fixed-size square and the points are translated so that their centroid is at the origin. These transformations effectively normalize the pen stroke so that comparisons are scale and position invariant. In the second step, the templated candidate stroke is compared to a number of training templates in order to find its best match. In order to be rotation invariant, \$1 employs golden section search during this step to determine the best angular alignment of the candidate with each template.

Rubine [106] developed one of the earliest handwritten gesture recognizers in his seminal paper. This technique begins by representing a pen stroke using 13 geometric features, such as the sum of the curvature values at each point of the stroke, the angle at the first point of the gesture, and the size of the stroke's bounding box. Gestures are recognized using a simple linear classifier, which is trained on these features. While this approach achieved good accuracy, it was not rotation invariant.

The Image-Based Recognizer [72] applies popular image recognition techniques. The Image-Based Recognizer converts each candidate stroke into a fixed-size binary bitmap. The distance between two bitmaps is computed using four traditional bitmap distances, e.g., the Hausdorff distance. Similar to §1, when a candidate template is compared to a training template, a search is employed to find the angular orientation that minimizes the distance to that training template. While this technique is typically more accurate than §1, it is much more complex in its implementation and is far less efficient; operations on bitmaps are typically $O(N^2)$.

These techniques have several features in common. Most importantly, they are all template-based. The unknown candidate must be compared to each of the templates in the training corpus to determine the best match. Also, §1 and Rubine’s method can recognize only single-stroke gestures. For example, they cannot recognize the letter “t” if it is drawn with two strokes.

While the Image-Based recognizer can recognize multi-stroke symbols, the set of strokes comprising the symbol must be distinguished from the other strokes in the sketch by some other process. Sezgin and Davis [108] developed a technique which uses Hidden Markov Models (HMMs) to identify all of the symbols in a sketch without preprocessing. By considering sequences of pen strokes, the HMMs naturally identify multi-stroke symbols. This approach begins by processing strokes using the Toolkit from [109]. This toolkit segments pen strokes into their constituent geometric primitives, such as lines and arcs. The segmented output is then converted into a sequence of discrete observations. A separate HMM is trained for each symbol class. Candidate sequences are classified according to the HMM that best recognizes the candidate sequence, i.e.,

returns the highest probability using the Forward-Backward algorithm.

While machine learning approaches such as linear classifiers and HMMs, are frequently used for sketch understanding techniques, time series classification techniques [22, 49, 22] have not been widely used. This is surprising as pen strokes are a time series of (x,y) values. Xi et al. [134] have shown that simple first-nearest-neighbor (1NN) Euclidean distance and Dynamic Time Warping (DTW) approaches are often more accurate than more complex time series classification techniques, with DTW being the more accurate of the two. Tappert [121] used DTW to implement a handwriting recognition technique which was used as a benchmark for \$1. It was found that \$1 typically performed as well as DTW and was much more efficient and easier to implement. However, DTW is much more expensive than Euclidean 1NN approaches. Our goal is develop an efficient gesture recognizer using a simple, Euclidean 1NN approach.

The key to using a Euclidean 1NN approach is developing a one-dimensional time series that allows for accurate recognition. The representation used in this chapter is based on that found in [135] which was used to recognize digitized hieroglyphs. We have significantly modified this representation to handle handwritten gestures. The hieroglyphs always formed a closed contour, which is rarely the case for handwritten gestures. Additionally, while the hieroglyph bitmaps contain no timing information, handwritten gestures do, and this may be leveraged to achieve greater accuracy. Furthermore, the one-dimensional hieroglyph representation is intrinsically sensitive to angular orientation, and thus, just as with \$1, search is required to determine the best angular orientation. As we show in Section 10.3.3, our representation is rotation invariant, eliminating the need for such a search.

10.3.3 Approach

1^c comprises two major steps. In the first step, the raw pen stroke data is transformed into a fixed-length, one-dimensional vector which we call a template. In the second step, distances are computed between the candidate and training templates to find the closest match.

10.3.3.1 Resample

A simple Euclidean distance comparison of two time series requires that both series be the same length. Thus, the first step in templating a pen stroke is to resample it to a fixed number of points, N , such that the distance between all pairs of successive points is equal. We accomplish this using the same approach used in [129]. To be consistent with the implementation of §1 we use $N = 64$ in this chapter, although as we show in the results section, any value of N between 16 and 128 is acceptable.

10.3.3.2 One Dimensional Representation

Next, we transform the three dimensional $(\mathbf{x}, \mathbf{y}, \mathbf{t})$ data of each stroke into a one-dimensional vector, \mathbf{d} . We begin by computing the centroid, c , of the points comprising the stroke, i.e., (μ_x, μ_y) . We then compute the Euclidean distance from c to each resampled point in the stroke. The resulting series of distances, \mathbf{d} , ordered by time, characterizes the stroke:

$$\mathbf{d} = \{d_1, \dots, d_N\} \text{ s.t. } d_i = \|p_i - c\| \quad (10.1)$$

Because this representation is based on the centroid of the gesture, it is intrinsically rotation invariant. Consider the example in Figure 10.2: a rotation transformation does not change a point’s distance to the centroid. This allows our technique to avoid expensive search that other methods require to find the optimal angular alignment. Additionally, this transformation is intrinsically position invariant, allowing for strokes drawn anywhere on a page to be compared to each other.

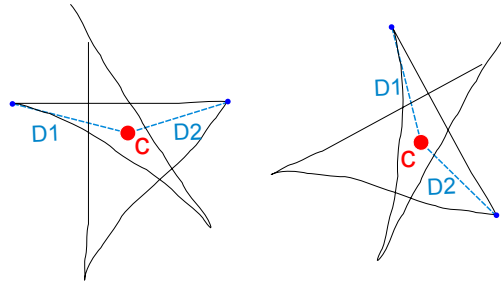


Figure 10.2: Rotation invariance of our one-dimensional pen stroke representation. Rotating the star gesture does not change the distance from the centroid (point “C”) to points on the stroke such as points D1 and D2.

Next, we z -normalize \mathbf{d} to ensure scale invariance:

$$\mathbf{z} = \{z_1, \dots, z_N\} \text{ s.t. } z_i = \frac{d_i - \mu_d}{\sigma_d} \quad (10.2)$$

where μ_d is the average of the stroke’s d_i values, and σ_d is the standard deviation.

This z -normalized distance vector, \mathbf{z} , is the template used to represent the candidate pen stroke. Figure 10.3 illustrates each of the major steps of this templating process.

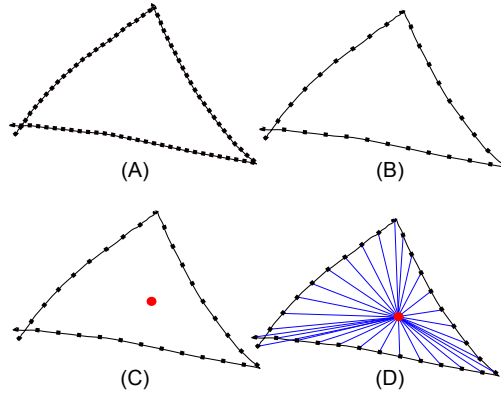


Figure 10.3: The templating process. The raw candidate pen stroke (A) is first resampled into N points (B). Next the centroid (red point) of the resampled points is computed (C) and the distance from it to every resampled point on the stroke is computed (blue lines). This sequence of distances is z-normalized and the resulting vector is the template for the candidate stroke.

10.3.3.3 Comparing Templates

Because all distance vectors are the same length and z-normalized, they may be directly compared. We use the L^2 distance to compare vectors. More formally, the distance between two vectors, \mathbf{v}_1 and \mathbf{v}_2 is defined as:

$$L^2(\mathbf{v}_1, \mathbf{v}_2) = \|\mathbf{v}_1 - \mathbf{v}_2\|^2 \quad (10.3)$$

Using this distance, we apply a simple 1NN approach to recognize gestures. A templated candidate (unknown) stroke, U , is compared to each training template, T , in the training set. The matching training template, T^* , is the one which minimizes the L^2 distance:

$$T^* = \arg \min_T L^2(U, T) \quad (10.4)$$

n	\$1 Tot.	1 ^c Tot.	\$1 Ave.	1 ^c Ave.
1	5.20	0.16	2.22	0.07
2	10.20	0.21	4.35	0.09
3	15.22	0.25	6.50	0.11
4	20.07	0.30	8.57	0.13
5	25.28	0.38	10.80	0.16
6	29.96	0.41	12.80	0.17
7	35.14	0.47	15.01	0.20
8	40.39	0.54	17.26	0.23
9	45.53	0.57	19.45	0.24

Table 10.1: Computation times for \$1 and 1^c as the number of examples (n) of each gesture used for training is varied. The second and third columns are the total time required by \$1 and 1^c to recognize all 2340 test gestures (seconds). The last two columns are the average times to recognize a single gesture (milliseconds).

10.3.4 Results

We present here an analysis of the efficiency and accuracy of both \$1 and 1^c. Two separate validation schemes were used, a user-independent scheme and a user-dependent one. In each fold of the user-independent scheme, n examples of each gesture from one participant were used for training, and all gestures from the other participants were used for testing. As there were 14 participants, we performed 14-folds of cross-validation for varying values of n . In each case, 2,340 gestures (13 participants x 180 gestures) were used for testing. The accuracy achieved by both recognizers for varying values of n are shown in Figure 10.4. Recognition timing results are shown in Table 10.1.

In the user-dependent scheme, a separate cross-validation is performed for each participant, with testing and training data from only that participant. We performed

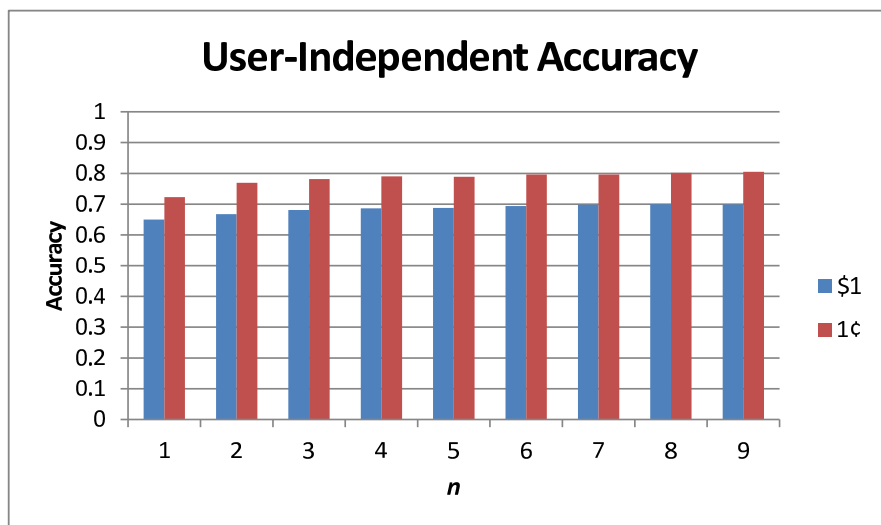


Figure 10.4: Average accuracy as a function of the number of training examples (n) for each gesture type for both \$1 and 1¢. In each case the difference between the groups is significant ($p < 0.01$)

18-folds of cross-validation for each participant. In each fold, one example of each gesture type is used for training and the others are used for testing (each gesture was used for training exactly one time). Because each participant drew 180 strokes, there were 170 strokes tested in each fold. Performing 18 folds resulted in 3,060 test recognitions. We averaged the accuracy for each participant across their 18-folds, producing the results in Figure 10.5.

Figure 10.6 presents the accuracy for both \$1 and 1¢ for each gesture type. Here accuracy is computed with the user-independent scheme with nine training templates for each gesture type ($n = 9$). The per-shape confusion matrix of the user-independent evaluation for 1¢ and \$1 are shown in Table 10.2 and Table 10.3 respectively. These tables show how often each shape was recognized as each other shape, which is useful

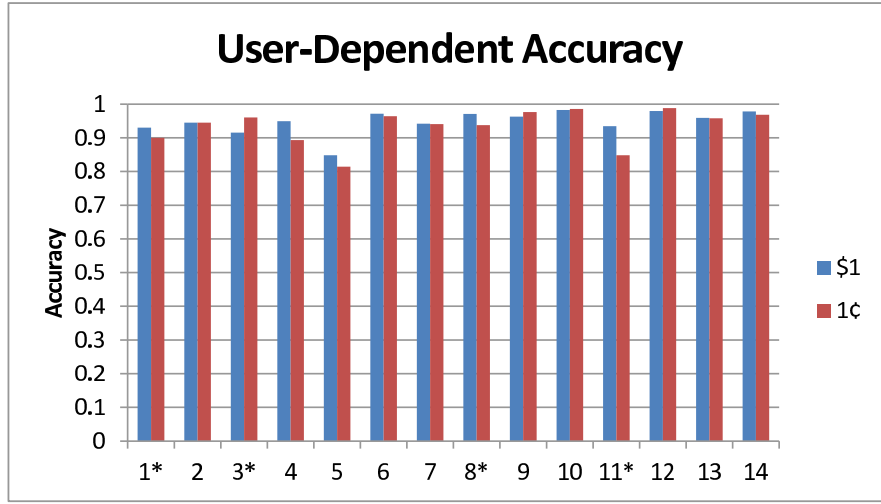


Figure 10.5: Accuracy results for each participant under the user-dependent scheme. Accuracy for a participant is averaged across 18-folds of cross-validation. An asterisk indicates that the difference in accuracies of the two methods for a particular participant is statistically significant ($p < 0.05$).

	Induc.	Not	Omega	Rect.	Resist.	Sigma	Sqrt	Star	Stoop	Triang.
Induc.	3245	0	0	0	5	0	26	0	0	0
Not	0	2089	96	435	0	133	22	51	161	289
Omega	3	3	2791	0	238	0	37	0	73	131
Rect.	0	144	33	2822	0	0	5	149	84	39
Resist.	319	0	0	0	2952	0	5	0	0	0
Sigma	153	0	3	0	4	3073	41	0	0	2
Sqrt	533	0	2	0	27	2	2712	0	0	0
Star	0	59	15	325	34	5	0	2776	43	19
Stoop	0	111	1052	374	0	186	11	53	1192	297
Triang.	0	143	182	45	0	104	0	7	63	2732

Table 10.2: Per-shape confusion matrix of 1^c on the user-independent scheme.

in understanding cases where the two recognizers perform poorly.

Table 10.4 presents the accuracy of each method across varying values of N . The difference between the accuracy of 1^c for $N = 16$ and $N = 128$ is not significant ($p = 0.61$). Similarly, the difference between the accuracy of \$1 for $N = 16$ and $N = 128$ is not significant ($p = 0.94$). Additionally, the difference in accuracy of the two methods

	Induc.	Not	Omega	Rect.	Resist.	Sigma	Sqrt	Star	Stoop	Triang.
Induc.	3043	0	0	0	27	42	164	0	0	0
Not	33	1217	60	861	23	503	210	59	126	184
Omega	82	36	2755	305	0	13	53	18	1	13
Rect.	0	3	214	2185	0	333	38	0	253	250
Resist.	154	0	0	0	3109	12	1	0	0	0
Sigma	5	1	0	0	0	3260	9	0	0	1
Sqrt	150	73	118	118	0	169	2635	0	0	13
Star	1	680	34	10	191	319	7	1997	25	12
Stoop	0	6	509	1144	0	213	13	6	798	587
Triang.	0	1	264	444	0	338	5	0	316	1908

Table 10.3: Per-shape confusion matrix of \$1 on the user-independent scheme.

	\$1	1 ^c	<i>p</i>
<i>N</i> = 16	94.8	94.2	0.68
<i>N</i> = 32	94.8	93.8	0.55
<i>N</i> = 64	94.7	93.4	0.43
<i>N</i> = 128	94.7	93.2	0.38

Table 10.4: Average accuracy of each technique across varying values of *N*, the resampling size, using the user-dependent method. The *p* column represents the p-value of a student t-test of the accuracies of the two techniques.

for each value of *N* is never significant ($p > 0.38$). These results demonstrate that the accuracy of either method is insensitive to this parameter and thus any value within this range is acceptable.

We repeated all of the above analyses, rotating the testing templates before recognition for increasing angles of rotation. We found in all such experiments that the rotation angle never affected the accuracy of either \$1 or 1^c, demonstrating a theoretical rotation invariance. A user study in which participants are asked to draw symbols at varying angles would be required to determine if, in practice, these methods were intolerant of additional transformations caused by drawing gestures at an angle.

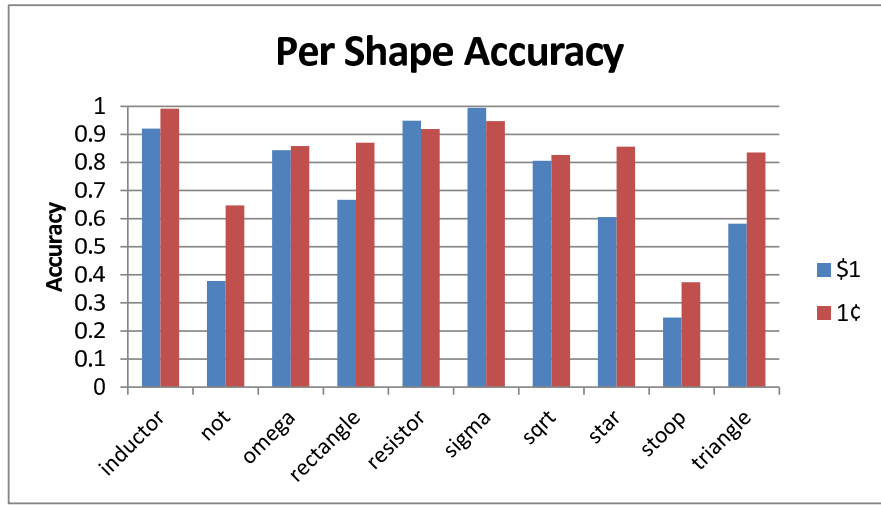


Figure 10.6: Average per shape accuracy results from the user-independent scheme with $n = 9$ training templates per gesture type. The difference in accuracy between the two methods is significant for each shape ($p < 0.03$).

10.3.5 Discussion

The user-independent scheme simulates off-the-shelf performance of each recognizer. One might imagine supplying a number of training examples with these recognizers so that a user can deploy them without needing to supply her own training examples. The user-independent evaluation considers the kind of performance that would be achieved in this circumstance. As Figure 10.4 indicates, user-independent accuracy increases slightly with the number of training templates used. Furthermore, 1^c performs significantly better than \$1 for this scheme.

Figure 10.6 provides additional insight into the difference in accuracy of the two recognizers when using the user-dependent scheme. For some gestures, the two recognizers achieve similar accuracy. On others, such as the not and star gestures, 1^c performs much better than \$1. Empirically, it seems that \$1 often misinterprets gestures

that form closed contours as other closed contour gestures, whereas 1^c is intrinsically well suited for recognizing such shapes.

Table 10.1 shows the expected result that computation time linearly increases with the number of training templates. Most importantly though, the computation time of 1^c is consistently two orders of magnitude less than that of $\$1$.

The user-dependent scheme simulates user-optimized performance for both recognizers. Figure 10.5 suggests that the two recognizers perform nearly equally well; in only four cases is the accuracy between the two techniques significant, and in three of those cases, $\$1$ performs better than 1^c .

10.3.6 Conclusion

We have presented 1^c , a fast, accurate, and easy-to-implement handwritten gesture recognizer. Our technique applies simple, yet effective techniques adapted from the time series classification literature. The key enabling component of our algorithm is our novel one-dimensional representation of handwritten strokes. This representation is intrinsically rotation invariant, eliminating the need for the expensive search-based angular alignment computation previous template-based gesture recognition techniques have required.

We have evaluated 1^c on a large database containing 2,520 strokes and compared its accuracy to that of $\$1$. By applying both user-dependent and user-independent cross-validation schemes, we evaluated accuracy on 75,600 test cases. Our results show that in the user-dependent scheme, both recognizers performed nearly as well as each other. In the user-independent scheme, 1^c always performs significantly better than

\$1. These results suggest that both recognizers are able to take advantage of users' individual drawing styles but 1^c is more tolerant of drawing styles not encountered in the training corpus. Additionally, our results show that 1^c is consistently two orders of magnitude faster than \$1. Just as important for developers, 1^c requires considerably less effort to implement than \$1; 1^c comprises 37 lines of pseudocode while \$1 comprises 72.

10.4 ClassySeg

10.4.1 Introduction

Automatic pen stroke segmentation is the process by which a digital pen stroke is segmented into its constituent lines and arcs². For example, stroke segmentation would decompose a hand-drawn triangle into the three straight lines that comprise it. The challenge in this process is determining which bumps and bends in the stroke are intended corners and which are not. It has been shown that curvature information alone is an unreliable indicator of segmentation [109], and thus a more sophisticated approach is required.

Stroke segmentation is an essential first step in shape recognition [80, 102] and thus is a crucial part of sketch-based interfaces. Decomposing a stroke into its constituent geometric primitives also facilitates beautification, in which the hand-drawn primitives are replaced by mathematically precise shapes to produce a neater final result [70, 68].

²This work was originally published in [53]

Existing segmentation techniques typically rely on heuristic algorithms and empirically determined parameters. ClassySeg provides greater extensibility and generality than these previous methods by employing general machine learning techniques to identify the segment points in a stroke. ClassySeg begins by identifying a set of candidate points consisting of all curvature maxima. Next, speed, curvature, and other geometric features are computed for each candidate point. These features are taken from previous segmentation approaches, effectively combining their strengths. The features are used to train a statistical classifier which determines which candidate points are true segment points and which are not. To optimize performance, a beam search was used to identify the subset of features that produces the most accurate classifier. ClassySeg was evaluated on a large data set of pen strokes from [59] and is more accurate than previous techniques for user-independent training conditions. Just as important, ClassySeg can be easily extended to include other features and is highly tunable. For example, it can be optimized for different kinds of shapes and can be tuned for individual users and various drawing hardware.

The next section places our approach in the context of previous work. Next, the dataset used to evaluate ClassySeg and benchmark it against prior techniques is described. This is followed by a discussion of the main components of the ClassySeg approach, including candidate point selection, feature computation, classifier training, and feature subset selection. Finally, ClassySeg’s accuracy is compared to that of three previous segmentation approaches, and to a baseline, naïve segmentation approach.

10.4.2 Related Work

Pen stroke segmentation is a well researched topic, and numerous methods have been developed. Yu and Cai's [141] technique first attempts to fit a stroke with a single primitive. If the fit is poor, the stroke is segmented at the point of highest curvature, and the two pieces are then recursively processed. Segments are merged in a post-processing phase, but the criteria for doing this are not specified.

The technique of Sezgin et al.[109], which we call SSD, uses speed and curvature to segment pen strokes. Segment points are located at points of minimum speed and maximum curvature. This work demonstrated the usefulness of pen speed data for segmentation, and showed that curvature data alone is inadequate. SSD is suitable for segmenting pen strokes into sequences of line segments, but cannot handle arcs.

Wolin et al. [130] developed ShortStraw, which begins by resampling the pen stroke, and then computes the "straw value" for each point, which gives an indication of the local curvature. All points with a straw value below an empirically determined threshold are considered candidate segment points. A top-down phase then examines the segments between each pair of consecutive candidate points to evaluate the quality of the line fit. If the fit is poor, segment points are added.

Xiong and La Viola [137] developed iStraw, which improves upon the ShortStraw approach by including timing information and curvature detection. iStraw achieves better accuracy than ShortStraw and is able to handle curve and arc segments, which ShortStraw cannot.

Wolin et. al's [131] Sort, Merge, Repeat (SMR) technique begins just as SSD

does, by locating candidate segment points at speed minima and curvature maxima. The algorithm then finds the shortest segment and merges it with one of its neighboring segments in an attempt to remove false positives. This process is repeated until the line and arc fit errors of the segments are below an empirically determined threshold.

Recently, Herold and Stahovich [59] presented SpeedSeg. This approach also identifies the initial candidate segment points at speed minima and curvature maxima. A set of heuristics are then used to both merge and split the initial segmentation to produce a more accurate final result. The heuristics employ several geometric and speed-based features with empirical thresholds. These threshold can be optimized to improve performance.

Nearly all of these approaches rely on heuristics and empirical parameters, which limit their extensibility. In many cases, there is no automated procedure for selecting optimal parameter values. By contrast, ClassySeg uses a general purpose machine learning approach that naturally extends to incorporate any number of features. Here, we use the approach with a collection of features derived from multiple existing segmentation techniques, but other features can be directly added. Furthermore, ClassySeg is highly optimizable. It can both determine the optimal subset of features to use and identify optimal parameter values (via a trained classifier). As a result, ClassySeg can be easily tuned for specific users, specific classes of shapes, and specific drawing hardware.

10.4.3 Approach

ClassySeg uses a C4.5 decision tree to determine which of the initial candidate segment points are true segment points. The candidates consist of the curvature maxima. The decision tree is trained using 48 features, many of which are taken from existing segmentation techniques. To improve performance, a beam search is used to identify the subset of features that results in the best-performing decision tree.

The sections that follow describe the approach in more detail, including the identification of candidate segment points, the features used for classification, and the approach used for training.

10.4.3.1 Candidate Point Selection

In the data set described in Section 10.2, only 2.61% of the data points are segment points. Training a classifier to detect such rare cases can be difficult [128]. There are a number of common ways to address this kind of problem, such as resampling schemes [69] and the use of sophisticated objective functions for training the classifier [77].

However, for the segmentation task, there is a more direct approach to overcome this problem: most segment points are maxima of curvature. In the data set we consider, only 0.20% of actual segment points are not maxima. Furthermore, 19.91% of the maxima are true segment points, which is a sufficiently high frequency to enable accurate classification.

ClassySeg identifies candidate segment points using the curvature computation in Section 10.4.3.3. As described there, this computation actually computes the absolute

value of curvature. Thus, all curvature extrema, regardless of the direction of concavity are selected as initial candidate segment points.

ClassySeg currently uses 48 features. Some describe basic geometric properties of the pen stroke such as arc length and the quality of fit of the line and arc segments. Others are adapted from SSD [109], Kim and Kim’s method [74], ShortStraw [130], iStraw [137], and SpeedSeg [59]. The details of these features are described below.

10.4.3.2 Arc Length

The ends of a pen stroke often require special handling because of the presence of hooks [59]. Thus, we characterize each candidate segment point by a number of arc length features. The arc length, d_i , of a point, p_i , is defined as:

$$d_i = \sum_{j=1}^i \left\| \vec{P}_j - \vec{P}_{j-1} \right\| \quad (10.5)$$

where \vec{P}_j is the coordinates of the j^{th} data point. The first data point has index $j = 0$ and $d_0 = 0$.

The arc length is computed between each candidate point and both the next candidate point (f_{aln}) and the previous candidate point (f_{alp}). The ratio of f_{aln} to f_{alp} is included as feature f_{alnpr} . A short arc length between consecutive candidate segment points may indicate that one, or both, is not a true segment point. To enable the classifier to properly handle the ends of a stroke, the arc length from the candidate point to both the the start point (f_{als}) and end point of the stroke (f_{ale}) are also computed.

To obtain arc length features that are independent of the size of the pen stroke, a second set of features is obtained by normalizing the above features by the arc length

of the stroke. The resulting four features are denoted by the addition of an “n” to the subscript: f_{alnn} , f_{alpn} , f_{alsn} , f_{alen} . (f_{alnpr} is already normalized.)

10.4.3.3 Curvature

We compute curvature using the approach from [59]. As all extrema of curvature are potential corners, ClassySeg works with the absolute value of curvature. In this way, all extrema are detected simply as maxima. For convenience, we use term “curvature” to mean the “absolute value of the curvature.”

The curvature, c , is computed as the derivative of the tangent angle, θ , with respect to arc length:

$$c = \left| \frac{\delta\theta}{\delta s} \right| \tag{10.6}$$

The feature f_c is defined to be equal to c

To construct the tangent angle at a point, a least squares line is constructed using the point and the five points on each side. This method naturally smooths the often noisy curvature data. If the line fit is inaccurate, i.e., if the average distance from the 11 points to the least squares line is greater than 10% of the arc length of the window of 11 points, a least squares circle is instead used to establish the tangent. The derivative of the tangent angle is also computed using a least squares line fit to the graph of the tangent angle vs. the arc length. The slope of this line gives the curvature in units of radians per pixel.

To provide the classifier with additional context about the curvature near a candidate segment point, the curvature of the points immediately preceding and pro-

ceeding the candidate point are used as features f_{cp} and f_{cn} .

The curvature at a corner can be arbitrarily large. Thus, to better characterize the shape, we compute two normalized curvature features. For the first (f_{cmm}), the curvature is normalized by the minimum and maximum curvature along the stroke:

$$f_{cmm} = \frac{(f_c - c_{min})}{c_{max}} \quad (10.7)$$

where c_{min} is the minimum curvature value over all points in the stroke, and c_{max} is the maximum. This normalized curvature value for the points immediately preceding and proceeding each candidate point are included as features: f_{cmm_p} and f_{cmm_n} . Likewise, the feature f_{ca} is obtained by normalizing the curvature by the average curvature of the stroke, c_{ave} :

$$f_{ca} = \frac{f_c}{c_{ave}} \quad (10.8)$$

Again, the average-normalized curvature of the points immediately preceding and proceeding each candidate point are also included as features: f_{cap} and f_{can} .

10.4.3.4 Straw

The ShortStraw algorithm of Wolin et al. [130] uses the “straw value”, which is analogous to curvature, to identify corners. The straw value at a point, p_i , is equal to the Euclidean distance between the points p_{i-w} and p_{i+w} :

$$straw_i = |p_{i-w}, p_{i+w}| \quad (10.9)$$

where w is a constant. Figure 10.7 shows an example of the straw value of both a segment point and a non-segment point. A small straw value is indicative of a true segment point.

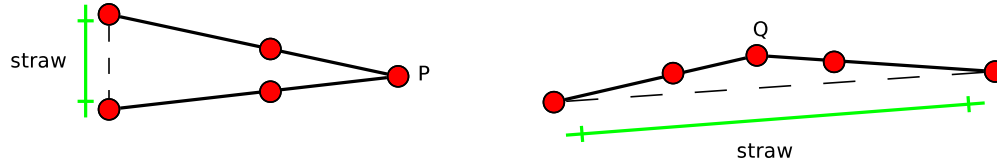


Figure 10.7: Straw values at a point P which is a true segment point (left) and a point Q which is not (right). ($w = 2$.) The straw value for a true segment point is much smaller than for a non-segment point.

With ShortStraw, the straw value is not well-behaved at the ends of the stroke.

The iStraw algorithm [137] uses a more general definition of the straw to remedy this:

$$\begin{aligned}
 straw_0 &= 0 \\
 straw_{N-1} &= 0 \\
 straw_1 &= |p_0, p_{1+w}| \times \frac{2w}{(w+1)} \\
 straw_2 &= |p_0, p_{2+w}| \times \frac{2w}{(w+2)} \\
 straw_{N-2} &= |p_{N-1}, p_{N-2-w}| \times \frac{2w}{(w+1)} \\
 straw_{N-3} &= |p_{N-1}, p_{N-3-w}| \times \frac{2w}{(w+2)} \tag{10.10}
 \end{aligned}$$

We compute the straw value using the iStraw approach. Before computing the straw value, both ShortStraw and iStraw resample the ink to produce equally spaced data points. We compute straw values two ways, first using the original data points, then using resampled data points. Additionally, we compute straw values using four

different values of w ($w = 1, 2, 3$, and 4), resulting in four straw features based on the original data points (f_{st1} , f_{st2} , f_{st3} , and f_{st4}) and four based on the resampled data points (f_{str1} , f_{str2} , f_{str3} , and f_{str4}).

Both ShortStraw and iStraw work with resampled data and do not classify the original data points. By contrast, our goal is to identify which of the original data points are true segment points. Thus, when resampling the ink to compute the straw value, we do a local computation that preserves the data point in question, and computes equally-spaced sample points on either side of it. We use the resampling technique used in [129].

10.4.3.5 Alpha and Beta

iStraw [137] uses two angles, α and β , to help distinguish true corners from smooth curves. Like the straw value, these angles are computed using a pair of resampled points, one on each side of the point in question. Angle β is computed using data points near the point in question, while angle α is computed using points that are farther.

Our features f_α and f_β are similar to angles α and β , except that our features are computed from the actual data points, rather than resampled ones. Also, we use a different offset for selecting the points that define the angles. As shown in Figure 10.8, we compute f_β as the angle between the line segment joining p_i and p_{i-1} and the line segment joining p_i and p_{i+1} . Similarly, we compute f_α as the angle between the line segment joining p_i and p_{i-2} and the line segment joining p_i and p_{i+2} .

In iStraw, the difference between α and β is used to distinguish corners. The

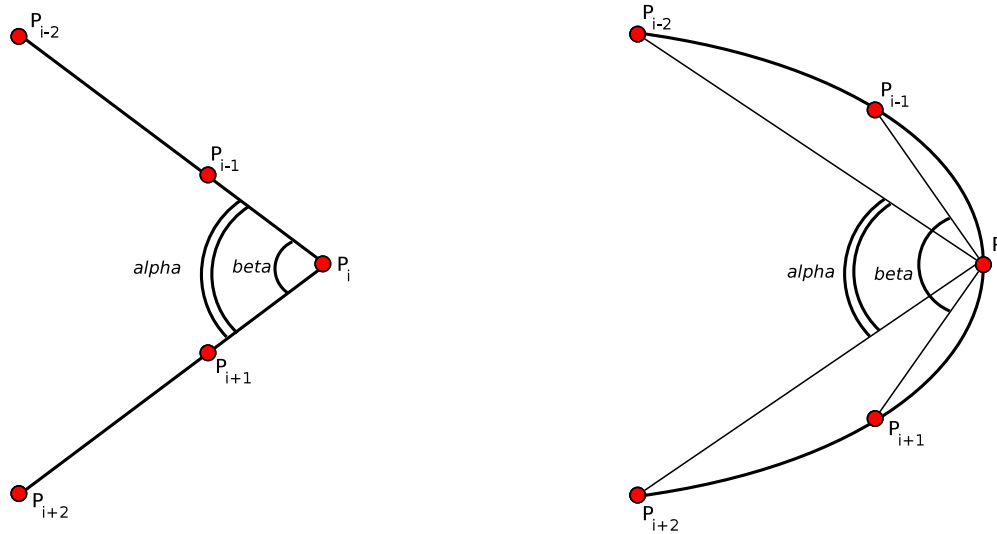


Figure 10.8: α and β values of a point, P_i , in the event of a corner (left) and curve (right). The difference between α and β is near zero in the case of a corner and relatively large in the case of a curve.

difference between α and β tends to be small in the case of true corners, and large in the case of curves. For this purpose, we use the feature $f_{\alpha\beta}$, which is defined as the difference between f_α and f_β .

We compute a second set of angle features ($f_{\alpha r}$, $f_{\beta r}$, and $f_{\alpha\beta r}$) using resampled ink. We resample the ink just as we do to compute the straw features in Section 10.4.3.4.

10.4.3.6 Region of Support

In their stroke segmentation approach, Kim and Kim [74] use a curvature estimation technique which combines the signed curvature values of the k -nearest neighbors of a point — called region of support — to determine the curvature at that point.

We define the region of support for a candidate point as the set of neighboring points that are locally convex to it. More precisely, we compute two region of support

features as illustrated in Figure 10.9: f_{asp} is the arc length from the candidate point to the previous curvature minima, while f_{asn} is the arc length from the candidate point to the next curvature minima. The larger these features, the more likely it is that the candidate point is a segment point.

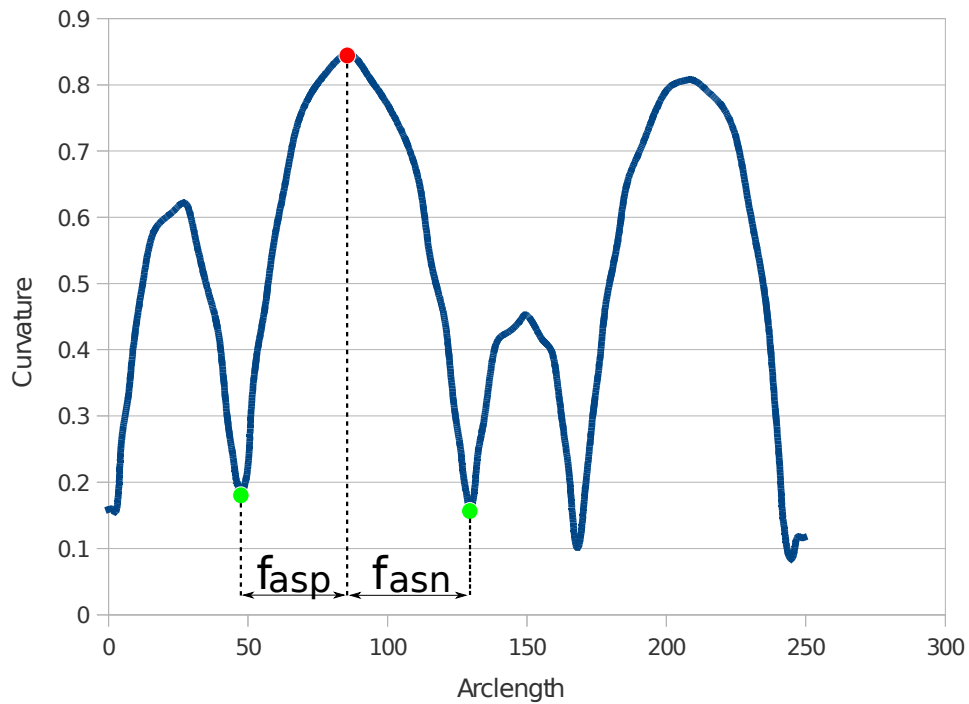


Figure 10.9: Region of support features for the red candidate point: f_{asp} and f_{asn} are the arc length to the next and previous curvature minima (green points).

10.4.3.7 Speed

People typically decrease pen speed when making deliberate corners [109]. Thus, low speed at a candidate point is often good evidence that the point is a true segment point.

We compute pen speed using a centered, finite difference approach from [59].

The speed at point p_i is initially computed as:

$$s_i = \frac{d_{i+1} - d_{i-1}}{t_{i+1} - t_{i-1}} \quad (10.11)$$

where, t_i and d_i are the time stamp and arc length, respectively, of point p_i . Because the speed data is noisy, we apply a simple filter: the speed at each point is averaged with that of the two points on either side. The feature f_s is defined as the value of the smoothed pen speed.

Just as with the curvature features, we also compute several normalized speed features. First, speed is normalized by the minimum and maximum pen speed of the stroke: f_{smm} . The minimum/maximum-normalized speed of the points immediately preceding and proceeding each candidate point are also included as features: f_{smmmp} and f_{smmmn} . Likewise, the feature f_{sa} is obtained by normalizing the speed by the average speed of the stroke. Finally, the average-normalized speed of the points immediately preceding and proceeding each candidate point are included as features: f_{sap} and f_{san} .

10.4.3.8 Arc and Line Fit

One measure of the quality of the segmentation is the degree to which the segments can be fit by geometric primitives. The approach in [141], for instance, recursively splits the segmentation until a good a good fit is achieved. Thus, we compute least squares arcs for the segments on either side of a candidate segment point. The features f_{afp} and f_{afn} are defined as the error of fit between these arcs and the actual data points. The features f_{lfp} and f_{lfn} are defined analogously, but represent the error of fit to least squares lines. If the next and previous segments form a single contiguous

line, the candidate segment point may be a false segment point. Thus, we compute a single least squares line fit to the ink spanning from the previous to the next candidate segment points. If the error of fit of this combined segment (f_{lfw}) is small, the candidate is likely a false segment point.

10.4.3.9 Classifier Training

The features and manually assigned labels of each candidate point were used to train a C4.5 decision tree classifier [104] implemented in WEKA [44]. The trained classifier is then used to predict if a given candidate point is a true segment point.

Using the full set of features may not yield the best accuracy. For example, if two features contain redundant information, this can adversely affect the classifier's accuracy. While a decision tree automatically determines which feature values best separate true positives from false positives, a separate process is required to determine the subset of features most useful in training the classifier. To approximate the optimal subset of features, we used a traditional beam search approach [10, 29] with a beam width of 10.

Beam search [142] is a modified best-first search used here to explore the large space of possible feature subsets to determine the one that produces the best classification accuracy. The search begins by training and evaluating all possible single-feature classifiers. The 10 single-feature subsets producing the highest accuracy are then expanded to create a set of two-feature classifiers. The 10 best of these are then expanded to create a set of three-feature classifiers, and so on, until the set of all features is reached. The best-performing subset is used as the optimal feature subset.

Two cross-validation schemes were used to evaluate ClassySeg. The first is a *user agnostic* scheme which evaluates ClassySeg’s performance under conditions in which the training set excludes data from the subject used for the testing set. The second is a *user optimized* scheme in which the training and testing data are from the same subject, but are still distinct. In both schemes, a beam search is performed in each testing-training fold. The classifiers created in the beam search are both trained and evaluated using the full training set of that fold.

In the user agnostic scheme, a 14-fold, “user-holdout” cross-validation was performed. In each fold, the data from one subject was selected for testing, and the data from the 13 other subjects was used for training. To begin each fold, a beam search was performed to determine the optimal feature set. The optimal features were then used to train the final C4.5 decision tree on the complete training set.

In the user optimized scheme, a 60-fold, “stroke-holdout” cross-validation was performed for each of the 14 subjects. In each fold of “stroke-holdout,” one of the subject’s strokes was used for testing, and the other 59 were used for training. Here again, a beam search was performed using the training set of each fold to determine the optimal feature set prior to training of the final C4.5 decision tree.

In both schemes, beam search found subsets containing eight to sixteen features. The following features occur in a majority of the feature subsets: f_{lfw} , f_{smmp} , f_{cn} , f_{str1} . The following features appear in none of the feature subsets: f_{lfp} , f_{alnpr} , f_{β} , f_{alsn} , f_{st1} , f_{st2} , f_{st3} , f_{str3} , f_{st4} , f_{str4} , f_{α} . These latter features may be redundant with other features, or may not be useful.

10.4.4 Results

We report accuracy using a number of measures common for other segmentation tasks (e.g., speech segmentation [41]): precision, recall, f-measure, and all-or-nothing accuracy. Precision, P , is the fraction of the predicted segment points that are true segment points:

$$P = \frac{\text{true positives}}{(\text{true positives} + \text{false positives})} \quad (10.12)$$

Here, true positives are true segment points that were classified as such by ClassySeg, while false positives are points that were erroneously classified as segment points.

Recall, R , is the fraction of the true segment points that were correctly identified:

$$R = \frac{\text{true positives}}{(\text{true positives} + \text{false negatives})} \quad (10.13)$$

Here, false negatives are true segment points that were incorrectly classified as not being segment points.

F-Measure, F , combines the precision and recall values; it is the harmonic mean of the two:

$$F = \frac{2PR}{(P + R)} \quad (10.14)$$

Finally, all-or-nothing accuracy, AoN , is the fraction of the pen strokes that have been perfectly segmented:

$$AoN = \frac{\text{perfect strokes}}{\text{total strokes}} \quad (10.15)$$

To benchmark ClassySeg’s performance, we compared it to five other techniques. Because these techniques may vary in where they place the segment point for a given corner, we used a 20 pixel threshold in evaluating the accuracy of each method – if a technique identified a segment point within 20 pixels of the correct location, as determined by manual labeling of the data, the segment point was considered to be correct. Using a threshold was particularly important for iStraw, which segments a resampled version of the stroke, rather than the original stroke data points.

The results are summarized in Table 10.5. The first technique is a naïve one in which every candidate segment point is considered to be a true segment point. The other four methods are more sophisticated and include: the SSD implementation from the authors of [130], the iStraw implementation from the authors of [137] (<http://www.eecs.ucf.edu/isuelab/>), and the SpeedSeg implementation from the authors of [59]. (We also implemented the algorithm of Yu and Cai [141]. However, it achieved poor performance on our data set: $AoN = 4.13\%$. This may be due to the omission of implementation details in the article.) All techniques were evaluated using the data described in Section 10.2. SpeedSeg’s performance is reported for both default parameter values and parameter values optimized for each individual subject.

The high recall of the naïve method demonstrates that the set of candidate points contains nearly all of the true segment points. The low precision, on the other hand, reveals that the candidate segment points include many false positives, so many

Method	P	R	F	AoN
Naïve	19.9%	99.8%	0.33	0%
SSD	84.7%	97.7%	0.91	27.8%
iStraw	93.2%	98.6%	0.96	69.9%
SpeedSeg (Default)	97.5%	96.4%	0.96	78.2%
SpeedSeg (Tuned)				88.6%
ClassySeg (UA)	95.6%	94.7%	0.95	74.6%
ClassySeg (UO)	99.0%	96.3%	0.98	85.7%

Table 10.5: Segmentation accuracy for various segmentation techniques. SpeedSeg Default and Tuned are the accuracies of SpeedSeg using default parameter values and user-optimized parameter values, respectively. ([59] reported only all-or-nothing accuracy for SpeedSeg Tuned.) ClassySeg UA and UO are the accuracies of ClassySeg for user-agnostic and user optimized conditions. P = precision; R = recall; F = f-measure; AoN = all-or-nothing accuracy.

in fact, that the naïve method did not correctly segment even a single stroke (AoN=0%).

The remaining segmentation methods achieve much higher precision than the naïve approach, with only a small decrease in recall. It is interesting to note the apparent logarithmic relationship between f-measure and all-or-nothing accuracy. For low values of f-measure, a large change in f-measure results in only a small increase in all-or-nothing accuracy: going from an f-measure of 0.33 to 0.91 improves all-or-nothing accuracy by only 27.8%. By contrast, at higher values, a small change in f-measure results in a large change in all-or-nothing accuracy: going from an f-measure of 0.91 to 0.98 improves all-or-nothing accuracy by 57.9%.

We performed an analysis of variance (ANOVA) to determine if there was a significant difference between the all-or-nothing accuracy of ClassySeg and each of the other segmentation techniques. This analysis focuses on all-or-nothing accuracy, as it is the most stringent of the performance metrics. When trained in a user-independent

fashion, ClassySeg (UA) achieves an all-or-nothing accuracy that is: significantly higher than that of SSD ($p < 0.001$); not statistically different than that of iStraw and SpeedSeg Default ($p = 0.26$ and $p = 0.31$ respectively); and significantly lower than that of SpeedSeg Tuned ($p < 0.001$). In the latter comparison, ClassySeg is at a disadvantage because SpeedSeg Tuned included user-specific training, while ClassySeg did not.

When trained on user-specific data, ClassySeg (UO) achieves an all-or-nothing accuracy that is significantly higher than that of SSD, iStraw, and SpeedSeg Default ($p < 0.001$, $p = 0.003$, and $p = 0.03$ respectively) and not significantly different than that of SpeedSeg Tuned ($p = 0.28$). Here, only the latter comparison considers comparable user-optimized evaluation conditions.

10.4.5 Conclusion

In contrast to previous stroke segmentation approaches which usually rely on empirically determined parameters and heuristics, we present ClassySeg, an automatic stroke segmentation technique that employs general-purpose machine learning techniques. ClassySeg begins by selecting all curvature maxima as candidate segment points. It then computes features for each candidate based on speed, curvature, and other geometric properties. These features are adapted from numerous prior segmentation approaches, effectively combining their strengths. These features are then used as input to a statistical classifier which is trained to distinguish true segment points from false ones. To improve performance, beam search is used to select the optimal subset of features to train the classifier.

When trained in a user-independent fashion, ClassySeg (UA) performed at

least as well as existing state-of-the-art algorithms also trained in a user-independent fashion. With user-specific training, ClassySeg (UO) performed better than existing algorithms with user independent training, and as good as the best existing algorithm with user-specific training.

While ClassySeg performs accurately, perhaps its most important property is its generality. The approach can be naturally extended to include any number of features. It is likely that even better performance can be achieved using features beyond the initial set considered here. Additionally, because the approach utilized a statistical classifier, it can be easily trained to optimize performance for specific users, specific classes of shapes, and specific drawing hardware. Finally, this approach is simpler to implement than techniques based on heuristic procedures.

10.5 Speech-Sketch Alignment

10.5.1 Introduction

Designers often communicate design concepts to each other with informal sketches, speech, and gestures.³ While the importance of such communication has long been recognized by designers [123], traditional design tools do not support this in any substantive way. Our long term goal is to remedy this by creating computational techniques to enable collaborative design tools that support natural multimodal communication.

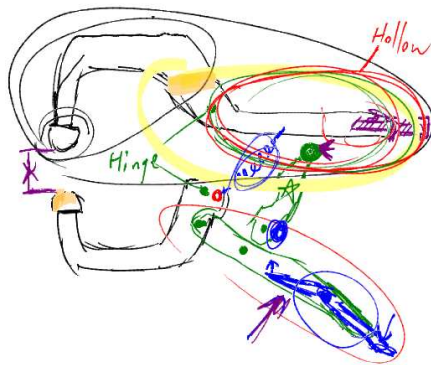
In previous work [10], we conducted a study to examine the nature of multi-

³This work originally appears in [54]

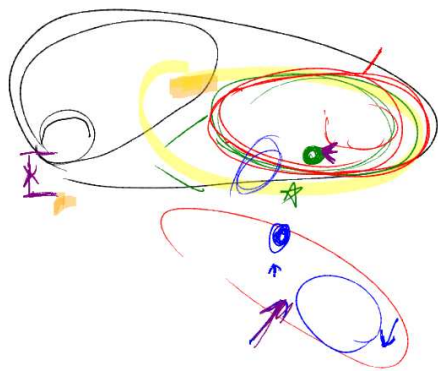
modal communication in collaborative design. Specifically, we examined how designers use natural free-form sketching and speaking to describe the structure and behavior of a mechanical device. We found that both the sketch and speech are essential to such descriptions, and that typically neither modality can be understood without the other. Additionally, the wide variety of information contained in the sketches makes them particularly challenging to interpret. While many of the pen strokes portray device structure, others are gestures, such as arrows used to indicate motion, or circles used to single out a component being discussed. Figure 10.10(a), which depicts a pair of C-clamp vise-grip pliers, is a typical sketch from the study. Consider the challenge such a drawing poses for any sketch-understanding software. To understand this sketch, it is first necessary to distinguish the *gesture strokes* (Figure 10.10(b)) from the *object strokes* representing device structure or handwritten text (Figure 10.10(c)).

Separating strokes in this way is valuable beyond the obvious purpose of facilitating sketch recognition. Most, if not all, gesture strokes have only temporary value. For example, gestures resolving deictic references or indicating the motion of a part may be superfluous once the discussion has moved on to a new topic. But over time, such gestures accumulate (e.g., Figure 10.10(a)), obscuring the sketch and hindering discussion. Detecting these and removing them from view when they are no longer needed may enable more efficient communication.

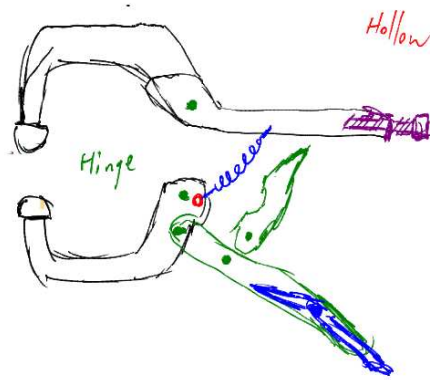
As part of our work in [10], we developed a technique for distinguishing gesture strokes from object strokes. The technique employs a statistical classifier that uses features of both the sketch and speech. The sketch features compute geometric properties



(a)



(b)



(c)

Figure 10.10: (a) Sketch of C-clamp vise-grip pliers. (b) Gesture pen strokes (c) Pen strokes representing device structure and text.

of the strokes, and the spatial and temporal relationships between them. The speech features compute statistical properties of the speech aligned with each stroke. Experiments with the technique indicated that the speech modality is more important than the sketch modality for gesture/object classification: the single most effective feature for classification was a speech feature.

The importance of speech for gesture/object classification suggests that the accuracy of the speech-sketch alignment process is critical to gesture classification. The work in [10], used a “three-second” alignment technique, in which the speech and sketch input were aligned based on temporal correlation. Each stroke was associated with the words that at least partially coincided with a temporal window extending three seconds on either side of the stroke. Our present work is focused on measuring the performance of this three-second alignment technique, and developing a new alignment technique to overcome some of its limitations, thus enabling more accurate gesture/object classification.

To evaluate the three-second alignment technique, we began by manually aligning the speech and sketch from the study in [10], as illustrated in Figure 10.11. We did this by first segmenting the speech primarily into clauses, and then aligning these with the strokes to which they refer. Comparison of the three-second and manual alignment revealed that the former has substantial room for improvement. For example, for 41% of pen strokes, there was no intersection between the three-second alignment and the correct (manual) alignment.

Consequently, we sought to develop an improved alignment technique, which

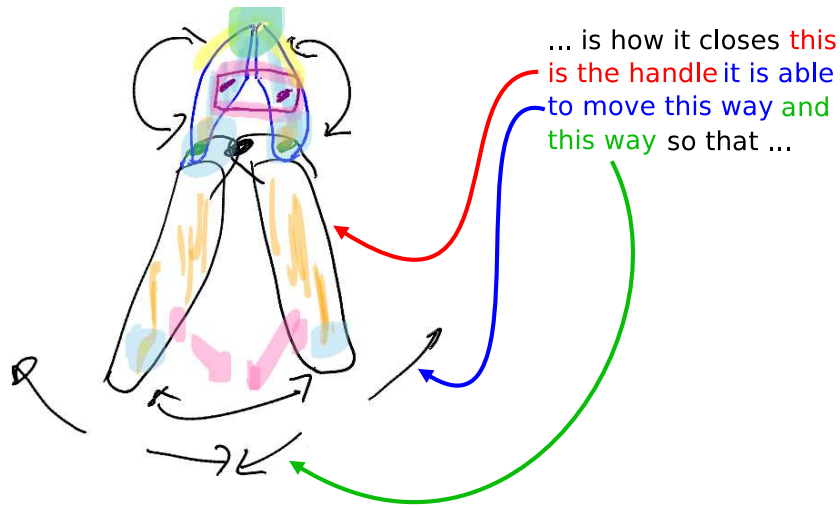


Figure 10.11: Example of the alignment of strokes with the speech that refers to them. The bold arrows link the word groups with the associated pen strokes.

we modeled on our manual alignment process. The new technique employs an explicit speech segmentation process, followed by a segment-stroke alignment process. Because both processes employ statistical classifiers, we call our technique “classifier-based alignment” (“CBA”). Evaluation of the new technique demonstrated that it produces considerably more accurate alignment than the three-second technique. More importantly, however, it results in substantially better gesture classification accuracy.

This work makes several contributions. First, we developed a technique for segmenting speech into meaningful clauses. The technique is well suited to the ungrammatical speech characteristic of multimodal dialog. The technique is effective, in part, because it uses information from the sketch input to help process the speech. Second, we developed a novel technique for aligning the segmented speech with the pen strokes to which it refers. These two efforts combine to produce an effective and accurate speech-sketch alignment technique for multimodal dialog. Finally, we demonstrated that the

new alignment technique enables accurate classification of gesture and object strokes in a multimodal dialog.

The next section places this work in the context of related work. This is followed in Section 10.5.3 by a description of our study from [10] and the gesture classification technique we developed in that work. Section 10.5.5 describes our manual speech-sketch alignment process and presents an evaluation of the alignment accuracy of the three-second technique. Next, Section 10.5.8 describes our classifier-based speech-sketch alignment technique, including the speech segmentation technique it employs. Section 10.5.10 presents the gesture classification accuracy obtained using the two speech-sketch alignment techniques and compares this to the accuracy achieved via manual alignment. After a discussion of these results in Section 10.3.5, conclusions are presented in Section 10.3.6.

10.5.2 Related Work

Multimodal systems date back at least to the work of Brown et al. [17], with subsequent early multimodal systems incorporating typed language and pointing with mouse or light-pen [126, 133]. Bolt's Put-that-there system [13] was the first to incorporate early speech and 3D pointing recognizers. Quickset [25] explores a general architecture for multimodal fusion, but unlike our work, QuickSet is a command-based system, i.e., the utterances are used as verbal replacements for mouse/menu commands. The iMap system handles free-hand gestures in a map-control user interface, using prosody cues to improve gesture recognition [76]. The system in [71] provides a speech and pen interface to restaurant and subway information for New York City, but it is not a

sketching system and has only text recognition and basic circling and pointing gestures for the graphical input modality.

Other applications of speaking and sketching include an early effort that used a diagram and written English text [94], interesting in part because it used a blackboard to help establish the reference relationships between the graphical and text entities. BBN’s Portable Voice Assistant [96] uses pen and voice input to enter and retrieve information on the World Wide Web. Their system integrates simultaneous speech and gesture inputs using a frame-based system. The Human-Centric Word Processor [98] enables radiologists to use pen-based selection gestures and command-based speech for post-dictation correction of transcriptions. nuSketch COA Creator [38] is designed as a general-purpose multimodal architecture, allowing users to sketch and talk to add symbols to a military map using commands like “add severely restricted terrain.” This system too uses command-based speech, and is focused on issues of reasoning about the content of the sketch rather than on recognition — the user assigns symbolic labels to the sketched objects.

Many systems have benefited from the series of empirical studies of multimodal communication in [99]. Cassell was among the first to argue that natural, free-hand gestures can be relevant to human computer interaction, and presented a helpful framework for gestural interaction in [19]. Oviatt et al. [100] have demonstrated advantages of multimodal interfaces, noting that multimodal input simplifies the users’ vocabulary and improves accuracy with accented speakers.

Our work is grounded in insights about how people use multimodal explanations to describe devices. Ullman [123] found that engineers commonly use five different

categories of pen strokes in a sketch. His “support” and “draw” strokes are analogous to our categories of gesture and object strokes. Heiser [50] concluded that when there are numerous arrow gestures in a sketch, students can more easily understand the functionality of a device, illustrating the importance of gestures in a design sketch.

Much of the previous work in understanding descriptions of mechanical devices has focused solely on sketching of structure (e.g., [12, 86]). By contrast, GIDeS++ [111] is a multimodal system specifically designed to understand descriptions of mechanical devices, but it uses pen strokes to replace mouse functionality rather than attempting to maintain a natural sketching environment. Likewise, ASSISTANCE [95] incorporates spoken behavioral descriptions to supplement the understanding of mechanical device sketches. However, it relies on limited vocabularies of speech patterns that must be explicitly identified in advance, where our system can adapt to new patterns via user-provided training data.

Hand and arm gestures have long been a topic of research. Kendon [73] provides an overview of the study of gesture, dating back to work by Quintilianus (circa the first century) in which he details how an orator ought to use gesture in discourse. More relevant to our work, Kendon explores the organization of speech and gestures. He finds that speech is organized into “idea units” marked by prosodic features, such as pitch level and loudness, rather than by lexical properties. Similarly, gestures are organized into “gesture units.” This suggests the need to segment our speech prior to aligning it with pen strokes. However, we segment speech based primarily on lexical considerations, and align each pen stroke with at most one speech segment.

Efron [32] classifies hand/arm gestures across three dimensions: the trajectory

of the gesture, whether the gesture involves the listener, and whether the gesture inherently contains semantic information. In our domain, gestures do not directly involve a listener, but they do contain semantic information which is frequently conveyed through shape.

We find parallels between the pen stroke gestures considered in our work and the hand and arm gestures studied by McNeill [88]. In McNeill’s classification scheme, hand/arm gestures describing objects or actions are called *imagistic*, while those that do not evoke imagery are *non-imagistic*. *Imagistic* gestures are further subdivided into *iconic* or *metaphoric* gestures. The former represent concrete concepts, such as a speaker illustrating how they threw a baseball by mimicking the action of throwing. The latter present abstract imagery, such a person balling their fists and then quickly spreading their fingers to convey a metaphoric “explosion,” illustrating frustration about the topic of discussion [88]. *Non-imagistic* gestures are also divided into two categories: *deictic* and *beats*. The former are pointing gestures, while the latter are typically involuntary movements of the hands made while speaking, and which carry no meaning.

Gestures are often understood in the context of accompanying speech. Oviatt [99] studied humans interacting with dynamic mapping software, quantifying the likelihood that speaking or sketching would occur first, or that they would start simultaneously. This work was extended by Adler [1] for design descriptions, who found consistent time delay patterns between when a pen stroke was drawn and when the related speech was spoken. The three-second speech-sketch alignment technique in [10] builds on this.

The findings of Oviatt and Adler are at odds with findings of McNeill [88]

which suggest that speech always co-occurs with its referent gesture. This discrepancy is likely due to the differences in the domains considered. Oviatt and Adler consider speaking and drawing, while McNeill [88] considers hand and arm gestures made during typical conversation. While hand/arm gestures are often made with minimal effort or concentration, drawing can often require enough concentration so as to interrupt speaking. Likewise, drawing is inherently slower than hand/arm gesturing, which may contribute significantly to the differences in gesture/speech alignment between the two domains.

Work in [20, 21] sets interaction in the context of a dialogue, using context, semantics and linguistic principles to resolve gestural references. Our task is different in that we must differentiate between gesture and object pen strokes, and we consider ungrammatical, disfluent speech, while they assume the speech is unambiguous. Furthermore, they interpret interaction in the context of a predefined image, while we consider an incrementally created sketch whose meaning is not known in advance.

There have been several prior efforts focused on segmenting speech into phrases and sentences. For example, Nakai and Shimodaira [93] describe a method that uses prosodic features to segment speech into accent phrases. A least squares approach is used to find the optimum match between the speech and pitch pattern templates. 97% segmentation accuracy is reported for a case in which the 30 best candidate segmentations are considered.

Most current techniques for identifying sentence boundaries in speech transcriptions are based on a hidden Markov model (HMM) [118, 119, 41, 66]. An n-gram language model is used to describe the joint distribution of words and sentence bound-

aries, which are modeled as events that occur between words. Many methods also use prosodic features for locating sentence boundaries. For example, Gotoh and Renals [41] combine their n-gram language model with a prosodic model based on pause duration. Likewise, Kim and Woodland [66] use a prosodic model based on 10 features. Stolcke and Shriberg [118] included part-of-speech information in an n-gram language model, and found that this improves accuracy. In later work, Stolcke et al. [119] augmented their n-gram language model with turn boundaries (change in speaker) and long pauses. All of these methods for locating sentence boundaries have been applied to telephone conversation and news broadcasts, while we consider a multimodal context with both speaking and sketching. Also, while these methods classify interword events as boundaries and non-boundaries, we classify words according to their position in a speech segment. As described in Section 10.5.8.1, this allows us to take advantage of the frequent occurrence of single-word segments.

Sentence boundary detection methods vary in the way they combine the language and prosodic models. Stolcke et al. [119] explore a variety of combination techniques including model interpolation, independent model combination, and joint modeling. In the latter case, a decision tree is used to combine posterior probabilities from the language model with prosodic features. Similarly, Liu et al. [83] use a maximum entropy model to combine prosodic and word-level features. We do not use an explicit language model, but instead use a single classifier (Ada-boosted C4.5 decision tree) to directly combine word-based, prosodic (pause), and sketch-based features.

Our gesture classifier [10] is related to the work of Patel [101] and Bishop [11] on separating text strokes from non-text strokes. These works differ from ours in

considering only features from the sketch, where we examine the accompanying speech. Additionally, in their work, text consists of a consistent set of letter and number glyphs, where the gestures in our domain are often unique, and frequently have the same shapes as object strokes.

Our gesture classifier also builds on work in shape recognition by using the kinds of features used by feature-based recognizers, as for example in [106, 101]. Our system relies on some of the features these systems use, but also extracts new features to address the special nature of identifying free-form gestures.

In examining properties of the accompanying speech, our gesture classifier does not try to understand it, but instead simply identify it as that which accompanies either a gesture or object stroke. We do this with Bayesian Filters (as in [42]) and Markovian Filters (as in [140]).

In summary, our work differs from much of the work in multi-modal interfaces in that we consider free-form speech and sketching, rather than a predefined vocabulary. Similarly, while most multi-modal systems use speech and sketch input as a substitute for mouse/menu commands, we consider the task of classifying sketch input as gesture and object strokes. While many speech segmentation techniques exist, ours is novel in that it uses information from the sketch modality. Also, it uses a single classifier to directly combine word-based, prosodic (pause), and sketch-based features. Finally, our speech-sketch alignment technique is novel in that it works from segmented speech and uses classifiers to detect and repair common alignment errors.

10.5.3 Background: Distinguishing Gesture and Object Strokes

As described in [10], we conducted a study to characterize how designers use natural free-form sketching, speaking, and gesturing to communicate design descriptions to each other.⁴ The study involved descriptions of four devices: C-clamp vise-grip pliers, bolt cutters, an air pump for inflating balls, and a door lock (Figures 10.10, 10.11, and 10.12). The participants were 16 graduate and senior undergraduate mechanical engineering students at UC Riverside. Four were female. English was the primary language for nine participants, but the speech of ten participants was indistinguishable from that of native English speakers. Eleven participants received their engineering instruction in English. There were only four participants that both did not have English as their primary language and did not receive engineering instruction in English. Fourteen participants had previously taken a course in engineering drawing, and seven had completed a team-based project-design course.

Each study session involved a pair of participants placed in separate rooms and allowed to communicate using Tablet PCs, microphones, and headphones. The tablets provided a shared drawing environment with a pen, highlighter, and eraser, and the ability to select from several ink colors. The audio and drawing were recorded with timestamps.

During a session, one participant was asked to describe a device to his or her partner, who could ask clarifying questions. At the end of the description, both participants were asked survey questions about the structure and behavior of the device.

⁴This section presents an overview of work from [10]. For complete details, refer to [10].

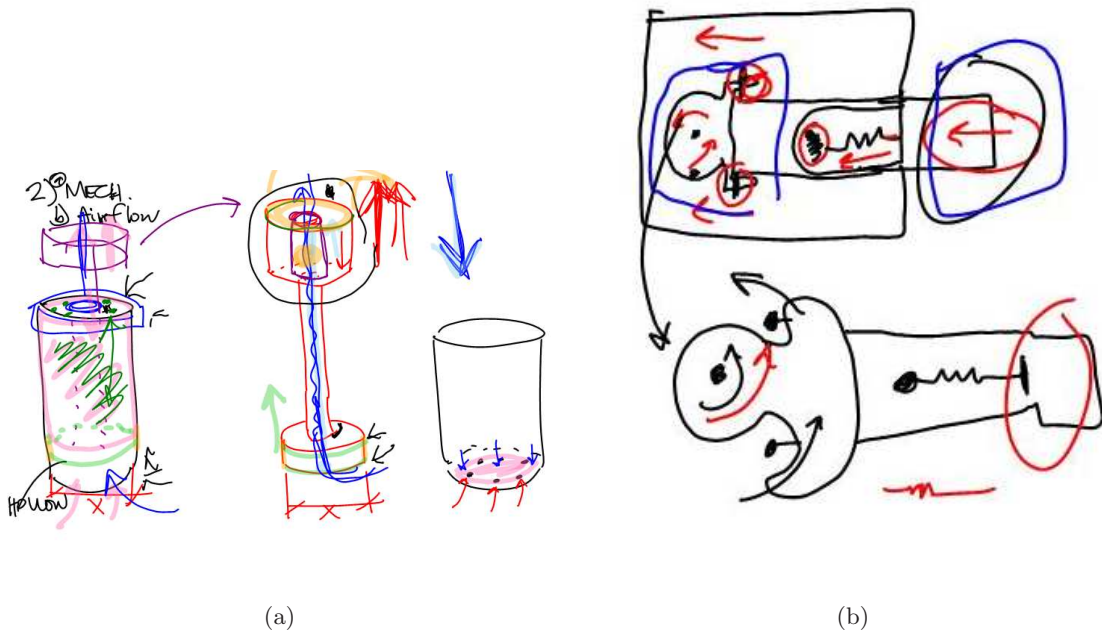


Figure 10.12: Two of the devices from the study. (a) Air pump for inflating balls. (b) Door lock.

To motivate effective dialog, the participants were informed that their compensation would be based on the accuracy of their answers. (All participants were in fact given the maximum compensation.) The two participants repeated this process three times, switching roles, so that each participant described two devices. In all, a total of 48 device descriptions were collected.

Figures 10.10, 10.11, and 10.12 show typical examples of sketches collected in the study. As discussed above, these sketches contain two types of pen strokes: *object strokes* and *gesture strokes*. The former depict device structure or comprise text. The latter can be classified into two categories, adopted from the terminology developed by McNeill [88]. Strokes that demonstrate an action, such as an arrow illustrating the direction in which the handles of a pair of vice grips may move, are iconic gestures.

Similarly, strokes that resolve deictic references from the speech modality are deictic gestures. These gestures may take many forms, such as tapping, circling, highlighting, and tracing. Object strokes could be considered iconic gestures, as they provide a representation of an object. However, we distinguish between object strokes and other iconic gesture strokes as our goal is to separate the representation of a device's structure from the description of its behavior.

10.5.3.1 Classifier Design

As Figure 10.10 illustrates, there can be a comparable number of gesture and object strokes in a sketch, making it challenging to understand the final image. There is a clear need for techniques to separate the two types of strokes. This would at first appear to be a shape recognition problem solvable with standard shape recognizers such as those in [72, 129]. However, this problem is not amenable to such approaches for several reasons. First, gesture and object strokes can have arbitrary shapes, but shape recognizers require a predefined set of shapes. Second, gesture and object strokes may be identical, and thus shape alone does not distinguish between the two classes of strokes. For example, a common selection gesture consists of tracing the shape of an object.

For these reasons, the gesture/object classifier described in [10] does not explicitly consider the shape of the pen stroke. Instead, each pen stroke is represented by features that are computed from both the sketch and speech input. The sketch features describe properties of the pen strokes, and the spatial and temporal relationships between them. The speech features describe properties of the speech aligned with each stroke. These features serve as inputs to a neural network which classifies a stroke as a

gesture or object stroke.

10.5.3.2 Sketch Features

The complete set of sketch features used for gesture/object classification is listed in Table 10.6. The first six features concern individual strokes. D_{SL} is the length of the pen stroke, while D_{SED} is the distance between its first and last points. D_{AC} is the sum of the absolute value of the curvature along a stroke. D_{DC} is similar to curvature, but is biased toward diagonal drawing directions. The ink density, D_{ID} , is a measure of the compactness of the stroke. The highlighter feature, D_{HL} , has a value of one if the stroke was made with a highlighter rather than an ordinary pen, and is zero otherwise.

The remaining 10 features describe the temporal and spatial relationships between strokes. D_{DPS} and D_{DNS} are the distance to the previous and next strokes. Similarly, D_{TPS} and D_{TNS} are the time between the stroke and the previous and next strokes. D_{TCS} is the time between the stroke and the closest previously-drawn stroke, while D_{ET} is the total elapsed time. The underlying color similarity, D_{UCS} , measures the extent to which earlier nearby strokes have the same color as the stroke. Underlying ink density, D_{UID} , is the density of the ink from other earlier pen strokes in the neighborhood (expanded bounding box) of the stroke. The two Hausdorff features [72] measure the extent to which a stroke traces underlying strokes. For each point on the stroke, the closest distance to a point on another earlier stroke is computed. D_{MHD} is the maximum of these closest distances, while D_{AHD} is the average.

Name	Description	Units
D_{SL}	Stroke length	Pixel
D_{SED}	Start to end distance	Pixel
D_{AC}	Total absolute curvature	Radian
D_{DC}	Diagonally-biased curvature	Radian
D_{ID}	Ink density	%
D_{HL}	Highlighter	Boolean
D_{DPS}	Distance to previous stroke	Pixel
D_{DNS}	Distance to the next stroke	Pixel
D_{TPS}	Time to the previous stroke	ms
D_{TNS}	The time to the next stroke	ms
D_{TCS}	Time to closest prior stroke	ms
D_{ET}	Total elapsed time	ms
D_{UCS}	Underlying color similarity	%
D_{UID}	Underlying ink density	%
D_{MHD}	Max. Hausdorff distance to underlying ink	Pixel
D_{AHD}	Ave. Hausdorff distance to underlying ink	Pixel
W_{TPS}	Time to previous speaker	ms
W_{WC}	No. of words in temporal window	Word
W_{BF}	Bayesian filter	Probability
W_{TBF}	Thesaurus Bayesian filter	Probability
W_{MF}	Markovian filter	Probability

Table 10.6: Features for gesture/object classification: D_x = sketch (drawing) feature; W_x = speech (word) feature.

10.5.3.3 Speech Features

To compute the speech features, it was first necessary to align the speech and sketch input, i.e., determine which words are associated with each pen stroke. The “three-second” alignment technique presented in [10] was grounded in observations in [1] and [99] suggesting that there is a strong temporal correlation between speaking and drawing. This technique employs a temporal window extending three seconds before and after the stroke. It is assumed that any words falling at least partially within this window are associated with the stroke. It is possible that a word may be associated with more than one stroke, or that a stroke may have no words associated with it.

The speech features associated with a pen stroke (Table 10.6) are computed from the speech aligned with it. To avoid inaccuracies inherent in current state-of-the-art speech-to-text tools, the speech was manually transcribed and then Sphinx [65] was used to align the text with the recorded audio to find timestamps for the words. The words were also labeled with the identity of the speaker. Using manual transcriptions provides an upper bound on the contribution of the speech content to gesture classification. However, the speech may contain other valuable information, such as prosody, which was not considered.

The simplest speech features are the time to the previous speaker, W_{TPS} , and the number of words aligned with the stroke, W_{WC} . The other speech features concern the words themselves. Understanding grammatically correct speech is difficult enough; the speech considered here is ungrammatical, filled with pauses, repetitions, and disflu-

encies like “um” and “ah.” Trying to perform semantic analysis on such ungrammatical text is intractable at present. As an alternative, statistical models are used to predict whether a set of words corresponds to a gesture or object stroke.

The first statistical speech feature, W_{BF} , is based on a Bayesian filter, a form of naïve Bayesian classifier that has had some success in spam recognition [42]. To construct the Bayesian filter, it is necessary to learn the conditional probability that a stroke is a gesture, given a specific word, w_i . $Pr(Gesture | w_i)$ can be estimated from training data using Bayes’ Theorem:

$$p_i = \frac{Pr(w_i | Gesture) \cdot Pr(Gesture)}{Pr(w_i)} \quad (10.16)$$

where $Pr(w_i | Gesture)$ is the conditional probability that word w_i will be observed, given that a gesture stroke is observed; $Pr(Gesture)$ is the prior probability of observing a gesture; and $Pr(w_i)$ is the prior probability of observing word w_i .

Participants in the study used a varied vocabulary to describe the same objects and gestures. If the Bayesian filter encounters a word that was not in the training corpus, it is unable to produce a probability. The Thesaurus Bayesian filter feature, W_{TBF} , provides a remedy for this situation. It is computed much like W_{BF} , except that a thesaurus is used to generalize the training data. One strength of these two features is that they learn which words are most likely to coincide with gesture or object strokes. However, these features do not consider word order. The Markovian filter feature, M_{MF} , is analogous to the Bayesian filter features, but considers word sequences rather than individual words.

10.5.4 Results: Gesture/Object Classification Accuracy

In [10], a form of holdout-validation was used to evaluate the accuracy of the gesture classifier. The holdout set was comprised of 39 randomly selected sketches for training, and 10 for testing. A conventional beam search approach [3, 43, 29] was used to determine which sets of features are the most effective at classification. To provide additional insights about which features are the most important, this process was performed three times: once considering only sketch features, once considering only speech features, and once considering both.

The best single sketch-feature classifier used D_{TNS} and achieved 69.5% accuracy. The best sketch-only classifier achieved 76.2% accuracy using nine features: D_{SL} , D_{DPS} , D_{UID} , D_{MHD} , D_{AHD} , D_{HL} , D_{ET} , D_{AC} , and D_{DC} . The best single speech-feature classifier used W_{BF} and achieved 77.7%. The best speech-only classifier achieved 78.2% accuracy using three features: W_{BF} , W_{TBF} , and W_{WC} . The best classifier considering all features achieved 81.9% accuracy using six features: D_{TCS} , D_{MHD} , D_{HL} , D_{ET} , W_{BF} , W_{TPS} .

In [10], this process was actually performed for four holdout sets. For all sets, the results were similar: the single best feature in all cases was either W_{BF} or W_{TBF} . Similarly, for multi-feature classifiers employing speech features, the best feature sets always contained at least one of these two features.

10.5.5 Evaluation of Three-Second Technique

As the results in the previous section demonstrate, the speech modality plays an important role in identifying gestures. For example, the two Bayesian filter features were

the most important single features for classifying pen strokes as gesture or object strokes. The importance of speech suggests the need to examine the validity of the speech-sketch alignment technique which serves as the foundation for the speech features.

As the name suggests, the three-second alignment approach uses only temporal correlation to align the speech and sketch input. It is possible for this technique to associate speech with a stroke that is not logically related to it. To evaluate the performance of the three-second technique, we manually aligned the speech and pen strokes based on semantic information. We then compared the resulting alignment with that produced by the three-second technique. We also used the manually aligned speech to compute the speech features for our gesture classifier to determine if more accurate alignment would improve classification accuracy. The latter results are presented in Section 10.5.10.

10.5.6 Manual Alignment

We manually aligned the speech and sketch modalities using a two-step approach. We first segmented the speech into small, meaningful statements. We then aligned each statement with the pen strokes, if any, to which it referred. The segmentation step proved to be difficult because the speech was terribly ungrammatical and disfluent, as is commonly the case in multimodal descriptions [1]. Because of the nature of the speech, we could not use a simple segmentation strategy, such as decomposing the speech into grammatically correct clauses. Oviatt et al. [99] suggest that in multimodal interactions, spoken phrases often follow a subject-verb-object (SVO) pattern. We used this as the starting point for developing our manual segmentation approach. Our approach is similar to the Simple Metadata Annotation Specification [120], but considers

information from the sketch input.

Our manual segmentation is comprised of: single “clauses” consisting of a subject, verb, and object; multiple logically related, sequential clauses; partial clauses; and filled pauses such as “uh” and “um”. Note that filled pauses are identified purely on lexical grounds and are not prosodic features of the speech. Whenever possible, we segmented the speech into subject-verb-object “clauses.” However, if the speaker moved on to a new thought before completing a clause, we segmented the incomplete thought into a partial clause. Likewise, a change in speaker before the completion of a clause also resulted in a partial clause.

Filled pauses could either comprise an entire segment, or be included in a larger clause, depending on the circumstances. If the filled pause was in the middle of a set of words that otherwise formed a clause, the pause was grouped with that set of words. For example, “this handle uh moves here” is considered a single clause. Likewise, if the filled pause occurred immediately before the start of a clause, it was grouped with it. For example, “uh this handle moves here” would be segmented as a single clause if there were little delay between “uh” and “this.” In all other cases, filled pauses were considered to form their own segments.

There were two occasional exceptions to our segmentation strategy. Two or more clauses were joined if they referred to the same pen stroke. We did this so that each stroke would be aligned with at most one speech segment. Also, if a clause had multiple objects referring to different strokes, the objects were split into separate clauses. These two cases are the primary differences between our segmentation approach and that in [120].

Figure 10.13 shows an example of the manual segmentation results. The first segment consists of the filled pause “uh.” This pause was not combined with the subsequent clause because the time gap was too large. “it’s kind of for cutting stuff” is a typical clause with subject “it”, verb “is”, and object “for cutting stuff”. “when they uh attach” is also considered to be a clause with subject “they” and intransitive verb “attach”. Also, the filled pause “uh” is included in the segment because it occurs inside an otherwise valid clause. “uh huh” is a segment consisting of two filled pauses in close succession. The phrase “and both sides move the” is a partial clause; the speaker changed thoughts before completing it. The word “these” is again a partial clause representing a new idea. Finally, the phrase “this moves” is a clause with a subject and verb, but no object.

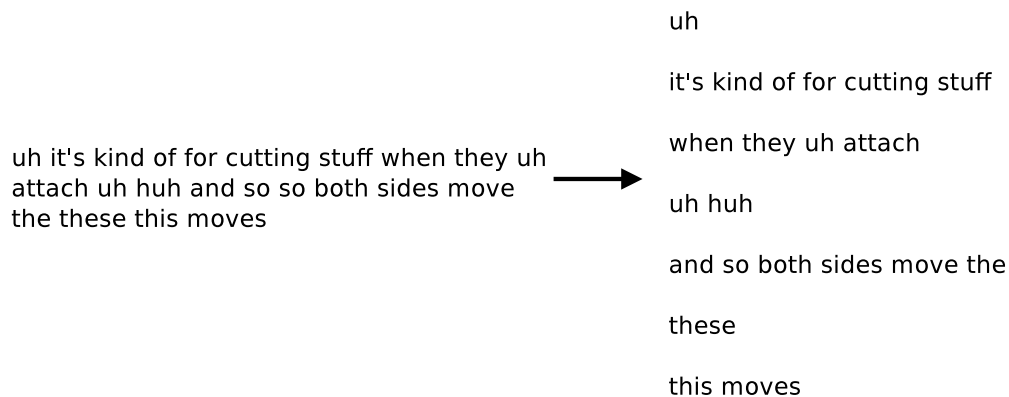


Figure 10.13: Example of manual segmentation. (left) Raw speech. (right) Segmented speech.

As this example illustrates, manually segmenting the speech required considerable judgment. The task was performed by two researchers. Each segmented one half of the speech and then verified the segmentation accuracy of the other half. Once the segmentation was completed, the two researchers then manually aligned the segments

Attribute	Count
Words	34,354
Segments	7,454
Strokes	6,470
Words/Segment, Ave.	4.6
Words/Associated Segment, Ave.	8.3
Segments Associated w/Strokes	22.5%
Strokes Associated w/Segments	78.8%

Table 10.7: Properties of manually aligned speech from user study.

and pen strokes. Each stroke was aligned with at most one speech segment. However, a speech segment could be aligned with multiple strokes. As with the segmentation, the researchers divided the task and verified each other’s work.

An alternative approach for annotating our data would have been for each researcher to annotate the entire corpus individually and then arbitrate the single, final annotation. This approach can lead to a more consistent annotation of the corpus than the cross-validation approach we used [6]. We opted for our approach in the interest of expediency, and note that any inconsistencies between the two halves of the annotation will only hamper the performance of our statistical classifier.

Table 10.7 tabulates the results of the manual segmentation and alignment process. The data from the study contained 34,354 words forming 7,454 speech segments. 78.8% of the 6,470 pen strokes were aligned with speech segments, but only 22.5% of the segments were aligned with strokes. On average there were 4.6 words per speech segment, but for segments aligned with strokes there was a much higher average of 8.3 words per segment.

10.5.7 Alignment Accuracy of Three-Second Technique

We evaluated the accuracy of the three-second alignment technique by direct comparison with the manual alignment, which constitutes the correct result. Specifically, we compared the set of words associated with each pen stroke in the two cases. Note that the three-second technique does not have an explicit segmentation step. Rather, any words that fall at least partially within the three-second temporal window of a stroke are associated with it. Thus, there is no notion of segmentation accuracy, and it is possible to evaluate accuracy only for those words that are associated with a pen stroke.

To illustrate the analysis, consider the speech and the accompanying gesture in Figure 10.14. The three-second approach has associated with this stroke the words “faces the other way this is uh like a”. The correct association determined by the manual alignment process is “this is uh like a handle”. In this case the three-second association begins too early and does not extend long enough. This situation occurred on average for 6% of the strokes (the average is computed over the 48 sketches).

There are a total of 14 possible relative arrangements of the three-second association and the correct (manual) association as shown in Figure 10.15. Each cell in the figure represents one of the possible arrangements. For example, cell 4 represents the arrangement from Figure 10.14. For clarity, the stroke itself is not represented in the various cells in Figure 10.15.

Cases 1 through 11 in Figure 10.15 are all cases in which the three-second approach associates words with strokes that should have associated words. Case 1 is

faces the other way this is uh like a handle it can move
faces the other way this is uh like a handle it can move



Figure 10.14: A gesture pen stroke and the speech associated with it by the three-second technique (top bar) and the manual alignment process (bottom bar).

when the three-second association exactly matches the correct result. This occurred on average for 2% of the strokes. Cases 2 through 9 are overlapping associations that do not perfectly match. These cases represent on average 56% of the strokes. Cases 10 and 11 are cases in which there is no overlap between the three-second association and the correct one. These cases represent on average 10% of the strokes. Case 12 describes strokes that should have no associated speech, but the three-second approach has made an association. This occurred on average for 26% of the strokes. Case 13 is the converse case in which there should be associated speech, but the three-second approach has associated none. This occurred on average for 5% of the strokes. Finally, case 14 describes situations in which the three-second approach has correctly associated no speech with a stroke. This case did not occur. On average, the three-second approach achieved the correct answer only 2% of the time (case 1), and 41% of the time the three-second association was completely disjoint from the correct result (cases 10 – 13).

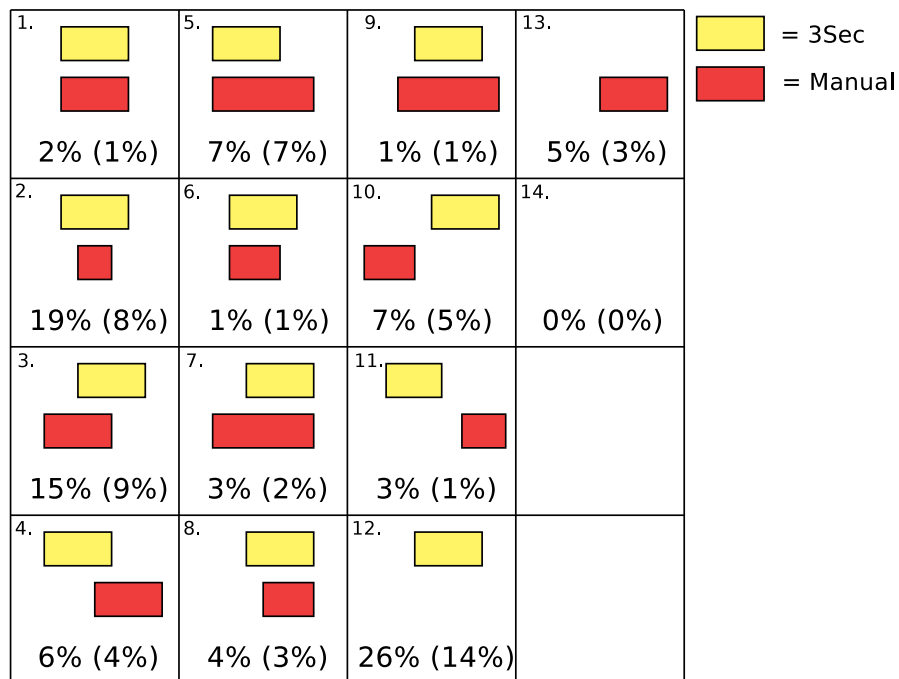


Figure 10.15: Accuracy of the three-second technique. In each cell, the top bar represents the speech associated with a stroke by the three-second technique, while the bottom bar represents the manual (correct) association. Each cell represents a distinct relative arrangement of the associations and the frequency with which it occurs. Results are averaged over the 48 sketches. Standard deviations are included in parenthesis.

10.5.8 Classifier-Based Alignment

As Figure 10.15 illustrates, the three-second technique does not accurately align the speech and sketch modalities. Consequently, we sought to develop an improved automatic alignment technique. We modeled the new technique on our manual process: the technique employs an explicit speech segmentation process, followed by a segment-stroke alignment process. Because both processes employ statistical classifiers, we call our alignment technique “classifier-based alignment” (CBA).

10.5.8.1 Speech Segmentation

Our approach to automatic segmentation uses a statistical classifier to classify words according to their position in a segment. We consider four classes of words: *start*, *middle*, *end*, and *only* words. As the names suggest, *start* and *end* words represent the start and end of a clause, respectively. All words in a clause other than these are defined as *middle* words. *Only* words are segments consisting of a single word, which is typically a filled pause. Figure 10.16 shows an example of the word classification for a passage of speech.

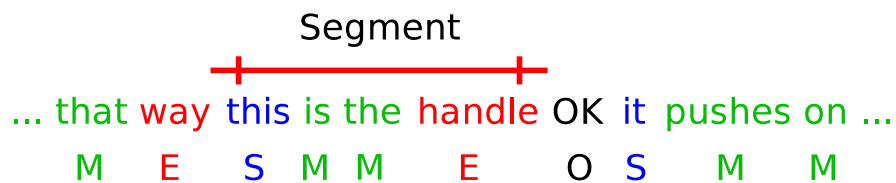


Figure 10.16: The classification of the words in a spoken passage. “S” = *start*, “M” = *middle*, “E” = *end*, and “O” = *only*.

The word classifications are used to directly construct the speech segmentation.

First, all valid segments are formed. Specifically, each *only* word is labeled as a segment. Likewise, each sequence of words that begins with a *start* word, ends with an *end* word, and has only *middle* words (if any) in between, is labeled as a segment.

Once all valid segments have been formed, a repair process is used to segment any remaining speech. The first and last words of the speech are always considered *start* and *end* words, respectively. Any unsegmented word immediately after a segment is treated as a *start* word, while any unsegmented word immediately before a segment is treated as an *end* word. A single word directly between two valid segments is considered an *only* word. After updating the word classifications in this fashion, any new valid segments are formed. The repair process is then repeated until all words have been segmented.

Consider a passage of speech that has been classified as: *start, middle, end, middle, middle, end*. In the initial segmentation pass, the first three words will be formed into a valid segment. Then, during the repair pass, the fourth word will be treated as a *start* word so that the last three words form a segment.

Our segmentation approach is based on four word classes. Many speech segmentation approaches such as [118, 119, 41, 66] classify interword boundaries as segment events or non-segment events. These approaches were developed for unimodal dialog such as the SWITCHBOARD corpus [40]. We consider multimodal dialog in which the speech is highly disfluent and filled pauses are common. We designed our four-class approach to take advantage of the discriminatory power of single-word segments. This approach is also consistent with work in [64, 84] suggesting that for some classification problems, decomposing a class into subclasses can result in higher accuracy.

10.5.8.2 Segmentation Classifier and Features

Our segmentation classifier is an Ada-boosted C4.5 decision tree computed with WEKA [44]. Each word is characterized by 25 features listed in Table 10.8. (The classifier considers the features of the word in question, as well as those of the word on either side.) The simplest feature is the word itself, W_s . Each word is also characterized by the parts of speech that it could possibly have in legal English usage, which is queried from the dictionary in the Stanford part-of-speech tagger [122]. The rationale for these features is that different parts of speech may be more likely to occur in particular locations within a speech segment. For example, a verb is unlikely to be the first word in a segment. We define nine boolean part-of-speech features indicating if the word could be a coordinating conjunction (W_{CCN}), determiner (W_{DET}), preposition (W_{PRP}), adjective (W_{ADJ}), personal pronoun (W_{PP}), adverb (W_{ADV}), verb (W_{VRB}), wh-determiner (W_{WHD}), or wh-adverb (W_{WHA}). Wh-determiners are the words “what” and “which” used as determiners. Wh-adverbs are the words “how,” “when,” “whence,” “where,” and “why” used as adverbs. Note that we use the possible parts of speech, rather than the actual part of speech, as the latter is difficult to determine because the speech is highly ungrammatical and the sentence boundaries are as yet unknown.

Four of the features compute temporal relationships between the words. W_{TNW} and W_{TPW} are the time to the next and previous words, respectively. To obtain a measure of the relative size of the time gap after a word, we compute the ratio of W_{TNW} to the sum of the values of W_{TNW} for the word and its two successors. We call this feature W_{TNR} . W_{TPR} is an analogous feature that concerns the relative size of the

gap before the word. A time gap that is large compared to the neighboring gaps (i.e., a large ratio) could indicate a segment boundary.

A change in speaker usually corresponds to a new segment. Thus, two features track changes in the “author” of the speech. W_{ACN} is a boolean feature that is true only when there is an author change immediately after the word. Similarly, W_{ACP} is true only when there is an author change immediately before the word.

A novel property of our segmentation technique is that we use information from the sketch modality. Specifically, we compute properties of the pen stroke drawn closest in time to the word. We refer to this as the “coincident stroke”, although the word and stroke may not actually overlap in time. We characterize this stroke with three intrinsic properties: its arc length (D_{SL}), start to end distance (D_{SED}), and duration (D_{DUR}). The first two of these features are the same as those used with the gesture/object classifier described in Section 10.5.3.1.

Four features describe the temporal relationships between the coincident stroke and the other strokes. These features are analogous to those used to describe the temporal relationships between the words. D_{TNS} and D_{TPS} are the time to the next and previous strokes, respectively. Again, to obtain a measure of the relative size of the time gap after the coincident stroke, we compute the ratio of D_{TNS} to the sum of the values of D_{TNS} for the stroke and its two successors. We call this feature D_{TNR} . D_{TPR} is an analogous feature that concerns the relative size of the time gap before the stroke.

The final two features track changes in the “author” of the pen strokes. D_{ACN} is a boolean feature that is true only when there is an author change immediately after the coincident stroke, while D_{ACP} is true only when there is an author change

Name	Description	Units
W_{TNW}	Time to Next Word	ms
W_{TPW}	Time to Previous Word	ms
W_{TNR}	Time to Next Ratio	*
W_{TPR}	Time to Previous Ratio	*
W_S	The Word	Text
W_{CCN}	Coordinating Conjunction	Boolean
W_{DET}	Determiner	Boolean
W_{PRP}	Preposition	Boolean
W_{ADJ}	Adjective	Boolean
W_{PP}	Personal Pronoun	Boolean
W_{ADV}	Adverb	Boolean
W_{VRB}	Verb	Boolean
W_{WHD}	Wh-Determiner	Boolean
W_{WHA}	Wh-Adverb	Boolean
W_{ACN}	Author Change Next	Boolean
W_{ACP}	Author Change Previous	Boolean
D_{TNS}	Time to Next Stroke	ms
D_{TPS}	Time to Previous Stroke	ms
D_{TNR}	Time to Next Ratio	*
D_{TPR}	Time to Previous Ratio	*
D_{ACN}	Author Change Next	Boolean
D_{ACP}	Author Change Previous	Boolean
D_{SL}	Stroke Length	Pixel
D_{SED}	Start to End Distance	Pixel
D_{DUR}	Stroke Duration	ms

Table 10.8: Features for speech segmenter: D_x = sketch (drawing) feature; W_x = speech (word) feature. * = dimensionless quantity.

immediately before it.

10.5.8.3 Segmentation Accuracy

We performed leave-one-out cross-validation to evaluate our speech segmenter.⁵

In each iteration of the cross-validation, the data from all but one sketch was used to train our classifier. We then used the trained classifier to predict the segment boundaries for the remaining sketch. The technique achieved an average accuracy of 92.7% at classifying words as *start*, *middle*, *end*, and *only* words.

To provide a more informative measure of accuracy, we directly compared our “classifier-based segmentation” (CBS) with the manual segmentation. Specifically, we computed the fraction of the classifier-based segments that matched the manual segments within a tolerance ranging from zero to three words. The results are shown in Figure 10.17. On average, about 33.9% of the classifier-based segments exactly matched a manual segment, and about 75.9% matched within three words. In the latter case, the errors could be distributed on both ends of the segment so long as the total number of errors did not exceed three. For example, compared to the manual segment, the classifier-based segment could be missing one word at the beginning, and have two extra words at the end, or vice versa.

⁵Cross-validation is a process of partitioning a dataset into complementary training and testing sets. Multiple alternative partitionings are considered, and the results from them are averaged.

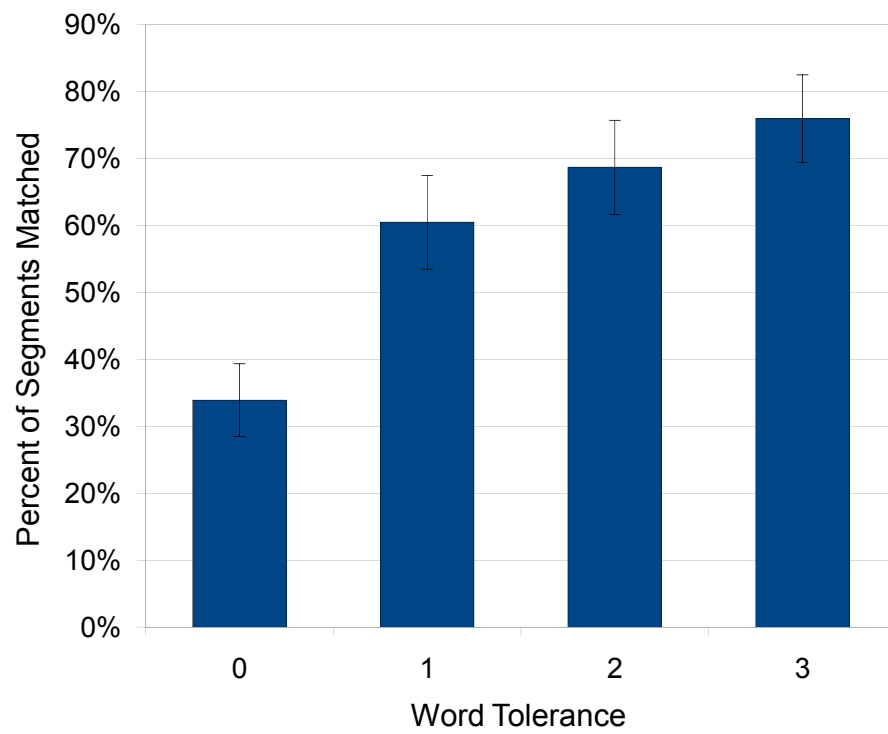


Figure 10.17: Percentage of classifier-based segments matching (correct) manual segments within a tolerance. Results averaged over the 48 sketches.

10.5.9 Stroke-Speech Alignment

Once the speech has been segmented, the next step is to align the segments with the pen strokes. We do this with a two-step process. First, segments are aligned with strokes based on simple temporal correlation. Then, we use a classifier to detect and repair two common alignment errors. The initial alignment borrows from the three-second approach. Each stroke is associated with the segment that has the greatest overlap with the stroke’s three-second temporal window, i.e., a window that extends three seconds before and after the stroke.

Using an analysis similar to that described in Figures 10.14 and 10.15, we computed the accuracy of the initial alignment to determine what improvements are necessary. The results are illustrated in Figure 10.18. The two most frequent problems are case 10 in which the initial association follows the correct association, and case 12 in which there is an association when there should be none. Case 10 occurs on average for 18% of pen strokes, while case 12 occurs for 26%.

Because of the prevalence of these two cases, we developed classifiers to detect them. The two classifiers are applied to each initial association. If a case-10 error is detected, the association of the pen stroke is changed to the next earlier segment. If a case-12 error is detected, the association for the stroke is removed. In this fashion, the classifiers enable an efficient approach to improving the initial alignment.

The “case-10” and “case-12” classifiers are Ada-boosted C4.5 decision trees computed with WEKA [44]. They consider features of both the speech segment and the initially associated pen stroke. There are a total of 11 features which are listed in

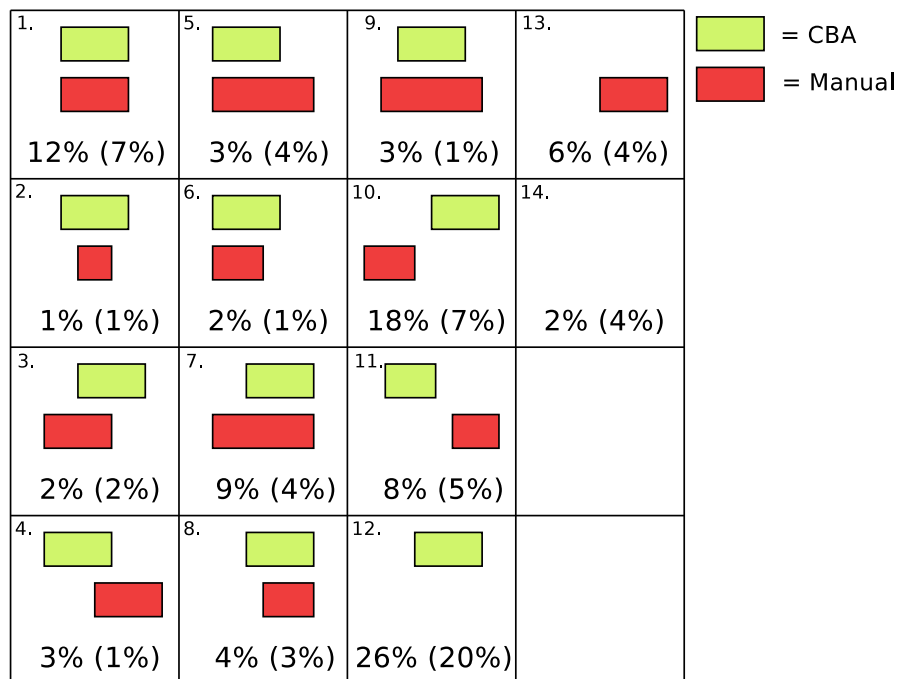


Figure 10.18: Alignment accuracy after the first step of classifier-based alignment (CBA), i.e., *before the final processing step*. In each cell, the top bar represents the speech associated with a stroke by the first step of CBA, while the bottom bar represents the manual (correct) association. Each cell represents a distinct relative arrangement of the associations and the frequency with which it occurs. Results are averaged over the 48 sketches. Standard deviations are included in parenthesis.

Table 10.9.

S_{WC} is the number of words in the segment, and S_{DUR} is its duration. S_{SC} is the number of strokes associated with the segment — the segment may also initially be associated with other strokes. S_{NV} is a boolean feature indicating if any of the words in the segment were tagged as a noun or verb by the Stanford part-of-speech tagger [122].⁶ The intuition is that segments containing no nouns or verbs are generally uninformative and are unlikely to refer to a stroke. The initially associated pen stroke is characterized by its arc length (D_{SL}), duration (D_{DUR}), and the time to the next stroke (D_{TNS}).

⁶At this point in the computation, the speech has been segmented into phrases, thus enabling the part-of-speech tagger to determine the actual part of speech of each word.

Name	Description	Units
D_{SL}	Arc Length	Pixel
D_{TNS}	Time to Next Stroke	ms
D_{DUR}	Stroke Duration	ms
S_{DUR}	Segment Duration	ms
S_{NV}	Contains Noun/Verb	Boolean
S_{WC}	Word Count	Word
S_{SC}	Stroke Count	Stroke
S_{SE}	Stroke Start End Duration	ms
S_{SS}	Stroke Start Start Duration	ms
S_{TN}	Time to Next Segment	ms
S_{TP}	Time to Previous Segment	ms

Table 10.9: Features used for segment-stroke alignment: D_x = sketch (drawing) feature; S_x = segment feature.

Four other features describe temporal relationships. S_{TN} is the time to the next segment, while S_{TP} is the time to the previous one. S_{SE} is the time between the start of the segment and the end of the associated stroke. Likewise, S_{SS} is the time from the start of the segment to the start of the associated stroke. Both of these features can have positive or negative values.

To train the case-10 classifier, all of the initial associations in the training set are labeled with a binary value indicating whether or not they are a case-10 error. An analogous approach is used to train the case-12 classifier.

10.5.9.1 Stroke-Speech Alignment Accuracy

To evaluate the performance of our two-step segment-stroke alignment technique, we again performed a leave-one-out cross-validation. In each iteration of the cross-validation, one sketch with speech was used for testing, while the others were used for training. We averaged the results across the 48 testing/training combinations.

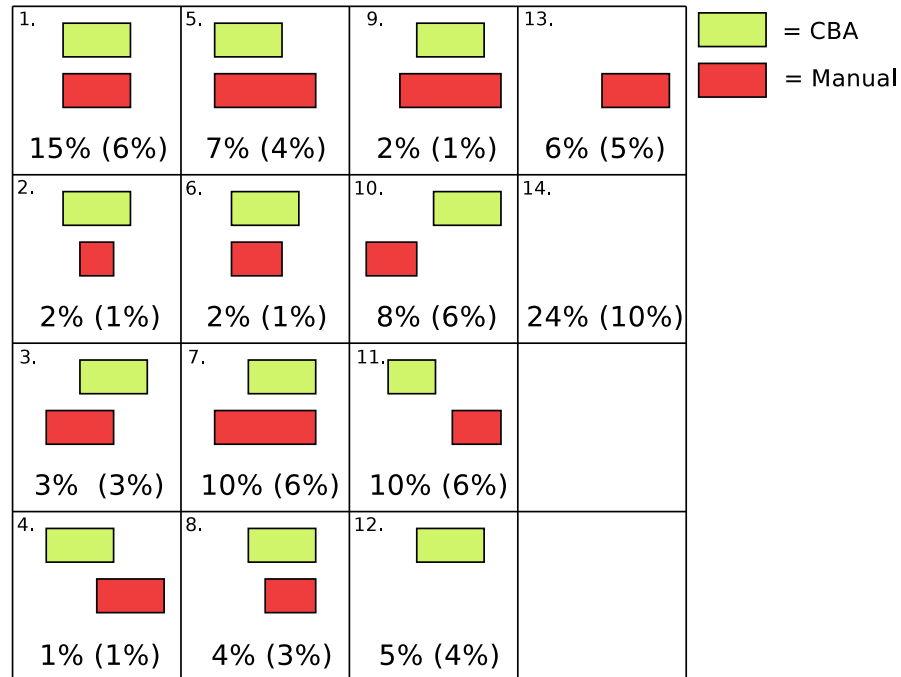


Figure 10.19: Accuracy of the classifier-based alignment (“CBA”) technique. In each cell, the top bar represents the speech associated with a stroke by CBA, while the bottom bar represents the manual (correct) association. Each cell represents a distinct relative arrangement of the associations and the frequency with which it occurs. Results averaged over the 48 sketches. Standard deviations are presented in parenthesis.

Figure 10.19 compares the final alignment to the correct (manual) alignment. The case-10 and case-12 classifiers were clearly effective. The case-10 errors have been reduced from an average of 18% in Figure 10.18 to an average of only 8%. Likewise, the case-12 errors have been reduced from an average of 26% to an average of only 5%. Overall, after the second step of alignment, an average of 39% of the associations are perfect (cases 1 and 14). Furthermore, on average only 29% of the associations are completely disjoint from the correct associations (cases 10 through 13).

To provide a more detailed evaluation of the alignment accuracy, we also computed the number of missing and extra words in each association. Extra words are those

associated with the pen stroke that should not have been. Conversely, missing words are those that should have been associated but were not. Consider the hypothetical example⁷ in Figure 10.20. The stroke is associated with the words “faces the other way this is uh like a”. The correct association (as determined by manual alignment) is the clause “this is uh like a handle.” In this case, the words “faces the other way” are extra words, and “handle” is a missing word.

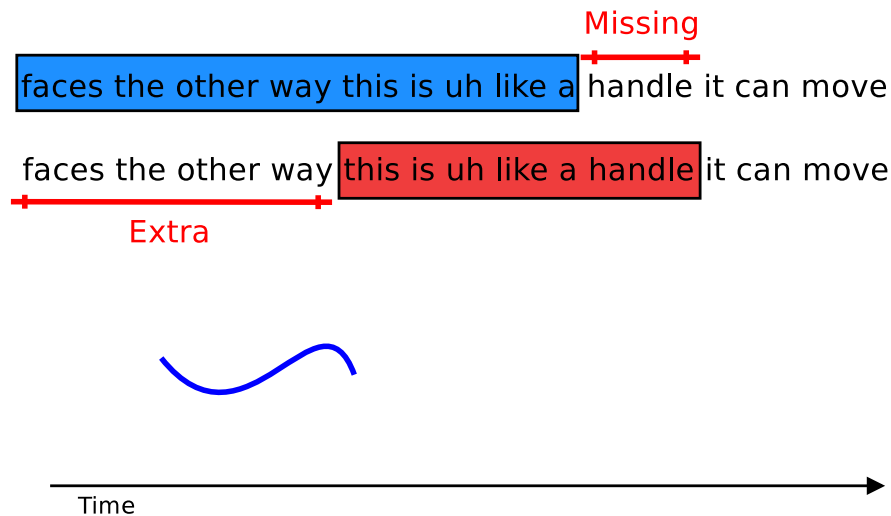


Figure 10.20: Missing and extra words in speech aligned with a pen stroke. The top bar indicates the words actually associated with the pen stroke; the bottom bar indicates the words that should have been associated.

Figure 10.21 presents the missing/extra accuracy of both the three-second and classifier-based alignment techniques. On average, the three-second approach has about 12 extra words and two missing words per stroke, while our classifier-based approach has only about two extra and four missing. Overall, the three-second approach has an average of 14 incorrect (missing plus extra) words per stroke, while our new approach

⁷The speech is taken directly from the user study data. The hypothetical stroke was designed to monotonically increase along the horizontal axis, thus suggesting a drawing process evolving in time.

has only six. This is a 57% reduction in errors.

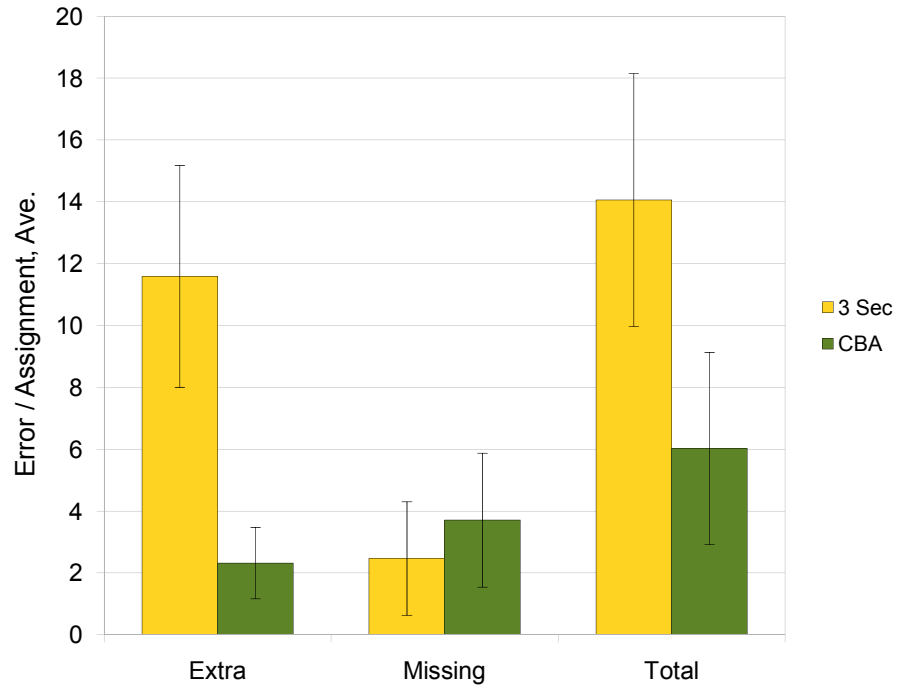


Figure 10.21: Average number of words incorrectly aligned with each stroke for the three-second and classifier-based alignment techniques. Averages are computed over the 48 sketches.

10.5.10 Gesture/Object Classification Accuracy

Our purpose in creating an improved technique for speech-sketch alignment is to enable more accurate identification of gesture pen strokes like those in Figure 10.10(b). Thus, to evaluate our classifier-based alignment technique, we computed the gesture classification accuracy using our technique and compared this to the accuracy achieved with the three-second alignment technique. We also computed the accuracy using manual alignment to obtain an upper bound on the achievable gesture classification accuracy.

For this analysis, we used the gesture classifier just as described in Section 10.5.3.1, except that we used an Ada-boosted C4.5 decision tree computed with WEKA [44] rather than using a neural network. The sketch features were computed as before, while the speech features were computed using the speech-sketch alignment technique in question.

We computed accuracy via leave-one-out cross-validation, with one sketch used for testing and the others used for training. Our results are the average across the 48 testing/training combinations. We evaluated classification accuracy for four sets of features: (a) the thesaurus Bayesian filter feature (W_{TBF}), (b) the five most important features, (c) the 10 most important features, and (d) all features. We determined the top five and top 10 features using an information gain algorithm [29, 136] as implemented by WEKA.⁸ The top five features include: the two Bayesian filter features (W_{TBF} , W_{BF}), the total elapsed time (D_{ET}), the time to the closest prior stroke (D_{TCS}), and the time to the next stroke (D_{TNS}). The top 10 features additionally include: the time to the previous stroke (D_{TPS}), the distance to previous stroke (D_{DPS}), the distance to the

⁸As discussed in Section 10.5.4, the work in [10] employed a beam search approach to determine the best features. While that approach may be more reliable, here we use information gain in the interest of expediency.

next stroke (D_{DNS}), the maximum Hausdorff distance to the underlying ink (D_{MHD}), and the average Hausdorff distance to the underlying ink (D_{AHD}).

The gesture/object classification results are shown in Figure 10.22. (The accuracy in Figure 10.22 differs from that in [10] because different classifiers were used, i.e., a neural network vs. Ada-boosted decision tree.) Typically, for a given set of features, the classifier-based alignment resulted in better accuracy than the three-second alignment, and the manual alignment resulted in the best accuracy. Likewise, using more features typically resulted in better accuracy. There was one exception. The three-second approach achieved nearly its best accuracy when only the thesaurus Bayesian filter feature was used. For this single-feature case, the three-second approach actually achieved better accuracy than even the manual alignment. This is discussed in the next section.

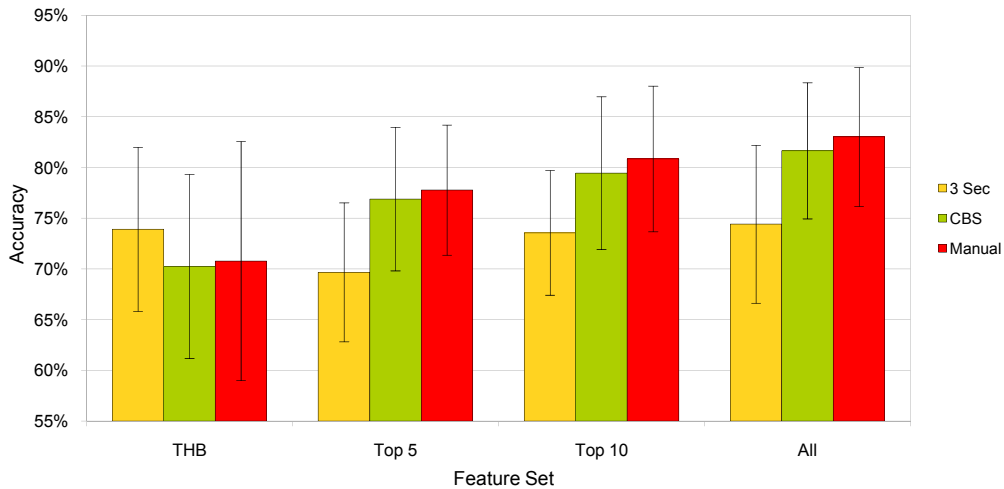


Figure 10.22: Gesture/object classification accuracy vs. speech-sketch alignment technique and number of features. “THB” = thesaurus Bayesian feature. Results are averaged over the 48 sketches.

10.5.11 Discussion

Our speech segmenter achieved an average accuracy of 92.7% in classifying words as *start*, *middle*, *end*, and *only* words. While the classification accuracy is high, a more important measure of accuracy is the segmentation accuracy: on average, 75.9% of the computed segments matched correct (manual) segments within a three-word tolerance.

Liu et al. [82] define a per-boundary-based metric for speech segmentation accuracy. This is defined as the sum of the false positive and false negative sentence boundaries normalized by the total number of interword boundaries. With their state-of-the-art technique based on a conditional random field model, they achieve a boundary-based accuracy of 95.4% on conversational telephone speech. (They achieve higher accuracy on broadcast news which is more grammatical than telephone conversation.) We achieve 86.8% accuracy on multimodal dialog that includes both speech and sketching. Our results compare favorably with theirs for several reasons. First, sentence boundaries occur less frequently in their data than in ours: in their dataset only 15.7% of interword boundaries are actually sentence boundaries, whereas in our dataset 21.8% are. Thus, a naïve classifier would perform better on their data than on ours. Also, their accuracy is likely to benefit from a greater amount of training data: they trained on 480,000 words, whereas we trained on about 33,000. Finally, we consider different kinds of speech: theirs is unimodal while ours is multimodal.

Our speech segmentation approach is unique in that it demonstrates the usefulness of sketch features for locating segment boundaries in multimodal dialog. Also, we

use a single classifier to directly combine word-based, prosodic (pause between words),⁹ and sketch-based features. Unlike many existing approaches (e.g., [118, 119, 41, 66]), we use a four-class (*start*, *middle*, *end*, and *only*) approach to locating segment boundaries. This approach was designed to take advantage of the discriminatory power of single-word segments. In future work, we plan to compare the performance of this approach to that of a more traditional approach in which interword boundaries are classified as either segment boundaries or non-boundaries. Also, unlike traditional approaches, we do not explicitly consider word sequence — we have no n-gram language model. When processing a given word, our classifier does consider the previous and next words, but we do not use a Markovian approach. In future work, we plan to combine our technique with an explicit language model, but this will likely require a much larger dataset. For example, to provide a benchmark for our results, we implemented the technique in [118] using a trigram language model. This approach performed poorly on our data: of the hypothesized segment boundaries, on average only 1.3% were true boundaries, while for our approach 79% were. It is likely that our corpus containing only about 34,000 words is too small to train the trigram model.

Our classification-based speech-sketch alignment technique performed significantly better than the three-second technique as indicated by multiple measures. On average, CBA aligned only six incorrect words (missing plus extra) per pen stroke, whereas the three-second approach had 14. Comparison of Figures 10.15 and 10.19 further illustrates the superiority of the CBA technique. For example, on average, CBA perfectly aligned the speech (cases 1 and 14) for 39% of pen strokes, whereas the three-

⁹The filled pauses discussed in Section 10.5.6 are not prosodic features but instead are a lexical concept. The time elapsed between words is the only prosodic feature we use.

second approach did this for only 2% of strokes. Similarly, for CBA an average of only 29% of the associations were completely disjoint from the correct associations (cases 10 – 13), whereas for the three-second approach 41% were. Likewise, for CBA an average of only 31% of the associations were partially disjoint (cases 3 – 9), whereas for the three-second approach 56% were. Note that for some particular completely or partially disjoint cases, the three-second approach did have fewer errors than CBA. However, on the whole, CBA had far fewer completely and partially disjoint cases, and thus overall is significantly more accurate than the three-second approach.

Leaving aside the case of the single-feature classifier, the results in Figure 10.22 support our hypothesis that better speech-sketch alignment leads to better accuracy for classifying pen strokes as gestures or object strokes. Our classification-based alignment technique resulted in much greater accuracy than the three-second approach, and performed nearly as well as the manual alignment.

The single-feature case, however, is an interesting anomaly. To understand why the three-second alignment technique outperformed even the manual alignment when the classifier used only the thesaurus Bayesian filter feature, we examined the distribution of the values of this feature for the three alignment methods as shown in Figure 10.23. Comparatively speaking, the three-second alignment results in a bimodal distribution in which each stroke is either a gesture (feature value of 1) or not (feature value of 0). The other two methods, by contrast, have a greater percentage of cases with a probability of 0.5, which indicates that a stroke is equally likely to be a gesture or object stroke. Thus, with more accurate alignment, the thesaurus Bayesian filter is

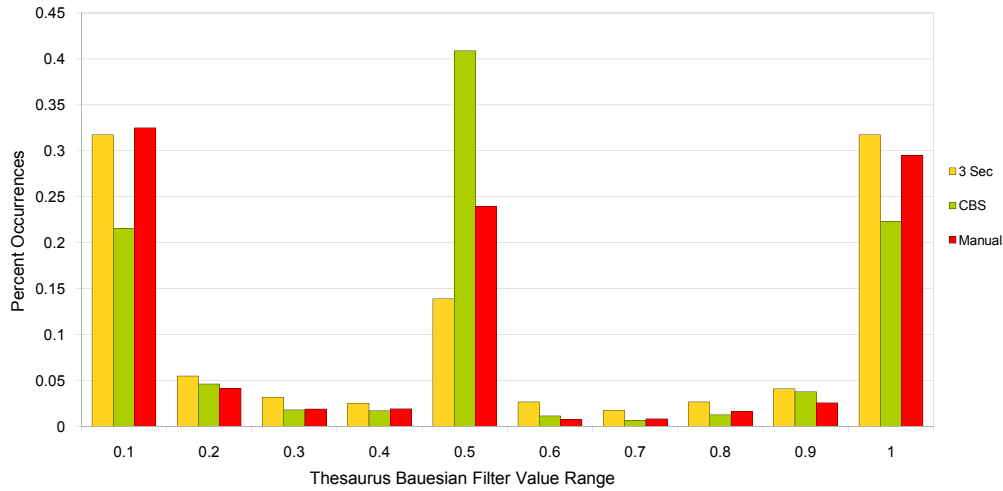


Figure 10.23: Histograms of the values of the thesaurus Bayesian filter feature for different alignment techniques.

able to make predictions about fewer strokes.

Examining Figure 10.21 gives some additional insight into this anomaly. The three-second alignment technique tends to align many extra words with each pen stroke. These extra words may allow the thesaurus Bayesian filter to make predictions for strokes that do not actually have associated speech. For strokes that do have associated speech, we would expect that better alignment would result in better classification accuracy. To test this hypothesis, we evaluated gesture/object classification accuracy for only those strokes with associated speech as determined by the manual alignment. Here again, we computed accuracy via leave-one-out cross-validation, with one sketch used for testing and the others used for training. However, in this case only strokes with associated speech were included in the testing and training sets.

The results are shown in Figure 10.24. For strokes with associated speech, improved alignment does result in improved accuracy, even when only the thesaurus Bayesian filter is used. It appears that the over-association of words by the three-

second approach is useful when only speech is used for gesture/object classification. However, the benefit is quickly lost as additional features are used. Apparently, the noi: 5.

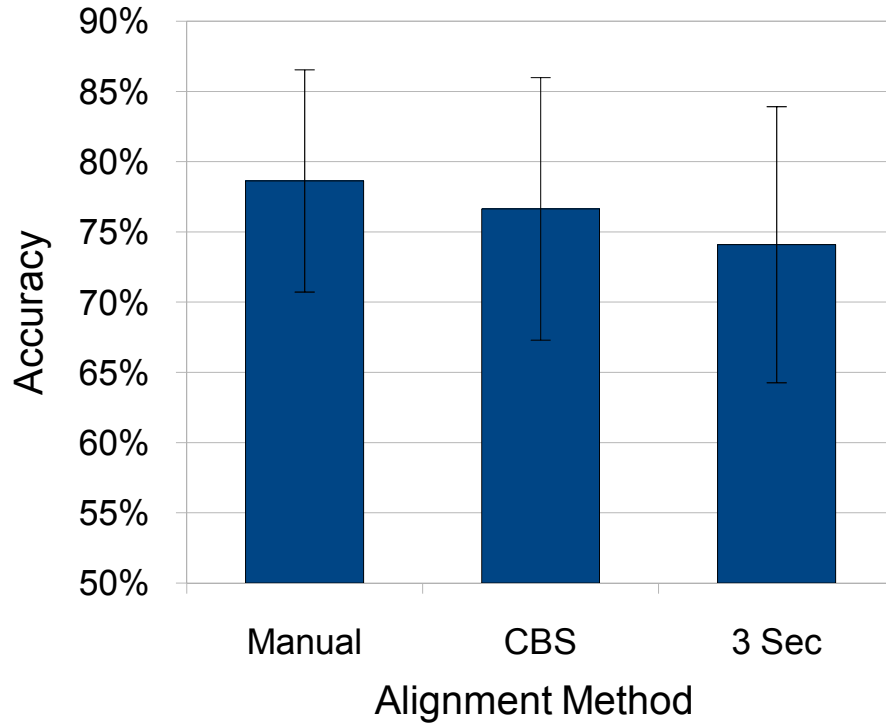


Figure 10.24: Gesture/object classification accuracy for strokes known to have associated speech. Classification based on only the thesaurus Bayesian filter feature.

Currently, our system is designed to be applied once the device description has been completed. An important next step will be to adapt our system to work in real-time so that strokes are classified as they are drawn. All of the features used for the various classifiers can be computed on the fly as they depend only on prior information. Thus, the primary challenge in creating a real-time system will be the problem of automatic speech recognition. The state-of-the-art Sphinx-4 speech recognition system [125] achieves a word error rate of 7% with a vocabulary of 5,000 words and a word

error rate of 19% with a vocabulary of 60,000 words. The errors inherent in automatic speech recognition will clearly present challenges. However, we may be able to compensate for these errors by using additional prosodic features (we currently use only pause duration).

We evaluated our techniques using a nearly user-independent approach. The training data used when testing on a particular sketch was comprised of 44 sketches by other authors, and only three sketches from the primary author of the test sketch. It is likely that increasing the amount of user-specific training data will increase the accuracy of the system. Such training data has proven beneficial for other recognition tasks, such as hand-drawn symbol recognition [36].

We have developed our techniques within the domain of collaborative engineering design, but they should generalize to many other domains. None of the features used by our classifiers are specific to mechanical devices or the task of designing, thus we believe our techniques should be suitable for any domain in which the task involves drawing a sketch or diagram and explaining its elements. Examples of such domains include giving driving directions, explaining the solution to a problem in a physics lecture, and explaining a sports play.

10.5.12 Conclusion

We have presented a new technique for aligning speech and sketch input in multimodal dialog. It is designed for use in classifying pen strokes as gesture and object strokes. The technique, which we call classifier-based alignment, employs a two-step process: the speech is first segmented into meaningful pieces (typically clauses), then

the segments are aligned with pen strokes. Our speech segmenter uses a statistical classifier to classify words according to their position in a segment. We consider four classes of words: *start*, *middle*, *end*, and *only* words. The word classifications are then used to form speech segments. The segment-stroke alignment step initially uses temporal correlation to align segments with pen strokes. Classifiers are then used to detect and correct two common alignment errors.

Our classification-based speech-sketch alignment technique performed significantly better than the existing “three-second” alignment technique, which is based solely on temporal correlation and has no explicit segmentation step. On average, our technique perfectly aligned the speech for 39% of pen strokes, whereas the three-second technique did this for only 2% of strokes. Furthermore, for our technique the aligned speech had no overlap with the correct alignment, on average, for only 29% of strokes. However, for the three-second technique there was no overlap for 41% of strokes. Finally, our technique had on average only six incorrectly aligned words (missing plus extra) per pen stroke, whereas the three-second approach had 14.

Our alignment technique is novel in that it uses information from the sketch modality for both the speech segmentation and alignment steps. Our results indeed demonstrate that features from the sketch input are valuable for segmenting speech.

Our purpose in developing an effective speech-sketch alignment technique was to enable accurate identification of gesture pen strokes in multimodal dialog. Our gesture classifier uses features of the pen strokes and the speech aligned with them. Experiments with this classifier demonstrated that, when multiple speech and sketch features are used for classification, better alignment accuracy does lead to more accurate gesture

classification. More precisely, when multiple features are used, our alignment technique resulted in much greater gesture classification accuracy than the three-second approach, and performed nearly as well as manual alignment. Inaccurate alignment was beneficial only when the gesture classifier used just a single statistical speech feature. In this case, the tendency of the three-second alignment technique to erroneously associate extra words with pen strokes allowed the gesture classifier to make predictions about pen strokes that in reality had no associated speech. Thus, in all but one unusual case, our new alignment technique enables substantially more accurate gesture classification than the prior technique.

Chapter 11

Conclusions

In this work, we have applied a breadth of machine learning and data mining techniques to students' ordinary, handwritten coursework. Educational Data Mining research requires digital instrumentation of students' learning processes. So far, prior work has typically considered educational data extracted from either Learning Content Management Systems or Intelligent Tutoring Systems. This has led to interesting discoveries in the ways students interact with these systems, but an investigation of students' problem-solving processes in their ordinary learning environment, i.e., working with pen and paper on their own schedule, has hitherto been unexplored.

In Chapter 3, we presented an in-depth description of our unique database of students' ordinary, handwritten coursework. For the past four years, students in an undergraduate Mechanical Engineering course at the University of California, Riverside have been given LiveScribe™ digital pens. These pens function as traditional ink pens, allowing students to write on paper. Additionally though, these pens digitize the work, generating a digital, time-stamped record of every pen stroke written. Students

were asked to complete all coursework each year with these pens, including homework, quizzes, midterms, and exams. By doing so, we have generated an electronic record of each students' problem-solving in the course. This record is, to our knowledge, the first of its kind. We have presented a systematic set of descriptive statistics and histograms which characterize the amount of time students spent on different parts of the coursework. This provides unique insights into how students solve problems in the context of this course. Furthermore, we have built experiments into some of the course offerings. Namely, during the 2011 and 2012 years, some students were provided with self-explanation prompts while others were not. Self-explanation has been demonstrated to positively impact students' learning gains, and by electronically capturing the handwritten work of those students who did and did not generate self-explanation we have begun to understand how this exercise affects students' problem-solving behavior.

We have found that not all self-explanation generated by the students was substantive. Thus, in Chapter 4, we applied an open information extraction technique to automatically determine if a student's self-explanation contained concepts that were also found in self-explanations generated by experts. This technique has proven to be quite reliable, achieving an accuracy of up to 97% at identifying the concepts on a particular explanation. This paves the way for an automated system which processes a student's self-explanation immediately after it is written and then prompts the student for further self-explanation if it is not substantive. It will be interesting in future work to study if the content of students' self-explanation correlates with performance.

Beyond identifying the content present within the transcripts of students' self-explanations, in Chapter 5, we sought to identify the ways in which students' problem-

solving behaviors were modified as a result of generating self-explanation. The same experts who generated self-explanations in Chapter 4 also solved the same homework assignments as the students. We found that these experts always completed problems in the order assigned, e.g., completing problem one before problem two, then problem three, and so on. Using this intuition, we investigated the problem number sequences of those students who did and did not generate self-explanations. We used an n-gram analysis, common in natural language processing tasks, to identify those problem number subsequences, or n-grams, which occurred more frequently in one group of students than in the other. This technique revealed that students who did not generate self-explanation were more likely to “bounce” back-and-forth between solving a particular problem and the previous problem. On the other hand, students who did generate self-explanation were more likely to solve problems in the order assigned, and thus, more like an expert. This study has demonstrated not only that self-explanation leads to improved learning gains, but also that considering the sequential aspect of students solutions can be used to distinguish between experimental groups. It will be interesting, in future work, to consider sequences beyond those that just consider problem number order and instead characterize other aspects of students’ problem solving, such as semantic content. This will enable studies to identify further differences in the ways students who do and do not generate self-explanations solve their problems. Furthermore, this technique is general enough that it may be applied to studies considering treatments other than self-explanation.

In Chapter 6 we expanded our investigation of sequences. We broadened our focus and considered not only the problem numbers that students worked on, but also

the semantic content of the writing and its duration. We developed an alphabet of canonical actions a student could make while solving a homework problem. Each action represented an uninterrupted period of problem-solving performed by the student and was characterized by duration, semantic content, and homework number. We have developed 49 unique actions. We then represented each student's solution to an entire homework problem as a sequence of these actions, called an *action sequence*. Instead of comparing the experimental groups, as in Chapter 5, here we compared the highest and lowest performing students. We applied a differential mining technique, previously developed by the educational data mining community, to identify those patterns that appear significantly more often in one group of students than in the other. This allowed us to capture behaviors exhibited by good- and poor-performing students. These behaviors provide insight into the problems and concepts that students may be having difficulty with. For example, we found that poor-performing students often repeatedly attempted the free body diagram portion of the first problem in a homework assignment. This may indicate that students who have difficulty completing their free body diagram, especially when it is the first in the assignment, are likely to not do well on that assignment, and thus should be targeted with instructional materials that guide them in their free body diagram construction. This work has shown that expanding the sequence information to consider semantic content and duration in addition to problem number provides greater insights into students' problem-solving behaviors. So far, the techniques described which seek out sequential patterns in students' work has been applied only to their homework solutions. It would be interesting in future work to consider sequences that appear in their exam work, and even more interesting to compare

the homework and exam sequences.

In related work, it was shown that the spatial organization students used when solving exam problems was indicative of their performance on those problems, but what are the ways in which students organize their solutions? To investigate this question in Chapter 7, we have applied an unsupervised learning technique to down-sampled bitmaps of the students' solutions. By greatly down-sampling, we abstract away minute differences between students' writing and capture the overall organization. We compute distances between these bitmaps using the Hausdorff distance, a common image-based metric. We were then able to then apply K-Means clustering to group the bitmaps by their distances to each other. Each group discovered with this algorithm can be considered its own distinct organizational category. We have compared the performance of students in each group and have found that often, there is a significant difference in their performance corresponding to the spatial organization of their solution. Furthermore, these groups characterize the different solution paths that students may take to solve particular problems and are the beginnings of a taxonomy of the solutions generated by students in the context of this course. This work demonstrates the regular nature of students' spatial organization on final exam problem. The techniques used in this study, K-Means clustering with the Hausdorff distance, can be improved in future work by applying more rigorous clustering methods, e.g., EM clustering, and more rigorous distance metrics, e.g., comparing the Z-order curves of the bitmaps. Furthermore, it would be interesting in future work to consider how different spatial organization groups perform in the class, by perhaps computing features based on the organizational groups that were used to train a statistical classifier to predict performance in the course.

In Chapter 6, we used the differential patterns to construct features to train a linear regression model which predicted students' performance in the course. We then used the underlying parameters of that model as indicators of which features, and subsequently patterns, were most indicative of good or bad performance. We continue with this idea of using model parameters to guide later analysis in Chapters 8 and 9. Instead of using sequential information, as in Chapter 6, in Chapters 8 and 9 we instead use numerical features. In particular, we use features which estimate the amount of effort expended by students on their homework assignments, namely, the total amount of ink written on different components of each homework assignment. In Chapter 8 we used the amount of ink written on each individual homework problem as its own feature. Furthermore, we computed four features which characterized how students distributed their effort during the problem-solving process. We used these features to train models which predicted both overall performance in the course as well as performance on individual exam problems. Again, we used the underlying parameters of those models to determine which features, and in turn homework problems, were most indicative of good and poor performance. This demonstrated, for example, that if students expended a great deal of effort on problem three of homework four, then they were more likely to do well on midterm one, problem one. By comparing these two problem descriptions, we were able to make conjectures about the types of knowledge transfer students made from particular homework problems to particular midterm problems that enabled them to do well on the midterm problems. This work is important, as it shows that the models computed from students handwritten coursework guide discoveries about the ways students' learn, and indicate which factors are most important to their success in the

course.

In Chapter 9, we used more fine-grained features than those used in Chapter 8 to predict more fine-grained performance metrics. The features investigated in Chapter 9 were the amount of free body diagram and equation ink written on each problem of each homework assignment. By considering these features, we are given a microscopic view into what students are spending their time and effort working on. Instead of predicting overall performance on exams, we use these features to predict if students will make specific errors on four of the midterm problems from the 2012 course offering. These errors are binary and indicate whether or not a student made that error in his or her solution. Furthermore, Chapter 9 expands upon Chapter 8 in the way that features are used to predict performance. In Chapter 8, a linear regression model that considered all effort-based features was computed, and then important homework problems were identified using the underlying parameters of that model. Instead in Chapter 9, we individually compute the information gained about each target error using each feature individually. In doing so, we lose information about interactions between these features, but our goal is to identify relationships between individual homework problems and midterm problems. This analysis revealed the features, and subsequently homework problems, that were most consistently predictive of student performance on each midterm problem. As a result, we have developed a technique for a microscopic study of which problems students should spend the most time working on and the characteristics of homework problems that instructors should employ in future offerings of the course. This work, and that of Chapter 8, has shown that the amount of ink spent writing on homework, which may be a rough estimation of effort, does indeed correlate with per-

formance in the course. Furthermore, this work shows that the more fine-grained the features and target metrics, the stronger the correlations between them. Future work should continue to consider the most fine-grain errors and numerical features possible. In this study, we have only considered amount-of-ink-based features. It will be interesting in future work to combine spatial, sequential, and temporal features and determine how well they predict performance when used together.

In this work, we have shown, for the first time, that one may infer a student's problem-solving process from his or her ordinary handwriting when captured with timestamps. We have demonstrated that scalable data mining and machine learning analyses of the sequential, spatial, and temporal characteristics of students' problem-solving process provide novel insights into the way students learn. By considering the problem number sequence of students' homework solutions, we were able to identify that generating self-explanations causes students to solve their assignments more like an expert. Additionally, by considering the action sequences of high- and low-performing students, we have identified problem-solving behaviors that both provide insights to the instructor and also may be used in an automated system to identify when students are having difficulty on their homework assignments. By applying an unsupervised learning technique to the spatial organization of students' exam solutions, we have revealed different solution organization styles which we have shown are indicative of performance. This is a low-cost method for predicting students' performance by considering the organization of the ink on the page. Lastly, by considering the amount of ink written on students' homework assignments, we have shown that effort spent on homework assignments correlates with performance on exam problems. Furthermore, this analysis revealed the

problems that most correlated with exam performance, in turn identifying the transfer that students made from homework to exams. In the future, this will enable a system to provide instructors with rapid feedback concerning the effectiveness of homework problems and the concepts students have learned.

Bibliography

- [1] A. Adler and R. Davis. Speech and sketching: an empirical study of multimodal interaction. In *SBIM '07: Proceedings of the 4th Eurographics workshop on Sketch-based interfaces and modeling*, pages 83–90, New York, NY, USA, 2007. ACM.
- [2] Rakesh Agrawal and Ramakrishnan Srikant. Mining sequential patterns. In *Data Engineering, 1995. Proceedings of the Eleventh International Conference on*, pages 3–14. IEEE, 1995.
- [3] David W. Aha and Richard L. Bankert. A comparative evaluation of sequential feature selection algorithms, 1994.
- [4] Hirotugu Akaike. A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*, 19(6):716–723, 1974.
- [5] J. L. Antonek. Interactive homework: Linking parents to the foreign language learning of their children. In *Dissertation Abstracts International*, volume 57, 1996.
- [6] Ron Artstein and Massimo Poesio. Bias decreases in proportion to the number of annotators. *Proceedings of FG-MoL*, pages 141–150, 2005.
- [7] Jay Ayres, Jason Flannick, Johannes Gehrke, Tomi Yiu, et al. Sequential pattern mining using a bitmap representation. In *Conference on Knowledge Discovery in Data: Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*, volume 23, pages 429–435, 2002.
- [8] S. Bai and S. Bai. The maximal frequent pattern mining of dna sequence. In *Granular Computing, 2009, GRC '09. IEEE International Conference on*, pages 23 –26, 2009.
- [9] Katerine Bielaczyc, Peter L. Pirolli, and Ann L. Brown. Training in self-explanation and self-regulation strategies: Investigating the effects of knowledge acquisition activities on problem solving. *Cognition and Instruction*, 13(2):221–252, 1995.

- [10] David Bischel, Thomas Stahovich, Eric Peterson, Randall Davis, and Aaron Adler. Combining speech and sketch to interpret unconstrained descriptions of mechanical devices. In *IJCAI'09: Proceedings of the 21st international joint conference on Artificial intelligence*, pages 1401–1406, San Francisco, CA, USA, 2009. Morgan Kaufmann Publishers Inc.
- [11] C.M. Bishop, M. Svensen, and G.E. Hinton. Distinguishing text from graphics in on-line handwritten ink. In *Proc. of the Int. Workshop on Frontiers in Handwriting Recognition*, pages 142–147, 2004.
- [12] K. Bloomenthal and R. Zeleznik. SKETCH-N-MAKE: Automated machining of CAD sketches. In *ASME DETC'98*, pages 1–11, 1998.
- [13] Richard A. Bolt. “put-that-there”: Voice and gesture at the graphics interface. In *SIGGRAPH '80: Proceedings of the 7th annual conference on Computer graphics and interactive techniques*, pages 262–270, New York, NY, USA, 1980. ACM.
- [14] N. K. Bowen and G. L. Bowen. The mediating role of educational meaning in the relationship between home academic culture and academic performance. In *Family Relations*, volume 47, pages 45–51, 1998.
- [15] Ann L. Brown. Knowing when, where, and how to remember: A problem of metacognition. Technical Report 47, University of Illinois at Urbana-Champaign, 1977.
- [16] Ann L. Brown and J.D. Day. Macrorules for summarizing texts: The development of expertise. *Journal of Verbal Learning and Verbal Behavior*, 22:1–14, 1983.
- [17] David C. Brown, Stanley C. Kwasny, B. Chandrasekaran, and Norman K. Sondheimer. An experimental graphics system with natural language input. *Computers & Graphics*, 4(1):13–22, 1979.
- [18] Peter F. Brown, Peter V. deSouza, Robert L. Mercer, Vincent J. Della Pietra, and Jenifer C. Lai. Class-based n-gram models of natural language. *Comput. Linguist.*, 18:467–479, December 1992.
- [19] Justine Cassell. A framework for gesture generation and interpretation. *Computer Vision in Human-Machine Interaction*, pages 191–215, 1998.
- [20] Joyce Y. Chai, Pengyu Hong, and Michelle X. Zhou. A probabilistic approach to reference resolution in multimodal user interfaces. In *IUI '04: Proceedings of the 9th international conference on Intelligent user interfaces*, pages 70–77, New York, NY, USA, 2004. ACM Press.
- [21] Joyce Y. Chai, Zahar Prasov, Joseph Blaim, and Rong Jin. Linguistic theories in efficient multimodal reference resolution: an empirical investigation. In *IUI '05: Proceedings of the 10th international conference on Intelligent user interfaces*, pages 43–50, New York, NY, USA, 2005. ACM Press.

- [22] Lei Chen and Mohamed S. Kamel. Design of multiple classifier systems for time series data. In *Proceedings of the 6th international conference on Multiple Classifier Systems*, MCS'05, pages 216–225, Berlin, Heidelberg, 2005. Springer-Verlag.
- [23] Michelene T. H. Chi, Nicholas De Leeuw, Mei hung Chiu, and Christian Lavancher. Eliciting self-explanations improves understanding. *Cognitive Science*, 18:439–477, 1994.
- [24] T. H. Michelene; Miriam Bassok; Matthew W. Lewis; Peter Reimann; Robert Glaser Chi. Self-explanations: How students study and use examples in learning to solve problems. In *Cognitive Science*, pages 145–182, 1989.
- [25] Philip R. Cohen, Michael Johnston, David McGee, Sharon Oviatt, Jay Pittman, Ira Smith, Liang Chen, and Josh Clow. Quickset: multimodal interaction for distributed applications. In *MULTIMEDIA '97: Proceedings of the fifth ACM international conference on Multimedia*, pages 31–40, New York, NY, USA, 1997. ACM.
- [26] H. Cooper, J. Robinson, and P. Erika. Does homework improve academic achievement? *Review of Educational Research*, 76, 2006.
- [27] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.
- [28] Aron Culotta. Integrating probabilistic extraction models and data mining to discover relations and patterns in text. In *In Proceedings of the HLT-NAACL-2006*, pages 296–303, 2006.
- [29] M. Dash and H. Liu. Feature selection for classification. *Intelligent Data Analysis*, 1:131–156, 1997.
- [30] Arthur P Dempster, Nan M Laird, and Donald B Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 1–38, 1977.
- [31] R. Deslandes, E. Royer, P. Potvin, and D. Leclerc. Patterns of home and school partnership for general and special education students at the secondary level. In *Exceptional Children*, volume 65, pages 496–506, 1999.
- [32] David Efron. *Gesture And Environment*. New York: King's Crown Press, 1941.
- [33] L. Eggli. Sketching with constraints. Master's thesis, University of Utah, 1994.
- [34] R. Zeleznik et al. SKETCH: an interface for sketching 3D scenes. In *Proceedings of SIGGRAPH'96*, pages 163–170, 1996.
- [35] Oren Etzioni, Michele Banko, Stephen Soderland, and Daniel S. Weld. Open information extraction from the web. *Commun. ACM*, 51:68–74, December 2008.

- [36] M. Field, S. Gordon, E. Peterson, R. Robinson, T. Stahovich, and C. Alvarado. Technical section: The effect of task on classification accuracy: Using gesture recognition techniques in free-sketch recognition. *Comput. Graph.*, 34:499–512, October 2010.
- [37] Manuel J. Fonseca and Joaquim A. Jorge. Using Fuzzy Logic to Recognize Geometric Shapes Interactively. In *Proceedings of the 9th Int. Conference on Fuzzy Systems (FUZZ-IEEE 2000)*, San Antonio, USA, May 2000.
- [38] Kenneth D. Forbus, Ronald W. Ferguson, and Jeffery M. Usher. Towards a computational model of sketching. In *Proceedings of the 6th International Conference on Intelligent User Interfaces*, pages 77–83. ACM Press, 2001.
- [39] Yoav Freund and Robert Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. In *Computational learning theory*, pages 23–37. Springer, 1995.
- [40] J. J. Godfrey, E. C. Holliman, and J. McDaniel. SWITCHBOARD: telephone speech corpus for research and development. In *Proc. ICASSP*, volume 1, pages 517–520 vol.1, 1992.
- [41] Yoshihiko Gotoh and Steve Renals. Sentence boundary detection in broadcast speech transcripts. In *in Proc. of ISCA Workshop: Automatic Speech Recognition: Challenges for the new Millennium ASR-2000*, pages 228–235, 2000.
- [42] P. Graham. *Hackers and Painters, Big Ideas from the Computer Age*. O’Reilly, 2004.
- [43] Puneet Gupta, David Doermann, and Daniel DeMenthon. Beam search for feature selection in automatic svm defect classification. *Pattern Recognition, International Conference on*, 2:20212, 2002.
- [44] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten. The weka data mining software: An update. In *SIGKDD Explorations*, 2009.
- [45] Mark A Hall. *Correlation-based feature selection for machine learning*. PhD thesis, The University of Waikato, 1999.
- [46] Simin Hall and Eric A. Vance. Improving self-efficacy in statistics: Role of self-explanation and feedback. *Journal of Statistics Education*, 18(3), 2010.
- [47] Tracy Hammond and Randall Davis. Ladder: A sketch recognition language. Technical report, IN MIT COMPUTER SCIENCE AND ARTIFICIAL INTELLIGENCE LABORATORY ANNUAL RESEARCH ABSTRACT. MIT CSAIL, 2004.
- [48] L. Han-lung, T. Stahovich, and J. Herold. Automatic handwritten statics solution classification and its applications in predicting student performance. In *Proceedings of the American Society for Engineering Education Annual Conference*, 2012.

- [49] Akira Hayashi, Yuko Mizuhara, and Nobuo Suematsu. Embedding time series data for classification. In *Proceedings of the 4th international conference on Machine Learning and Data Mining in Pattern Recognition*, MLDM'05, pages 356–365, Berlin, Heidelberg, 2005. Springer-Verlag.
- [50] J. Heiser and B. Tversky. Arrows in comprehending and producing mechanical diagrams. *Cognitive Science*, 30(3):581–592, 2006.
- [51] J. Herold, Zundel A., and T. F. Stahovich. Mining meaningful patterns from students' handwritten coursework. In *Proceedings of the Sixth International Conference on Educational Data Mining*, 2012.
- [52] J. Herold and T. Stahovich. Characterizing students' handwritten self-explanations. In *Proceedings of the American Society for Engineering Education Annual Conference*, 2012.
- [53] J. Herold and T. F. Stahovich. Classyseg: A machine learning approach to stroke segmentation. In *Proceedings of the 8th Eurographics workshop on Sketch-based interfaces and modeling*, 2011.
- [54] J. Herold and T. F. Stahovich. Using speect to identify gesture pen strokes in collaborative multimodal device descriptions. *Artificial Intelligence for Engineering Design, Analysis and Manufacturin*, 2011.
- [55] J. Herold and T. F. Stahovich. Automatically understanding student self-explanations. In *Proceedings of the 2012 American Society for Engineering Education Annual Conference and Exposition*, 2012.
- [56] J. Herold and T. F. Stahovich. Characterizing students' handwritten self-explanations. In *Proceedings of the 2012 American Society for Engineering Education Annual Conference and Exposition*, 2012.
- [57] J. Herold and T. F. Stahovich. Using educational data mining to identify correlations between homework effort and performance. In *Proceedings of the 2013 American Society for Engineering Education Annual Conference and Exposition*, 2013.
- [58] J. Herold, A. Zundel, and T. F. Stahovich. A data-driven approach to categorizing the spatial organization of homework solutions. In *Proceedings of the 2013 American Society for Engineering Education Pacific Southwest Section Conference*, 2013.
- [59] James Herold and Thomas F. Stahovich. Speedseg: A technique for segmenting pen strokes using pen speed. *Computers & Graphics*, 35(2):250 – 264, 2011.
- [60] T. F. Herold, J. and Stahovich. The 1¢recognizer: A fast, accurate, and easy-to-implement handwritten gesture recognition technique. In *Proceedings of the 9th Eurographics workshop on Sketch-based interfaces and modeling*, 2012.

- [61] Russell C Hibbeler and SC Fan. *Statics and mechanics of materials*, volume 2. Prentice Hall, 2004.
- [62] Joshua Ho, Lior Lukov, and Sanjay Chawla. Sequential pattern mining with constraints on large protein databases. In *Proceedings of the 12th International Conference on Management of Data (COMAD 2005b)*, pages 89–100, 2005.
- [63] Jerry R. Hobbs, Mabry Tyson, John Bear, and David Israel. Sri international: Description of the fastus system used for muc-4. In *In Proceedings of the Fourth Message Understanding Conference (MUC-4)*, pages 268–275. Morgan Kaufmann, 1992.
- [64] Achim G. Hoffmann, Rex Bing Hung Kwok, and Paul Compton. Using subclasses to improve classification learning. In *EMCL '01: Proceedings of the 12th European Conference on Machine Learning*, pages 203–213, London, UK, 2001. Springer-Verlag.
- [65] X. Huang, F. Alleva, and H. Hon. The SPHINX-II speech recognition system: An overview. *Computer, Speech and Language*, 7:137–148, 1993.
- [66] Ji hwan Kim and P. C. Woodland. The use of prosody in a combined system for punctuation generation and speech recognition. In *In Proc. EUROSPEECH*, pages 2757–2760, 2001.
- [67] Wayne Iba and Pat Langley. Induction of one-level decision trees. In *Proceedings of the Ninth International Conference on Machine Learning*, pages 233–240, 1992.
- [68] T. Igarashi, S. Matsuoka, S. Kawachiya, and H. Tanaka. Interactive beautification: A technique for rapid geometric design. In *UIST '97*, pages 105–114, 1997.
- [69] Nathalie Japkowicz. Concept-learning in the presence of between-class and within-class imbalances. In *Proceedings of the 14th Biennial Conference of the Canadian Society on Computational Studies of Intelligence: Advances in Artificial Intelligence, AI '01*, pages 67–77, London, UK, 2001. Springer-Verlag.
- [70] D. L. Jenkins and R. R. Martin. Applying constraints to enforce users' intentions in free-hand 2-D sketches. *Intelligent Systems Engineering*, 1(1), 1992.
- [71] M. Johnston, S. Bangalore, G. Vasireddy, A. Stent, P. Ehlen, M. Walker, S. Whittaker, and P. Maloor. MATCH: An architecture for multimodal dialogue systems. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, pages 276–383, 2002.
- [72] Levent Burak Kara. An image-based trainable symbol recognizer for sketch-based interfaces. In *in AAAI Fall Symposium Series 2004: Making Pen-Based Interaction Intelligent and Natural*, pages 99–105. AAAI Press, 2004.
- [73] Kendon, Adam. Gesture. *Annual Review of Anthropology*, 26(1):109–128, 1997.

- [74] Dae Hyun Kim and Myoung-Jun Kim. A curvature estimation for pen input segmentation in sketch-based modeling. *Comput. Aided Des.*, 38:238–248, March 2006.
- [75] J. Kinnebrew and G. Biswas. Identifying learning behaviors by contextualizing differential sequence mining with action features and performance evolution. In *Proceedings of the Fifth International Conference on Educational Data Mining*, 2012.
- [76] N. Krahnstoever, S. Kettebekov, M. Yeasin, and R. Sharma. A real-time framework for natural multimodal interaction with large screen displays. In *Proc. of Fourth Intl. Conference on Multimodal Interfaces (ICMI 2002)*, Pittsburgh, PA, USA, October 2002.
- [77] Miroslav Kubat, Robert Holte, and Stan Matwin. Machine learning for the detection of oil spills in satellite radar images. In *Machine Learning*, pages 195–215, 1998.
- [78] James A. Landay and Brad A. Myers. Sketching interfaces: Toward more human interface design. *IEEE Computer*, 34(3):56–64, 2001.
- [79] C. Lee, T. F. Stahovich, and R. C. Calfee. A pen-based statics tutoring system. In *Proceedings of the 2011 American Society for Engineering Education Annual Conference and Exposition*, 2011.
- [80] WeeSan Lee, Levent Burak Kara, and Thomas F. Stahovich. An efficient graph-based recognizer for hand-drawn symbols. *Comput. Graph.*, 31:554–567, August 2007.
- [81] V. C. Lin, D. C. Gossani, and R. A. Light. Variational geometry in computer-aided design. *Computer Graphics*, 15(3), 1981.
- [82] Yang Liu, Andreas Stolcke, Elizabeth Shriberg, and Mary Harper. Using conditional random fields for sentence boundary detection in speech. In *ACL '05: Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*, pages 451–458, Morristown, NJ, USA, 2005. Association for Computational Linguistics.
- [83] Yang Liu, Andreas Stolcke, Elizabeth Shriberg, and Mary P. Harper. Comparing and combining generative and posterior probability models: Some advances in sentence boundary detection in speech. In *Proceedings of the Empirical Methods in Natural Language Processing*, Barcelona Spain, 2004.
- [84] Yun Luo. Can subclasses help a multiclass learning problem? In *Intelligent Vehicles Symposium, 2008 IEEE*, pages 214–219, 4-6 2008.
- [85] James MacQueen et al. Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, volume 1, pages 281–297. California, USA, 1967.

- [86] M. Masry, D.J. Kang, and H. Lipson. A freehand sketching interface for progressive construction of 3D objects. *C&G*, 29(4):563–575, 2005.
- [87] Richard E Mayer. Cognitive, metacognitive, and motivational aspects of problem solving. *Instructional science*, 26(1-2):49–63, 1998.
- [88] McNeill, David. *Hand and Mind: What Gestures Reveal about Thought*. University Of Chicago Press, August 1992.
- [89] JL Meriam and LG Kraige. *Engineering mechanics: statics*. John Wiley & Son, 2009.
- [90] Danielle P Mersch, Alessandro Crespi, and Laurent Keller. Tracking individuals shows spatial fidelity is a key regulator of ant social organization. *Science*, 2013.
- [91] Scott Miller, Michael Crystal, Heidi Fox, Lance Ramshaw, Richard Schwartz, Rebecca Stone, Ralph Weischedel, and The Annotation Group BBN Technologies. BBN: Description of the Sift system as used for MUC-7. In *Proceedings of the Seventh Message Understanding Conference (MUC-7)*, 1998.
- [92] J. Mostow, J. Gonzalez-Brenes, and B. H. Tan. Learning classifiers from a relational database of tutor logs. In *Proceedings of the Fourth International Conference on Educational Data Mining*, 2011.
- [93] Mitsuru Nakai and Hiroshi Shimodaira. Accent phrase segmentation by finding n-best sequences of pitch pattern templates. In *Proceedings of the Third International Conference on Spoken Language Processing (ICSLP 94)*, pages 347–350, 1994.
- [94] Gordon S. Jr. Novak and William C. Bulko. Diagrams and text as computer input. *Journal of Visual Languages and Computing*, 4(2):161–175, 1993.
- [95] Michael Oltmans. Understanding naturally conveyed explanations of device behavior. Master’s thesis, Massachusetts Institute of Technology, 2000.
- [96] Sharon Oviatt. Taming recognition errors with a multimodal interface. *Commun. ACM*, 43(9):45–51, 2000.
- [97] Sharon Oviatt, Alex Arthur, and Julia Cohen. Quiet interfaces that help students think. In *Proceedings of the 19th annual ACM symposium on User interface software and technology*, pages 191–200, 2006.
- [98] Sharon Oviatt, Phil Cohen, Lizhong Wu, John Vergo, Lisbeth Duncan, Bernhard Suhm, Josh Bers, Thomas Holzman, Terry Winograd, James Landay, Jim Larson, and David Ferro. Designing the user interface for multimodal speech and pen-based gesture applications: state-of-the-art systems and future research directions. *Hum.-Comput. Interact.*, 15(4):263–322, 2000.

- [99] Sharon Oviatt, Antonella DeAngeli, and Karen Kuhn. Integration and synchronization of input modes during multimodal human-computer interaction. In *CHI '97: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 415–422, New York, NY, USA, 1997. ACM.
- [100] Sharon L. Oviatt. Mutual disambiguation of recognition errors in a multimodel architecture. In *Proceedings of the CHI 99 Conference on Human Factors in Computing Systems: the CHI is the Limit*, pages 576–583. ACM Press, 1999.
- [101] R. Patel, B. Plimmer, J. Grundy, and R. Ihaka. Ink features for diagram recognition. In *Proc. of SBIM '07*, pages 131–138, 2007.
- [102] Brandon Paulson and Tracy Hammond. Paleosketch: accurate primitive sketch recognition and beautification. In *Proceedings of the 13th international conference on Intelligent user interfaces, IUI '08*, pages 1–10, New York, NY, USA, 2008. ACM.
- [103] K. Pezdek, T. Berry, and P. A. Renno. Children’s mathematics achievement: The role of parents’ perceptions and their involvement in homework. In *Journal of Educational Psychology*, volume 94, pages 771–777, 2002.
- [104] J. Ross Quinlan. *C4.5: programs for machine learning*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1993.
- [105] C. Romero, J. Romero, J. Luna, and Ventura S. Mining rare association rules from e-learning data. In *Proceedings of the Third International Conference on Educational Data Mining*, 2010.
- [106] Dean Rubine. Specifying gestures by example. *Computer Graphics*, 25:329–337, July 1991.
- [107] Hiroto Saigo, Jean-Philippe Vert, Nobuhisa Ueda, and Tatsuya Akutsu. Protein homology detection using string alignment kernels. *Bioinformatics*, 20:1682–1689, July 2004.
- [108] Tevfik Metin Sezgin and Randall Davis. Hmm-based efficient sketch recognition. In *Proceedings of the 10th international conference on Intelligent user interfaces, IUI '05*, pages 281–283, New York, NY, USA, 2005. ACM.
- [109] Tevfik Metin Sezgin, Thomas Stahovich, and Randall Davis. Sketch based interfaces: early processing for sketch understanding. In *SIGGRAPH '06: ACM SIGGRAPH 2006 Courses*, page 22, New York, NY, USA, 2006. ACM.
- [110] M. Shpitalni and H. Lipson. Classification of sketch strokes and corner detection using conic sections and adaptive clustering. *ASME Journal of Mechanical Design*, 119(2):131–135, 1996.
- [111] N. Silva and T. Cardoso. GIDeS++ – using constraints to model scenes. Technical report, Information Society Technologies, 2004.

- [112] K. Singh, M. Granville, and S_ç Dika. Mathematics and science achievement: Effects of motivation, interest, and academic engagement. In *Journal of Educational Research*, volume 95, pages 323–332, 2002.
- [113] Stephen Soderland, Brendan Roof, Bo Qin, Mausam Shi Xu, and Oren Etzioni. Adapting open information extraction to domain-specific relations. *AI Magazine*, 31:93–102, 2010.
- [114] T.F. Stahovich. Segmentation of pen strokes using pen speed. In *AAAI Fall Symposium Series 2004: Making Pen-Based Interaction Intelligent and Natural*, 2004.
- [115] Paul S. Steif and John A. Dantzler. A statics concept inventory : development and psychometric analysis. *Journal of Engineering Education*, October 2005.
- [116] P.S. Steif, J. Lobue, A. L. Fay, and L. B. Kara. Improving problem solving performance by inducing talk about salient problem features. *Journal of Engineering Education*, pages 135–142, 2010.
- [117] Andreas Stolcke. SRILM – an extensible language modeling toolkit. In *Proceedings of ICSLP*, volume 2, pages 901–904, Denver, USA, 2002.
- [118] Andreas Stolcke and Elizabeth Shriberg. Automatic linguistic segmentation of conversational speech. In H.T. Bunnell and W. Idsardi, editors, *Proceedings of the International Conference on Spoken Language Processing*, volume 2, pages 1005–1008, Philadelphia, 1996.
- [119] Andreas Stolcke, Elizabeth Shriberg, Rebecca Bates, Mari Ostendorf, Dilek Hakkani, Madelaine Plauche, Gokhan Tur, and Yu Lu. Automatic detection of sentence boundaries and disfluencies based on recognized words. In R.H. Mannell and J. Robert-Ribes, editors, *Proceedings of the International Conference on Spoken Language Processing*, volume 5, pages 2247–2250, Sydney, 1998. Australian Speech Science and Technology Association.
- [120] Stephanie Strassel. Simple metadata annotation specification version 6.2. Linguistic Data Consortium, 2004.
- [121] C. C. Tappert. Cursive script recognition by elastic matching. *IBM J. Res. Dev.*, 26(6):765–771, November 1982.
- [122] Kristina Toutanova and Christopher D. Manning. Enriching the knowledge sources used in a maximum entropy part-of-speech tagger. In *Proceedings of the 2000 Joint SIGDAT conference on Empirical methods in natural language processing and very large corpora*, pages 63–70, Morristown, NJ, USA, 2000. Association for Computational Linguistics.
- [123] David G. Ullman, Stephen Wood, and David Craig. The importance of drawing in the mechanical design process. *Computers & Graphics*, 14(2):263–274, 1990.

- [124] T. Van Arsdale and T. Stahovich. Does neatness count? what the organization of student work says about understanding. In *Proceedings of the American Society for Engineering Education Annual Conference*, 2012.
- [125] Willie Walker, Paul Lamere, Philip Kwok, Bhiksha Raj, Rita Singh, Evandro Gouvea, Peter Wolf, and Joe Woelfel. Sphinx-4: A Flexible Open Source Framework for Speech Recognition. *Sun Microsystems Technical Report*, November 2004.
- [126] Kenneth Wauchope. Eucalyptus: Integrating natural language input with a graphical user interface. Technical report, US Naval Research Laboratory Report NRL/FR/5510-94-9711, 1994.
- [127] Amali Weerasinghe and Antonija Mitrovic. Supporting self-explanation in an open-ended domain. In Mircea Negoita, Robert Howlett, and Lakhmi Jain, editors, *Knowledge-Based Intelligent Information and Engineering Systems*, volume 3213 of *Lecture Notes in Computer Science*, pages 306–313. Springer Berlin, 2004.
- [128] Gary M. Weiss. Mining with rarity: a unifying framework. *SIGKDD Explor. Newsl.*, 6:7–19, June 2004.
- [129] Jacob O. Wobbrock, Andrew D. Wilson, and Yang Li. Gestures without libraries, toolkits or training: a \$1 recognizer for user interface prototypes. In *Proceedings of the 20th annual ACM symposium on User interface software and technology*, UIST '07, pages 159–168, New York, NY, USA, 2007. ACM.
- [130] A. Wolin, B. Eoff, and T. Hammond. Shortstraw: A simple and effective corner finder for polylines. In *Eurographics Workshop on Sketch-Based Interfaces and Modeling (SBIM'08)*, 2008.
- [131] A. Wolin, B. Paulson, and T. Hammond. Sort, merge, repeat: an algorithm for effectively finding corners in hand-sketched strokes. In *SBIM '09: Proceedings of the 6th Eurographics Symposium on Sketch-Based Interfaces and Modeling*, pages 93–99, New York, NY, USA, 2009. ACM.
- [132] Aaron Wolin, Devin Smith, and Christine Alvarado. A pen-based tool for efficient labeling of 2d sketches. In *Proceedings of the 4th Eurographics workshop on Sketch-based interfaces and modeling*, SBIM '07, pages 67–74, New York, NY, USA, 2007. ACM.
- [133] W.A Woods, L. Bates, R. Bobrow, R. Brachman, P. R. Cohen, and J. Klovstad. Research in natural language understanding. In *Cambridge, MA, Annual report 4274, Bolt Beranek and Newman*, 1979.
- [134] Xiaopeng Xi, Eamonn Keogh, Christian Shelton, Li Wei, and Chotirat Ann Ratanamahatana. Fast time series classification using numerosity reduction. In *Proceedings of the 23rd international conference on Machine learning*, ICML '06, pages 1033–1040, New York, NY, USA, 2006. ACM.

- [135] Xiaopeng Xi, Eamonn Keogh, Li Wei, and Agenor Mafra-Neto. Finding motifs in database of shapes. In *Proceedings of Siam International Conference on Data Mining (SDM'07)*, 2007.
- [136] Eric P. Xing, Michael I. Jordan, and Richard M. Karp. Feature selection for high-dimensional genomic microarray data. In *In Proceedings of the Eighteenth International Conference on Machine Learning*, pages 601–608. Morgan Kaufmann, 2001.
- [137] Yiyang Xiong and Joseph J. LaViola Jr. A shortstraw-based algorithm for corner finding in sketch-based interfaces. *Computers & Graphics*, 34(5):513 – 527, 2010.
- [138] Yiyang Xiong and Joseph J. LaViola, Jr. Revisiting shortstraw: improving corner finding in sketch-based interfaces. In *SBIM '09: Proceedings of the 6th Eurographics Symposium on Sketch-Based Interfaces and Modeling*, pages 101–108, New York, NY, USA, 2009. ACM.
- [139] Lexiang Ye and Eamonn Keogh. Time series shapelets: a new primitive for data mining. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 947–956. ACM, 2009.
- [140] W.S. Yezhov. The spam-filtering accuracy plateau at 99.9 percent accuracy and how to get past it. In *Proc. MIT Spam Conference 2004*, 2004.
- [141] Bo Yu and Shijie Cai. A domain-independent system for sketch recognition. In *GRAPHITE '03: Proceedings of the 1st international conference on Computer graphics and interactive techniques in Australasia and South East Asia*, pages 141–146. ACM, 2003.
- [142] Weixiong Zhang. *State-space search: Algorithms, complexity, extensions, and applications*. Springer, 1999.
- [143] Rui Zhao. Incremental recognition in gesture-based and syntax directed diagram editor. In *Proceedings of InterCHI'93*, pages 95–100, 1993.

Appendix - Histograms

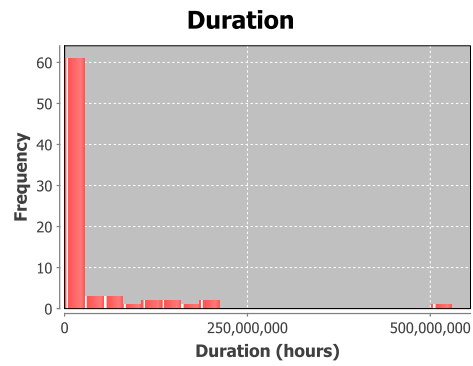


Figure .1: Histogram of the duration, time from first to last pen stroke for all students for homework two problem one for the 2010 course offering.

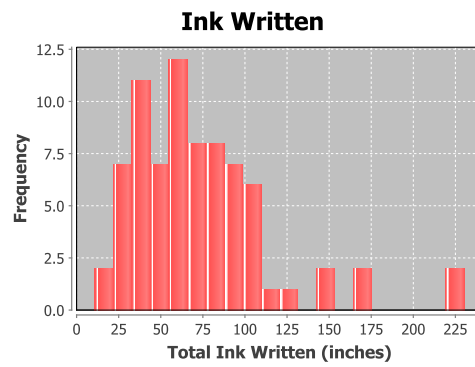


Figure .2: Histogram of the amount of ink, in inches, written, for each student on homework two problem one for the 2010 course offering.

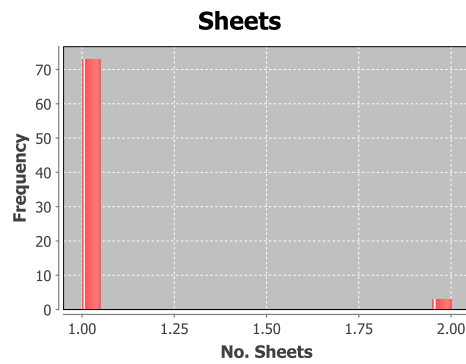


Figure .3: Histogram of the number of pages written on by each student to solve homework two problem one for the 2010 course offering.

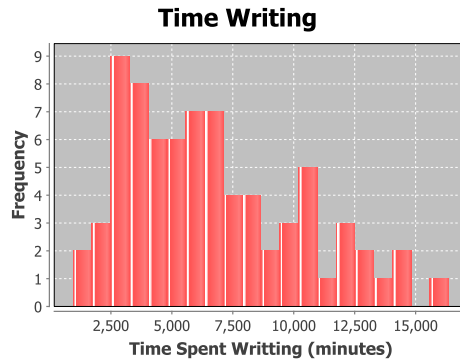


Figure .4: Histogram of the amount of time spent writing by each student on homework two problem one for the 2010 course offering.

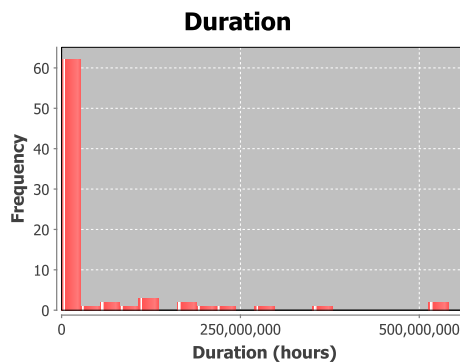


Figure .5: Histogram of the duration, time from first to last pen stroke for all students for homework two problem two for the 2010 course offering.

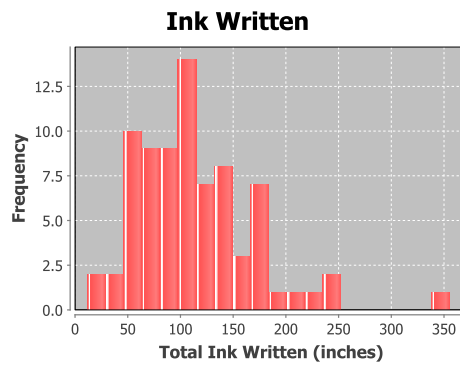


Figure .6: Histogram of the amount of ink, in inches, written, for each student on homework two problem two for the 2010 course offering.

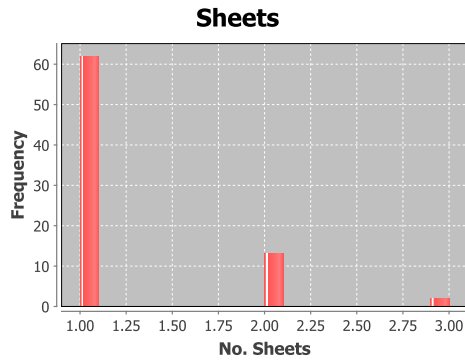


Figure .7: Histogram of the number of pages written on by each student to solve homework two problem two for the 2010 course offering.

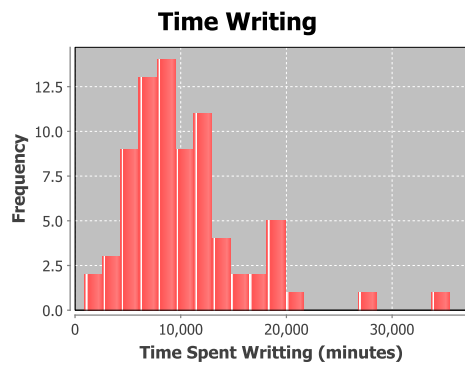


Figure .8: Histogram of the amount of time spent writing by each student on homework two problem two for the 2010 course offering.

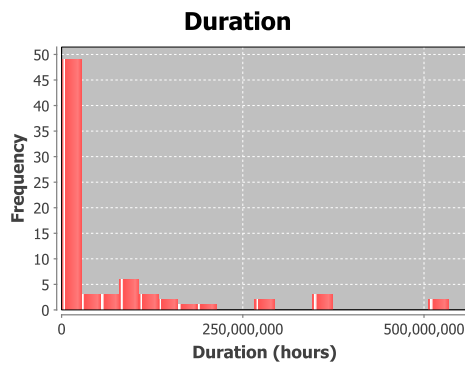


Figure .9: Histogram of the duration, time from first to last pen stroke for all students for homework two problem three for the 2010 course offering.

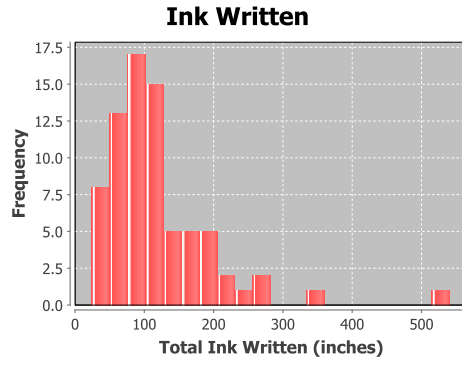


Figure .10: Histogram of the amount of ink, in inches, written, for each student on homework two problem three for the 2010 course offering.

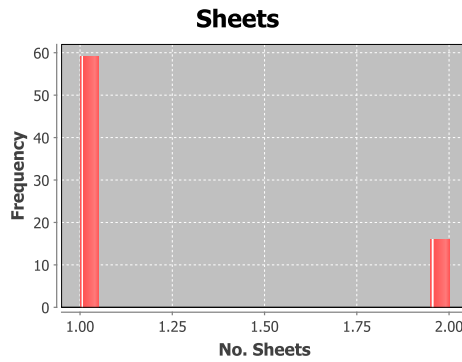


Figure .11: Histogram of the number of pages written on by each student to solve homework two problem three for the 2010 course offering.

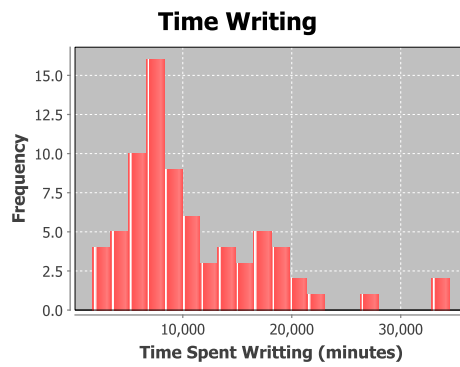


Figure .12: Histogram of the amount of time spent writing by each student on homework two problem three for the 2010 course offering.

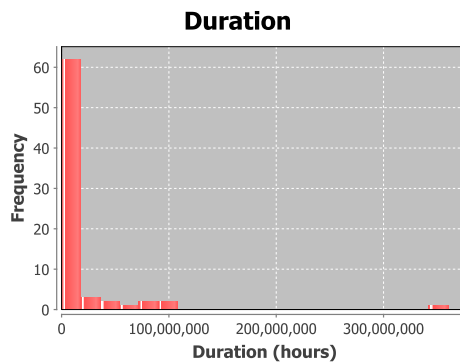


Figure .13: Histogram of the duration, time from first to last pen stroke for all students for homework two problem four for the 2010 course offering.

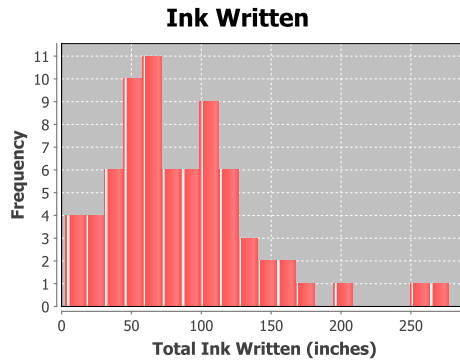


Figure .14: Histogram of the amount of ink, in inches, written, for each student on homework two problem four for the 2010 course offering.

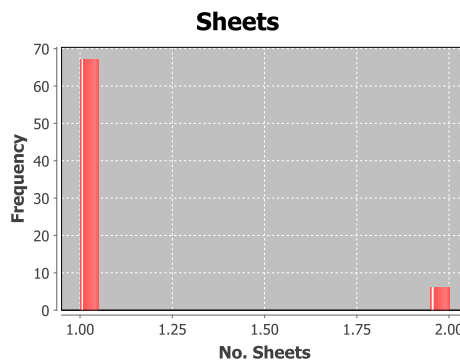


Figure .15: Histogram of the number of pages written on by each student to solve homework two problem four for the 2010 course offering.

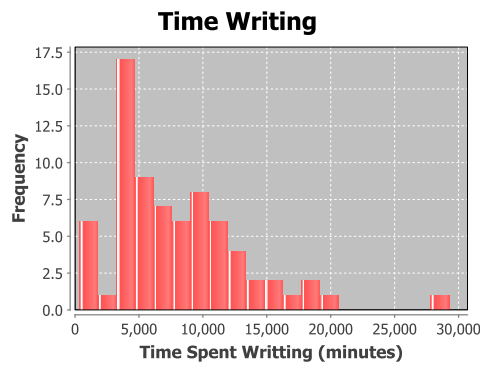


Figure .16: Histogram of the amount of time spent writing by each student on homework two problem four for the 2010 course offering.

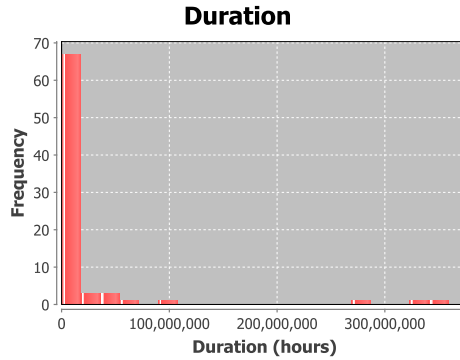


Figure .17: Histogram of the duration, time from first to last pen stroke for all students for homework two problem five for the 2010 course offering.

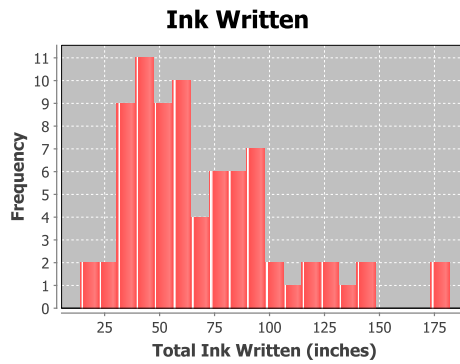


Figure .18: Histogram of the amount of ink, in inches, written, for each student on homework two problem five for the 2010 course offering.

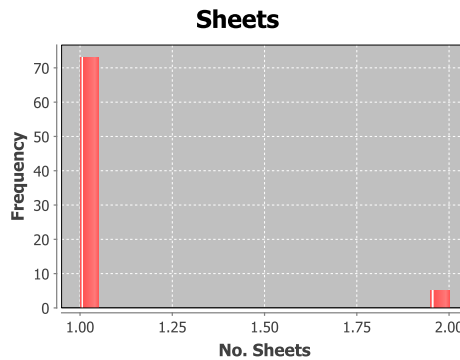


Figure .19: Histogram of the number of pages written on by each student to solve homework two problem five for the 2010 course offering.

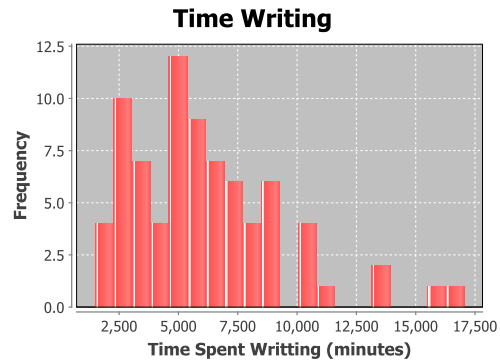


Figure .20: Histogram of the amount of time spent writing by each student on homework two problem five for the 2010 course offering.

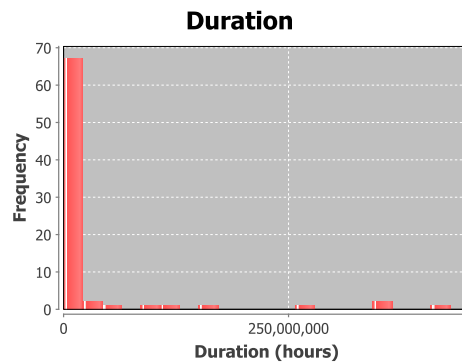


Figure .21: Histogram of the duration, time from first to last pen stroke for all students for homework two problem six for the 2010 course offering.

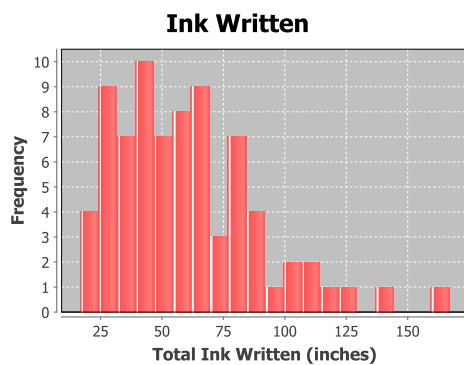


Figure .22: Histogram of the amount of ink, in inches, written, for each student on homework two problem six for the 2010 course offering.

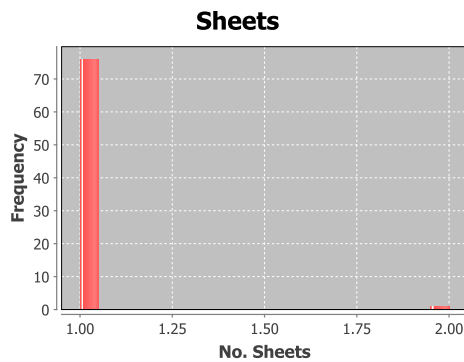


Figure .23: Histogram of the number of pages written on by each student to solve homework two problem six for the 2010 course offering.

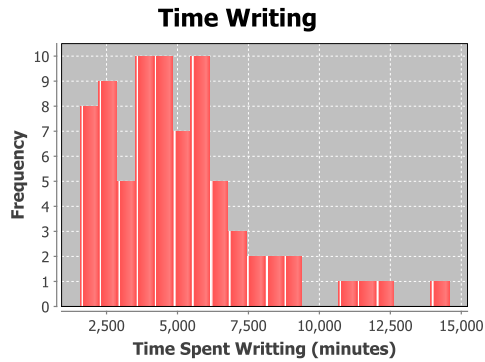


Figure .24: Histogram of the amount of time spent writing by each student on homework two problem six for the 2010 course offering.

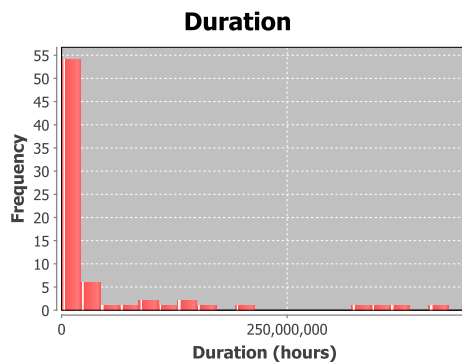


Figure .25: Histogram of the duration, time from first to last pen stroke for all students for homework two problem seven for the 2010 course offering.

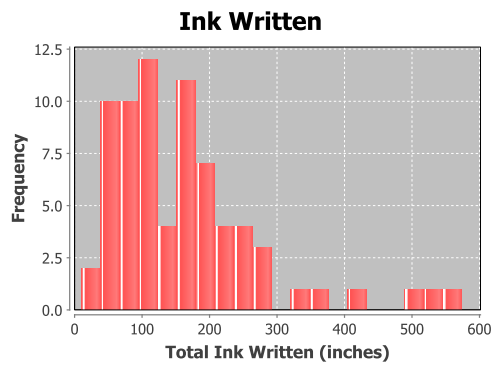


Figure .26: Histogram of the amount of ink, in inches, written, for each student on homework two problem seven for the 2010 course offering.

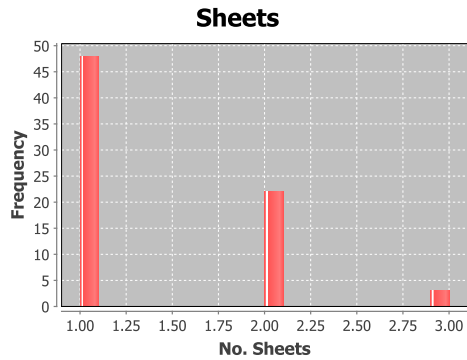


Figure .27: Histogram of the number of pages written on by each student to solve homework two problem seven for the 2010 course offering.

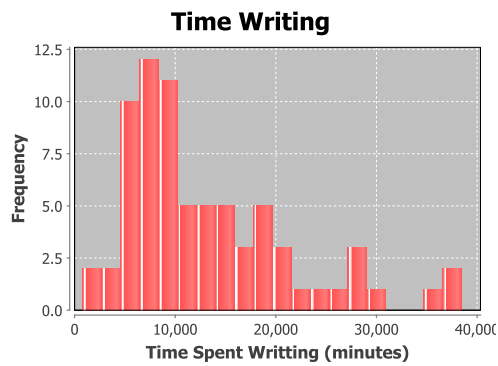


Figure .28: Histogram of the amount of time spent writing by each student on homework two problem seven for the 2010 course offering.

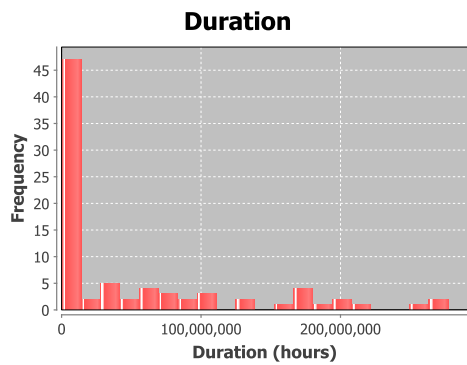


Figure .29: Histogram of the duration, time from first to last pen stroke for all students for homework three problem one for the 2010 course offering.

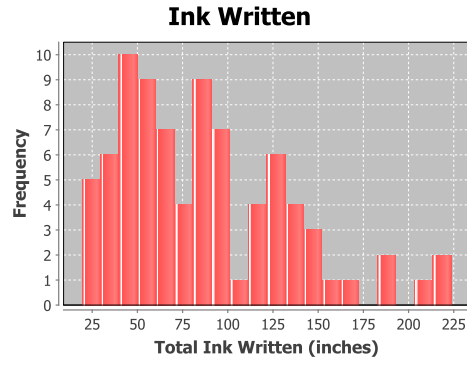


Figure .30: Histogram of the amount of ink, in inches, written, for each student on homework three problem one for the 2010 course offering.

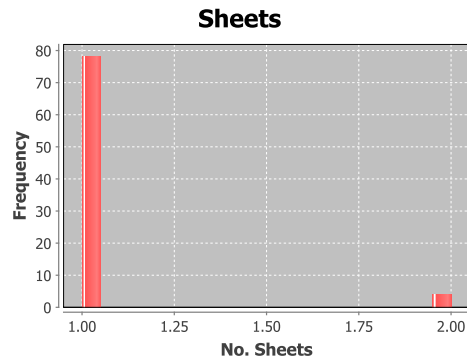


Figure .31: Histogram of the number of pages written on by each student to solve homework three problem one for the 2010 course offering.

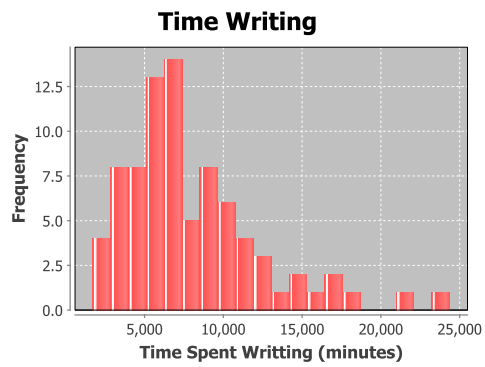


Figure .32: Histogram of the amount of time spent writing by each student on homework three problem one for the 2010 course offering.

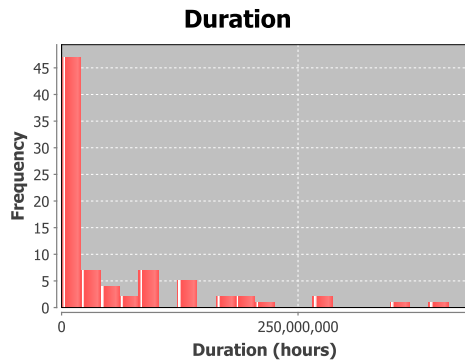


Figure .33: Histogram of the duration, time from first to last pen stroke for all students for homework three problem two for the 2010 course offering.

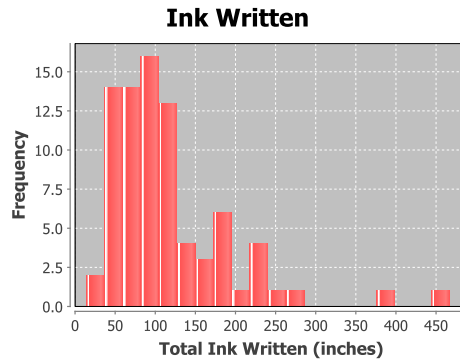


Figure .34: Histogram of the amount of ink, in inches, written, for each student on homework three problem two for the 2010 course offering.

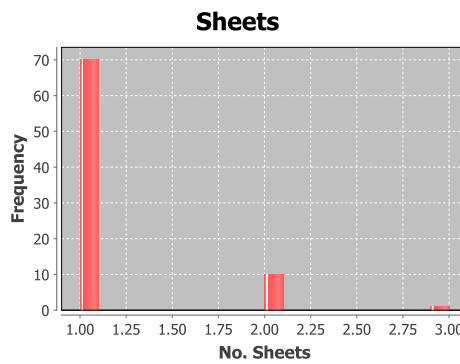


Figure .35: Histogram of the number of pages written on by each student to solve homework three problem two for the 2010 course offering.

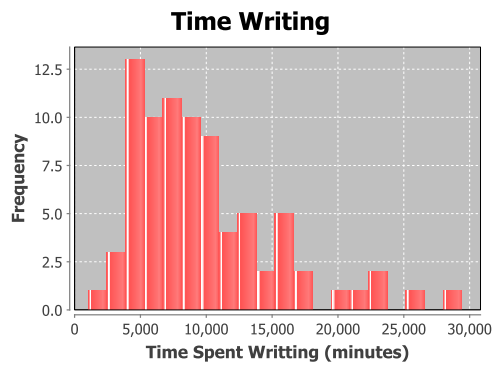


Figure .36: Histogram of the amount of time spent writing by each student on homework three problem two for the 2010 course offering.

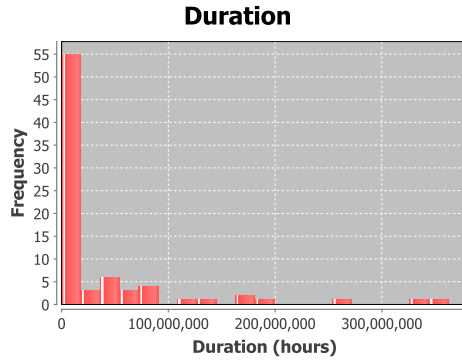


Figure .37: Histogram of the duration, time from first to last pen stroke for all students for homework three problem three for the 2010 course offering.

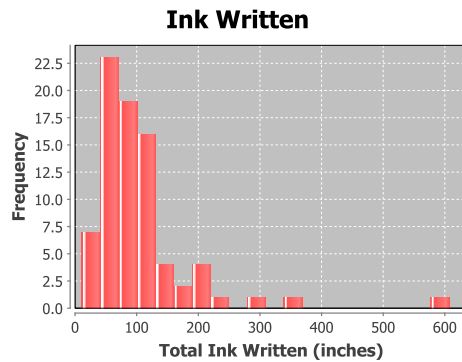


Figure .38: Histogram of the amount of ink, in inches, written, for each student on homework three problem three for the 2010 course offering.

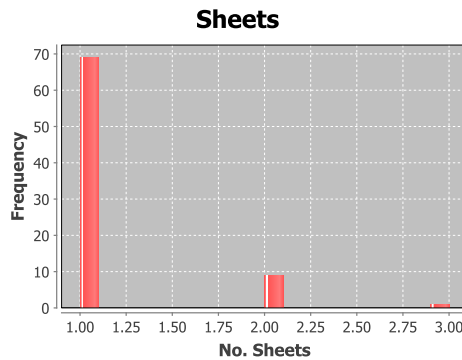


Figure .39: Histogram of the number of pages written on by each student to solve homework three problem three for the 2010 course offering.

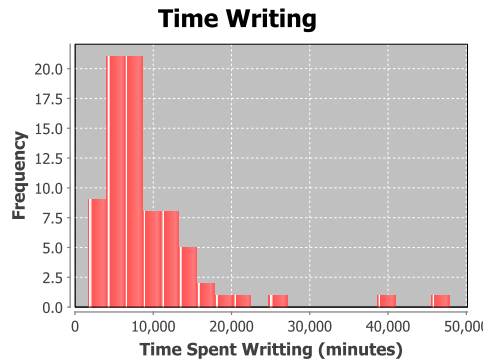


Figure .40: Histogram of the amount of time spent writing by each student on homework three problem three for the 2010 course offering.

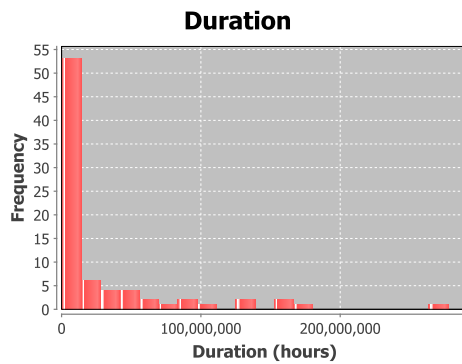


Figure .41: Histogram of the duration, time from first to last pen stroke for all students for homework three problem three for the 2010 course offering.

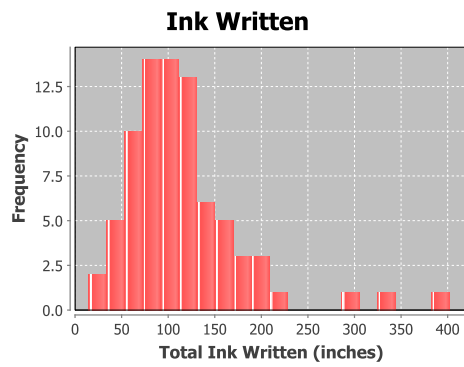


Figure .42: Histogram of the amount of ink, in inches, written, for each student on homework three problem four for the 2010 course offering.

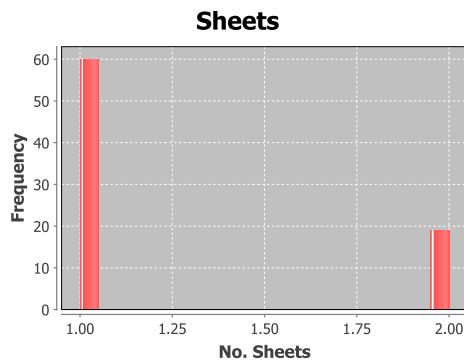


Figure .43: Histogram of the number of pages written on by each student to solve homework three problem four for the 2010 course offering.

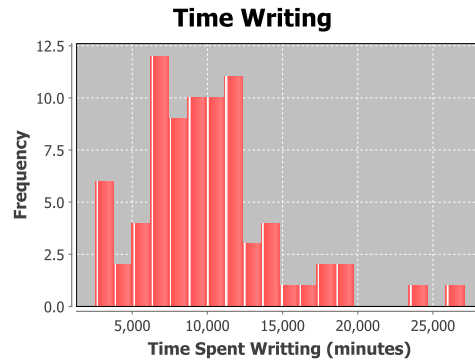


Figure .44: Histogram of the amount of time spent writing by each student on homework three problem four for the 2010 course offering.

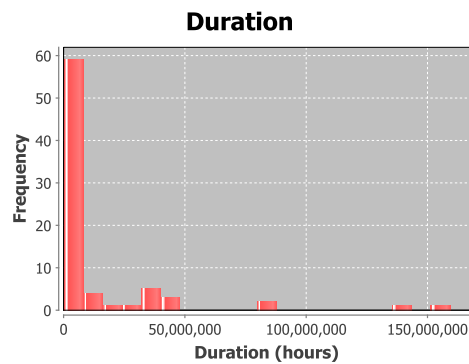


Figure .45: Histogram of the duration, time from first to last pen stroke for all students for homework three problem five for the 2010 course offering.

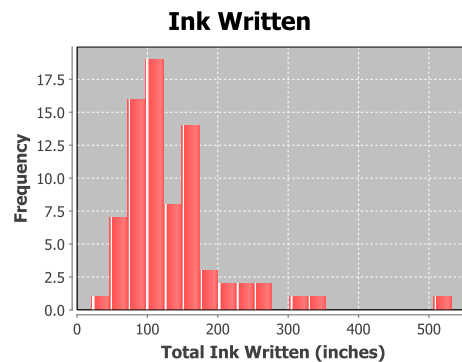


Figure .46: Histogram of the amount of ink, in inches, written, for each student on homework three problem five for the 2010 course offering.

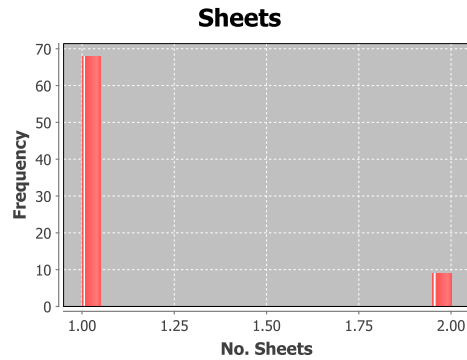


Figure .47: Histogram of the number of pages written on by each student to solve homework three problem five for the 2010 course offering.

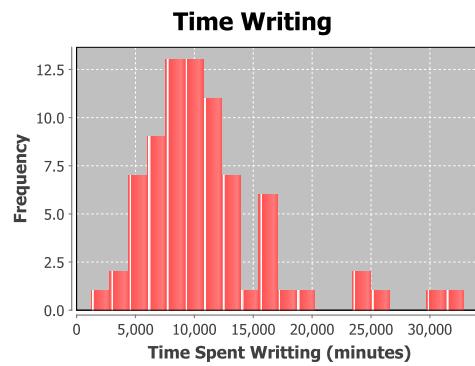


Figure .48: Histogram of the amount of time spent writing by each student on homework three problem five for the 2010 course offering.

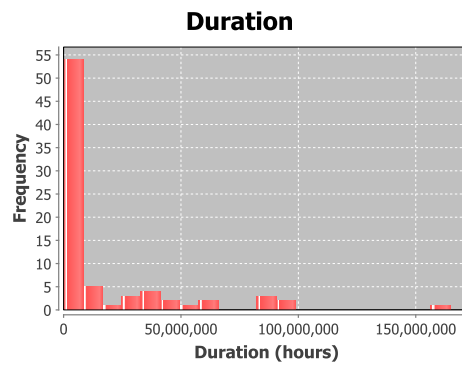


Figure .49: Histogram of the duration, time from first to last pen stroke for all students for homework three problem six for the 2010 course offering.

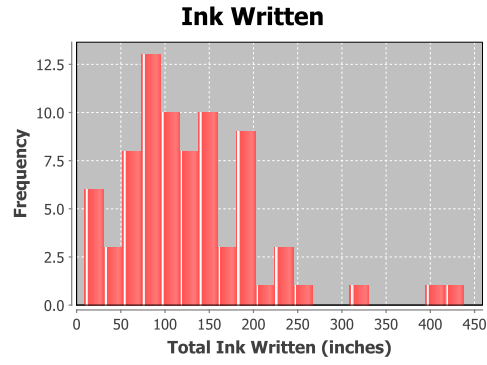


Figure .50: Histogram of the amount of ink, in inches, written, for each student on homework three problem six for the 2010 course offering.

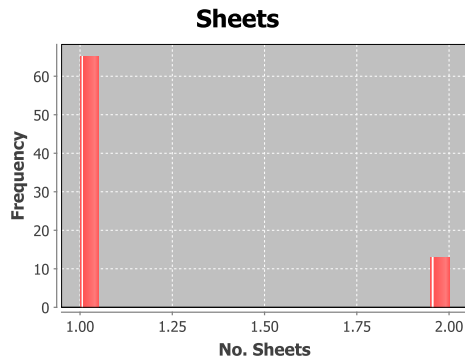


Figure .51: Histogram of the number of pages written on by each student to solve homework three problem six for the 2010 course offering.

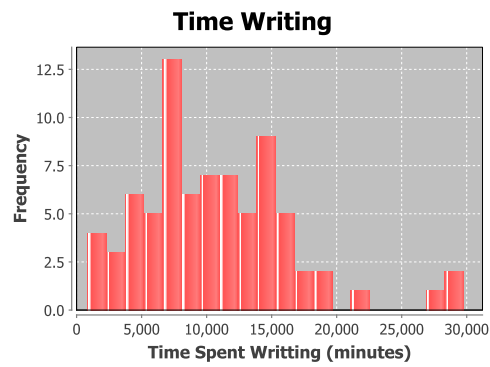


Figure .52: Histogram of the amount of time spent writing by each student on homework three problem six for the 2010 course offering.

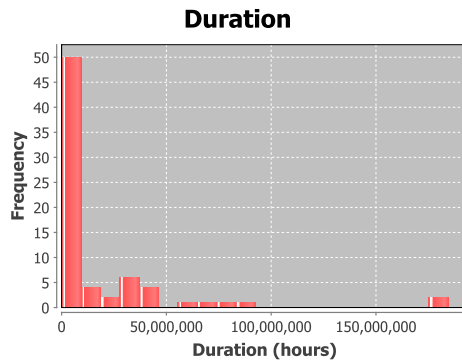


Figure .53: Histogram of the duration, time from first to last pen stroke for all students for homework three problem seven for the 2010 course offering.

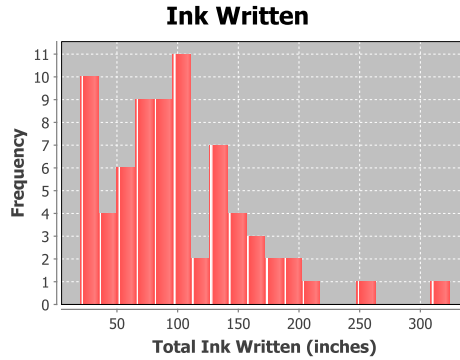


Figure .54: Histogram of the amount of ink, in inches, written, for each student on homework three problem seven for the 2010 course offering.

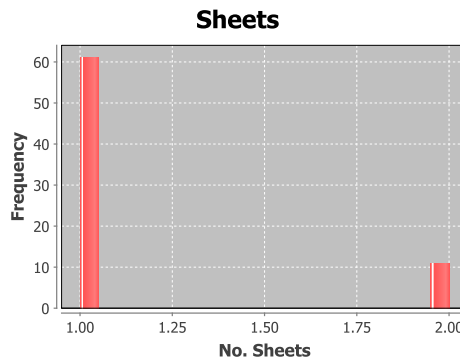


Figure .55: Histogram of the number of pages written on by each student to solve homework three problem seven for the 2010 course offering.

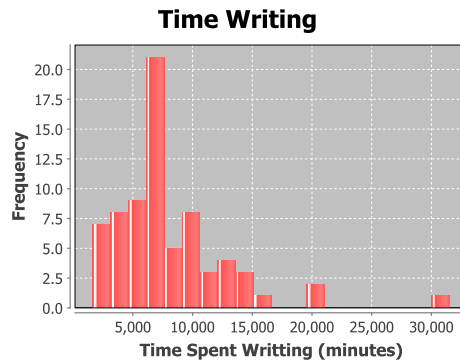


Figure .56: Histogram of the amount of time spent writing by each student on homework three problem seven for the 2010 course offering.

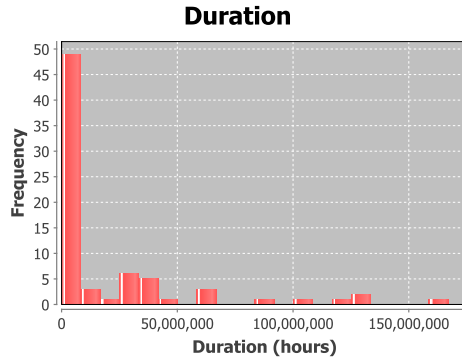


Figure .57: Histogram of the duration, time from first to last pen stroke for all students for homework three problem eight for the 2010 course offering.

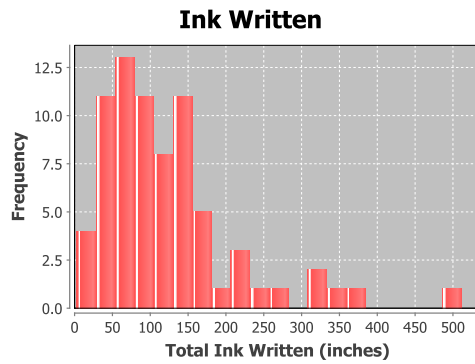


Figure .58: Histogram of the amount of ink, in inches, written, for each student on homework three problem eight for the 2010 course offering.

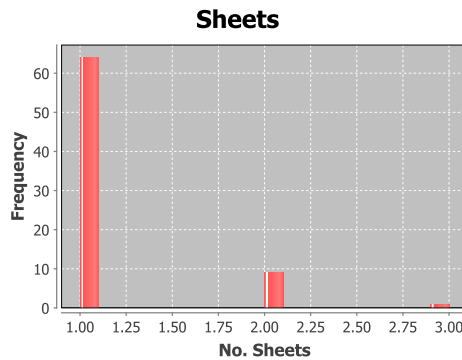


Figure .59: Histogram of the number of pages written on by each student to solve homework three problem eight for the 2010 course offering.

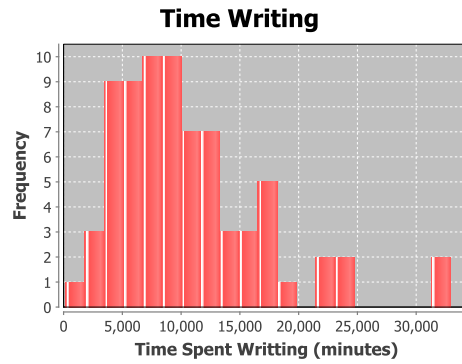


Figure .60: Histogram of the amount of time spent writing by each student on homework three problem eight for the 2010 course offering.

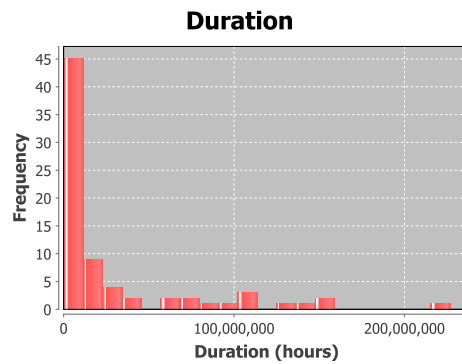


Figure .61: Histogram of the duration, time from first to last pen stroke for all students for homework four problem one for the 2010 course offering.

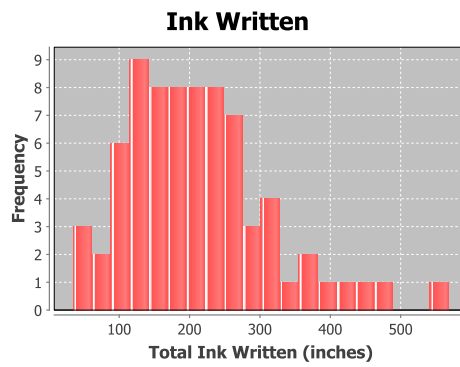


Figure .62: Histogram of the amount of ink, in inches, written, for each student on homework four problem one for the 2010 course offering.

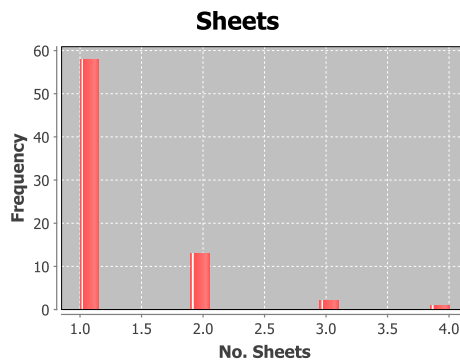


Figure .63: Histogram of the number of pages written on by each student to solve homework four problem one for the 2010 course offering.

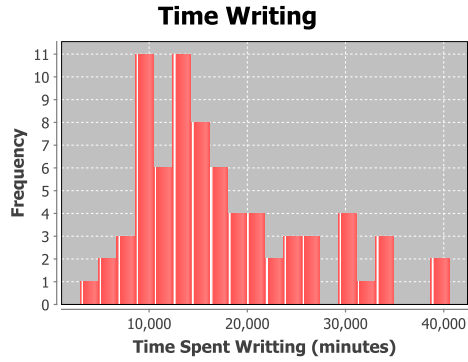


Figure .64: Histogram of the amount of time spent writing by each student on homework four problem one for the 2010 course offering.

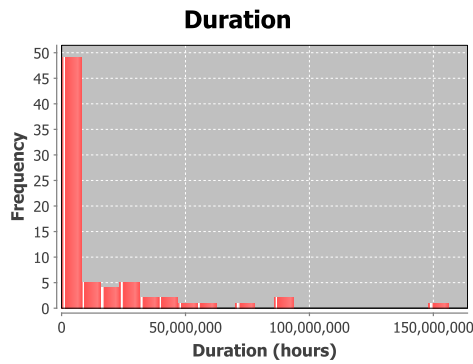


Figure .65: Histogram of the duration, time from first to last pen stroke for all students for homework four problem two for the 2010 course offering.

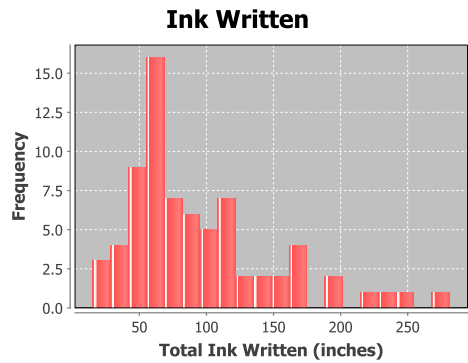


Figure .66: Histogram of the amount of ink, in inches, written, for each student on homework four problem two for the 2010 course offering.

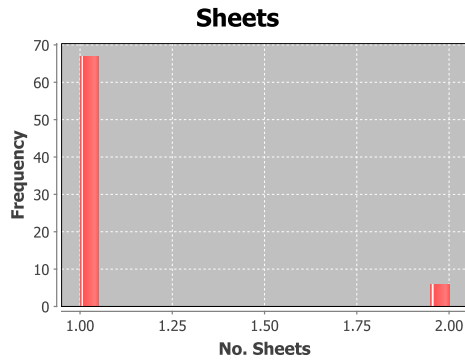


Figure .67: Histogram of the number of pages written on by each student to solve homework four problem two for the 2010 course offering.

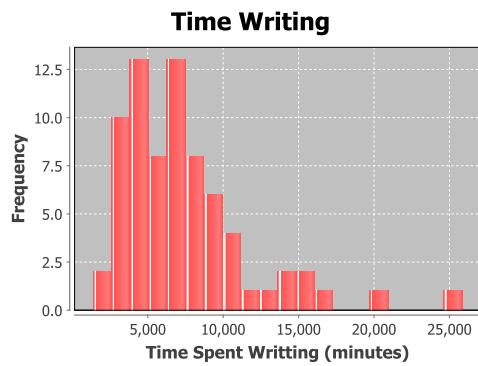


Figure .68: Histogram of the amount of time spent writing by each student on homework four problem two for the 2010 course offering.

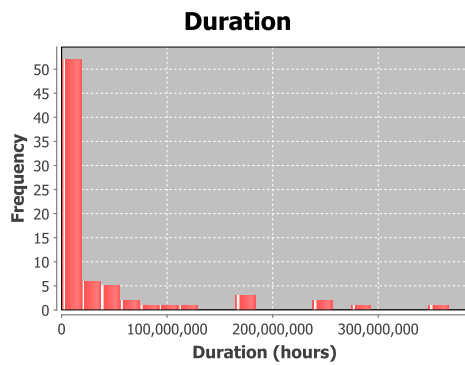


Figure .69: Histogram of the duration, time from first to last pen stroke for all students for homework five problem one for the 2010 course offering.

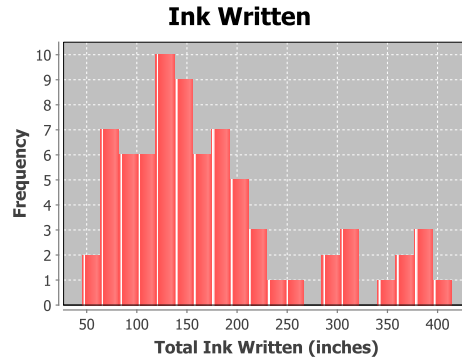


Figure .70: Histogram of the amount of ink, in inches, written, for each student on homework five problem one for the 2010 course offering.

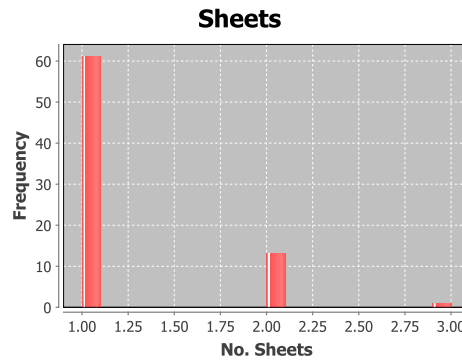


Figure .71: Histogram of the number of pages written on by each student to solve homework five problem one for the 2010 course offering.

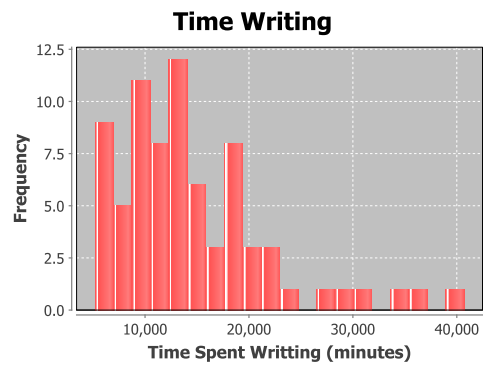


Figure .72: Histogram of the amount of time spent writing by each student on homework five problem one for the 2010 course offering.

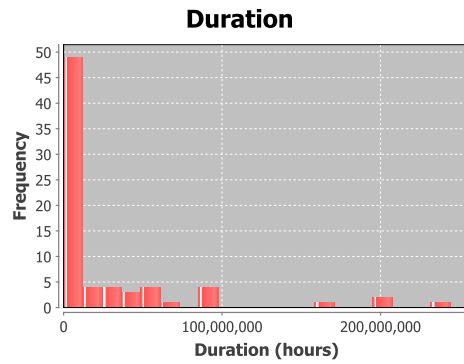


Figure .73: Histogram of the duration, time from first to last pen stroke for all students for homework five problem two for the 2010 course offering.

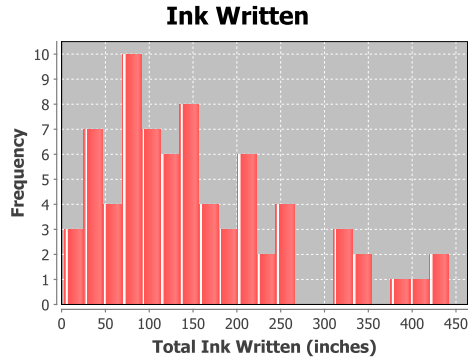


Figure .74: Histogram of the amount of ink, in inches, written, for each student on homework five problem two for the 2010 course offering.

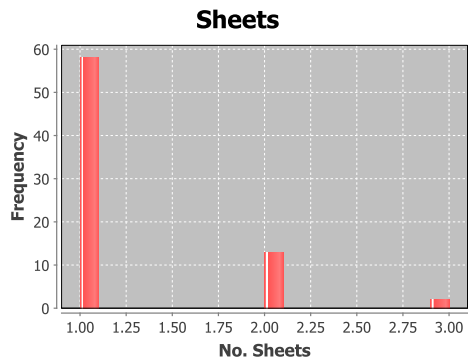


Figure .75: Histogram of the number of pages written on by each student to solve homework five problem two for the 2010 course offering.

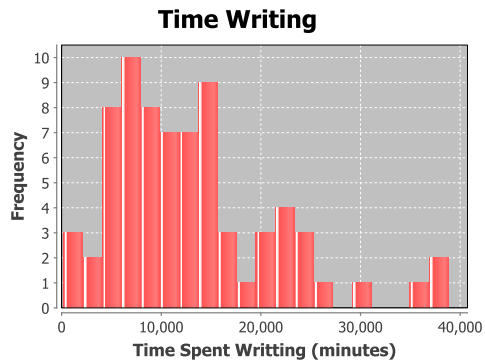


Figure .76: Histogram of the amount of time spent writing by each student on homework five problem two for the 2010 course offering.

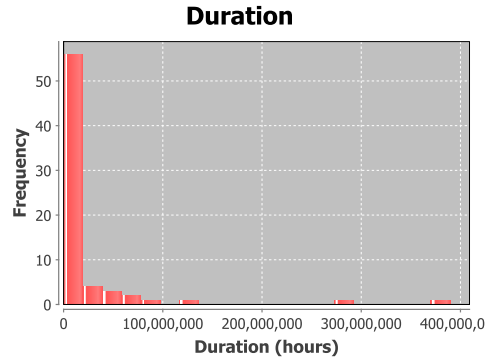


Figure .77: Histogram of the duration, time from first to last pen stroke for all students for homework five problem three for the 2010 course offering.

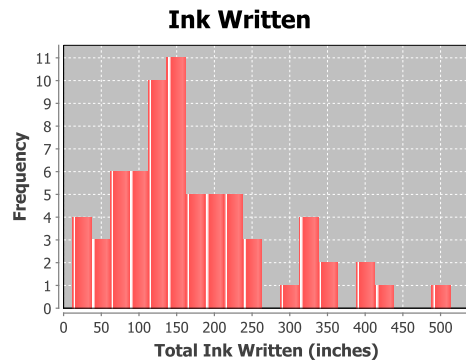


Figure .78: Histogram of the amount of ink, in inches, written, for each student on homework five problem three for the 2010 course offering.

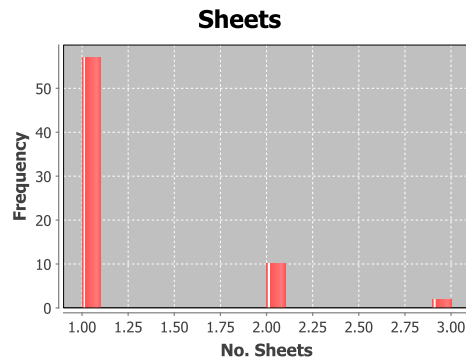


Figure .79: Histogram of the number of pages written on by each student to solve homework five problem three for the 2010 course offering.

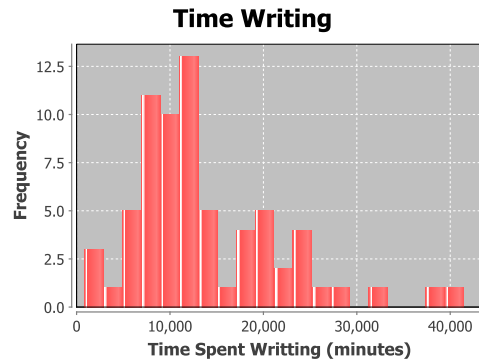


Figure .80: Histogram of the amount of time spent writing by each student on homework five problem three for the 2010 course offering.

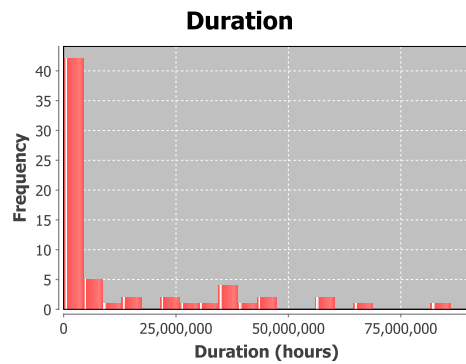


Figure .81: Histogram of the duration, time from first to last pen stroke for all students for homework five problem four for the 2010 course offering.

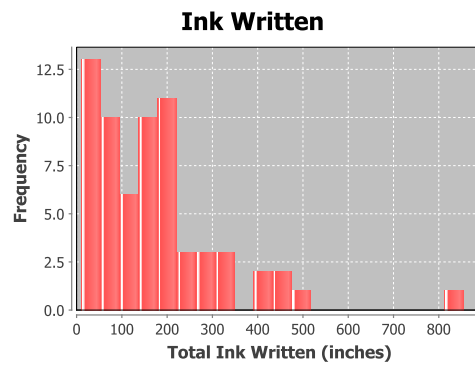


Figure .82: Histogram of the amount of ink, in inches, written, for each student on homework five problem four for the 2010 course offering.

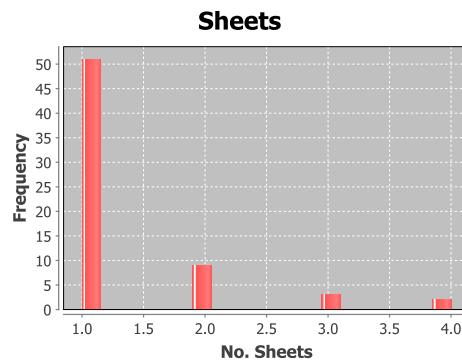


Figure .83: Histogram of the number of pages written on by each student to solve homework five problem four for the 2010 course offering.

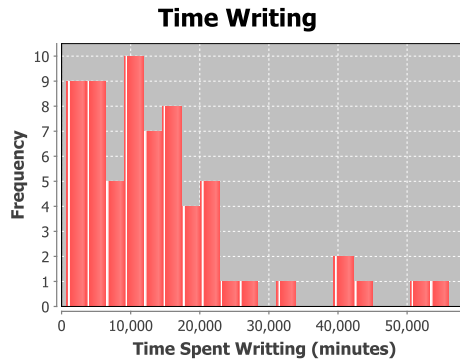


Figure .84: Histogram of the amount of time spent writing by each student on homework five problem four for the 2010 course offering.

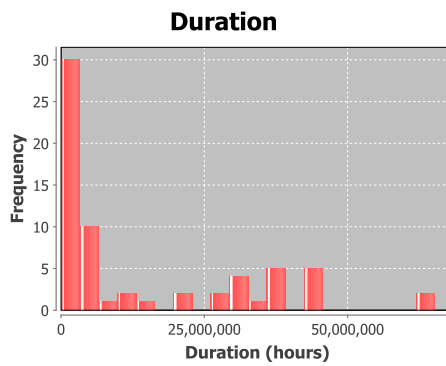


Figure .85: Histogram of the duration, time from first to last pen stroke for all students for homework five problem five for the 2010 course offering.

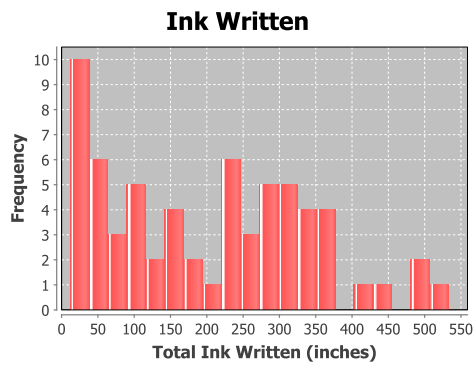


Figure .86: Histogram of the amount of ink, in inches, written, for each student on homework five problem five for the 2010 course offering.

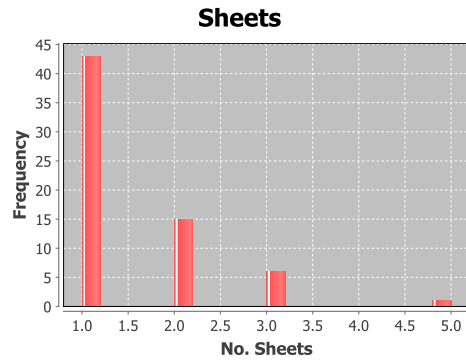


Figure .87: Histogram of the number of pages written on by each student to solve homework five problem five for the 2010 course offering.

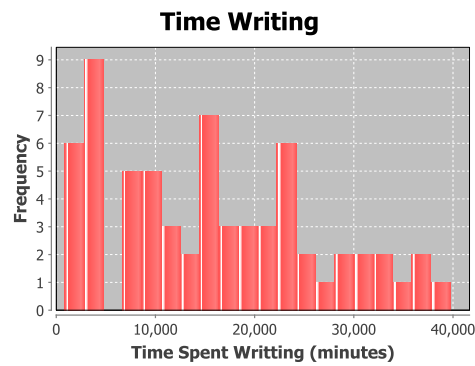


Figure .88: Histogram of the amount of time spent writing by each student on homework five problem five for the 2010 course offering.

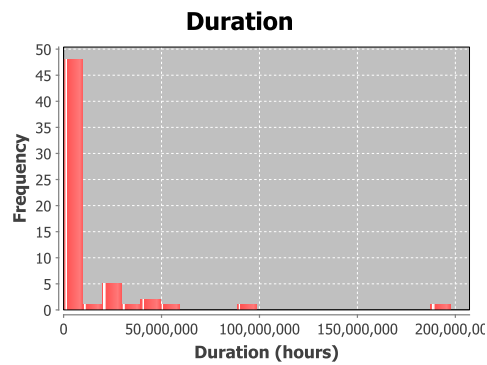


Figure .89: Histogram of the duration, time from first to last pen stroke for all students for homework five problem six for the 2010 course offering.

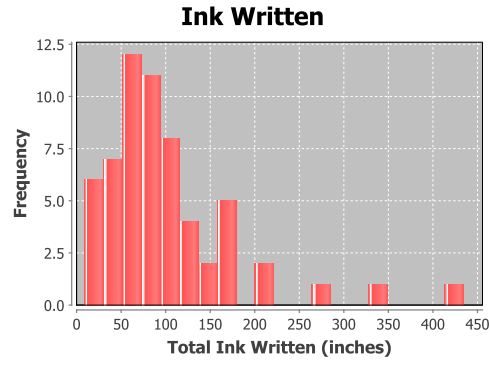


Figure .90: Histogram of the amount of ink, in inches, written, for each student on homework five problem six for the 2010 course offering.

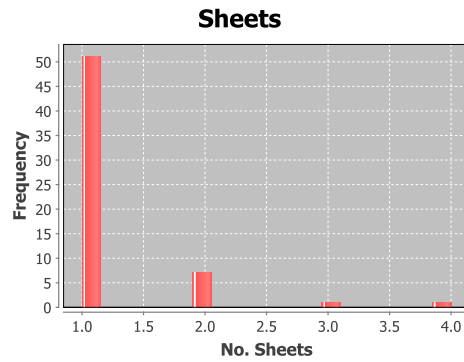


Figure .91: Histogram of the number of pages written on by each student to solve homework five problem six for the 2010 course offering.

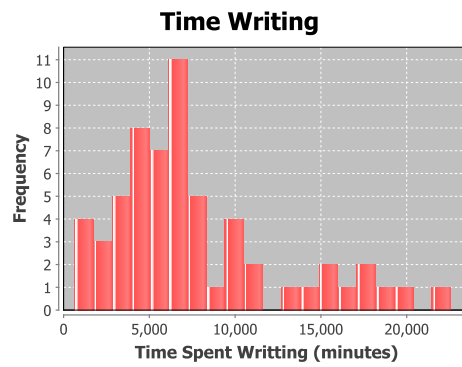


Figure .92: Histogram of the amount of time spent writing by each student on homework five problem six for the 2010 course offering.

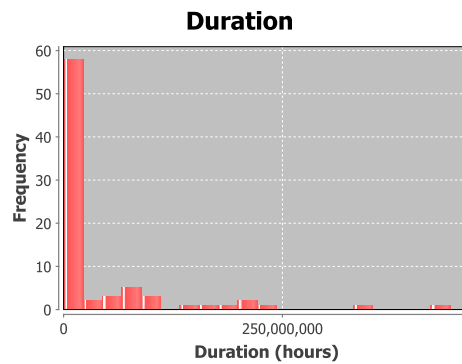


Figure .93: Histogram of the duration, time from first to last pen stroke for all students for homework six problem one for the 2010 course offering.

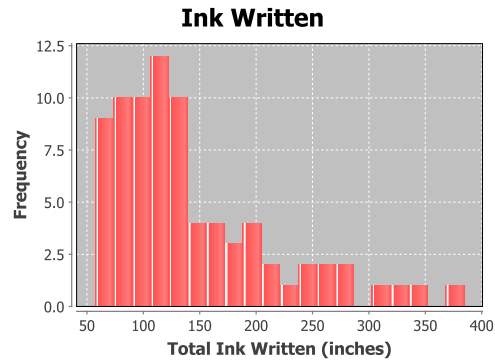


Figure .94: Histogram of the amount of ink, in inches, written, for each student on homework six problem one for the 2010 course offering.

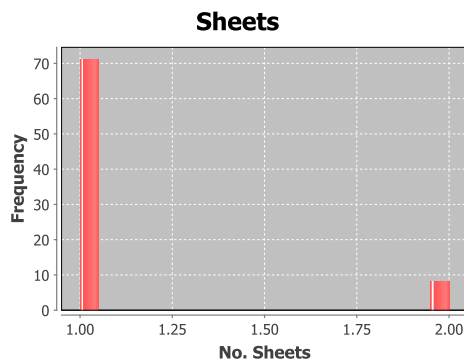


Figure .95: Histogram of the number of pages written on by each student to solve homework six problem one for the 2010 course offering.

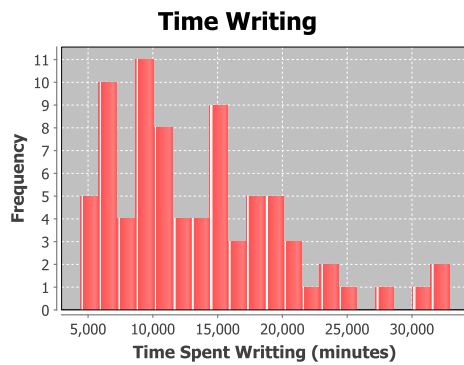


Figure .96: Histogram of the amount of time spent writing by each student on homework six problem one for the 2010 course offering.

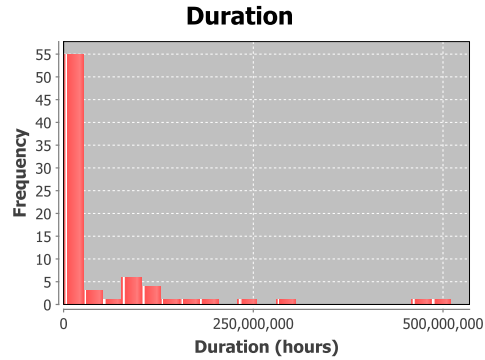


Figure .97: Histogram of the duration, time from first to last pen stroke for all students for homework six problem two for the 2010 course offering.

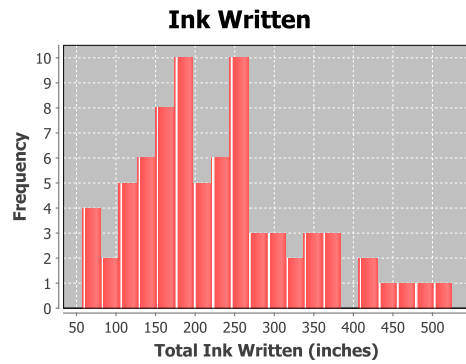


Figure .98: Histogram of the amount of ink, in inches, written, for each student on homework six problem two for the 2010 course offering.

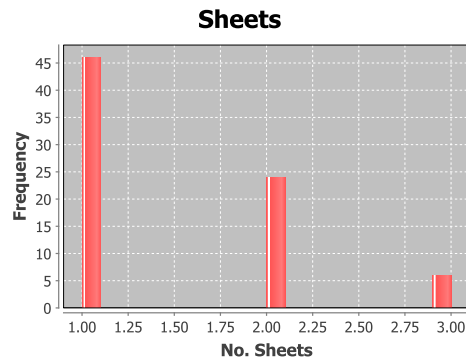


Figure .99: Histogram of the number of pages written on by each student to solve homework six problem two for the 2010 course offering.

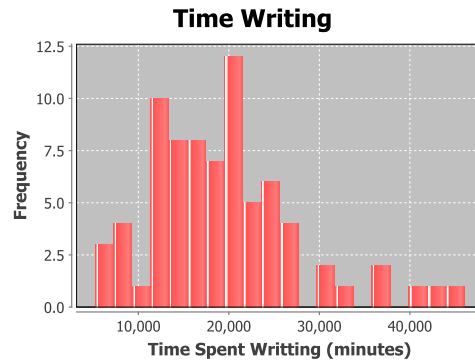


Figure .100: Histogram of the amount of time spent writing by each student on homework six problem two for the 2010 course offering.

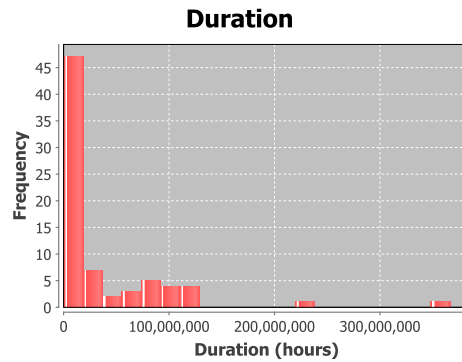


Figure .101: Histogram of the duration, time from first to last pen stroke for all students for homework six problem three for the 2010 course offering.

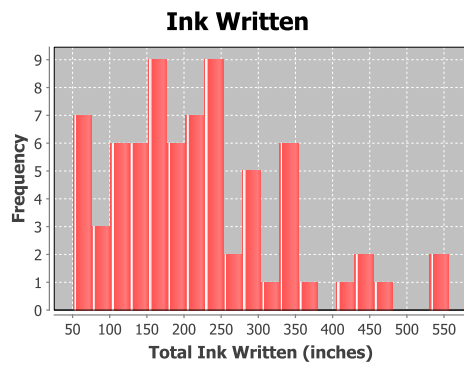


Figure .102: Histogram of the amount of ink, in inches, written, for each student on homework six problem three for the 2010 course offering.

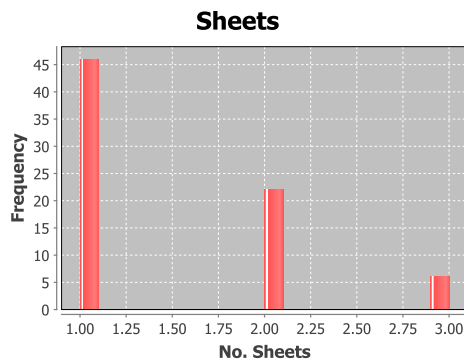


Figure .103: Histogram of the number of pages written on by each student to solve homework six problem three for the 2010 course offering.

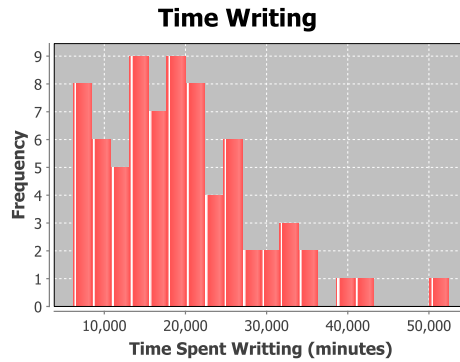


Figure .104: Histogram of the amount of time spent writing by each student on homework six problem three for the 2010 course offering.

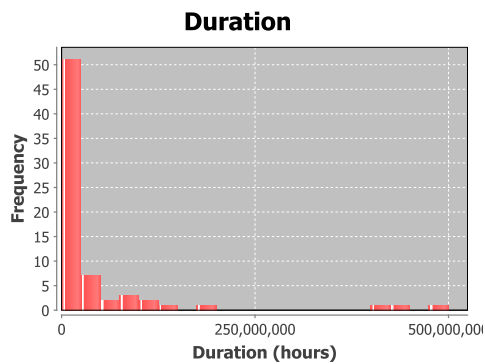


Figure .105: Histogram of the duration, time from first to last pen stroke for all students for homework six problem four for the 2010 course offering.

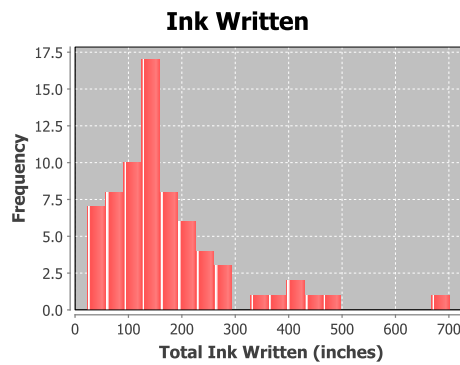


Figure .106: Histogram of the amount of ink, in inches, written, for each student on homework six problem four for the 2010 course offering.

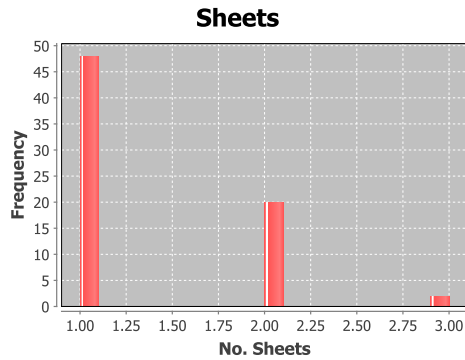


Figure .107: Histogram of the number of pages written on by each student to solve homework six problem four for the 2010 course offering.

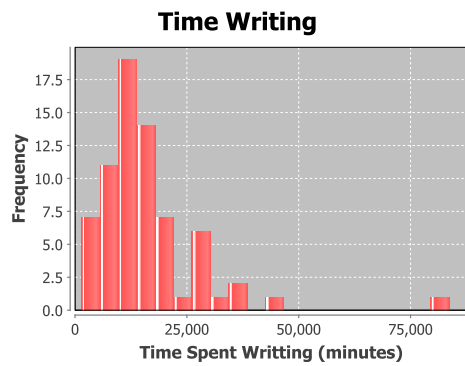


Figure .108: Histogram of the amount of time spent writing by each student on homework six problem four for the 2010 course offering.

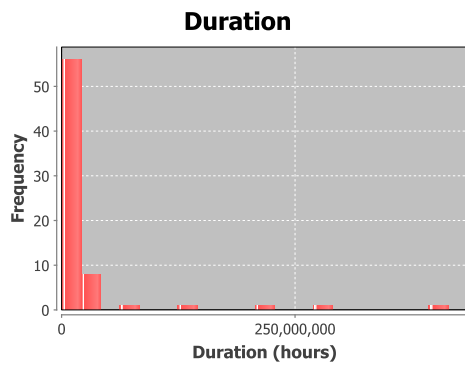


Figure .109: Histogram of the duration, time from first to last pen stroke for all students for homework six problem five for the 2010 course offering.

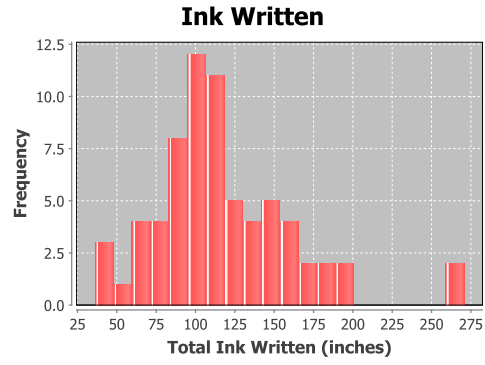


Figure .110: Histogram of the amount of ink, in inches, written, for each student on homework six problem five for the 2010 course offering.

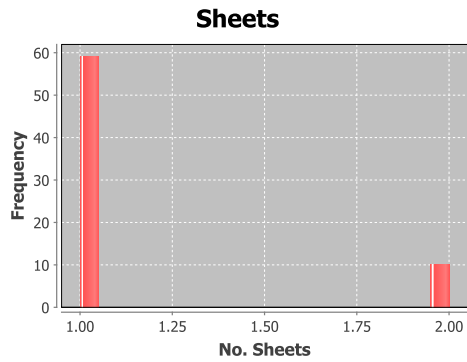


Figure .111: Histogram of the number of pages written on by each student to solve homework six problem five for the 2010 course offering.

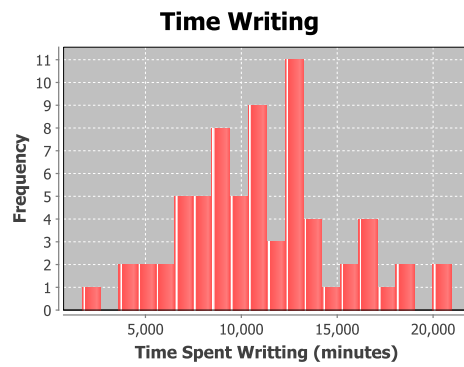


Figure .112: Histogram of the amount of time spent writing by each student on homework six problem five for the 2010 course offering.

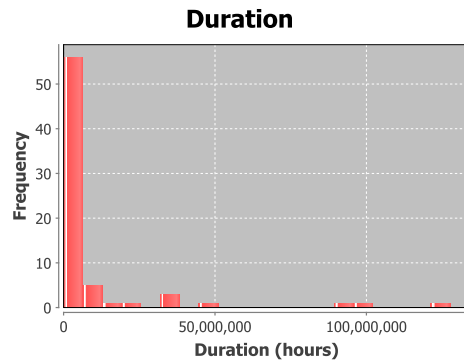


Figure .113: Histogram of the duration, time from first to last pen stroke for all students for homework six problem six for the 2010 course offering.

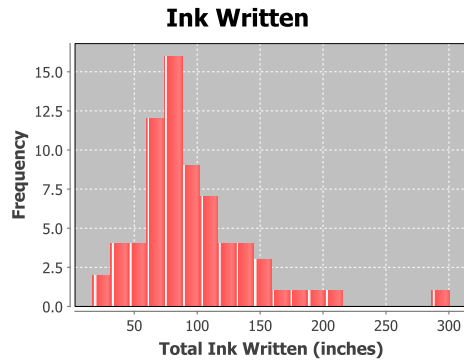


Figure .114: Histogram of the amount of ink, in inches, written, for each student on homework six problem six for the 2010 course offering.

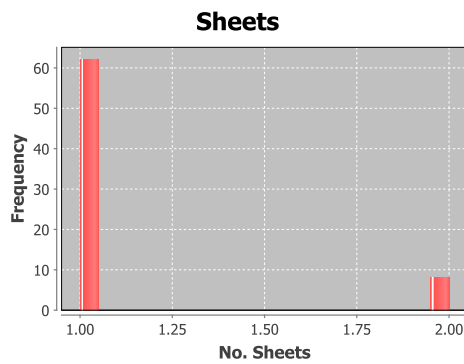


Figure .115: Histogram of the number of pages written on by each student to solve homework six problem six for the 2010 course offering.

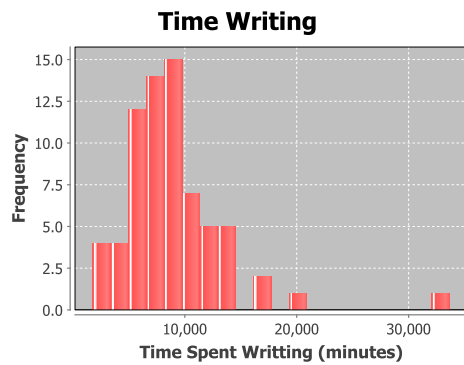


Figure .116: Histogram of the amount of time spent writing by each student on homework six problem six for the 2010 course offering.

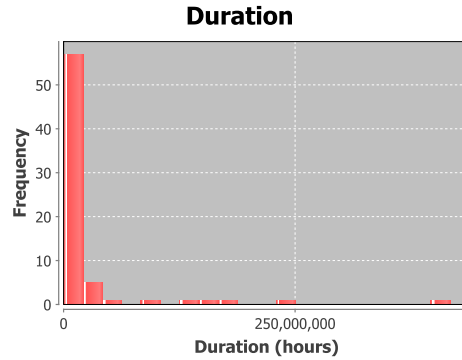


Figure .117: Histogram of the duration, time from first to last pen stroke for all students for homework six problem seven for the 2010 course offering.

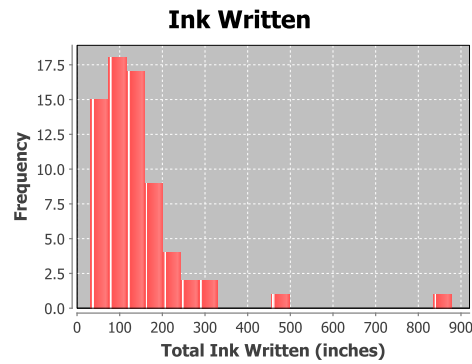


Figure .118: Histogram of the amount of ink, in inches, written, for each student on homework six problem seven for the 2010 course offering.

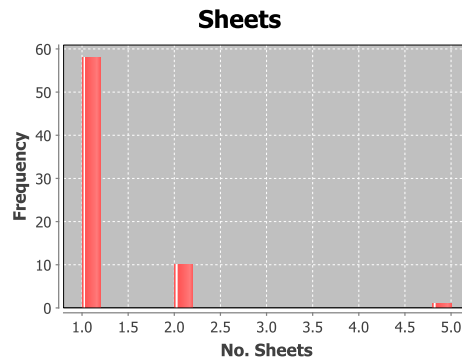


Figure .119: Histogram of the number of pages written on by each student to solve homework six problem seven for the 2010 course offering.

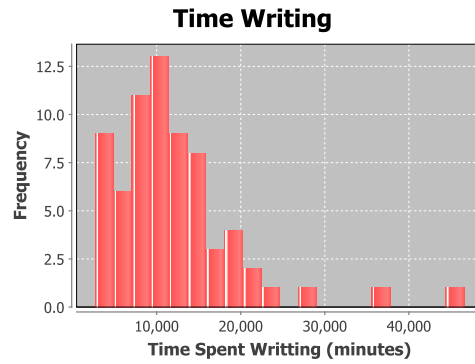


Figure .120: Histogram of the amount of time spent writing by each student on homework six problem seven for the 2010 course offering.

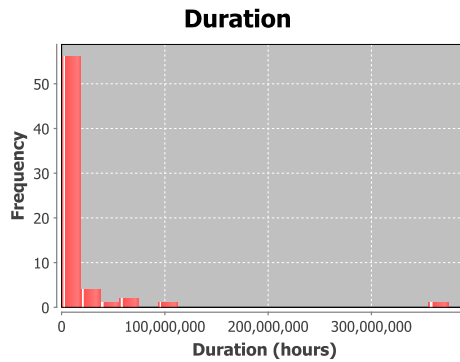


Figure .121: Histogram of the duration, time from first to last pen stroke for all students for homework six problem eight for the 2010 course offering.

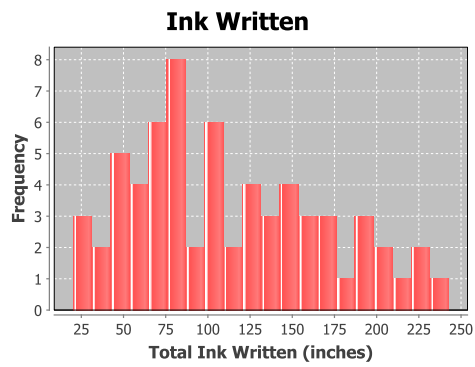


Figure .122: Histogram of the amount of ink, in inches, written, for each student on homework six problem eight for the 2010 course offering.

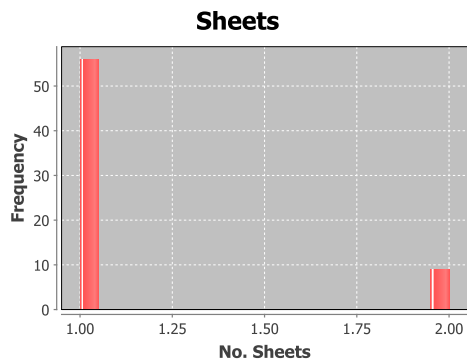


Figure .123: Histogram of the number of pages written on by each student to solve homework six problem eight for the 2010 course offering.

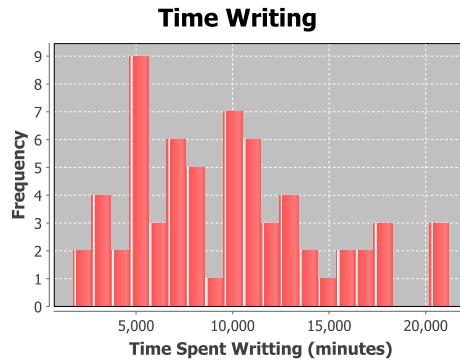


Figure .124: Histogram of the amount of time spent writing by each student on homework six problem eight for the 2010 course offering.

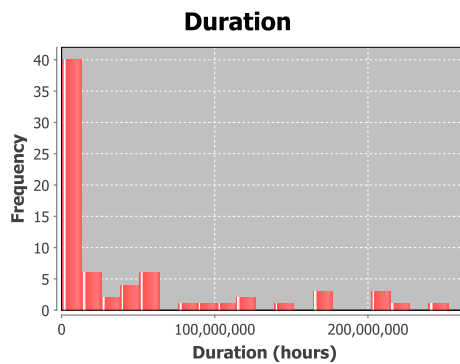


Figure .125: Histogram of the duration, time from first to last pen stroke for all students for homework seven problem one for the 2010 course offering.

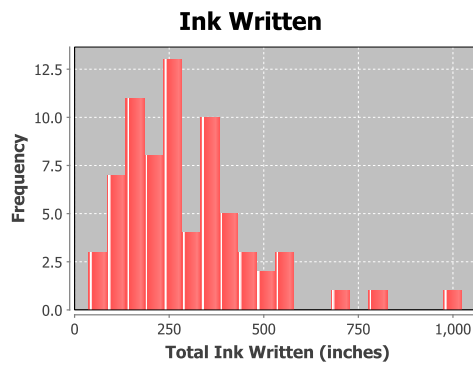


Figure .126: Histogram of the amount of ink, in inches, written, for each student on homework seven problem one for the 2010 course offering.

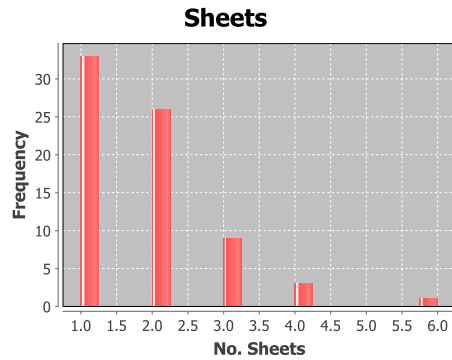


Figure .127: Histogram of the number of pages written on by each student to solve homework seven problem one for the 2010 course offering.

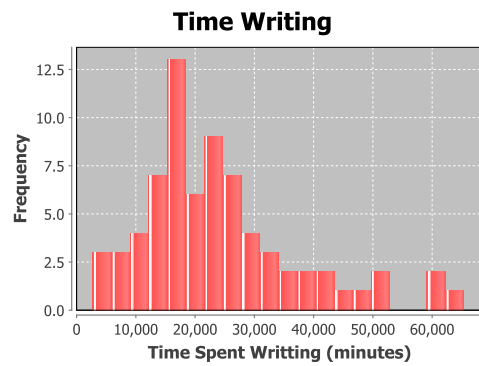


Figure .128: Histogram of the amount of time spent writing by each student on homework seven problem one for the 2010 course offering.

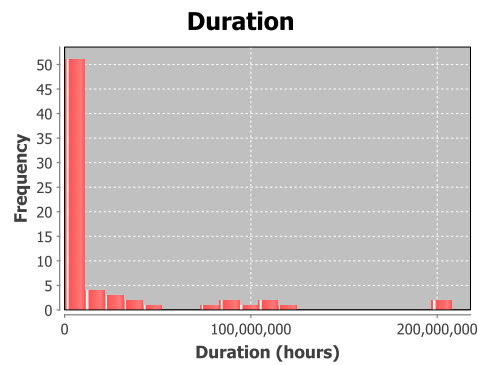


Figure .129: Histogram of the duration, time from first to last pen stroke for all students for homework seven problem two for the 2010 course offering.

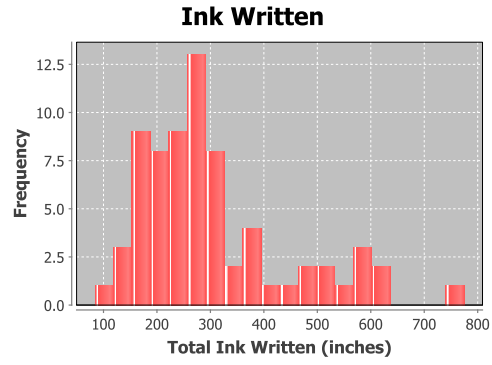


Figure .130: Histogram of the amount of ink, in inches, written, for each student on homework seven problem two for the 2010 course offering.

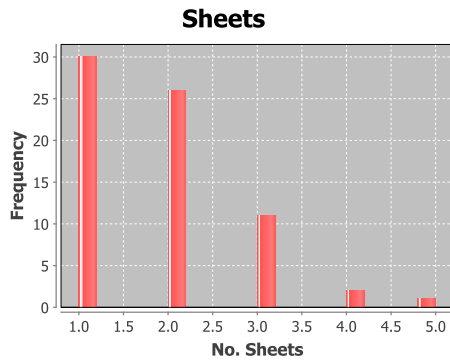


Figure .131: Histogram of the number of pages written on by each student to solve homework seven problem two for the 2010 course offering.

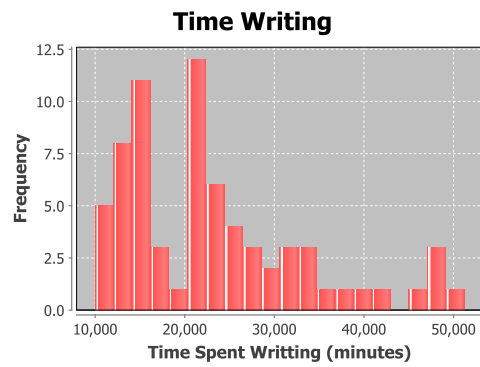


Figure .132: Histogram of the amount of time spent writing by each student on homework seven problem two for the 2010 course offering.

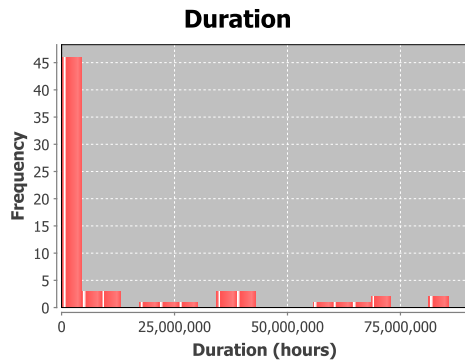


Figure .133: Histogram of the duration, time from first to last pen stroke for all students for homework seven problem three for the 2010 course offering.

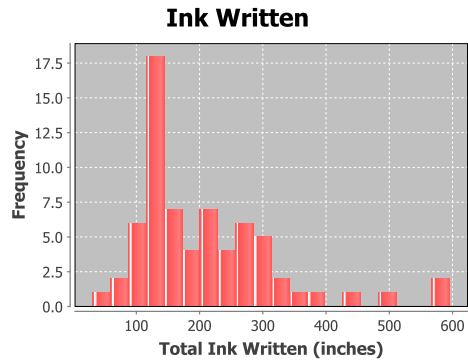


Figure .134: Histogram of the amount of ink, in inches, written, for each student on homework seven problem three for the 2010 course offering.

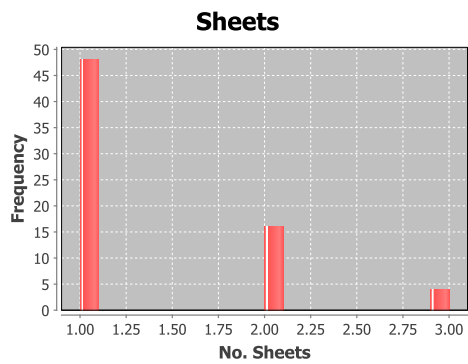


Figure .135: Histogram of the number of pages written on by each student to solve homework seven problem three for the 2010 course offering.

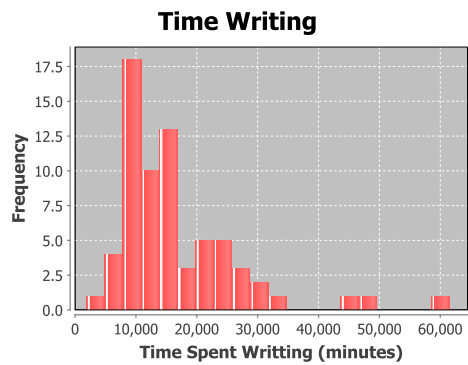


Figure .136: Histogram of the amount of time spent writing by each student on homework seven problem three for the 2010 course offering.

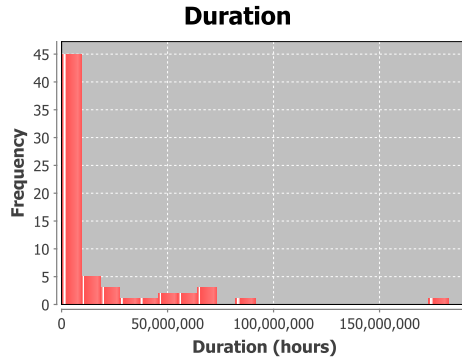


Figure .137: Histogram of the duration, time from first to last pen stroke for all students for homework seven problem four for the 2010 course offering.

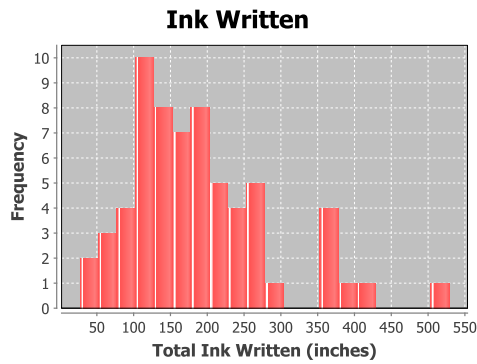


Figure .138: Histogram of the amount of ink, in inches, written, for each student on homework seven problem four for the 2010 course offering.

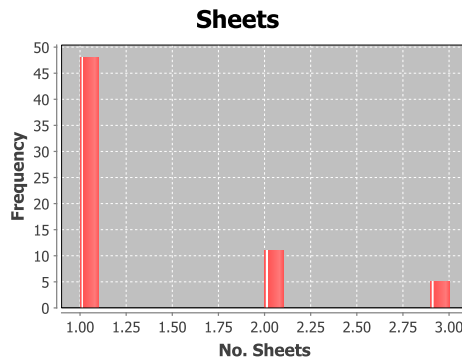


Figure .139: Histogram of the number of pages written on by each student to solve homework seven problem four for the 2010 course offering.

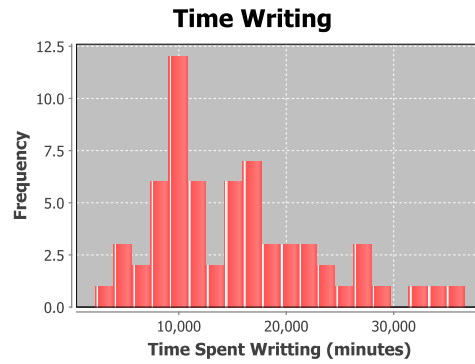


Figure .140: Histogram of the amount of time spent writing by each student on homework seven problem four for the 2010 course offering.

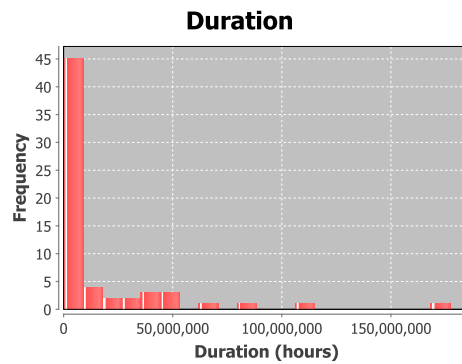


Figure .141: Histogram of the duration, time from first to last pen stroke for all students for homework seven problem five for the 2010 course offering.

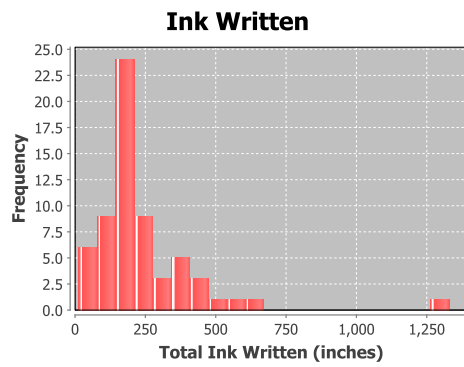


Figure .142: Histogram of the amount of ink, in inches, written, for each student on homework seven problem five for the 2010 course offering.

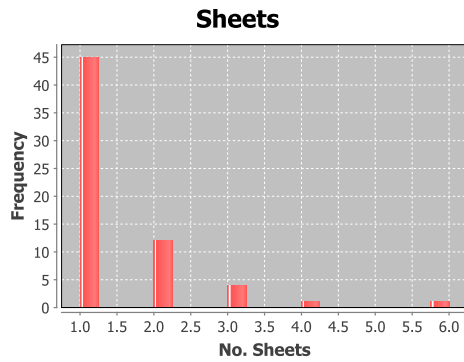


Figure .143: Histogram of the number of pages written on by each student to solve homework seven problem five for the 2010 course offering.

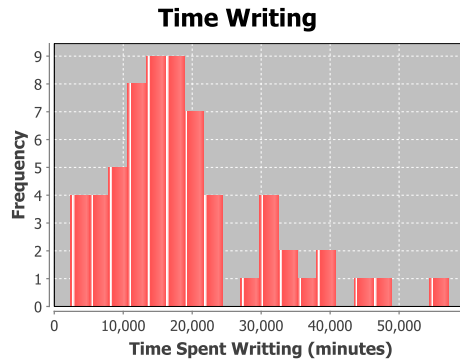


Figure .144: Histogram of the amount of time spent writing by each student on homework seven problem five for the 2010 course offering.

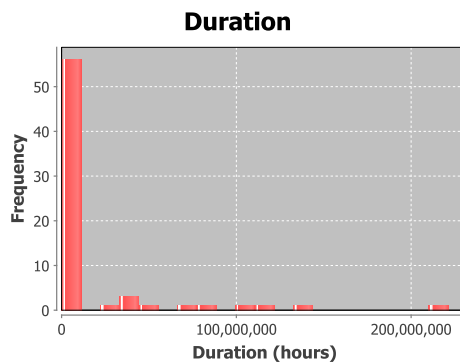


Figure .145: Histogram of the duration, time from first to last pen stroke for all students for homework eight problem one for the 2010 course offering.

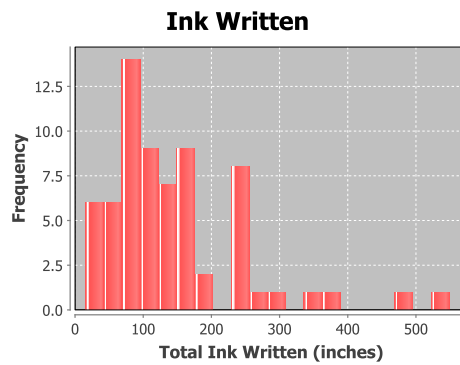


Figure .146: Histogram of the amount of ink, in inches, written, for each student on homework eight problem one for the 2010 course offering.

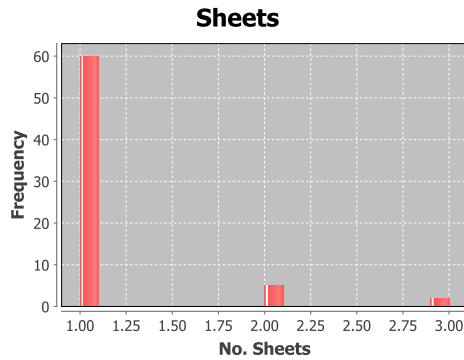


Figure .147: Histogram of the number of pages written on by each student to solve homework eight problem one for the 2010 course offering.

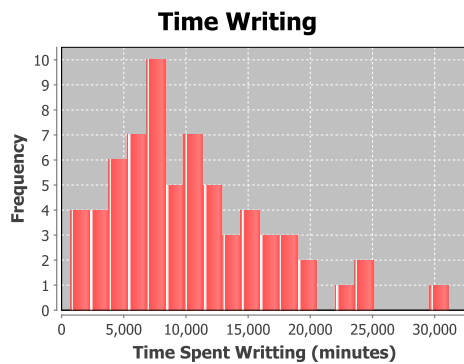


Figure .148: Histogram of the amount of time spent writing by each student on homework eight problem one for the 2010 course offering.

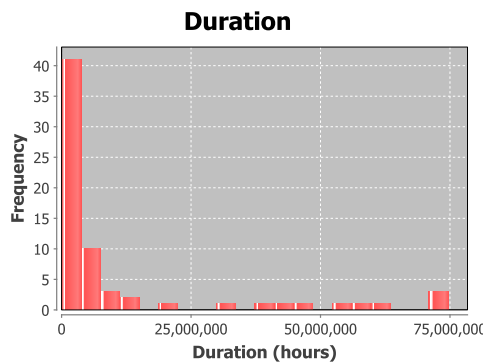


Figure .149: Histogram of the duration, time from first to last pen stroke for all students for homework eight problem two for the 2010 course offering.

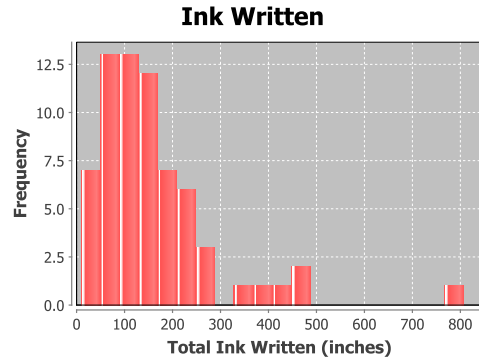


Figure .150: Histogram of the amount of ink, in inches, written, for each student on homework eight problem two for the 2010 course offering.

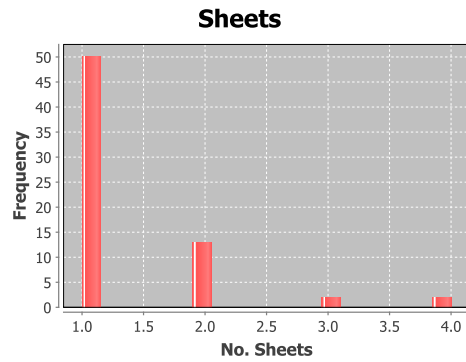


Figure .151: Histogram of the number of pages written on by each student to solve homework eight problem two for the 2010 course offering.

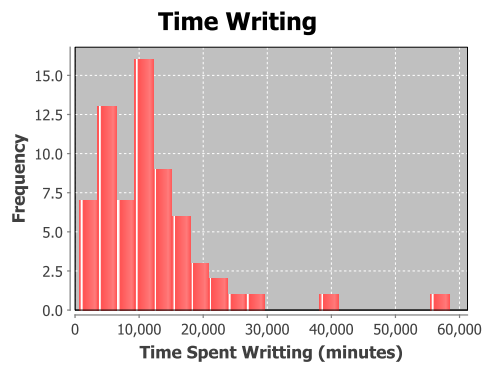


Figure .152: Histogram of the amount of time spent writing by each student on homework eight problem two for the 2010 course offering.

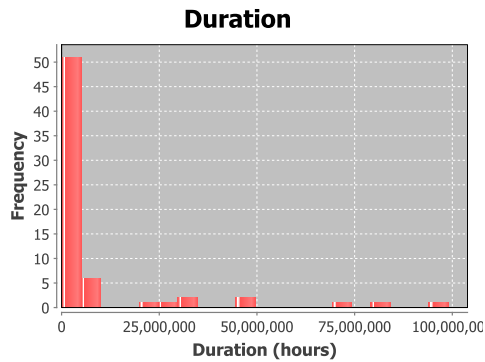


Figure .153: Histogram of the duration, time from first to last pen stroke for all students for homework eight problem three for the 2010 course offering.

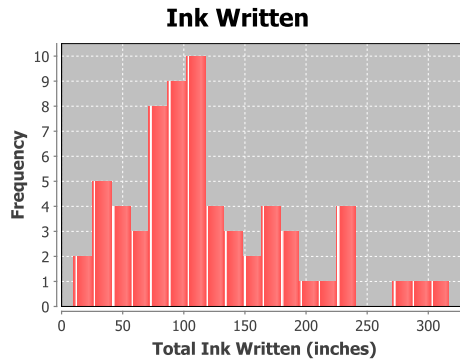


Figure .154: Histogram of the amount of ink, in inches, written, for each student on homework eight problem three for the 2010 course offering.

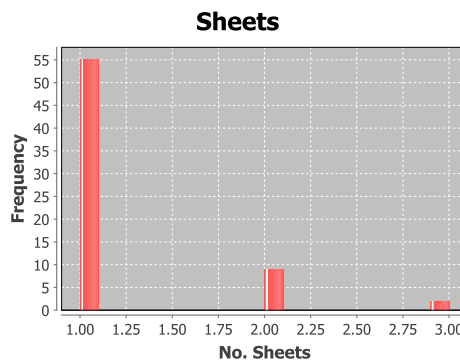


Figure .155: Histogram of the number of pages written on by each student to solve homework eight problem three for the 2010 course offering.

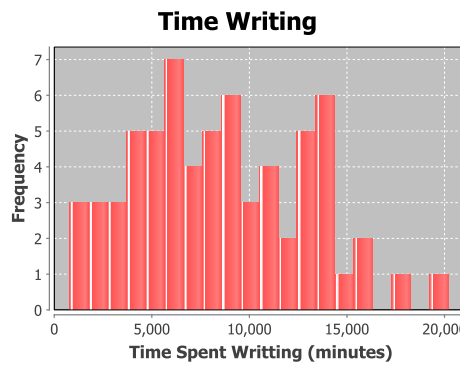


Figure .156: Histogram of the amount of time spent writing by each student on homework eight problem three for the 2010 course offering.

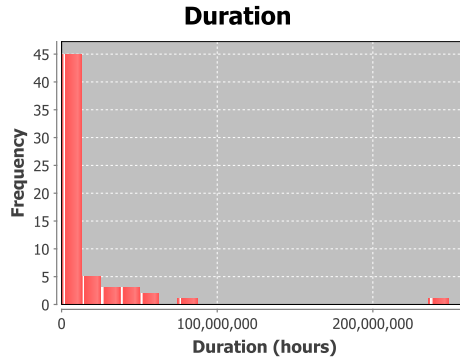


Figure .157: Histogram of the duration, time from first to last pen stroke for all students for homework eight problem four for the 2010 course offering.

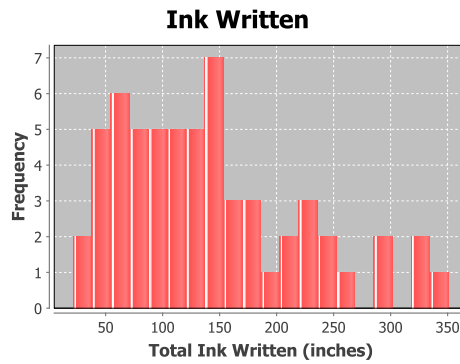


Figure .158: Histogram of the amount of ink, in inches, written, for each student on homework eight problem four for the 2010 course offering.

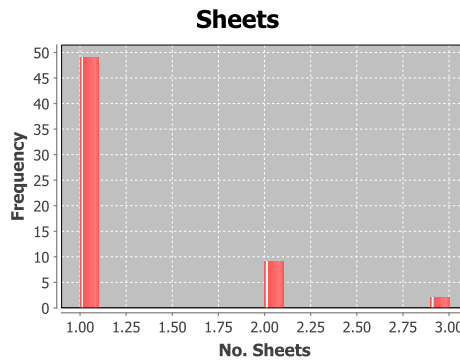


Figure .159: Histogram of the number of pages written on by each student to solve homework eight problem four for the 2010 course offering.

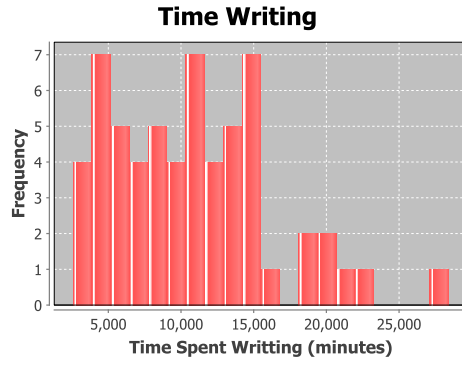


Figure .160: Histogram of the amount of time spent writing by each student on homework eight problem four for the 2010 course offering.

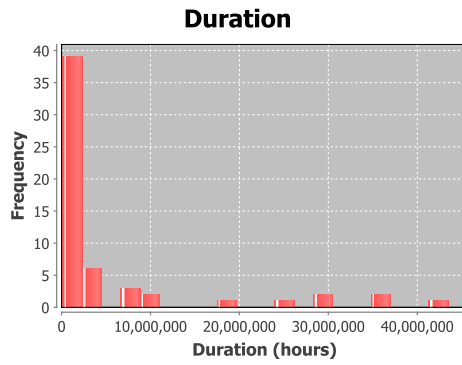


Figure .161: Histogram of the duration, time from first to last pen stroke for all students for homework eight problem five for the 2010 course offering.

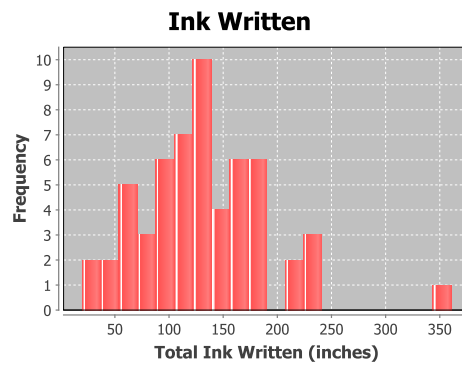


Figure .162: Histogram of the amount of ink, in inches, written, for each student on homework eight problem five for the 2010 course offering.

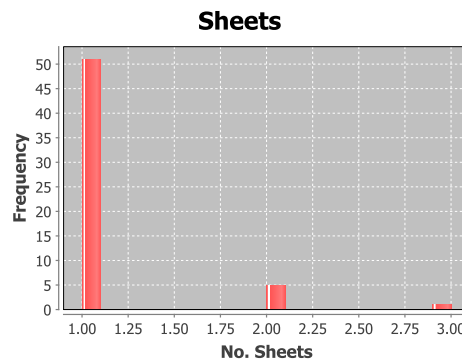


Figure .163: Histogram of the number of pages written on by each student to solve homework eight problem five for the 2010 course offering.

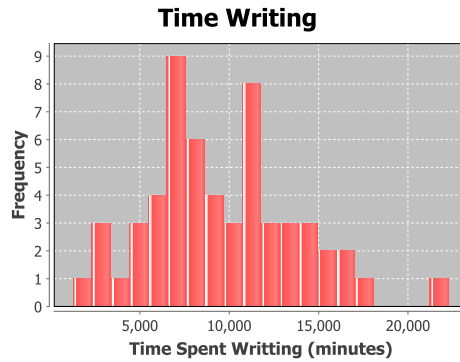


Figure .164: Histogram of the amount of time spent writing by each student on homework eight problem five for the 2010 course offering.

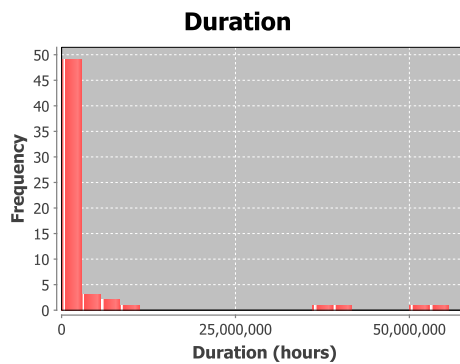


Figure .165: Histogram of the duration, time from first to last pen stroke for all students for homework eight problem six for the 2010 course offering.

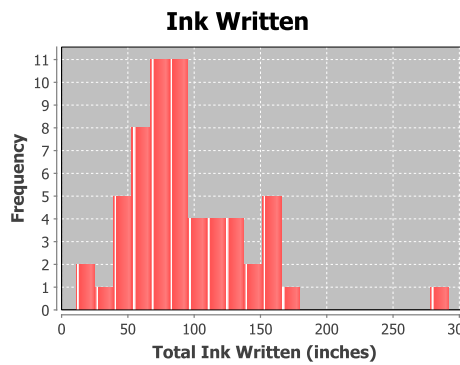


Figure .166: Histogram of the amount of ink, in inches, written, for each student on homework eight problem six for the 2010 course offering.

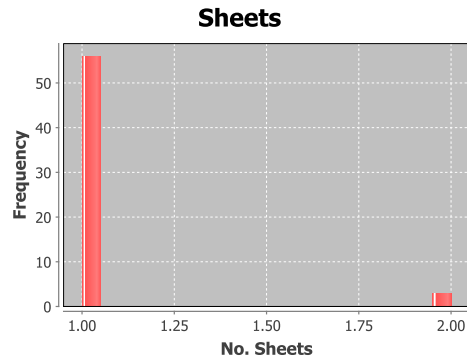


Figure .167: Histogram of the number of pages written on by each student to solve homework eight problem six for the 2010 course offering.

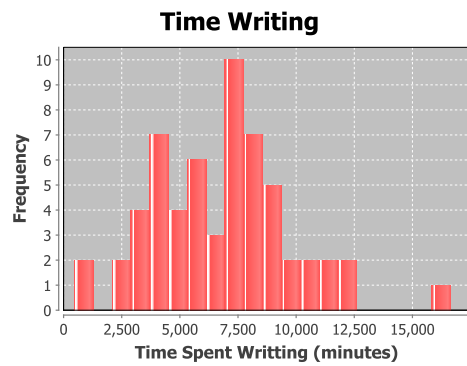


Figure .168: Histogram of the amount of time spent writing by each student on homework eight problem six for the 2010 course offering.

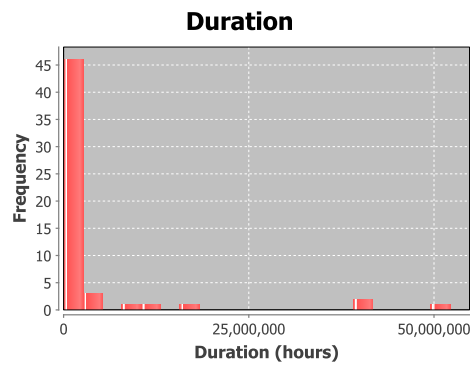


Figure .169: Histogram of the duration, time from first to last pen stroke for all students for homework eight problem seven for the 2010 course offering.

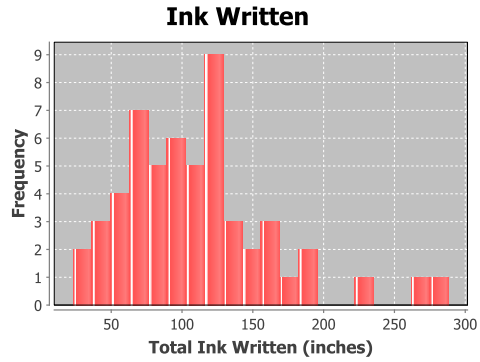


Figure .170: Histogram of the amount of ink, in inches, written, for each student on homework eight problem seven for the 2010 course offering.

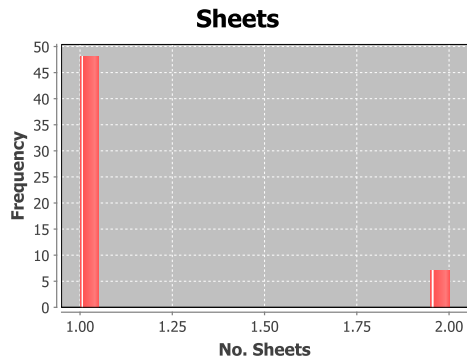


Figure .171: Histogram of the number of pages written on by each student to solve homework eight problem seven for the 2010 course offering.

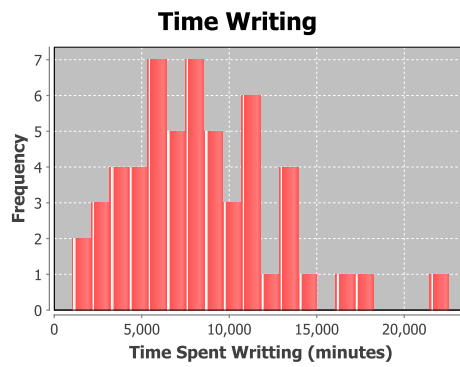


Figure .172: Histogram of the amount of time spent writing by each student on homework eight problem seven for the 2010 course offering.

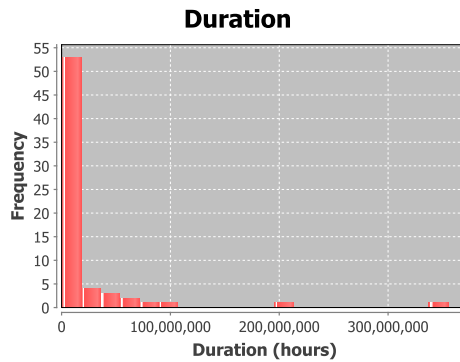


Figure .173: Histogram of the duration, time from first to last pen stroke for all students for homework nine problem one for the 2010 course offering.

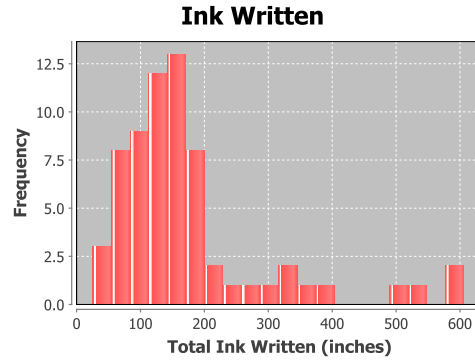


Figure .174: Histogram of the amount of ink, in inches, written, for each student on homework nine problem one for the 2010 course offering.

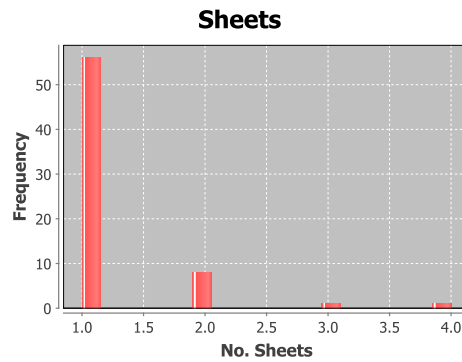


Figure .175: Histogram of the number of pages written on by each student to solve homework nine problem one for the 2010 course offering.

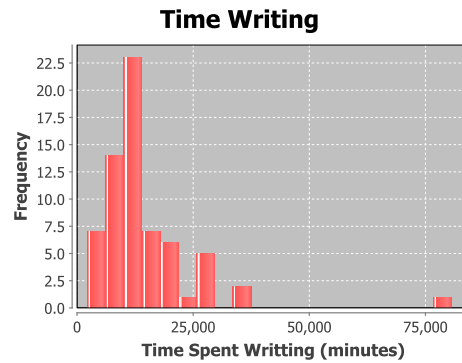


Figure .176: Histogram of the amount of time spent writing by each student on homework nine problem one for the 2010 course offering.

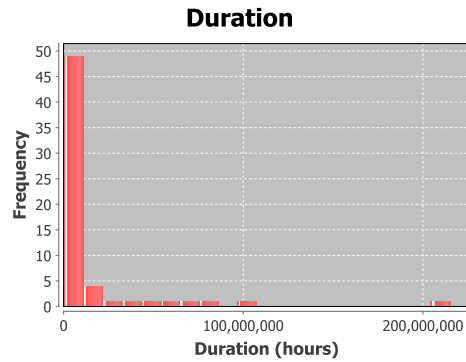


Figure .177: Histogram of the duration, time from first to last pen stroke for all students for homework nine problem two for the 2010 course offering.

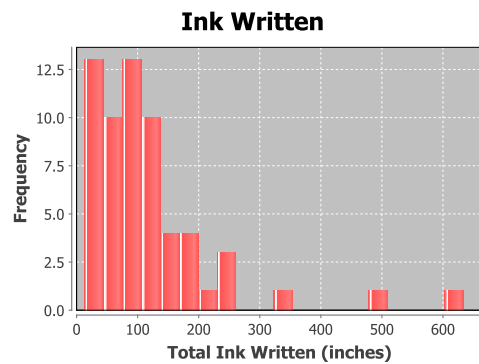


Figure .178: Histogram of the amount of ink, in inches, written, for each student on homework nine problem two for the 2010 course offering.

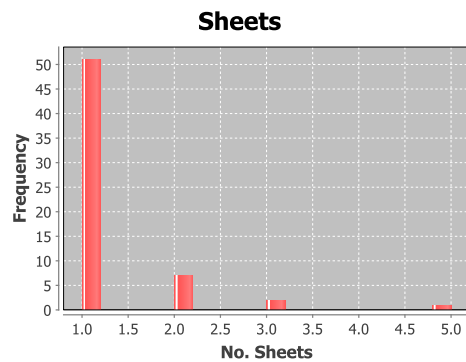


Figure .179: Histogram of the number of pages written on by each student to solve homework nine problem two for the 2010 course offering.

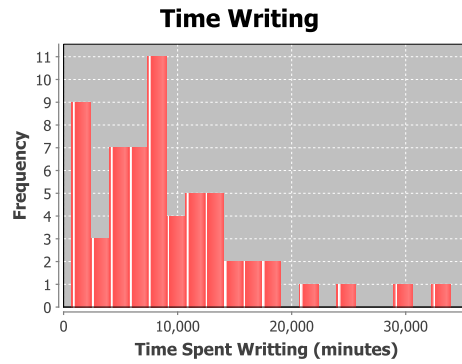


Figure .180: Histogram of the amount of time spent writing by each student on homework nine problem two for the 2010 course offering.

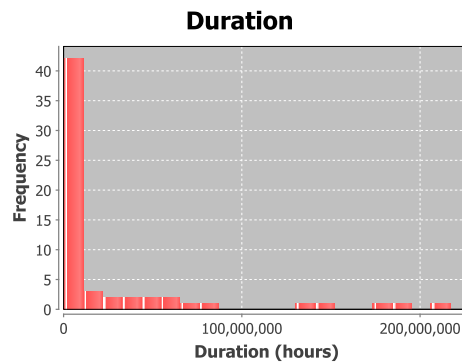


Figure .181: Histogram of the duration, time from first to last pen stroke for all students for homework nine problem three for the 2010 course offering.

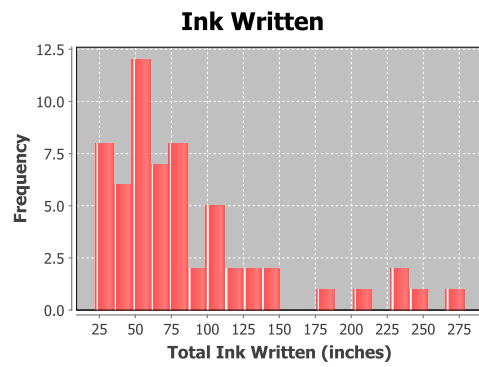


Figure .182: Histogram of the amount of ink, in inches, written, for each student on homework nine problem three for the 2010 course offering.

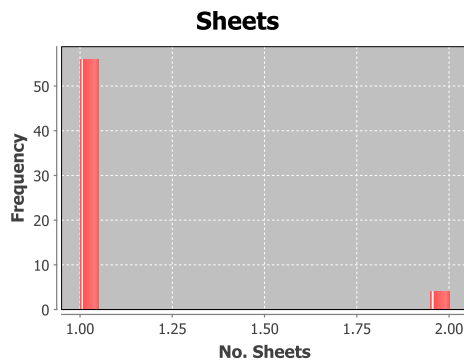


Figure .183: Histogram of the number of pages written on by each student to solve homework nine problem three for the 2010 course offering.

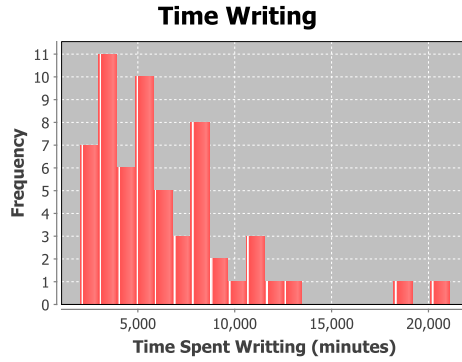


Figure .184: Histogram of the amount of time spent writing by each student on homework nine problem three for the 2010 course offering.

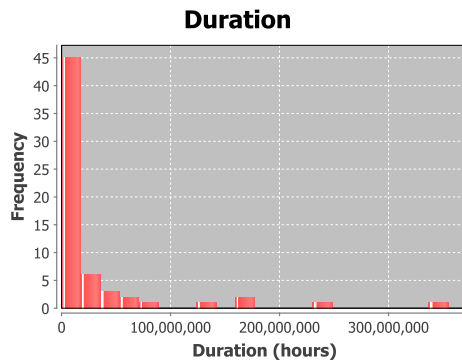


Figure .185: Histogram of the duration, time from first to last pen stroke for all students for homework nine problem four for the 2010 course offering.

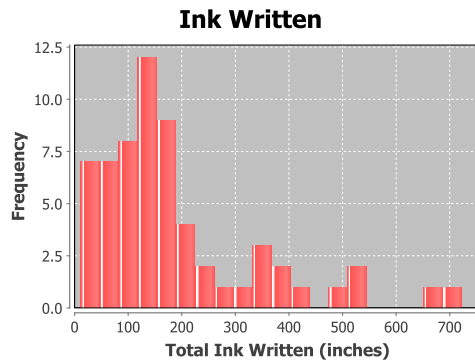


Figure .186: Histogram of the amount of ink, in inches, written, for each student on homework nine problem four for the 2010 course offering.

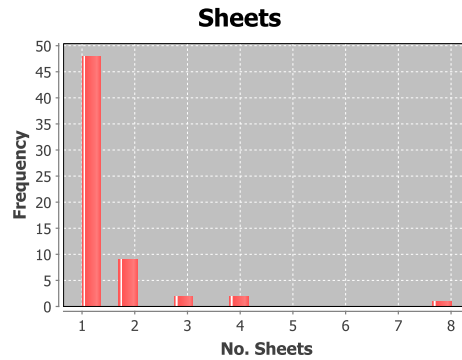


Figure .187: Histogram of the number of pages written on by each student to solve homework nine problem four for the 2010 course offering.

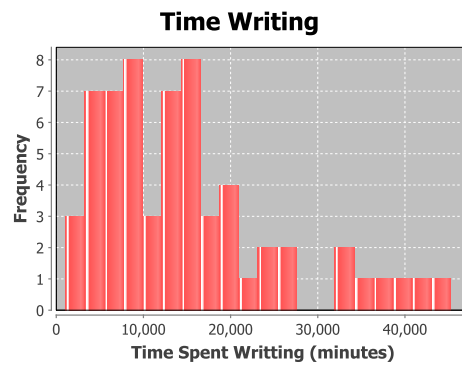


Figure .188: Histogram of the amount of time spent writing by each student on homework nine problem four for the 2010 course offering.

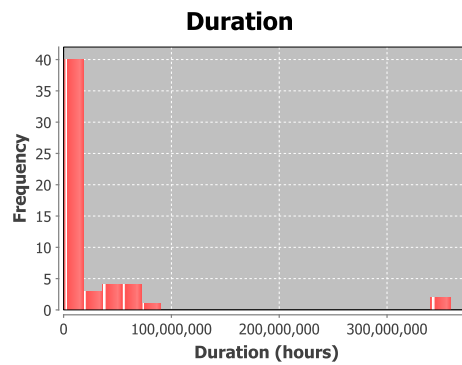


Figure .189: Histogram of the duration, time from first to last pen stroke for all students for homework nine problem five for the 2010 course offering.

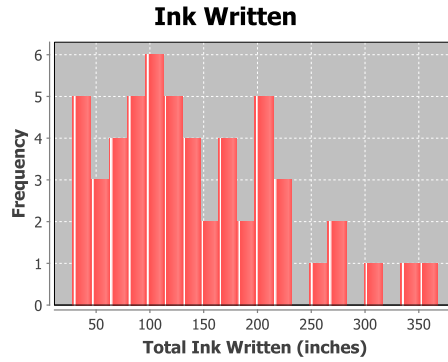


Figure .190: Histogram of the amount of ink, in inches, written, for each student on homework nine problem five for the 2010 course offering.

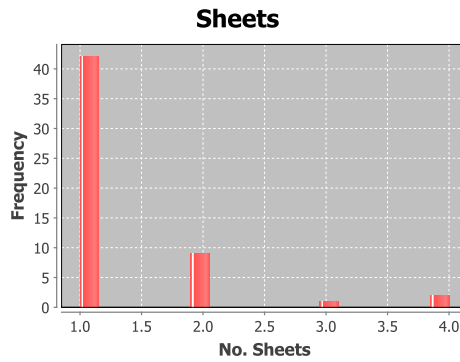


Figure .191: Histogram of the number of pages written on by each student to solve homework nine problem five for the 2010 course offering.

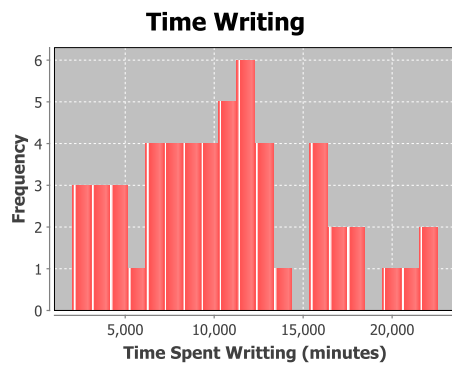


Figure .192: Histogram of the amount of time spent writing by each student on homework nine problem five for the 2010 course offering.

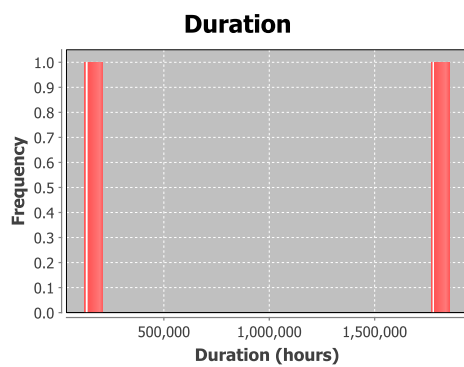


Figure .193: Histogram of the duration, time from first to last pen stroke for all students for homework one problem one for the 2011 course offering.

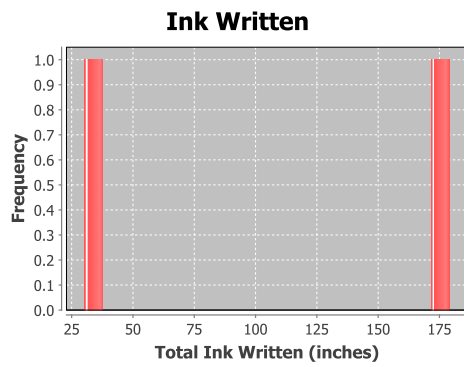


Figure .194: Histogram of the amount of ink, in inches, written, for each student on homework one problem one for the 2011 course offering.

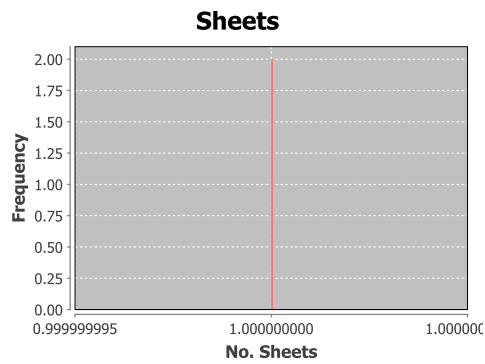


Figure .195: Histogram of the number of pages written on by each student to solve homework one problem one for the 2011 course offering.

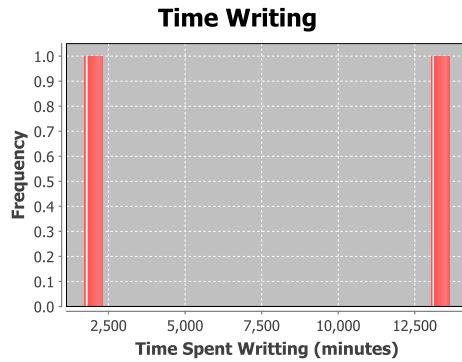


Figure .196: Histogram of the amount of time spent writing by each student on homework one problem one for the 2011 course offering.

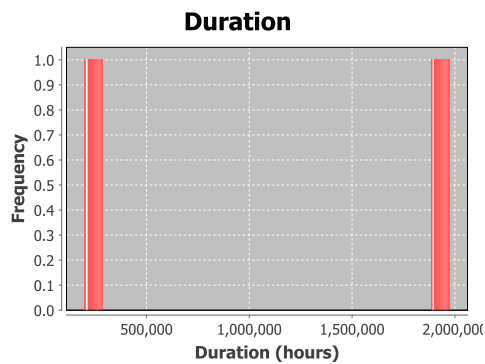


Figure .197: Histogram of the duration, time from first to last pen stroke for all students for homework one problem two for the 2011 course offering.

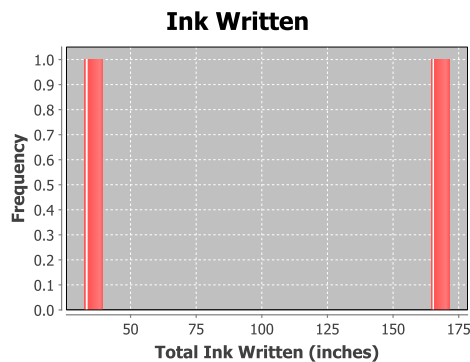


Figure .198: Histogram of the amount of ink, in inches, written, for each student on homework one problem two for the 2011 course offering.

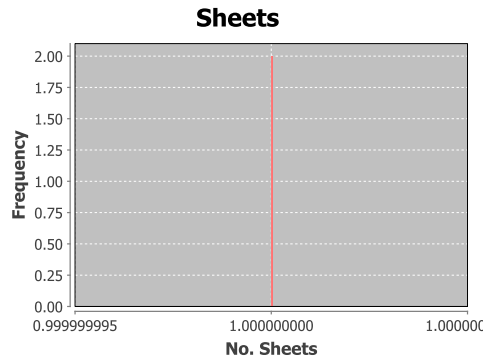


Figure .199: Histogram of the number of pages written on by each student to solve homework one problem two for the 2011 course offering.

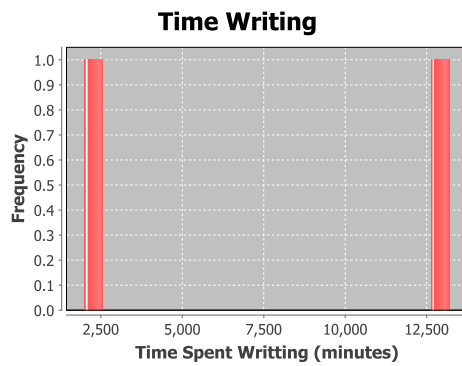


Figure .200: Histogram of the amount of time spent writing by each student on homework one problem two for the 2011 course offering.

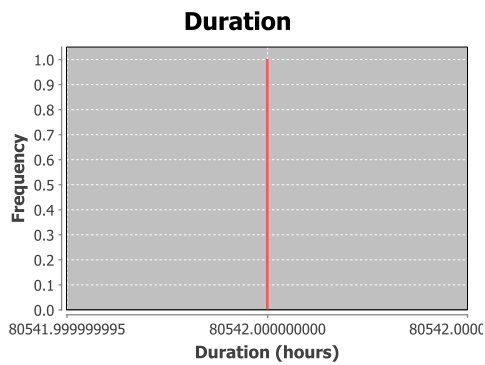


Figure .201: Histogram of the duration, time from first to last pen stroke for all students for homework one problem three for the 2011 course offering.

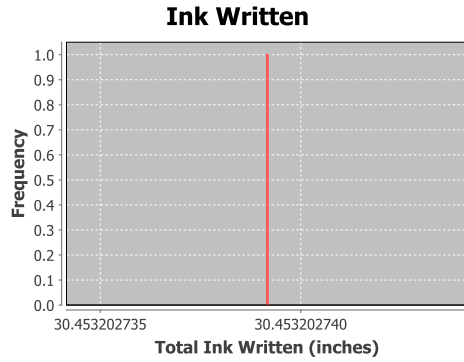


Figure .202: Histogram of the amount of ink, in inches, written, for each student on homework one problem three for the 2011 course offering.

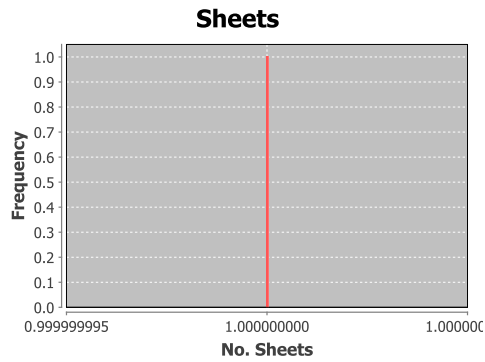


Figure .203: Histogram of the number of pages written on by each student to solve homework one problem two for the 2011 course offering.

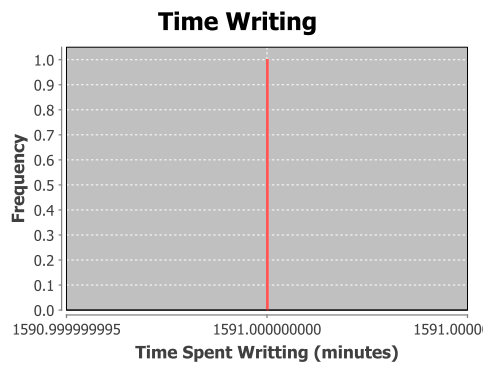


Figure .204: Histogram of the amount of time spent writing by each student on homework one problem three for the 2011 course offering.

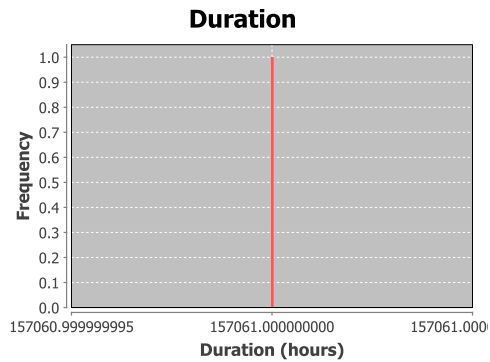


Figure .205: Histogram of the duration, time from first to last pen stroke for all students for homework one problem four for the 2011 course offering.

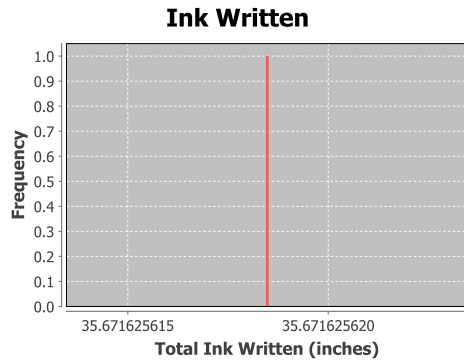


Figure .206: Histogram of the amount of ink, in inches, written, for each student on homework one problem four for the 2011 course offering.

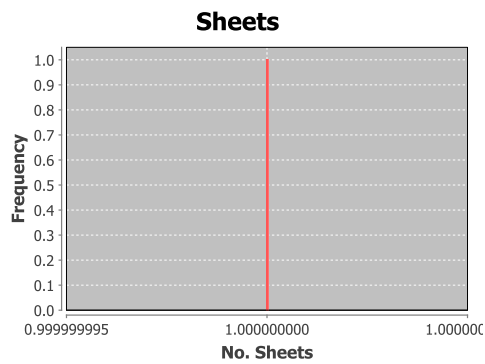


Figure .207: Histogram of the number of pages written on by each student to solve homework one problem four for the 2011 course offering.

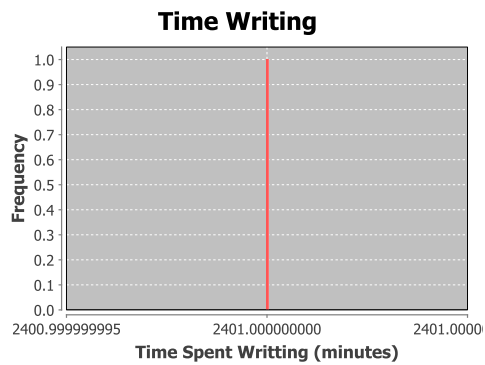


Figure .208: Histogram of the amount of time spent writing by each student on homework one problem four for the 2011 course offering.

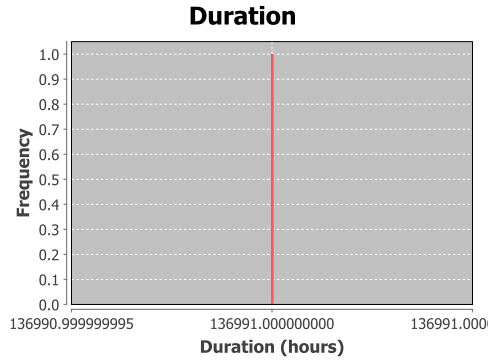


Figure .209: Histogram of the duration, time from first to last pen stroke for all students for homework one problem five for the 2011 course offering.

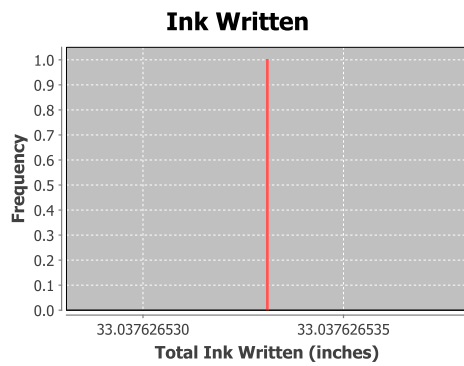


Figure .210: Histogram of the amount of ink, in inches, written, for each student on homework one problem five for the 2011 course offering.

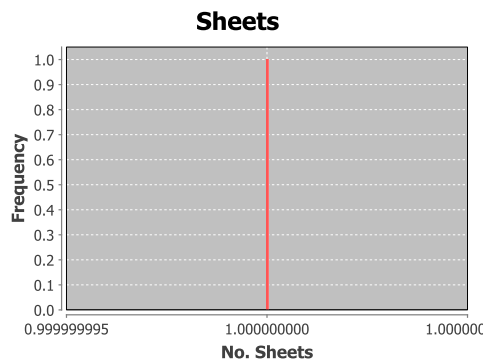


Figure .211: Histogram of the number of pages written on by each student to solve homework one problem five for the 2011 course offering.

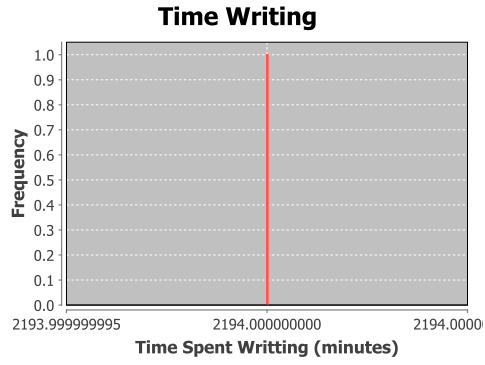


Figure .212: Histogram of the amount of time spent writing by each student on homework one problem five for the 2011 course offering.

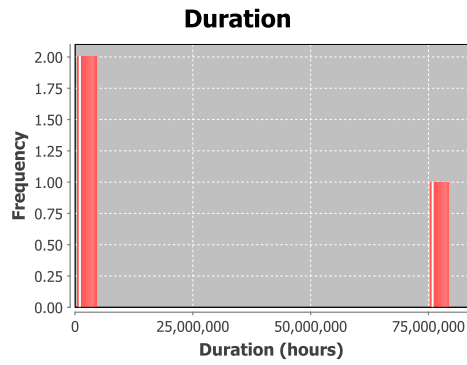


Figure .213: Histogram of the duration, time from first to last pen stroke for all students for homework two problem one for the 2011 course offering.

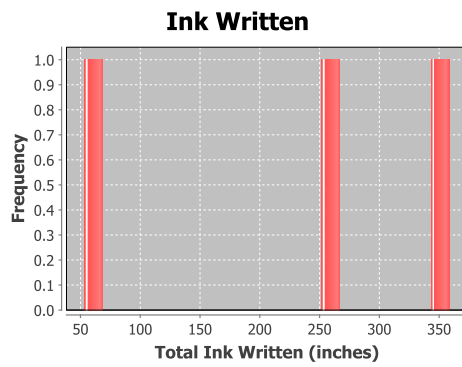


Figure .214: Histogram of the amount of ink, in inches, written, for each student on homework two problem one for the 2011 course offering.

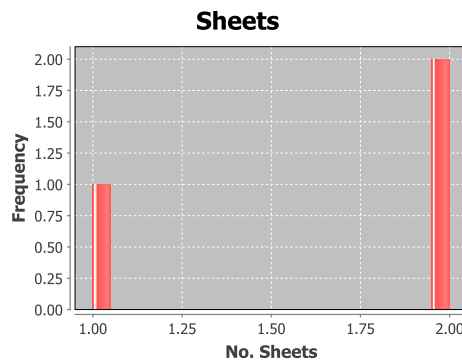


Figure .215: Histogram of the number of pages written on by each student to solve homework two problem one for the 2011 course offering.

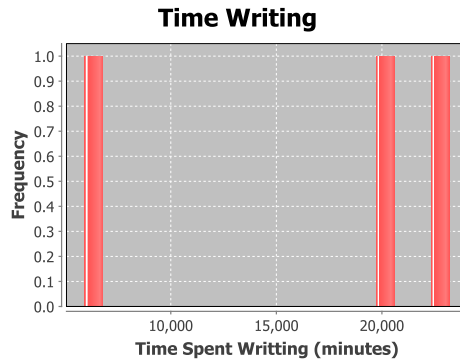


Figure .216: Histogram of the amount of time spent writing by each student on homework two problem one for the 2011 course offering.

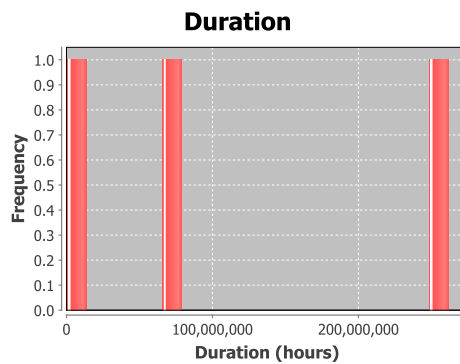


Figure .217: Histogram of the duration, time from first to last pen stroke for all students for homework two problem two for the 2011 course offering.

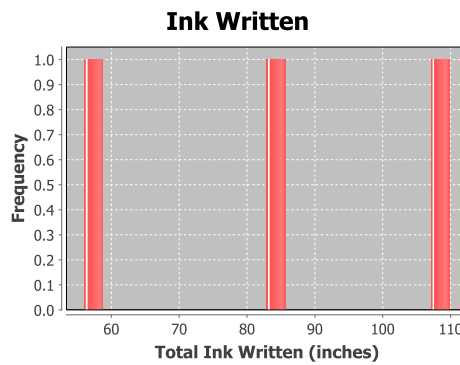


Figure .218: Histogram of the amount of ink, in inches, written, for each student on homework two problem two for the 2011 course offering.

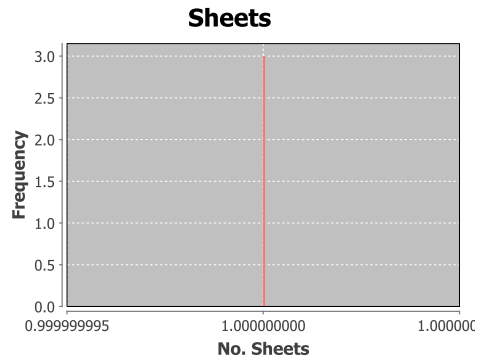


Figure .219: Histogram of the number of pages written on by each student to solve homework two problem two for the 2011 course offering.

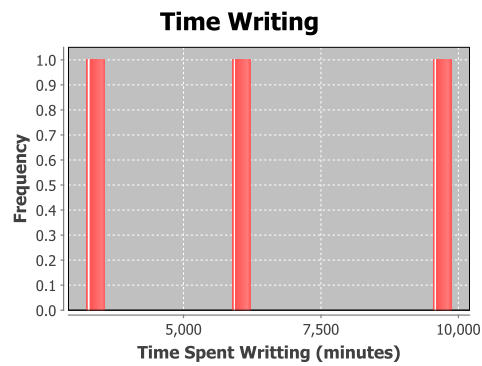


Figure .220: Histogram of the amount of time spent writing by each student on homework two problem two for the 2011 course offering.

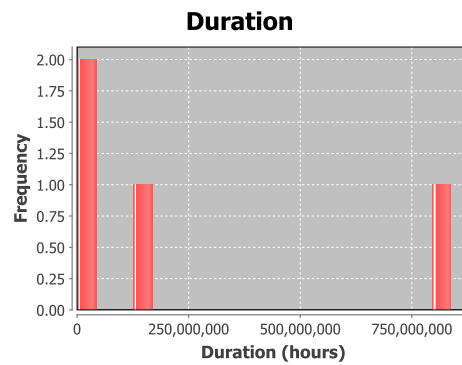


Figure .221: Histogram of the duration, time from first to last pen stroke for all students for homework two problem three for the 2011 course offering.

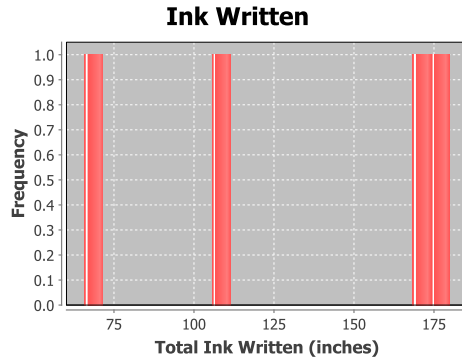


Figure .222: Histogram of the amount of ink, in inches, written, for each student on homework two problem three for the 2011 course offering.

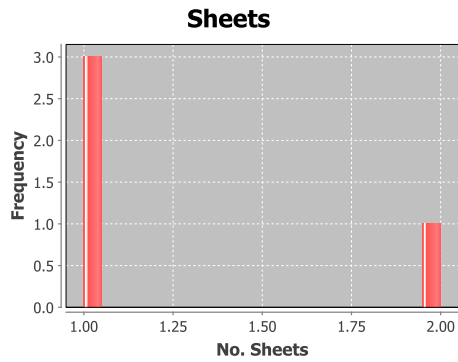


Figure .223: Histogram of the number of pages written on by each student to solve homework two problem three for the 2011 course offering.

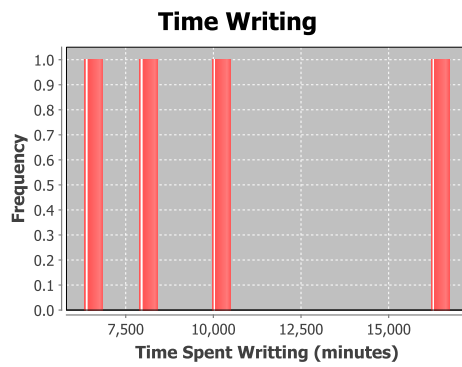


Figure .224: Histogram of the amount of time spent writing by each student on homework two problem three for the 2011 course offering.

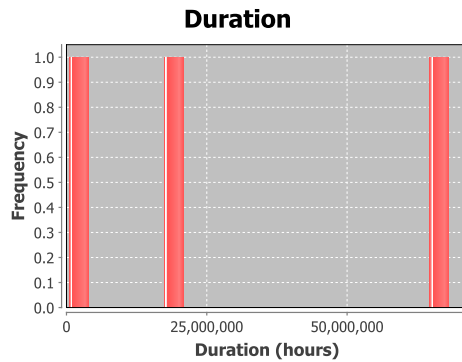


Figure .225: Histogram of the duration, time from first to last pen stroke for all students for homework two problem four for the 2011 course offering.

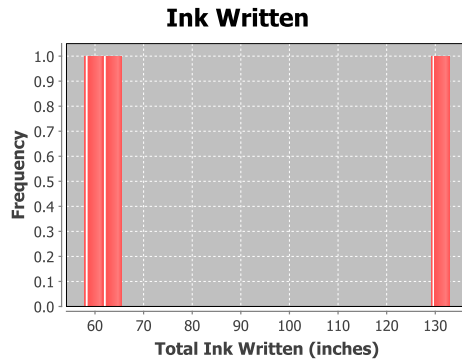


Figure .226: Histogram of the amount of ink, in inches, written, for each student on homework two problem four for the 2011 course offering.

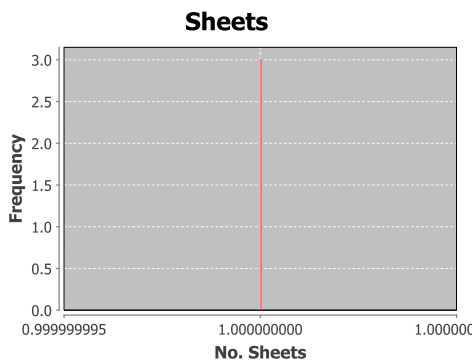


Figure .227: Histogram of the number of pages written on by each student to solve homework two problem four for the 2011 course offering.

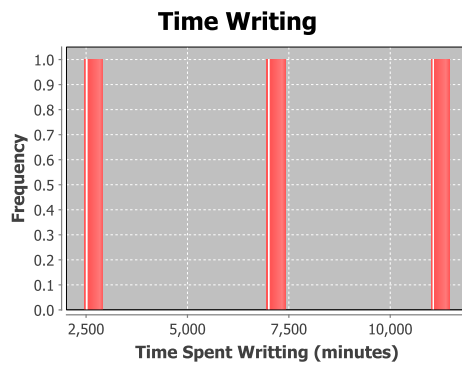


Figure .228: Histogram of the amount of time spent writing by each student on homework two problem four for the 2011 course offering.

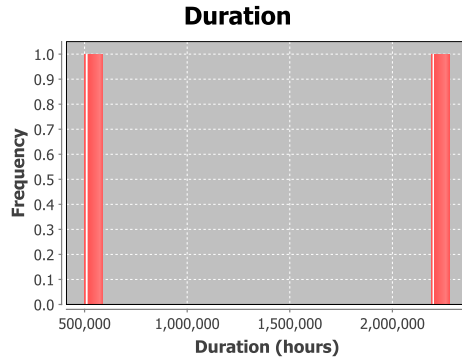


Figure .229: Histogram of the duration, time from first to last pen stroke for all students for homework two problem five for the 2011 course offering.

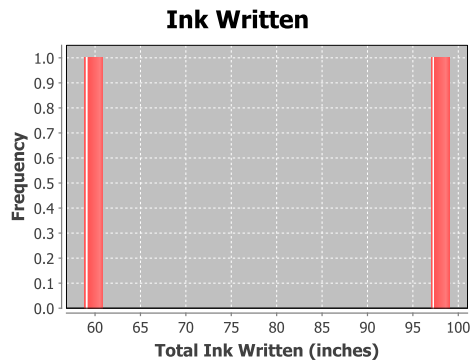


Figure .230: Histogram of the amount of ink, in inches, written, for each student on homework two problem five for the 2011 course offering.

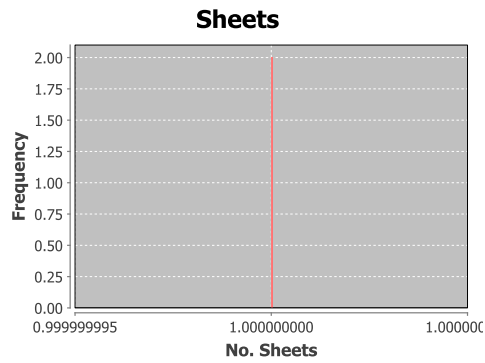


Figure .231: Histogram of the number of pages written on by each student to solve homework two problem five for the 2011 course offering.

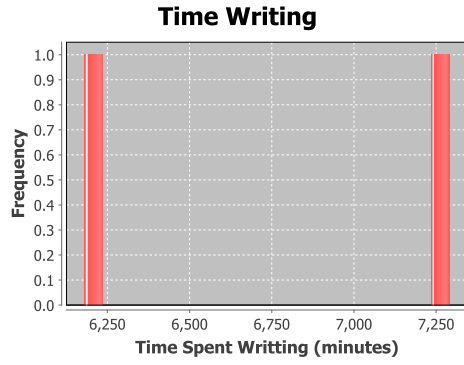


Figure .232: Histogram of the amount of time spent writing by each student on homework two problem five for the 2011 course offering.

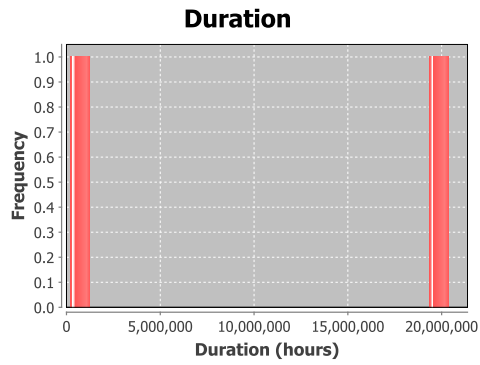


Figure .233: Histogram of the duration, time from first to last pen stroke for all students for homework two problem six for the 2011 course offering.

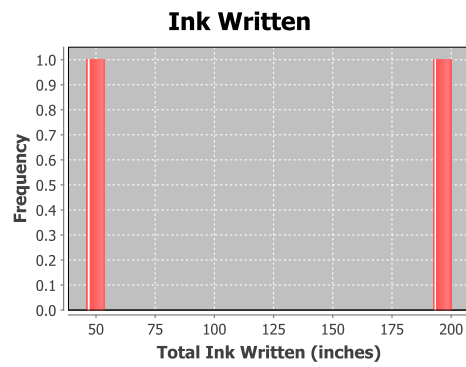


Figure .234: Histogram of the amount of ink, in inches, written, for each student on homework two problem six for the 2011 course offering.

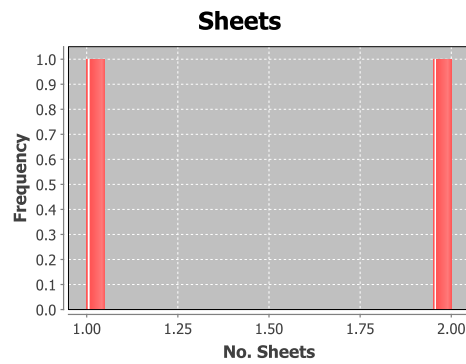


Figure .235: Histogram of the number of pages written on by each student to solve homework two problem six for the 2011 course offering.

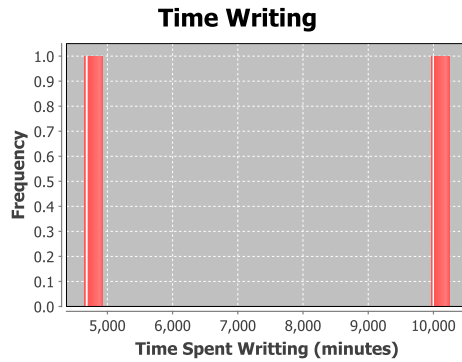


Figure .236: Histogram of the amount of time spent writing by each student on homework two problem six for the 2011 course offering.

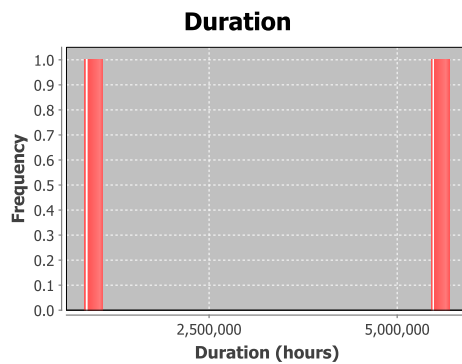


Figure .237: Histogram of the duration, time from first to last pen stroke for all students for homework two problem seven for the 2011 course offering.

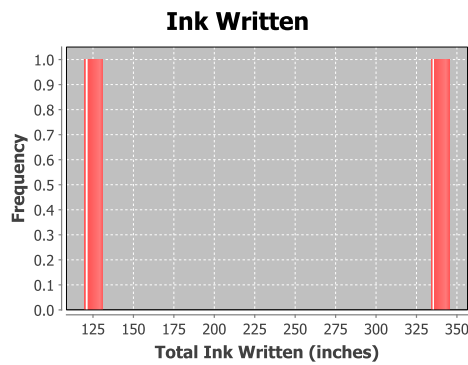


Figure .238: Histogram of the amount of ink, in inches, written, for each student on homework two problem seven for the 2011 course offering.

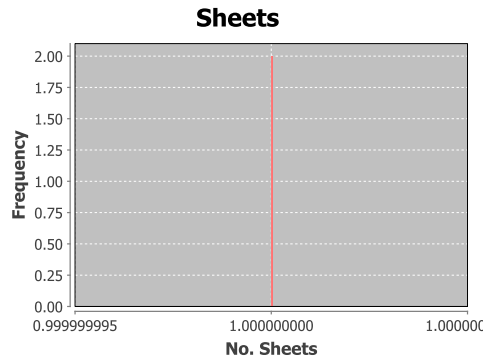


Figure .239: Histogram of the number of pages written on by each student to solve homework two problem seven for the 2011 course offering.

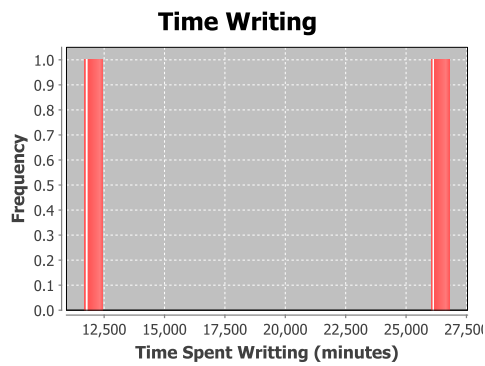


Figure .240: Histogram of the amount of time spent writing by each student on homework two problem seven for the 2011 course offering.

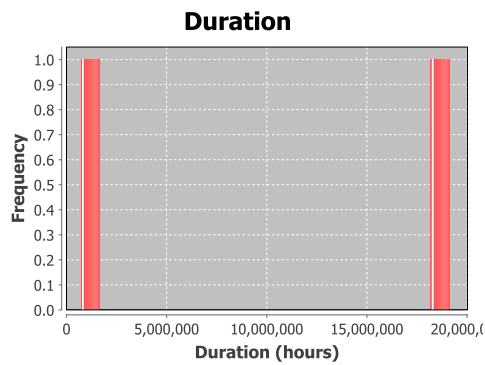


Figure .241: Histogram of the duration, time from first to last pen stroke for all students for homework two problem eight for the 2011 course offering.

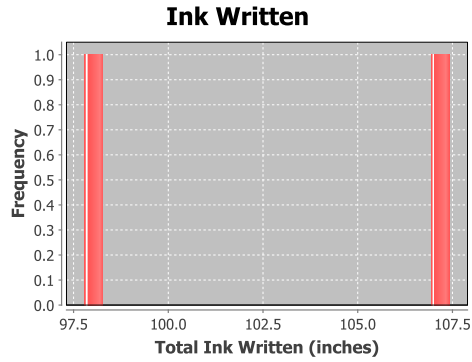


Figure .242: Histogram of the amount of ink, in inches, written, for each student on homework two problem eight for the 2011 course offering.

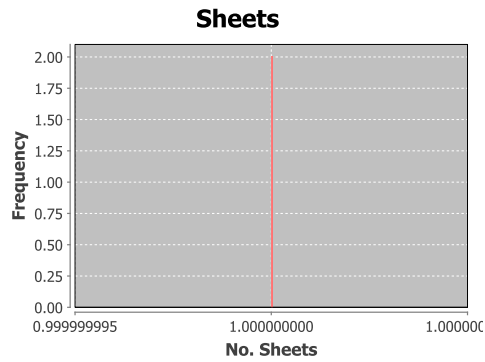


Figure .243: Histogram of the number of pages written on by each student to solve homework two problem eight for the 2011 course offering.

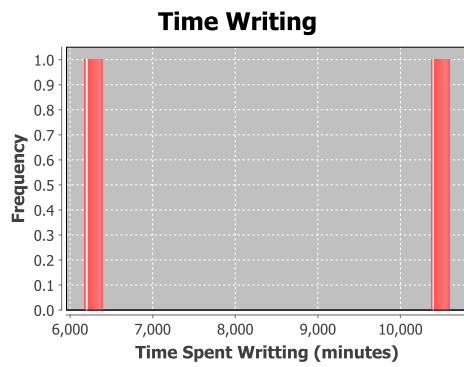


Figure .244: Histogram of the amount of time spent writing by each student on homework two problem eight for the 2011 course offering.

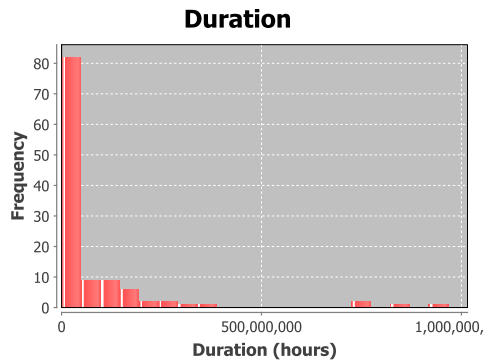


Figure .245: Histogram of the duration, time from first to last pen stroke for all students for homework three problem one for the 2011 course offering.

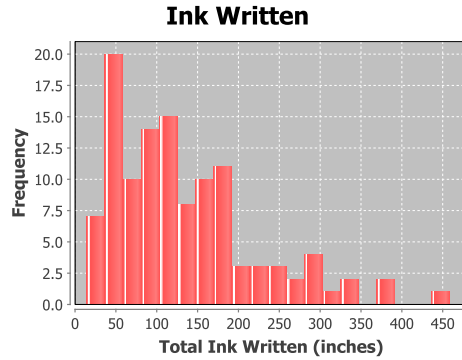


Figure .246: Histogram of the amount of ink, in inches, written, for each student on homework three problem one for the 2011 course offering.

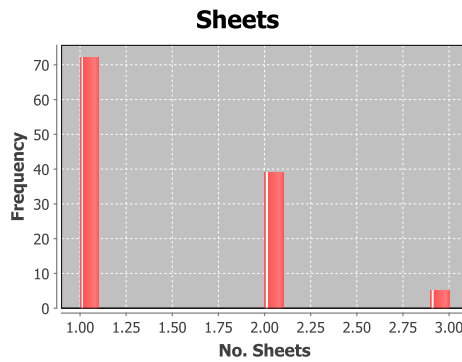


Figure .247: Histogram of the number of pages written on by each student to solve homework three problem one for the 2011 course offering.

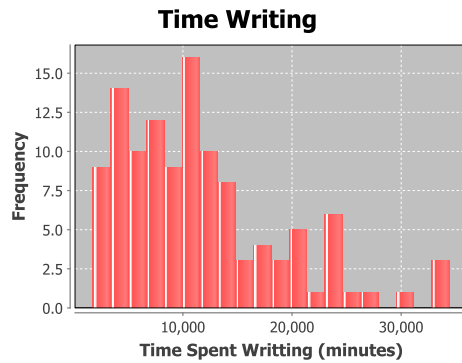


Figure .248: Histogram of the amount of time spent writing by each student on homework three problem one for the 2011 course offering.

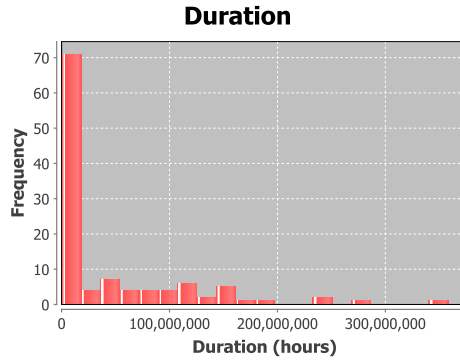


Figure .249: Histogram of the duration, time from first to last pen stroke for all students for homework three problem two for the 2011 course offering.

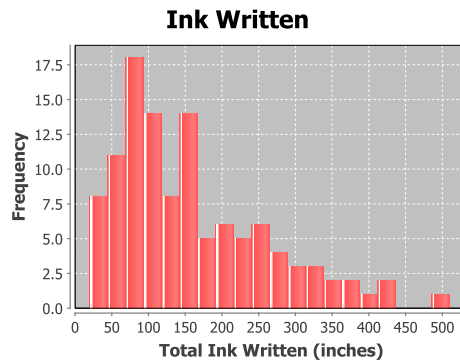


Figure .250: Histogram of the amount of ink, in inches, written, for each student on homework three problem two for the 2011 course offering.

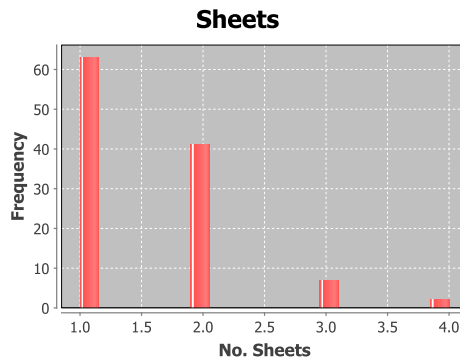


Figure .251: Histogram of the number of pages written on by each student to solve homework three problem two for the 2011 course offering.

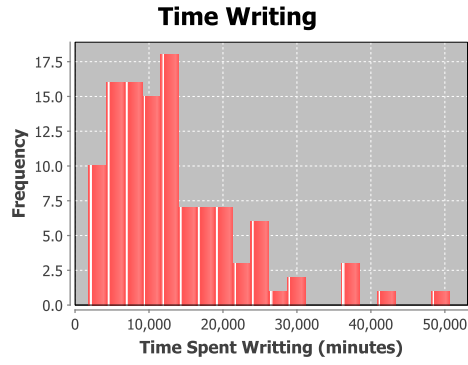


Figure .252: Histogram of the amount of time spent writing by each student on homework three problem two for the 2011 course offering.

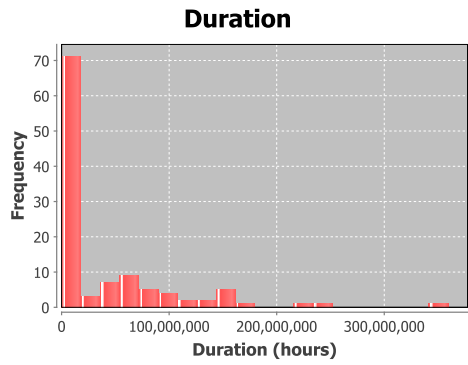


Figure .253: Histogram of the duration, time from first to last pen stroke for all students for homework three problem three for the 2011 course offering.

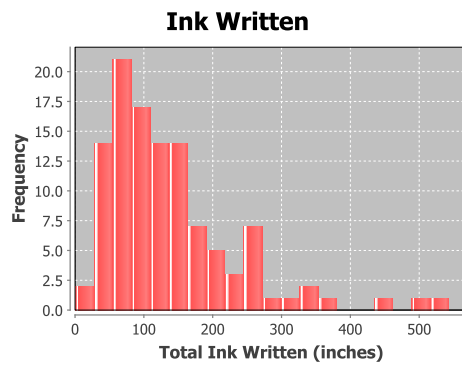


Figure .254: Histogram of the amount of ink, in inches, written, for each student on homework three problem three for the 2011 course offering.

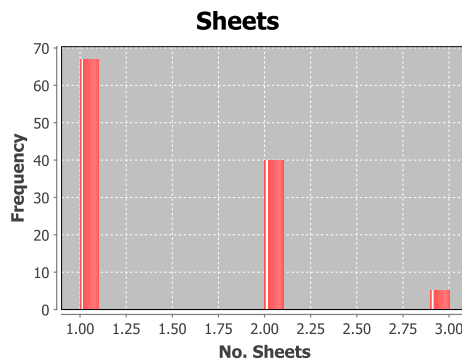


Figure .255: Histogram of the number of pages written on by each student to solve homework three problem three for the 2011 course offering.

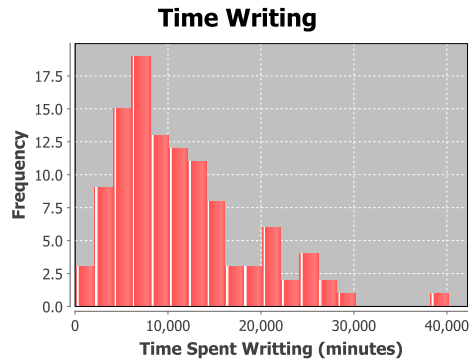


Figure .256: Histogram of the amount of time spent writing by each student on homework three problem three for the 2011 course offering.

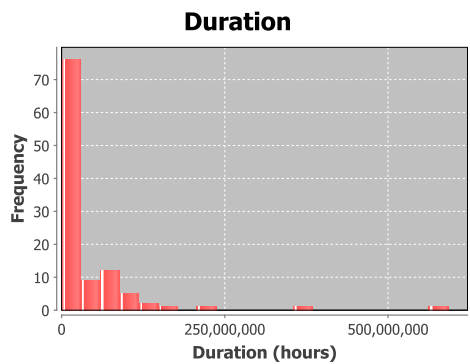


Figure .257: Histogram of the duration, time from first to last pen stroke for all students for homework three problem four for the 2011 course offering.

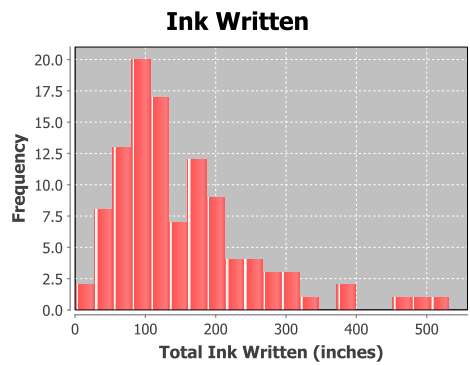


Figure .258: Histogram of the amount of ink, in inches, written, for each student on homework three problem four for the 2011 course offering.

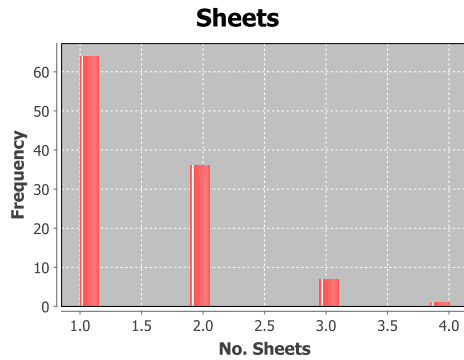


Figure .259: Histogram of the number of pages written on by each student to solve homework three problem four for the 2011 course offering.

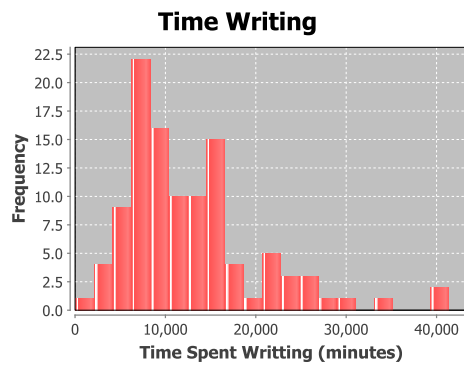


Figure .260: Histogram of the amount of time spent writing by each student on homework three problem four for the 2011 course offering.

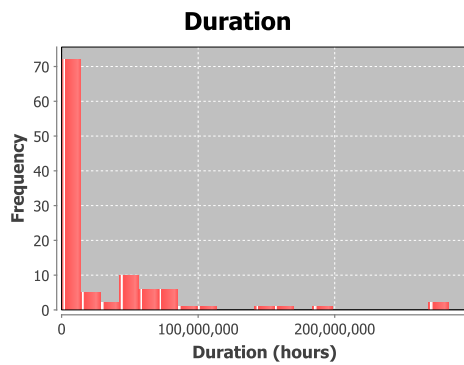


Figure .261: Histogram of the duration, time from first to last pen stroke for all students for homework three problem five for the 2011 course offering.

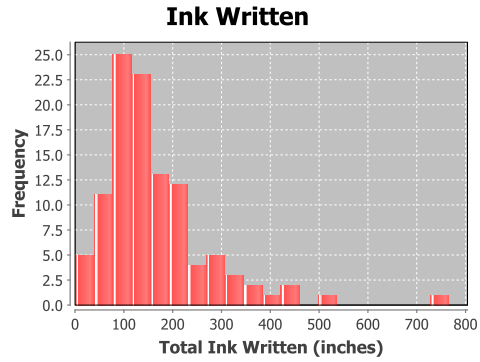


Figure .262: Histogram of the amount of ink, in inches, written, for each student on homework three problem five for the 2011 course offering.

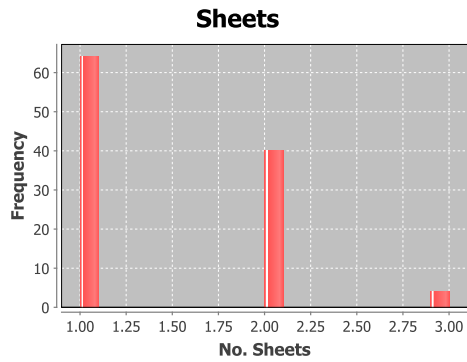


Figure .263: Histogram of the number of pages written on by each student to solve homework three problem five for the 2011 course offering.

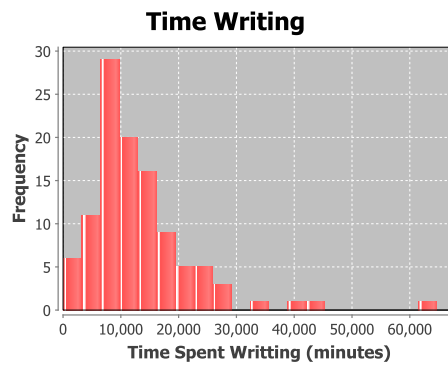


Figure .264: Histogram of the amount of time spent writing by each student on homework three problem five for the 2011 course offering.

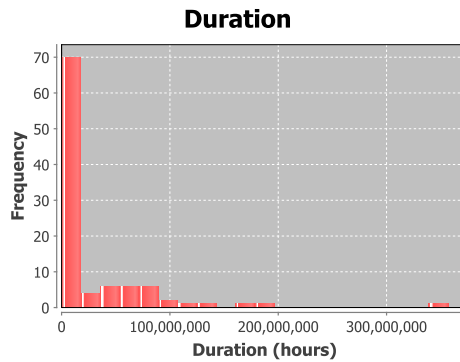


Figure .265: Histogram of the duration, time from first to last pen stroke for all students for homework three problem six for the 2011 course offering.

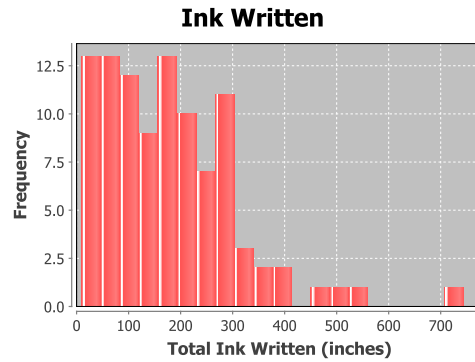


Figure .266: Histogram of the amount of ink, in inches, written, for each student on homework three problem six for the 2011 course offering.

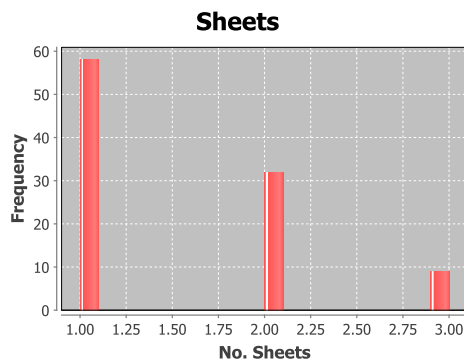


Figure .267: Histogram of the number of pages written on by each student to solve homework three problem six for the 2011 course offering.

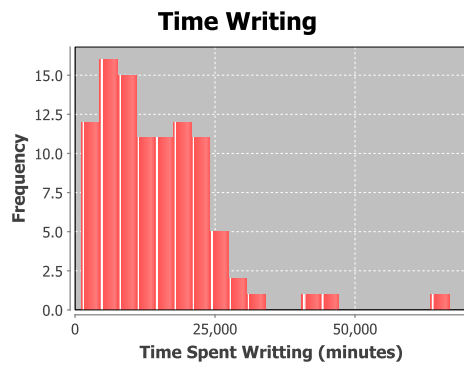


Figure .268: Histogram of the amount of time spent writing by each student on homework three problem six for the 2011 course offering.

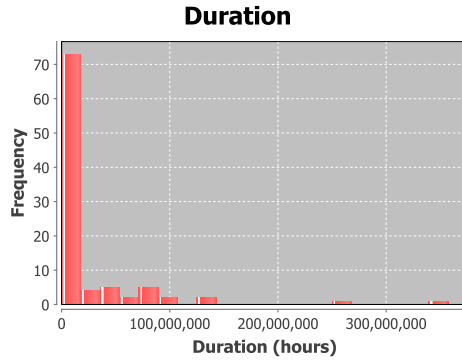


Figure .269: Histogram of the duration, time from first to last pen stroke for all students for homework three problem seven for the 2011 course offering.

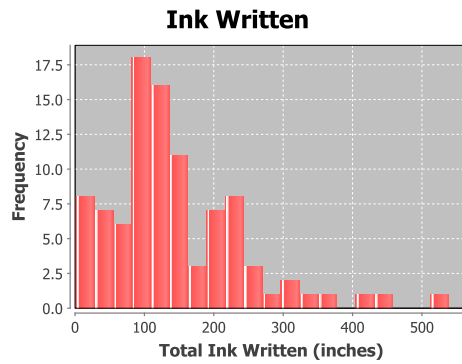


Figure .270: Histogram of the amount of ink, in inches, written, for each student on homework three problem seven for the 2011 course offering.

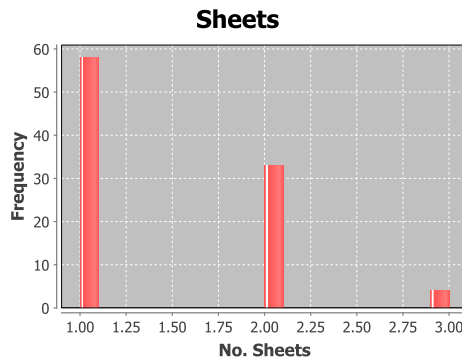


Figure .271: Histogram of the number of pages written on by each student to solve homework three problem seven for the 2011 course offering.

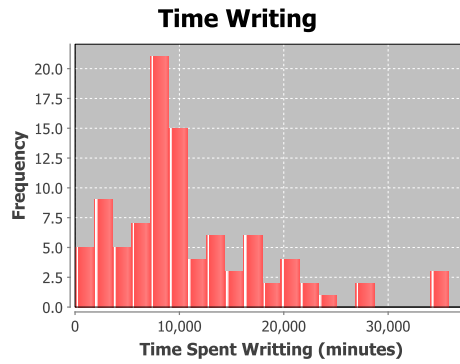


Figure .272: Histogram of the amount of time spent writing by each student on homework three problem seven for the 2011 course offering.

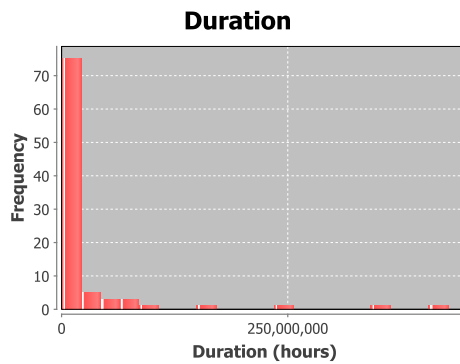


Figure .273: Histogram of the duration, time from first to last pen stroke for all students for homework three problem eight for the 2011 course offering.

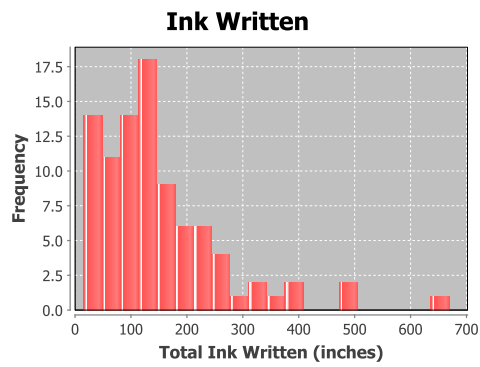


Figure .274: Histogram of the amount of ink, in inches, written, for each student on homework three problem eight for the 2011 course offering.

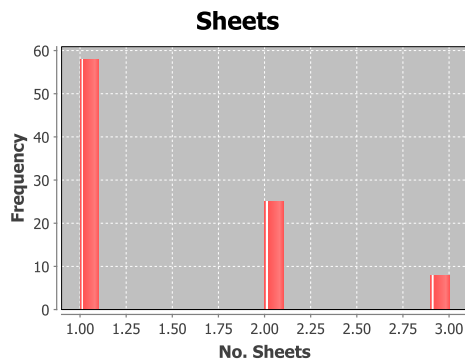


Figure .275: Histogram of the number of pages written on by each student to solve homework three problem eight for the 2011 course offering.

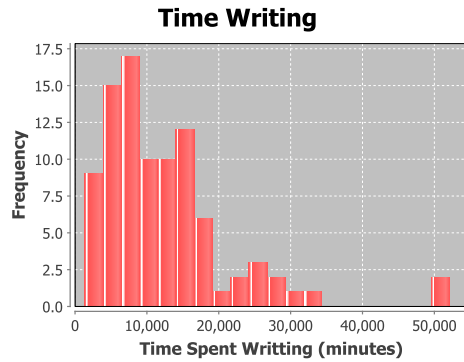


Figure .276: Histogram of the amount of time spent writing by each student on homework three problem eight for the 2011 course offering.

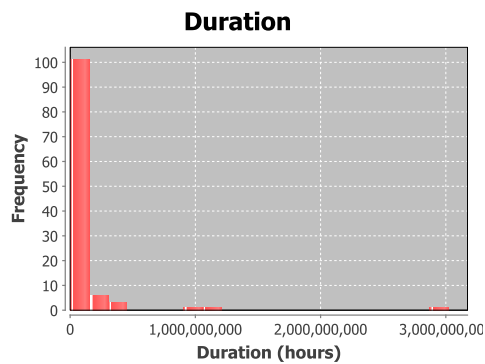


Figure .277: Histogram of the duration, time from first to last pen stroke for all students for homework four problem one for the 2011 course offering.

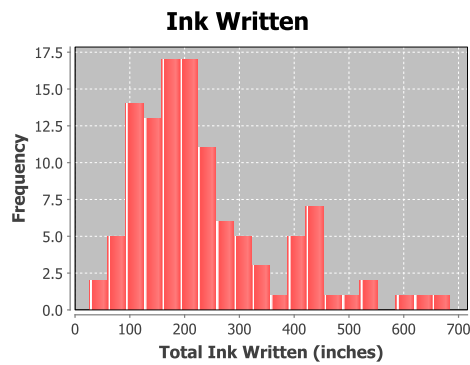


Figure .278: Histogram of the amount of ink, in inches, written, for each student on homework four problem one for the 2011 course offering.

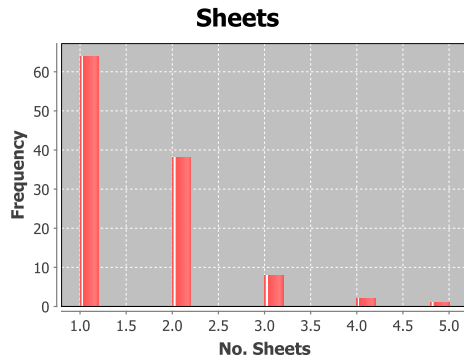


Figure .279: Histogram of the number of pages written on by each student to solve homework four problem one for the 2011 course offering.

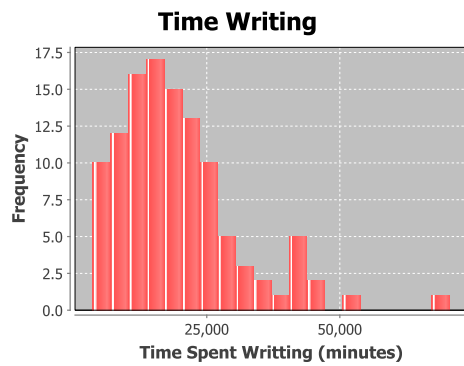


Figure .280: Histogram of the amount of time spent writing by each student on homework four problem one for the 2011 course offering.

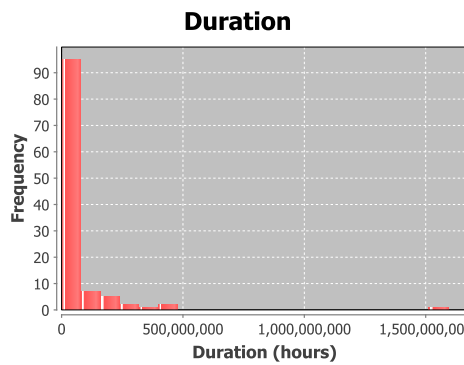


Figure .281: Histogram of the duration, time from first to last pen stroke for all students for homework four problem two for the 2011 course offering.

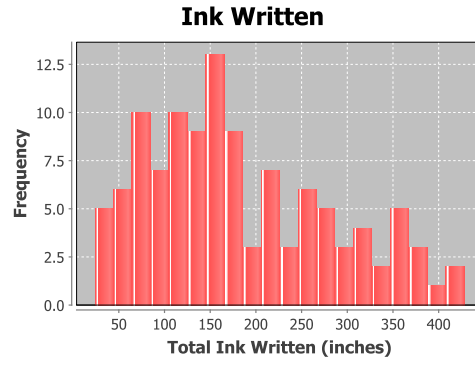


Figure .282: Histogram of the amount of ink, in inches, written, for each student on homework four problem two for the 2011 course offering.

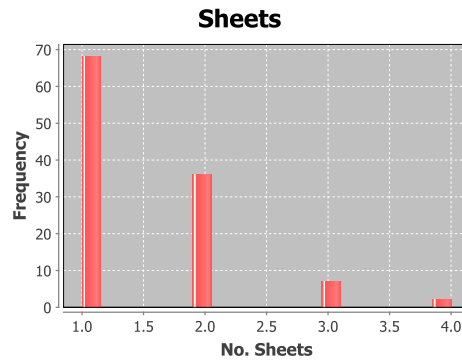


Figure .283: Histogram of the number of pages written on by each student to solve homework four problem two for the 2011 course offering.

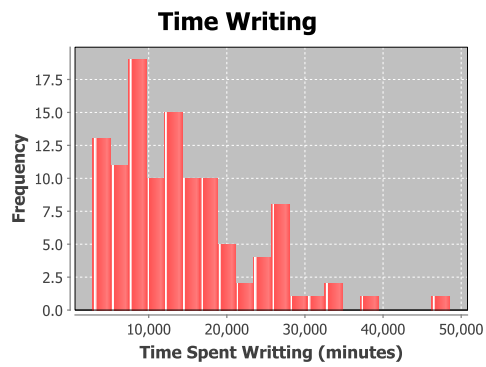


Figure .284: Histogram of the amount of time spent writing by each student on homework four problem two for the 2011 course offering.

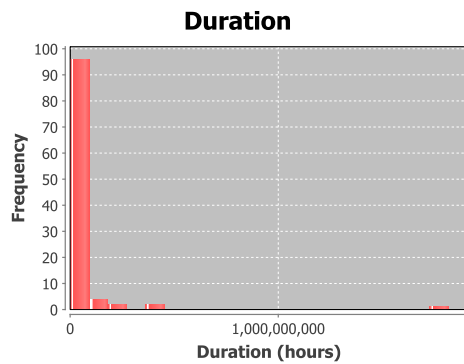


Figure .285: Histogram of the duration, time from first to last pen stroke for all students for homework four problem three for the 2011 course offering.

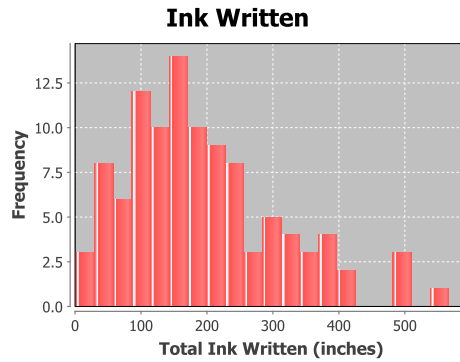


Figure .286: Histogram of the amount of ink, in inches, written, for each student on homework four problem three for the 2011 course offering.

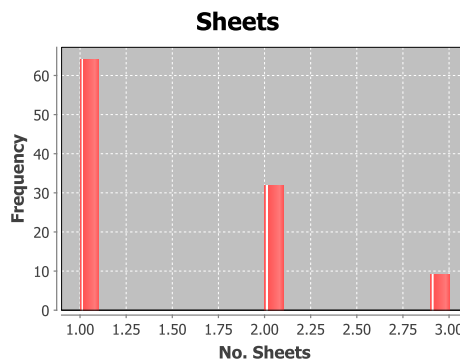


Figure .287: Histogram of the number of pages written on by each student to solve homework four problem three for the 2011 course offering.

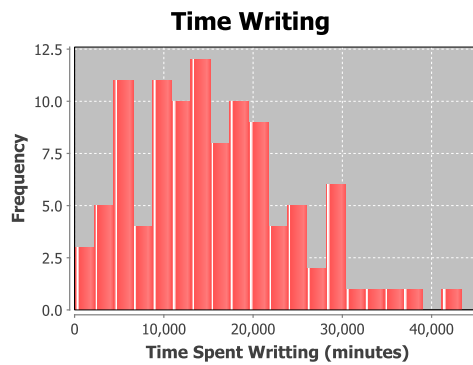


Figure .288: Histogram of the amount of time spent writing by each student on homework four problem three for the 2011 course offering.

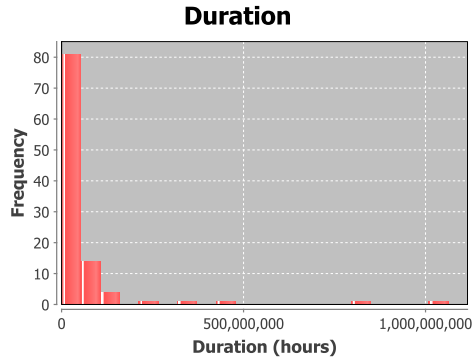


Figure .289: Histogram of the duration, time from first to last pen stroke for all students for homework four problem four for the 2011 course offering.

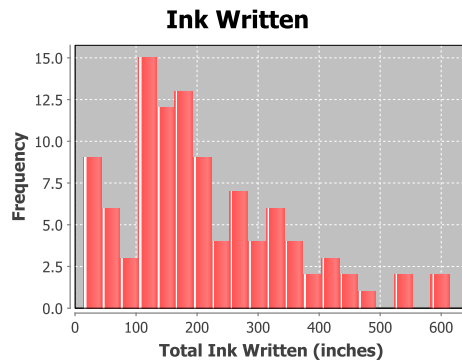


Figure .290: Histogram of the amount of ink, in inches, written, for each student on homework four problem four for the 2011 course offering.

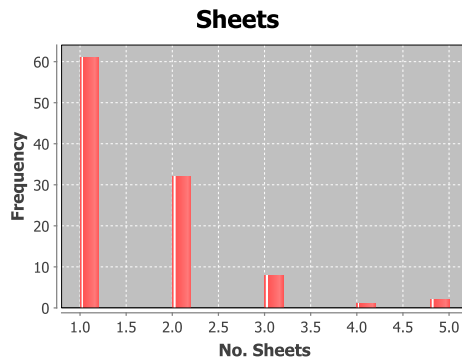


Figure .291: Histogram of the number of pages written on by each student to solve homework four problem four for the 2011 course offering.

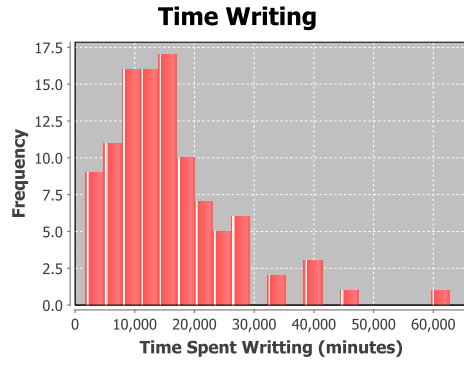


Figure .292: Histogram of the amount of time spent writing by each student on homework four problem four for the 2011 course offering.

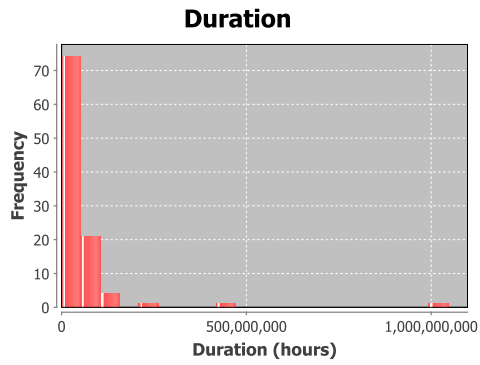


Figure .293: Histogram of the duration, time from first to last pen stroke for all students for homework four problem five for the 2011 course offering.

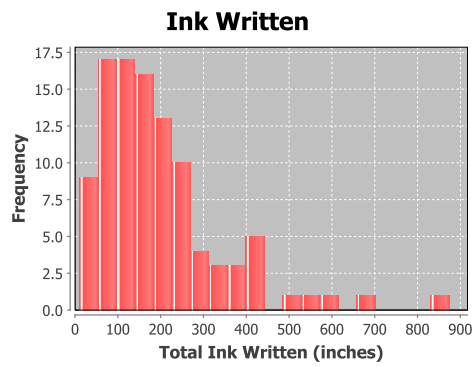


Figure .294: Histogram of the amount of ink, in inches, written, for each student on homework four problem five for the 2011 course offering.

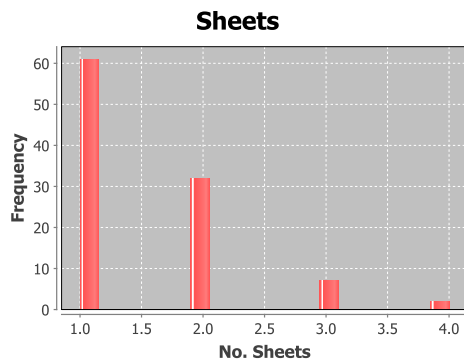


Figure .295: Histogram of the number of pages written on by each student to solve homework four problem five for the 2011 course offering.

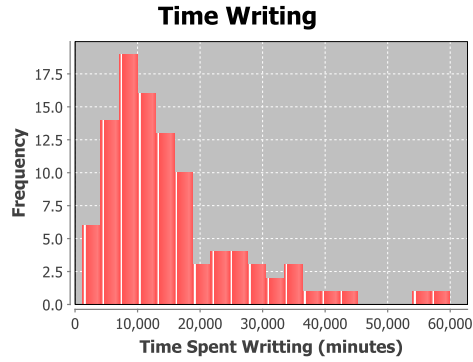


Figure .296: Histogram of the amount of time spent writing by each student on homework four problem five for the 2011 course offering.

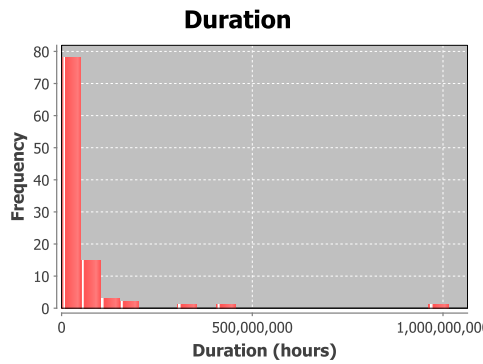


Figure .297: Histogram of the duration, time from first to last pen stroke for all students for homework four problem six for the 2011 course offering.

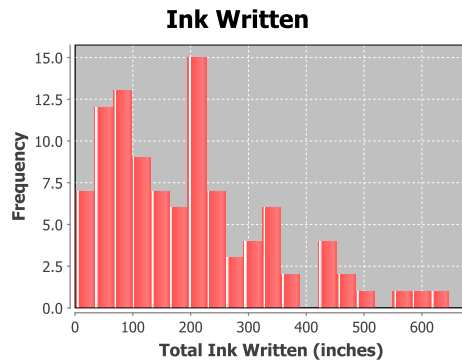


Figure .298: Histogram of the amount of ink, in inches, written, for each student on homework four problem six for the 2011 course offering.

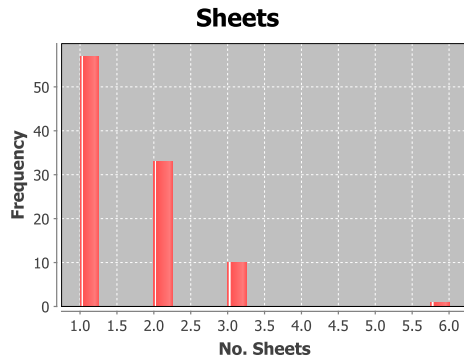


Figure .299: Histogram of the number of pages written on by each student to solve homework four problem six for the 2011 course offering.

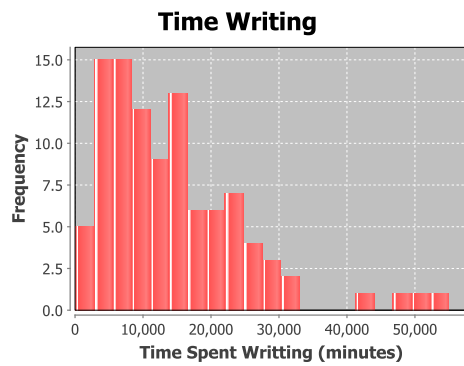


Figure .300: Histogram of the amount of time spent writing by each student on homework four problem six for the 2011 course offering.

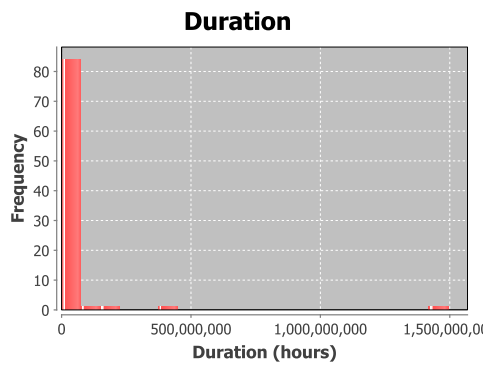


Figure .301: Histogram of the duration, time from first to last pen stroke for all students for homework four problem seven for the 2011 course offering.

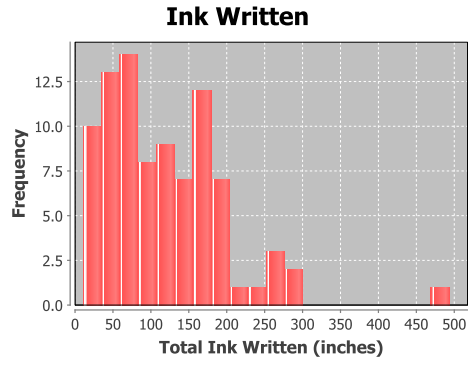


Figure .302: Histogram of the amount of ink, in inches, written, for each student on homework four problem seven for the 2011 course offering.

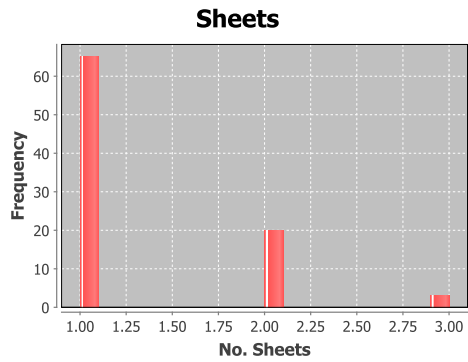


Figure .303: Histogram of the number of pages written on by each student to solve homework four problem seven for the 2011 course offering.

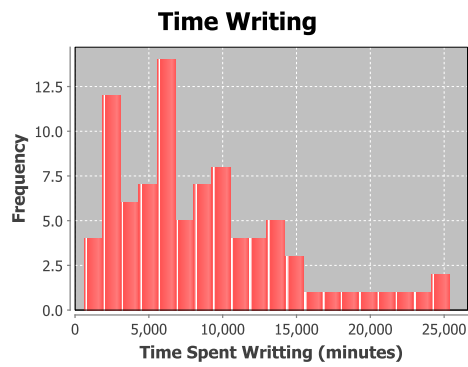


Figure .304: Histogram of the amount of time spent writing by each student on homework four problem seven for the 2011 course offering.

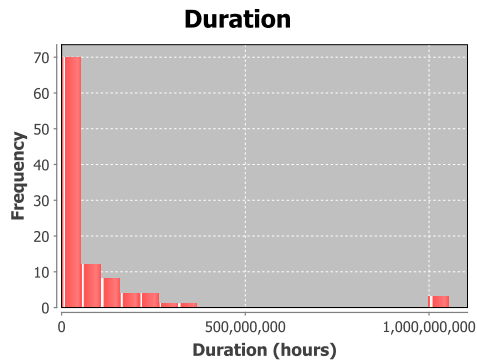


Figure .305: Histogram of the duration, time from first to last pen stroke for all students for homework five problem one for the 2011 course offering.

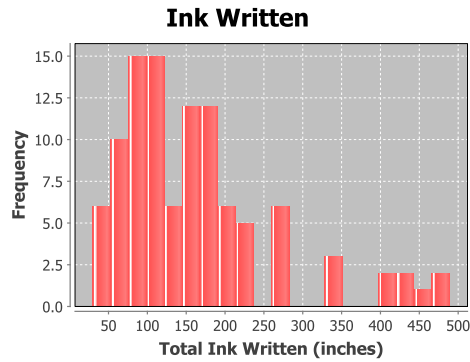


Figure .306: Histogram of the amount of ink, in inches, written, for each student on homework five problem one for the 2011 course offering.

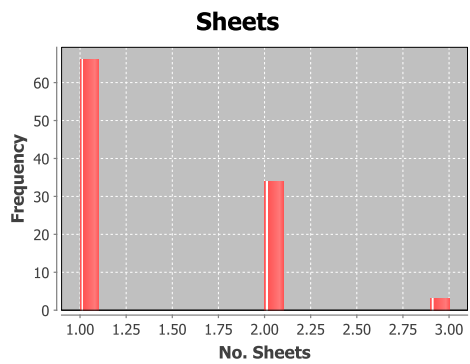


Figure .307: Histogram of the number of pages written on by each student to solve homework five problem one for the 2011 course offering.

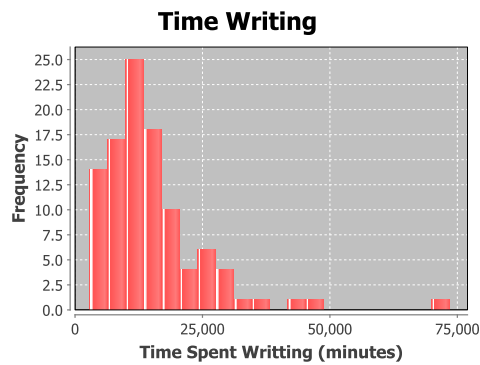


Figure .308: Histogram of the amount of time spent writing by each student on homework five problem one for the 2011 course offering.

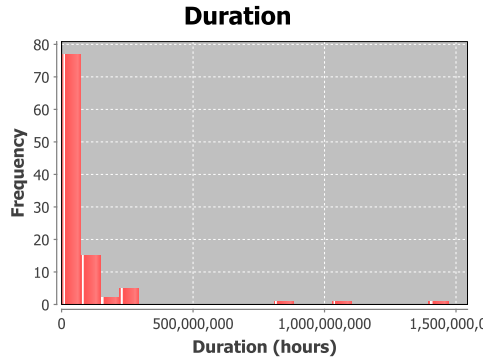


Figure .309: Histogram of the duration, time from first to last pen stroke for all students for homework five problem two for the 2011 course offering.

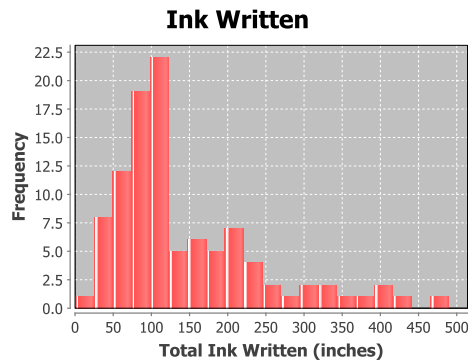


Figure .310: Histogram of the amount of ink, in inches, written, for each student on homework five problem two for the 2011 course offering.

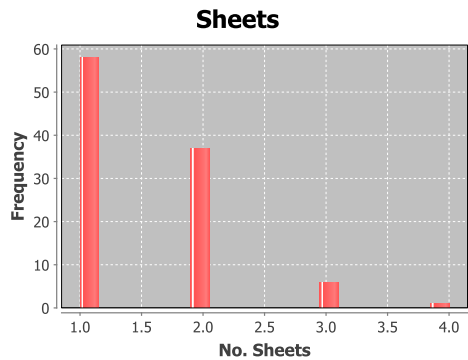


Figure .311: Histogram of the number of pages written on by each student to solve homework five problem two for the 2011 course offering.

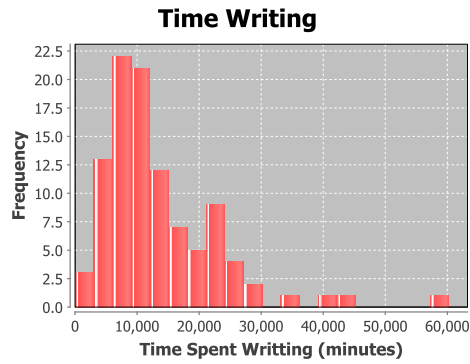


Figure .312: Histogram of the amount of time spent writing by each student on homework five problem two for the 2011 course offering.

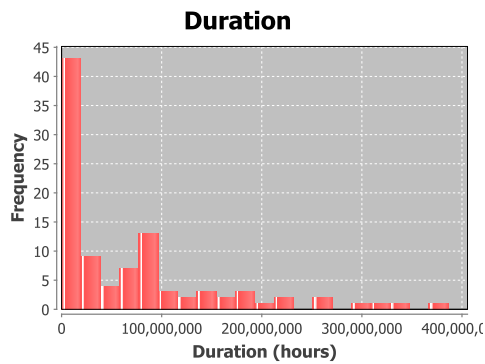


Figure .313: Histogram of the duration, time from first to last pen stroke for all students for homework five problem three for the 2011 course offering.

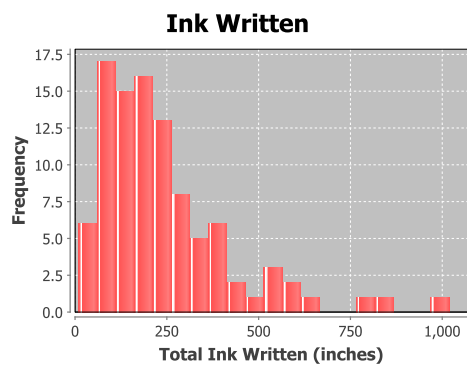


Figure .314: Histogram of the amount of ink, in inches, written, for each student on homework five problem three for the 2011 course offering.

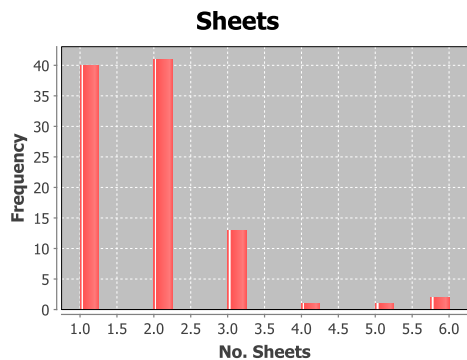


Figure .315: Histogram of the number of pages written on by each student to solve homework five problem three for the 2011 course offering.

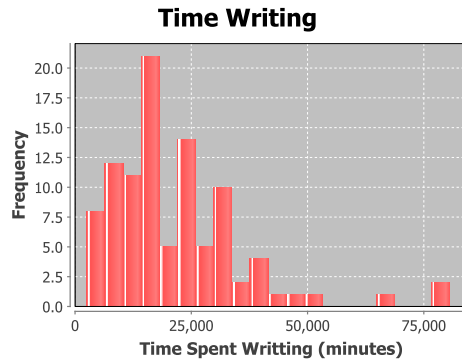


Figure .316: Histogram of the amount of time spent writing by each student on homework five problem three for the 2011 course offering.

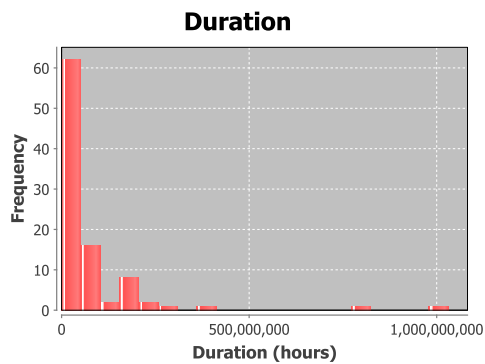


Figure .317: Histogram of the duration, time from first to last pen stroke for all students for homework five problem four for the 2011 course offering.

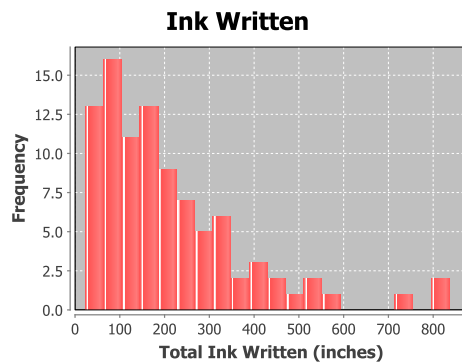


Figure .318: Histogram of the amount of ink, in inches, written, for each student on homework five problem four for the 2011 course offering.

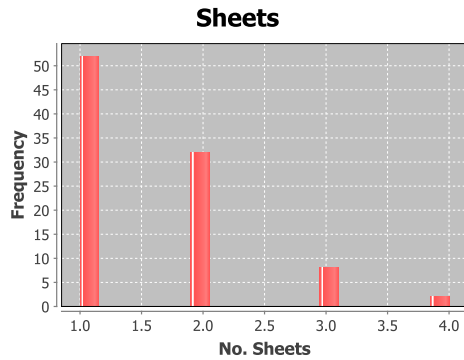


Figure .319: Histogram of the number of pages written on by each student to solve homework five problem four for the 2011 course offering.

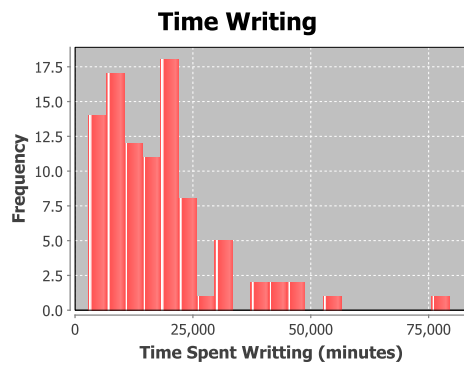


Figure .320: Histogram of the amount of time spent writing by each student on homework five problem four for the 2011 course offering.

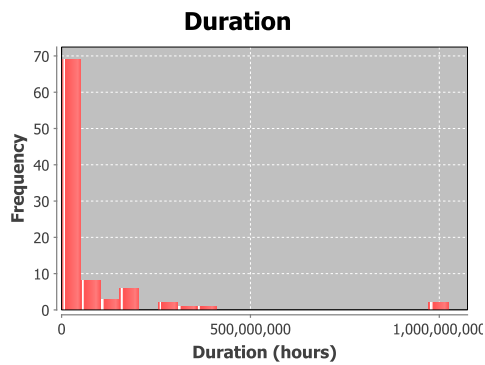


Figure .321: Histogram of the duration, time from first to last pen stroke for all students for homework five problem five for the 2011 course offering.

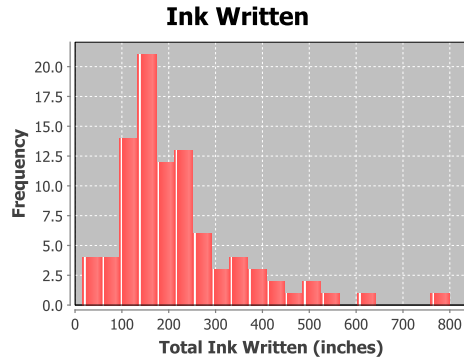


Figure .322: Histogram of the amount of ink, in inches, written, for each student on homework five problem five for the 2011 course offering.

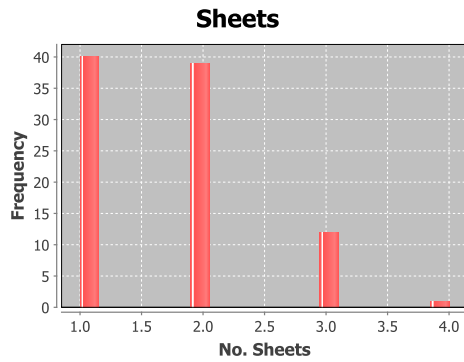


Figure .323: Histogram of the number of pages written on by each student to solve homework five problem five for the 2011 course offering.

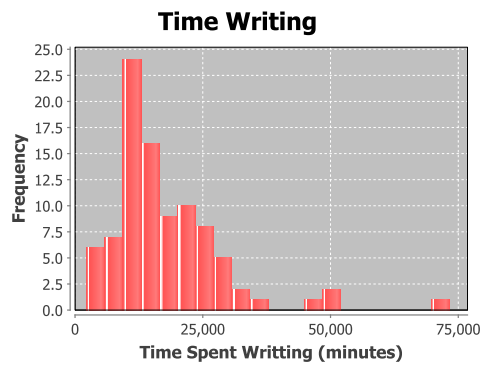


Figure .324: Histogram of the amount of time spent writing by each student on homework five problem five for the 2011 course offering.

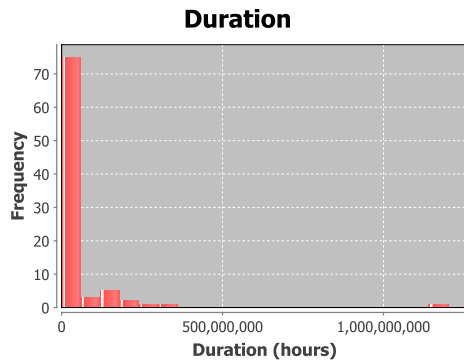


Figure .325: Histogram of the duration, time from first to last pen stroke for all students for homework five problem six for the 2011 course offering.

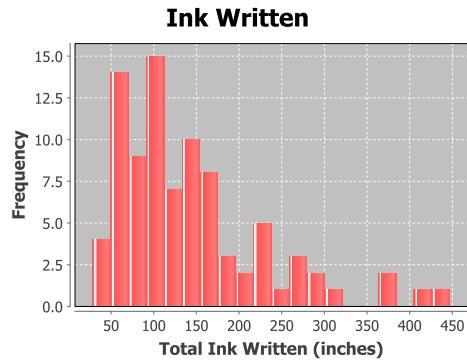


Figure .326: Histogram of the amount of ink, in inches, written, for each student on homework five problem six for the 2011 course offering.

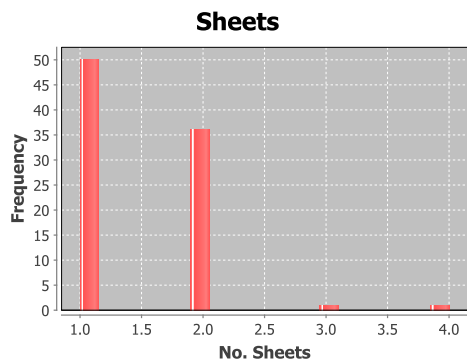


Figure .327: Histogram of the number of pages written on by each student to solve homework five problem six for the 2011 course offering.

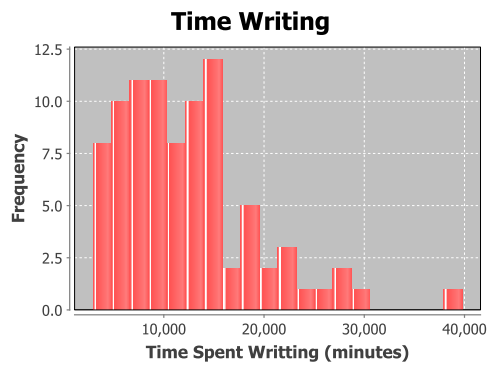


Figure .328: Histogram of the amount of time spent writing by each student on homework five problem six for the 2011 course offering.

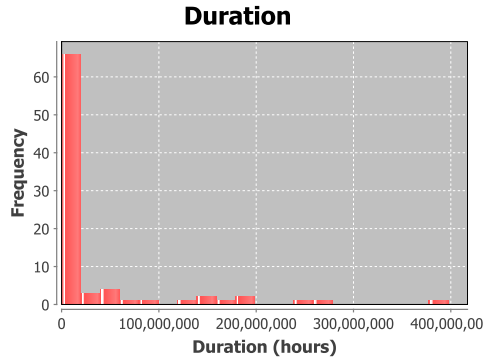


Figure .329: Histogram of the duration, time from first to last pen stroke for all students for homework five problem seven for the 2011 course offering.

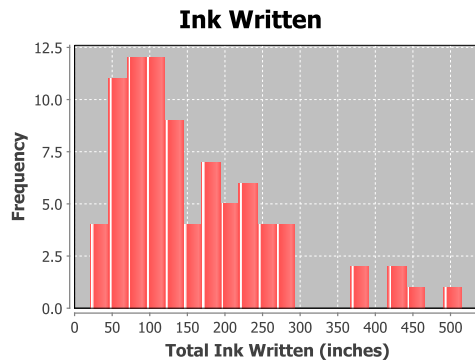


Figure .330: Histogram of the amount of ink, in inches, written, for each student on homework five problem seven for the 2011 course offering.

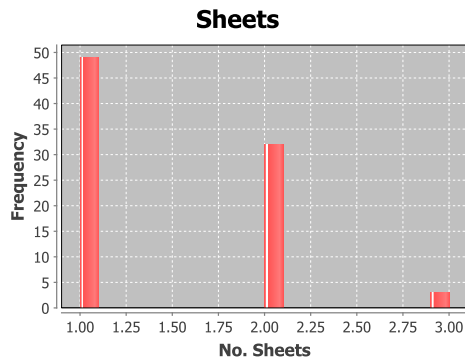


Figure .331: Histogram of the number of pages written on by each student to solve homework five problem seven for the 2011 course offering.

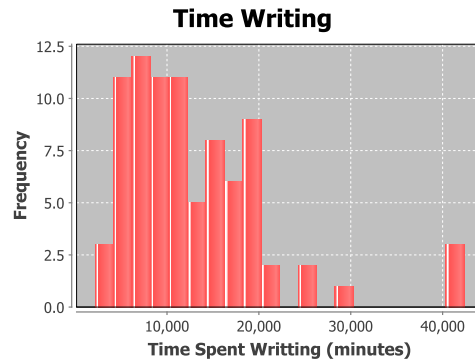


Figure .332: Histogram of the amount of time spent writing by each student on homework five problem seven for the 2011 course offering.

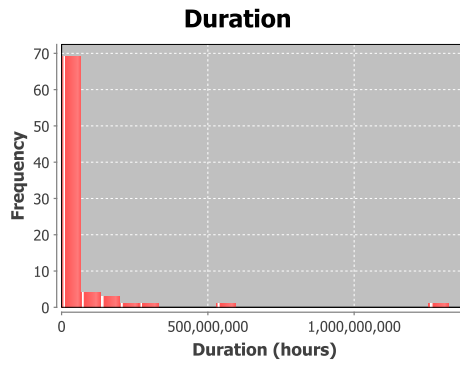


Figure .333: Histogram of the duration, time from first to last pen stroke for all students for homework five problem eight for the 2011 course offering.

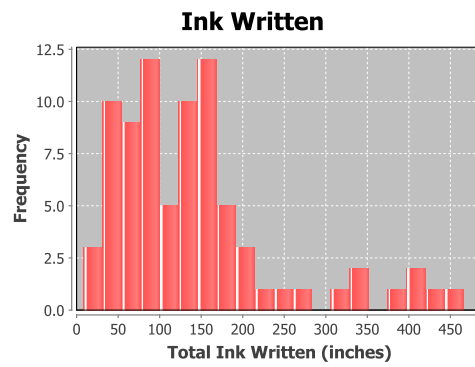


Figure .334: Histogram of the amount of ink, in inches, written, for each student on homework five problem eight for the 2011 course offering.

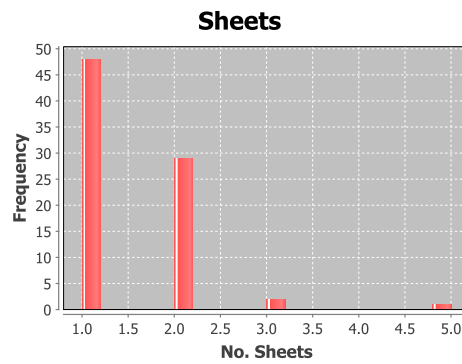


Figure .335: Histogram of the number of pages written on by each student to solve homework five problem eight for the 2011 course offering.

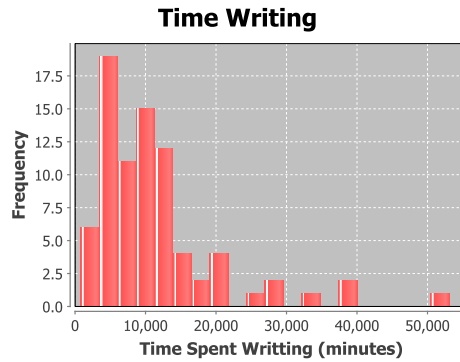


Figure .336: Histogram of the amount of time spent writing by each student on homework five problem eight for the 2011 course offering.

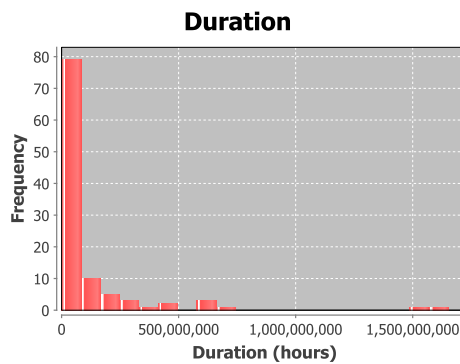


Figure .337: Histogram of the duration, time from first to last pen stroke for all students for homework six problem one for the 2011 course offering.

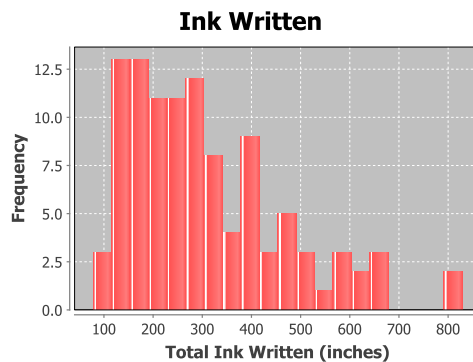


Figure .338: Histogram of the amount of ink, in inches, written, for each student on homework six problem one for the 2011 course offering.

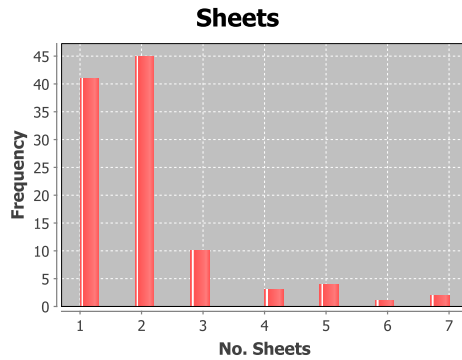


Figure .339: Histogram of the number of pages written on by each student to solve homework six problem one for the 2011 course offering.

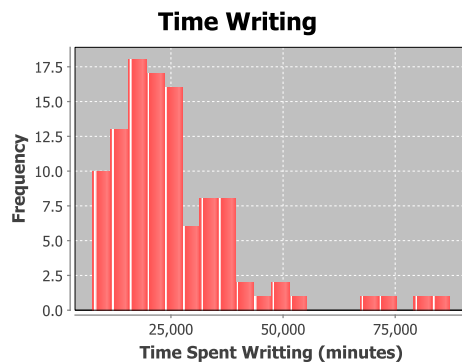


Figure .340: Histogram of the amount of time spent writing by each student on homework six problem one for the 2011 course offering.

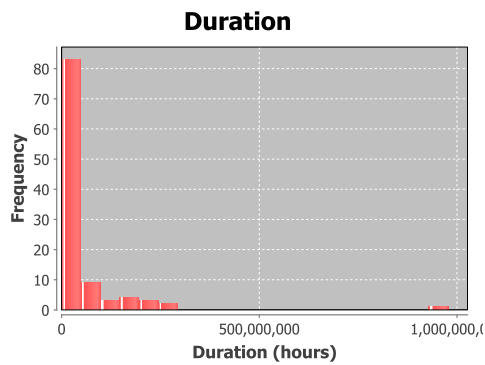


Figure .341: Histogram of the duration, time from first to last pen stroke for all students for homework six problem two for the 2011 course offering.

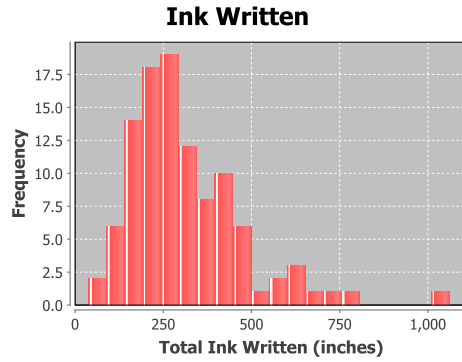


Figure .342: Histogram of the amount of ink, in inches, written, for each student on homework six problem two for the 2011 course offering.

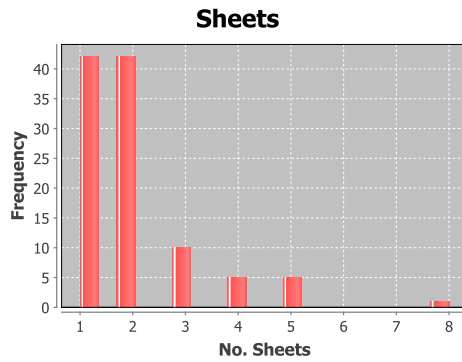


Figure .343: Histogram of the number of pages written on by each student to solve homework six problem two for the 2011 course offering.

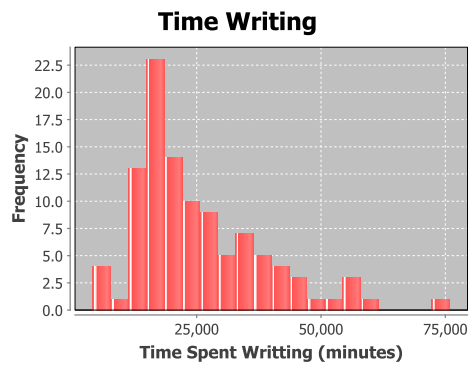


Figure .344: Histogram of the amount of time spent writing by each student on homework six problem two for the 2011 course offering.

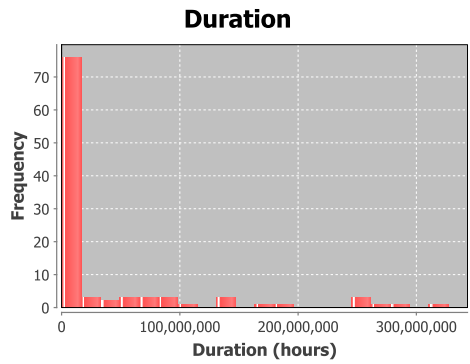


Figure .345: Histogram of the duration, time from first to last pen stroke for all students for homework six problem three for the 2011 course offering.

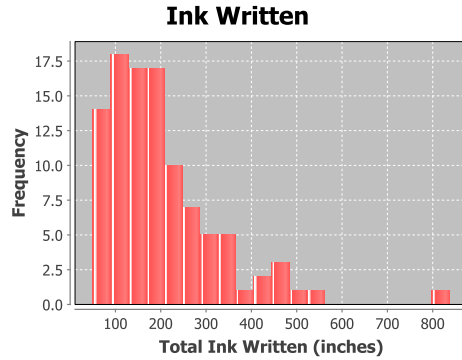


Figure .346: Histogram of the amount of ink, in inches, written, for each student on homework six problem three for the 2011 course offering.

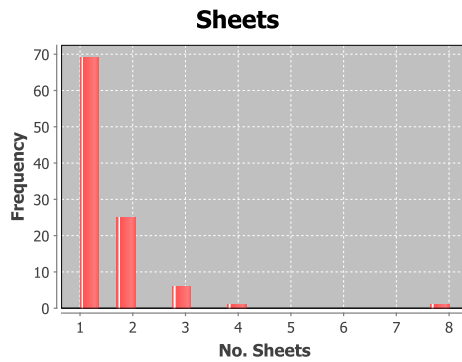


Figure .347: Histogram of the number of pages written on by each student to solve homework six problem three for the 2011 course offering.

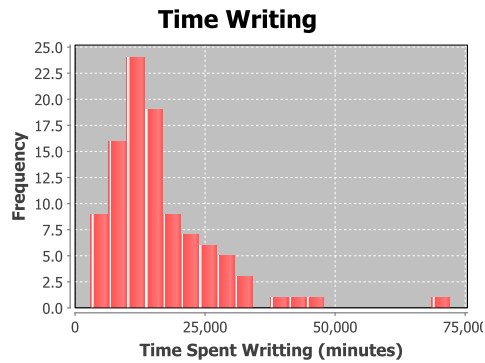


Figure .348: Histogram of the amount of time spent writing by each student on homework six problem three for the 2011 course offering.

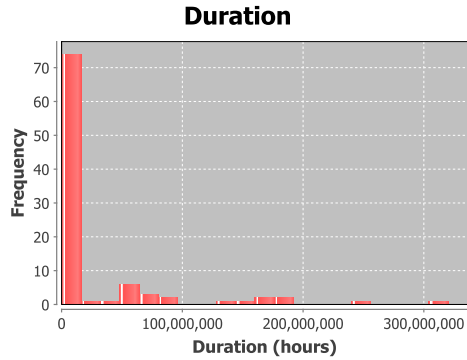


Figure .349: Histogram of the duration, time from first to last pen stroke for all students for homework six problem four for the 2011 course offering.

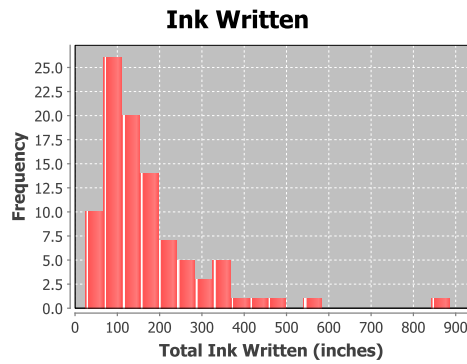


Figure .350: Histogram of the amount of ink, in inches, written, for each student on homework six problem four for the 2011 course offering.

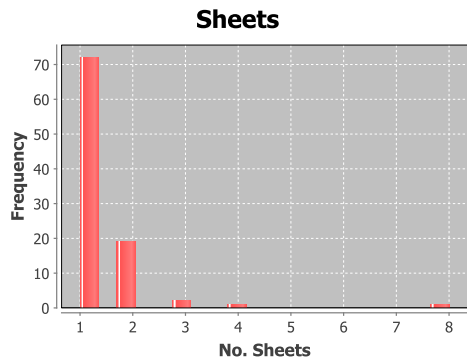


Figure .351: Histogram of the number of pages written on by each student to solve homework six problem four for the 2011 course offering.

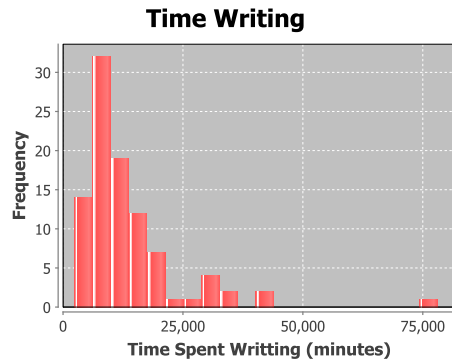


Figure .352: Histogram of the amount of time spent writing by each student on homework six problem four for the 2011 course offering.

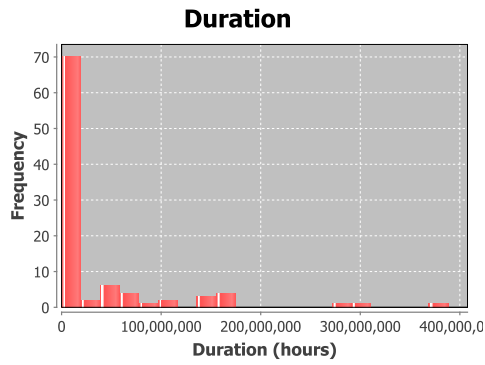


Figure .353: Histogram of the duration, time from first to last pen stroke for all students for homework six problem five for the 2011 course offering.

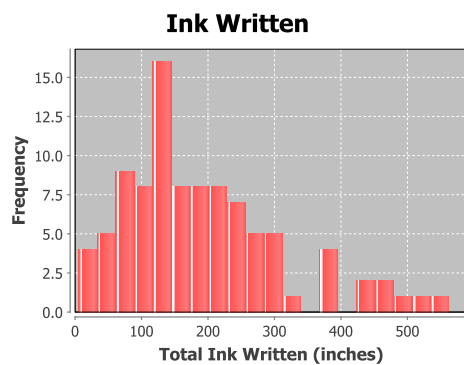


Figure .354: Histogram of the amount of ink, in inches, written, for each student on homework six problem five for the 2011 course offering.

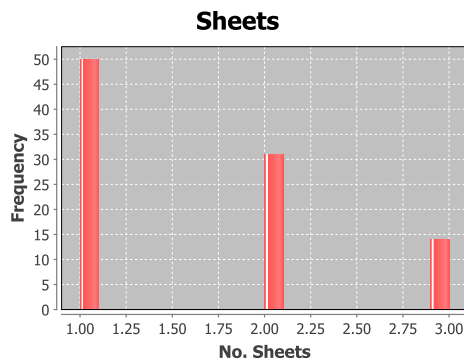


Figure .355: Histogram of the number of pages written on by each student to solve homework six problem five for the 2011 course offering.

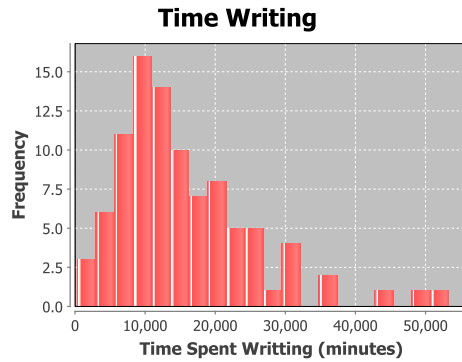


Figure .356: Histogram of the amount of time spent writing by each student on homework six problem five for the 2011 course offering.

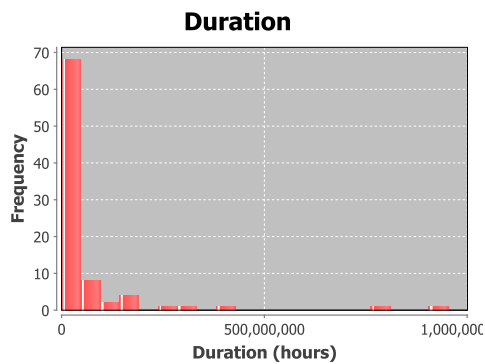


Figure .357: Histogram of the duration, time from first to last pen stroke for all students for homework six problem six for the 2011 course offering.

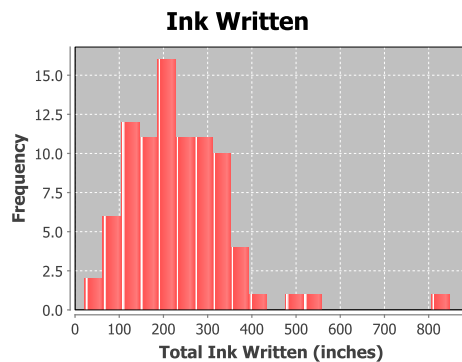


Figure .358: Histogram of the amount of ink, in inches, written, for each student on homework six problem six for the 2011 course offering.

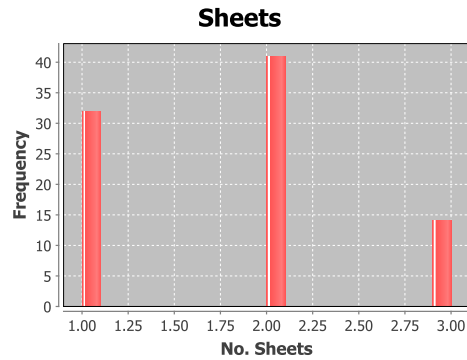


Figure .359: Histogram of the number of pages written on by each student to solve homework six problem six for the 2011 course offering.

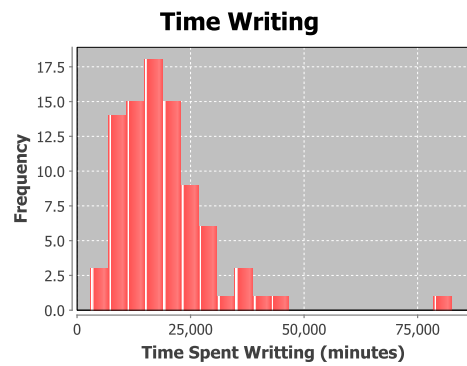


Figure .360: Histogram of the amount of time spent writing by each student on homework six problem six for the 2011 course offering.

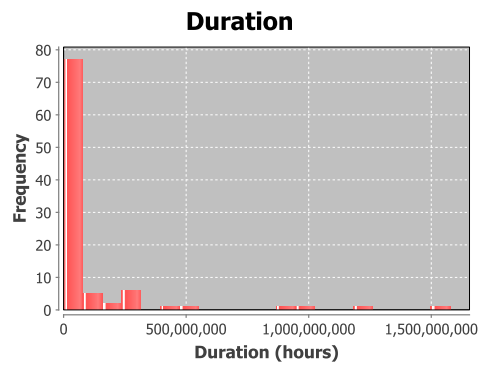


Figure .361: Histogram of the duration, time from first to last pen stroke for all students for homework seven problem one for the 2011 course offering.

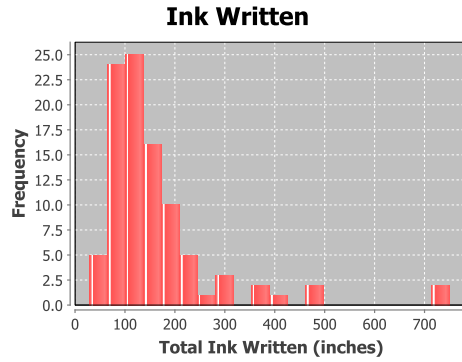


Figure .362: Histogram of the amount of ink, in inches, written, for each student on homework seven problem one for the 2011 course offering.

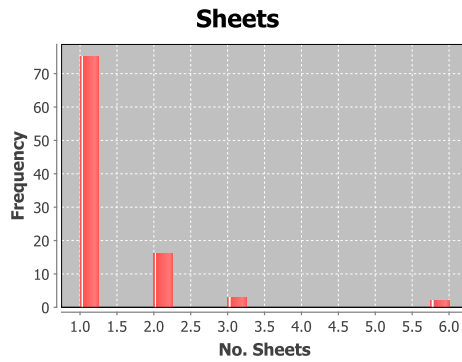


Figure .363: Histogram of the number of pages written on by each student to solve homework six problem seven for the 2011 course offering.

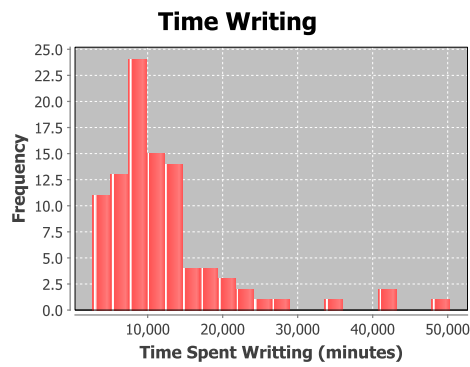


Figure .364: Histogram of the amount of time spent writing by each student on homework seven problem one for the 2011 course offering.

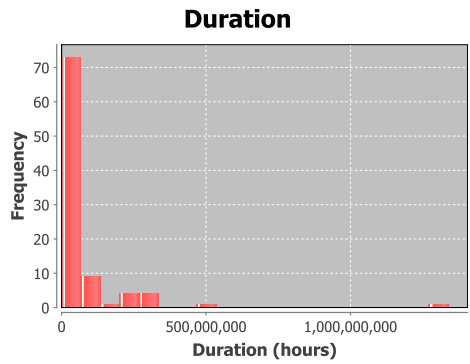


Figure .365: Histogram of the duration, time from first to last pen stroke for all students for homework seven problem two for the 2011 course offering.

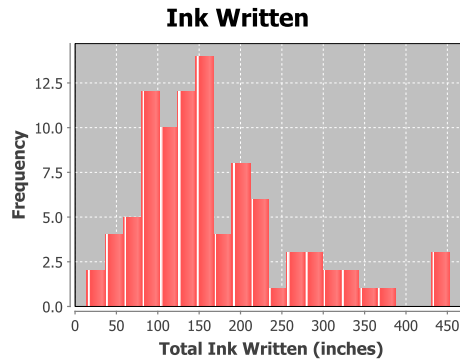


Figure .366: Histogram of the amount of ink, in inches, written, for each student on homework seven problem two for the 2011 course offering.

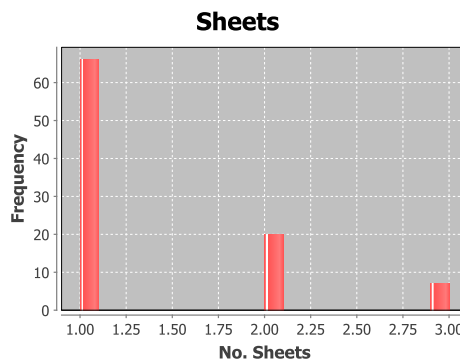


Figure .367: Histogram of the number of pages written on by each student to solve homework seven problem one for the 2011 course offering.

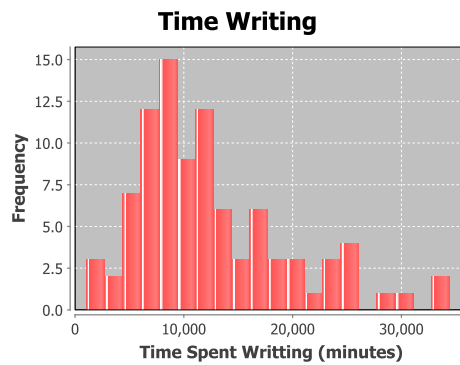


Figure .368: Histogram of the amount of time spent writing by each student on homework seven problem two for the 2011 course offering.

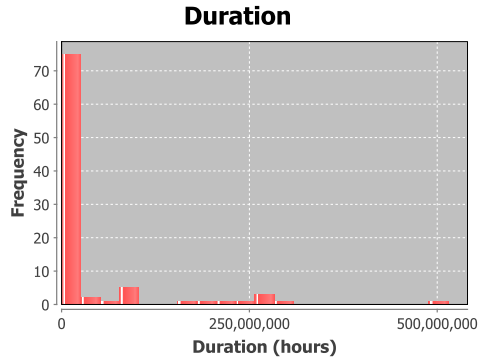


Figure .369: Histogram of the duration, time from first to last pen stroke for all students for homework seven problem three for the 2011 course offering.

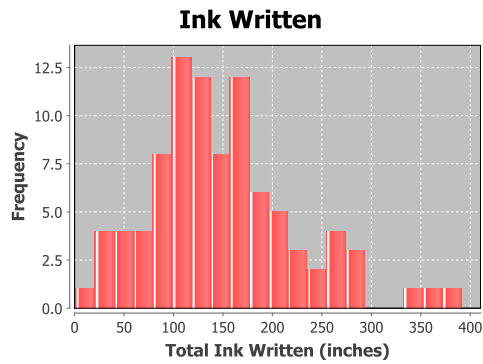


Figure .370: Histogram of the amount of ink, in inches, written, for each student on homework seven problem three for the 2011 course offering.

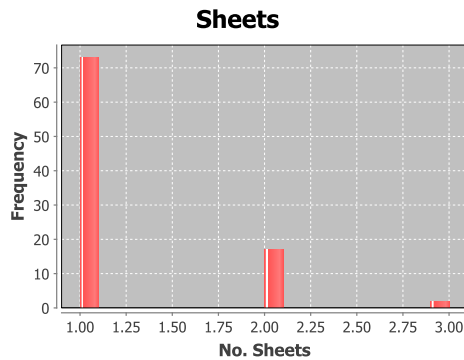


Figure .371: Histogram of the number of pages written on by each student to solve homework seven problem three for the 2011 course offering.

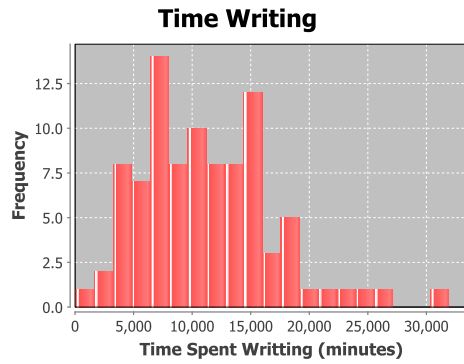


Figure .372: Histogram of the amount of time spent writing by each student on homework seven problem three for the 2011 course offering.

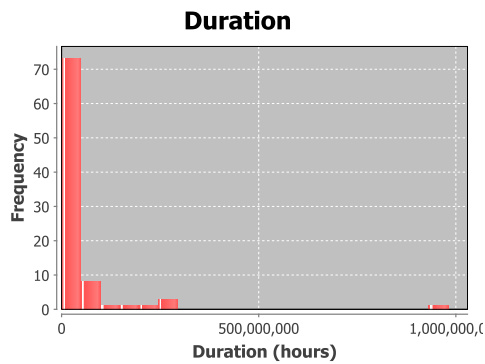


Figure .373: Histogram of the duration, time from first to last pen stroke for all students for homework seven problem four for the 2011 course offering.

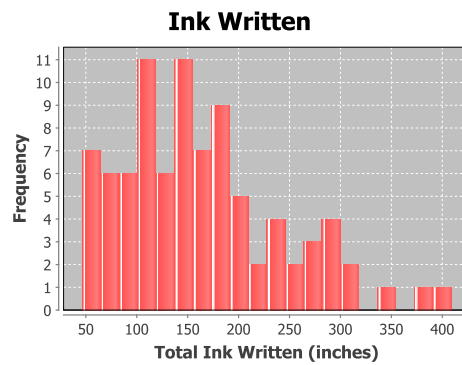


Figure .374: Histogram of the amount of ink, in inches, written, for each student on homework seven problem four for the 2011 course offering.

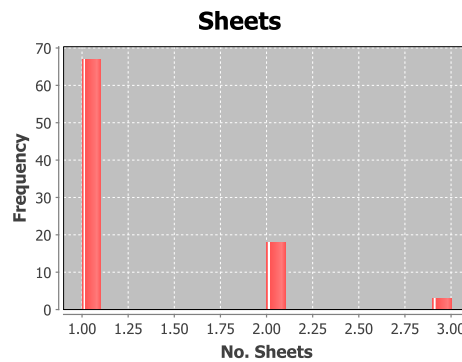


Figure .375: Histogram of the number of pages written on by each student to solve homework seven problem four for the 2011 course offering.

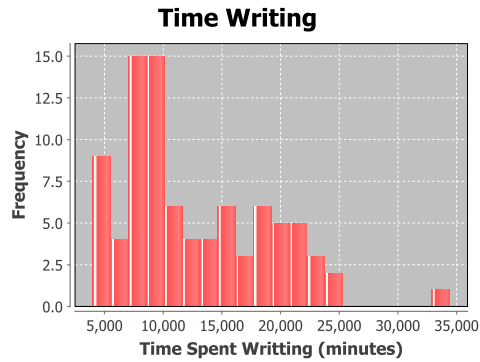


Figure .376: Histogram of the amount of time spent writing by each student on homework seven problem four for the 2011 course offering.

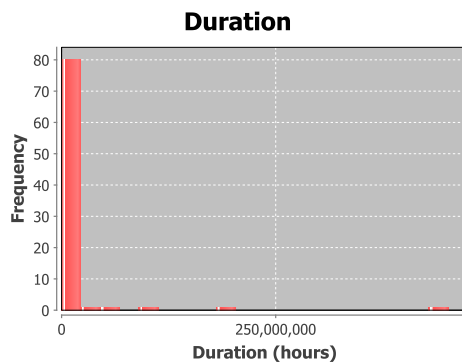


Figure .377: Histogram of the duration, time from first to last pen stroke for all students for homework seven problem five for the 2011 course offering.

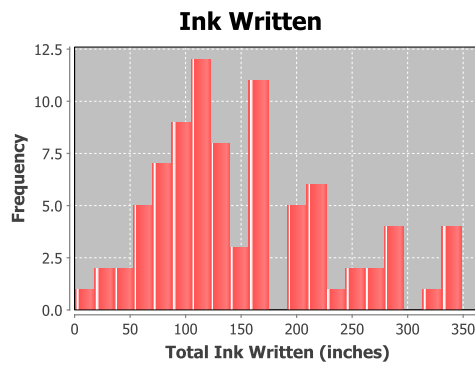


Figure .378: Histogram of the amount of ink, in inches, written, for each student on homework seven problem five for the 2011 course offering.

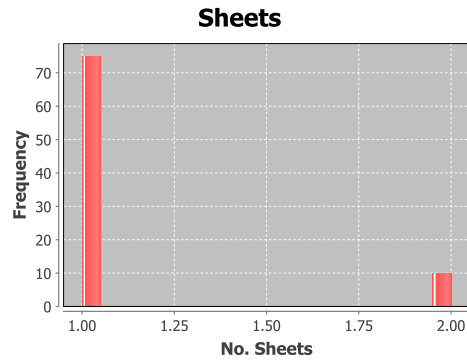


Figure .379: Histogram of the number of pages written on by each student to solve homework seven problem five for the 2011 course offering.

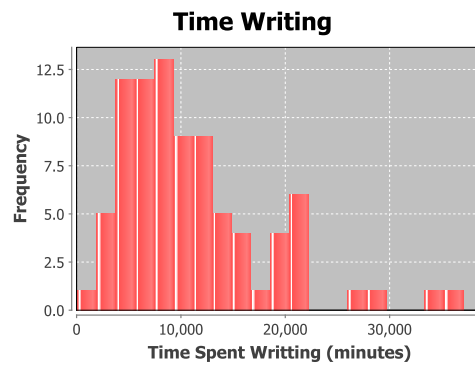


Figure .380: Histogram of the amount of time spent writing by each student on homework seven problem five for the 2011 course offering.

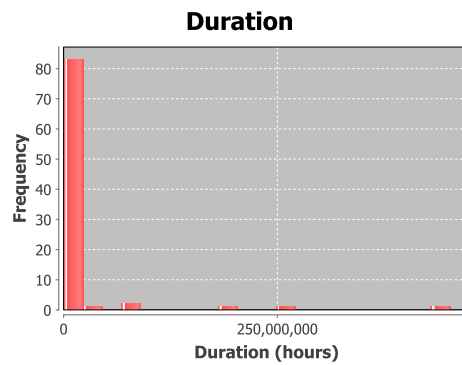


Figure .381: Histogram of the duration, time from first to last pen stroke for all students for homework seven problem six for the 2011 course offering.

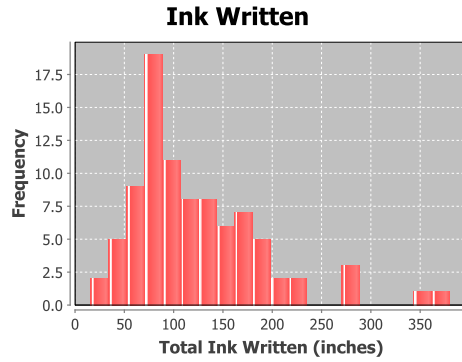


Figure .382: Histogram of the amount of ink, in inches, written, for each student on homework seven problem six for the 2011 course offering.

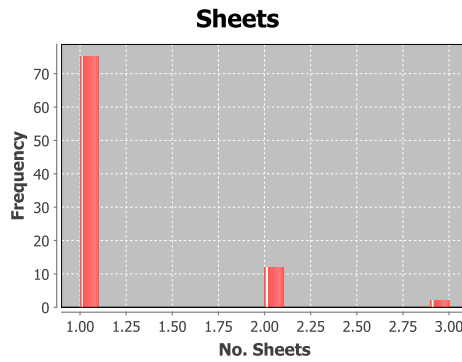


Figure .383: Histogram of the number of pages written on by each student to solve homework seven problem six for the 2011 course offering.

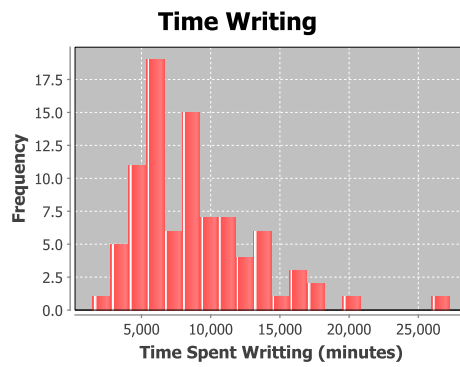


Figure .384: Histogram of the amount of time spent writing by each student on homework seven problem six for the 2011 course offering.

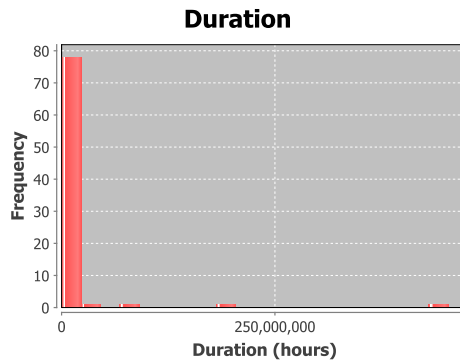


Figure .385: Histogram of the duration, time from first to last pen stroke for all students for homework seven problem seven for the 2011 course offering.

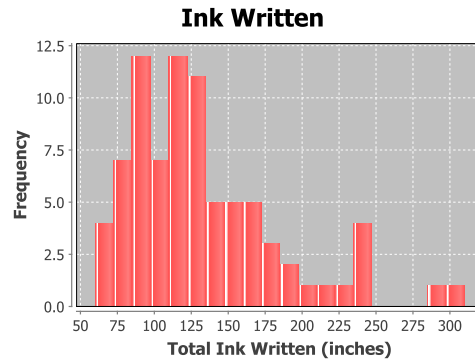


Figure .386: Histogram of the amount of ink, in inches, written, for each student on homework seven problem seven for the 2011 course offering.

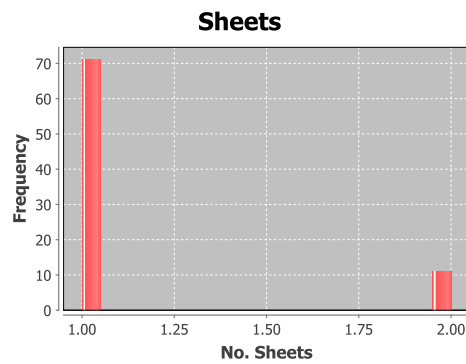


Figure .387: Histogram of the number of pages written on by each student to solve homework seven problem seven for the 2011 course offering.

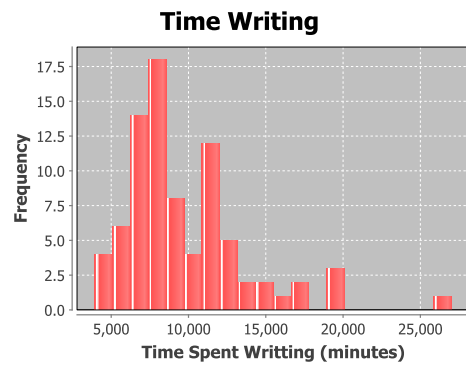


Figure .388: Histogram of the amount of time spent writing by each student on homework seven problem seven for the 2011 course offering.

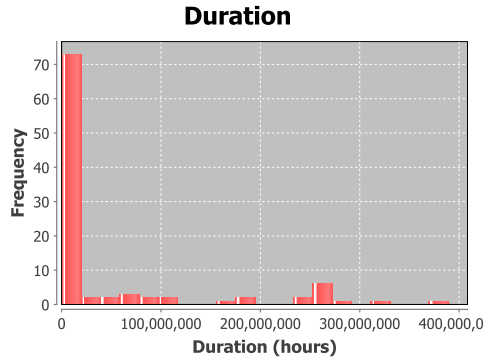


Figure .389: Histogram of the duration, time from first to last pen stroke for all students for homework eight problem one for the 2011 course offering.

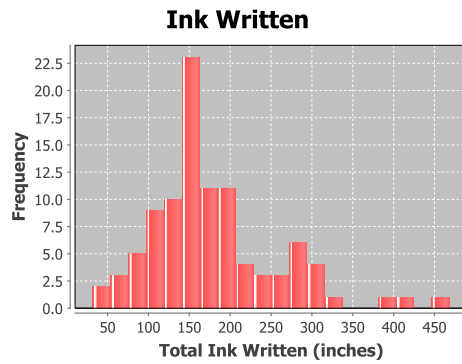


Figure .390: Histogram of the amount of ink, in inches, written, for each student on homework eight problem one for the 2011 course offering.

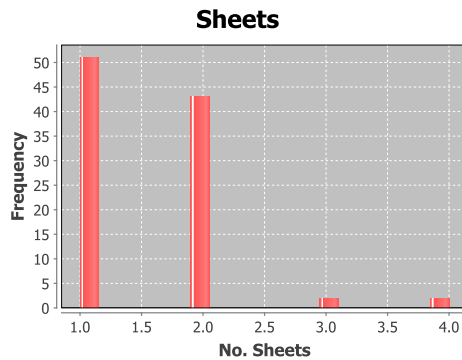


Figure .391: Histogram of the number of pages written on by each student to solve homework eight problem one for the 2011 course offering.

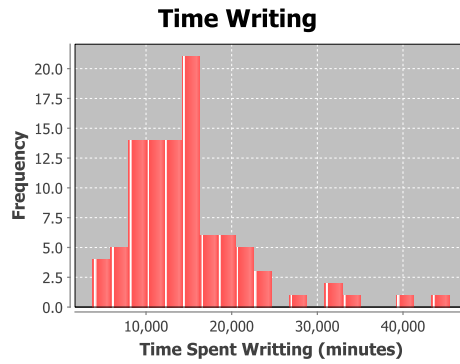


Figure .392: Histogram of the amount of time spent writing by each student on homework eight problem one for the 2011 course offering.

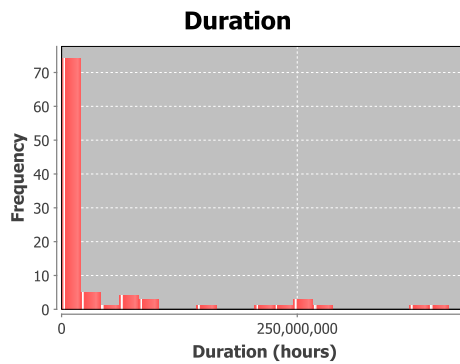


Figure .393: Histogram of the duration, time from first to last pen stroke for all students for homework eight problem two for the 2011 course offering.

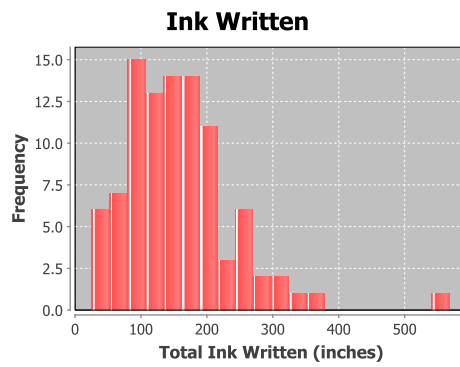


Figure .394: Histogram of the amount of ink, in inches, written, for each student on homework eight problem two for the 2011 course offering.

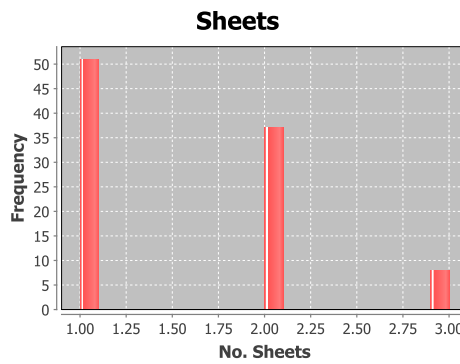


Figure .395: Histogram of the number of pages written on by each student to solve homework eight problem two for the 2011 course offering.

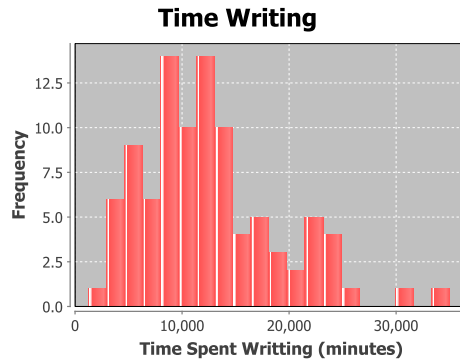


Figure .396: Histogram of the amount of time spent writing by each student on homework eight problem two for the 2011 course offering.

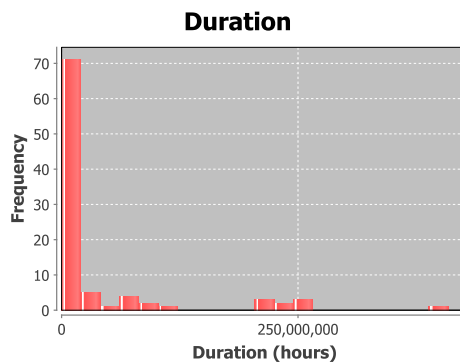


Figure .397: Histogram of the duration, time from first to last pen stroke for all students for homework eight problem three for the 2011 course offering.

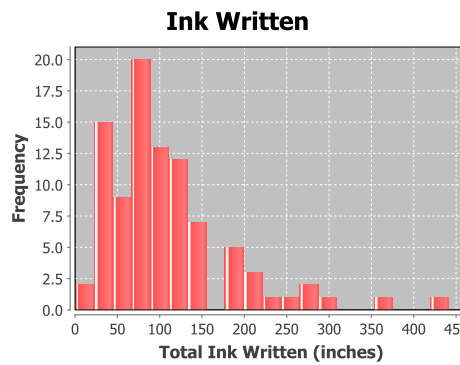


Figure .398: Histogram of the amount of ink, in inches, written, for each student on homework eight problem three for the 2011 course offering.

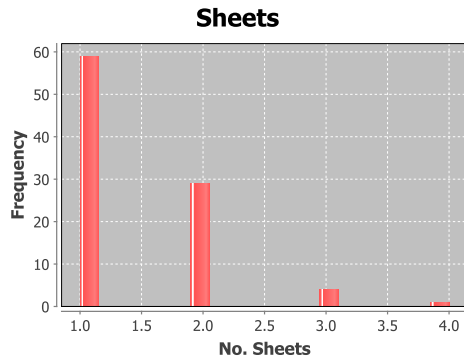


Figure .399: Histogram of the number of pages written on by each student to solve homework eight problem three for the 2011 course offering.

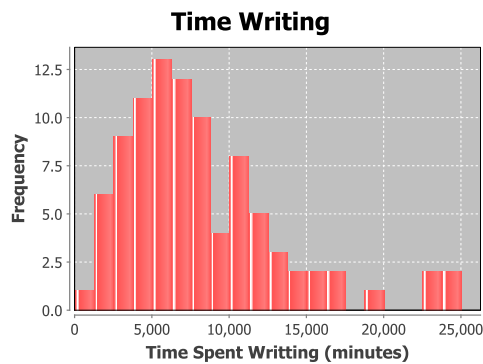


Figure .400: Histogram of the amount of time spent writing by each student on homework eight problem three for the 2011 course offering.

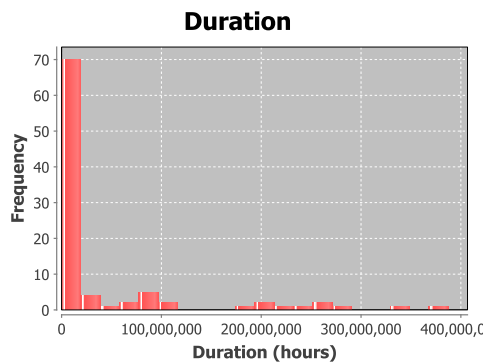


Figure .401: Histogram of the duration, time from first to last pen stroke for all students for homework eight problem four for the 2011 course offering.

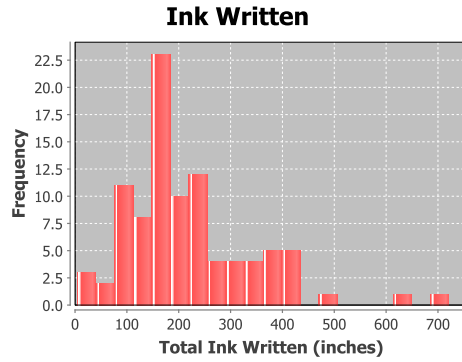


Figure .402: Histogram of the amount of ink, in inches, written, for each student on homework eight problem four for the 2011 course offering.

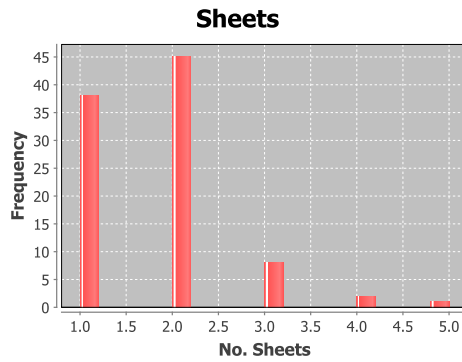


Figure .403: Histogram of the number of pages written on by each student to solve homework eight problem four for the 2011 course offering.

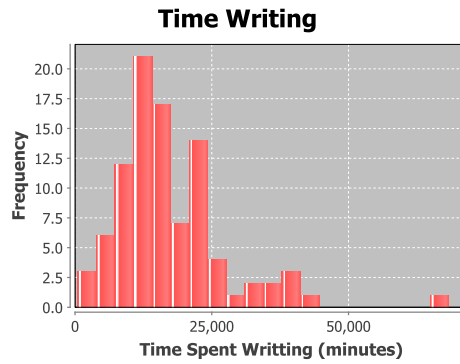


Figure .404: Histogram of the amount of time spent writing by each student on homework eight problem four for the 2011 course offering.

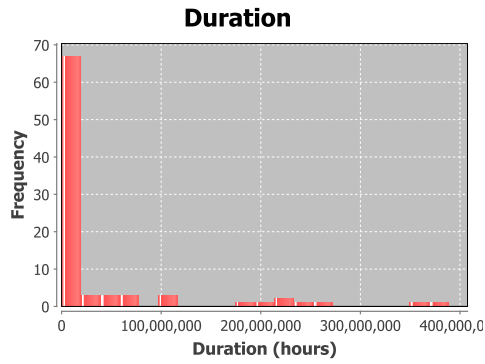


Figure .405: Histogram of the duration, time from first to last pen stroke for all students for homework eight problem five for the 2011 course offering.

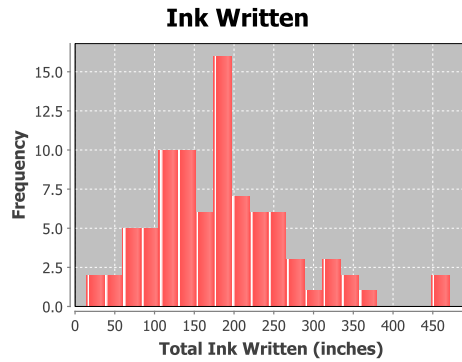


Figure .406: Histogram of the amount of ink, in inches, written, for each student on homework eight problem five for the 2011 course offering.

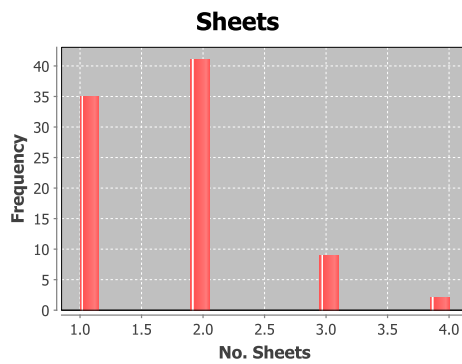


Figure .407: Histogram of the number of pages written on by each student to solve homework eight problem five for the 2011 course offering.

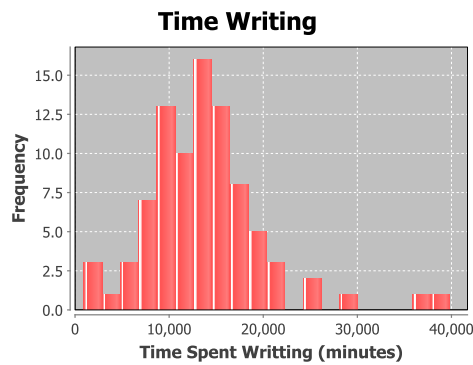


Figure .408: Histogram of the amount of time spent writing by each student on homework eight problem five for the 2011 course offering.

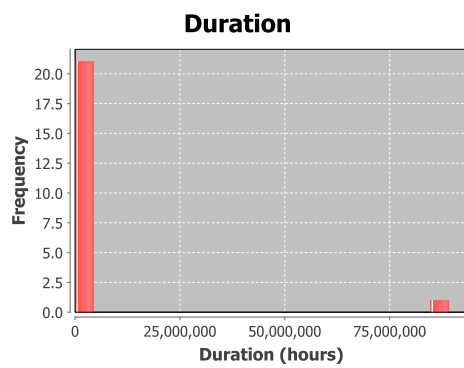


Figure .409: Histogram of the duration, time from first to last pen stroke for all students for homework one problem one for the 2012 course offering.

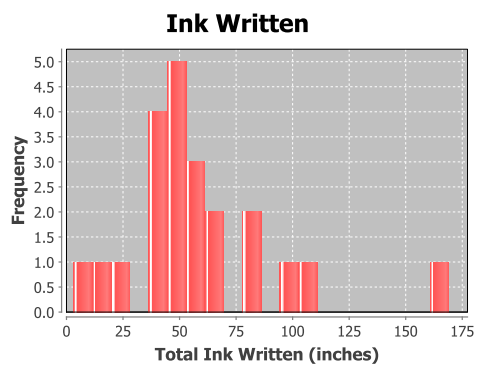


Figure .410: Histogram of the amount of ink, in inches, written, for each student on homework one problem one for the 2012 course offering.

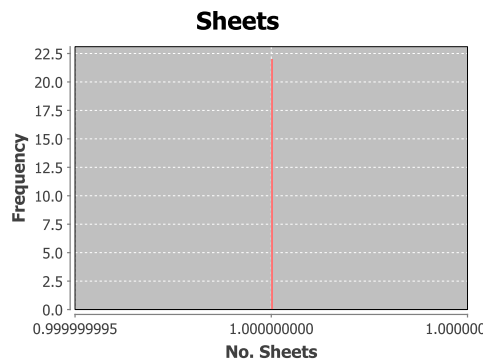


Figure .411: Histogram of the number of pages written on by each student to solve homework one problem one for the 2012 course offering.

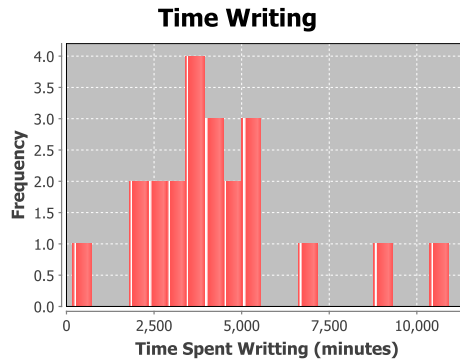


Figure .412: Histogram of the amount of time spent writing by each student on homework one problem one for the 2012 course offering.

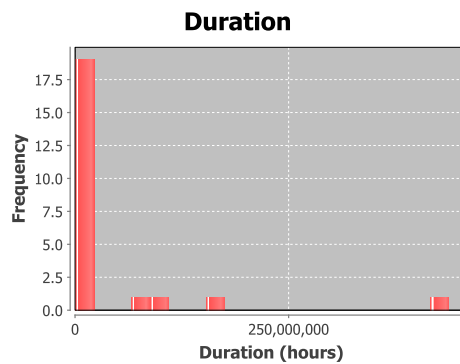


Figure .413: Histogram of the duration, time from first to last pen stroke for all students for homework one problem two for the 2012 course offering.

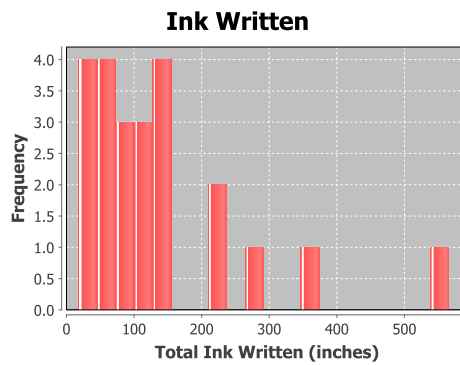


Figure .414: Histogram of the amount of ink, in inches, written, for each student on homework one problem two for the 2012 course offering.

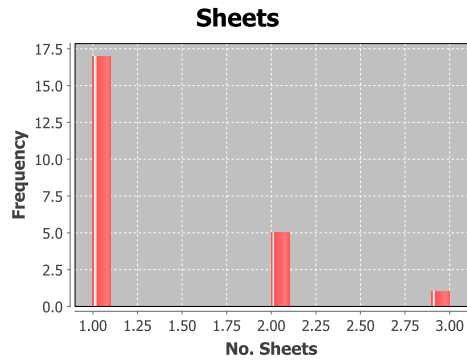


Figure .415: Histogram of the number of pages written on by each student to solve homework one problem two for the 2012 course offering.

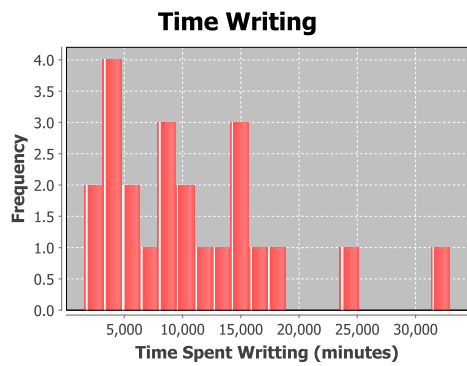


Figure .416: Histogram of the amount of time spent writing by each student on homework one problem two for the 2012 course offering.

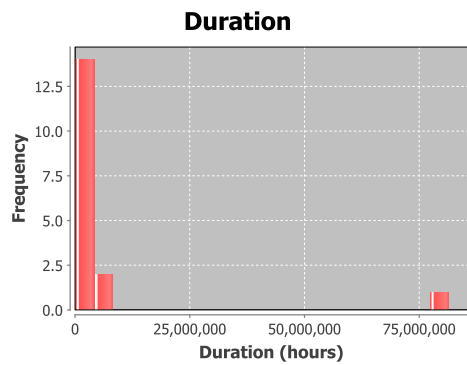


Figure .417: Histogram of the duration, time from first to last pen stroke for all students for homework one problem three for the 2012 course offering.

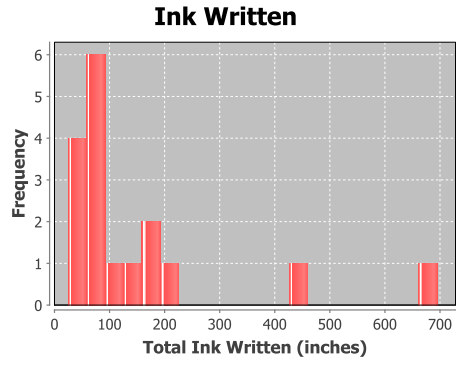


Figure .418: Histogram of the amount of ink, in inches, written, for each student on homework one problem three for the 2012 course offering.

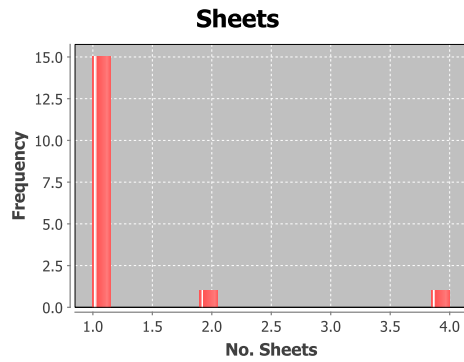


Figure .419: Histogram of the number of pages written on by each student to solve homework one problem three for the 2012 course offering.

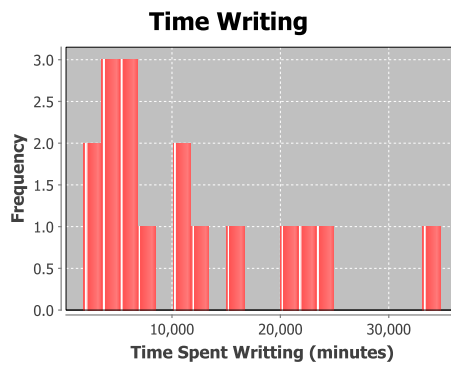


Figure .420: Histogram of the amount of time spent writing by each student on homework one problem three for the 2012 course offering.

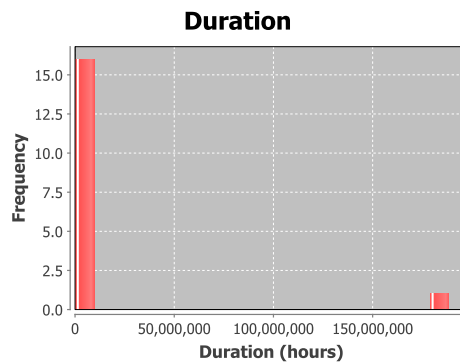


Figure .421: Histogram of the duration, time from first to last pen stroke for all students for homework one problem four for the 2012 course offering.

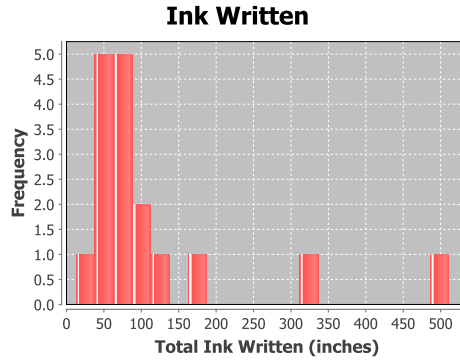


Figure .422: Histogram of the amount of ink, in inches, written, for each student on homework one problem four for the 2012 course offering.

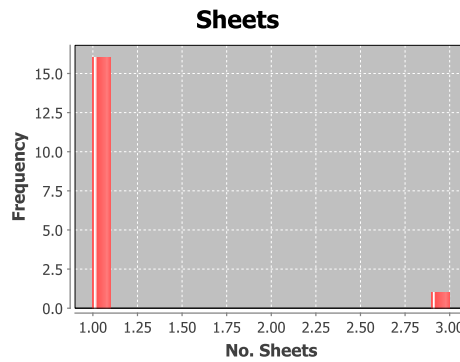


Figure .423: Histogram of the number of pages written on by each student to solve homework one problem four for the 2012 course offering.

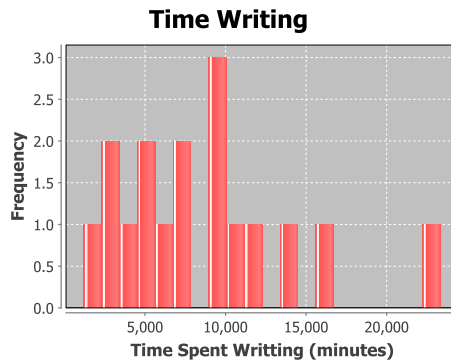


Figure .424: Histogram of the amount of time spent writing by each student on homework one problem four for the 2012 course offering.

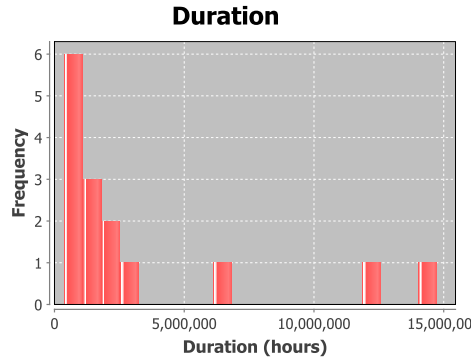


Figure .425: Histogram of the duration, time from first to last pen stroke for all students for homework one problem five for the 2012 course offering.

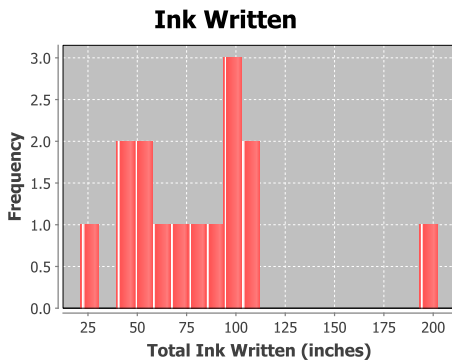


Figure .426: Histogram of the amount of ink, in inches, written, for each student on homework one problem five for the 2012 course offering.

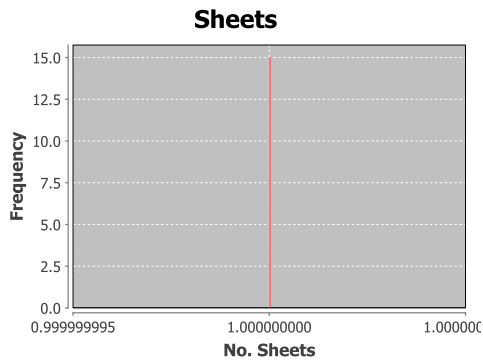


Figure .427: Histogram of the number of pages written on by each student to solve homework one problem five for the 2012 course offering.

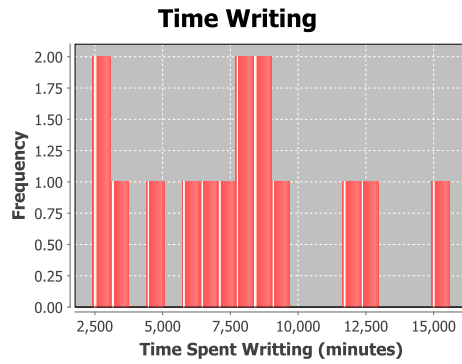


Figure .428: Histogram of the amount of time spent writing by each student on homework one problem five for the 2012 course offering.

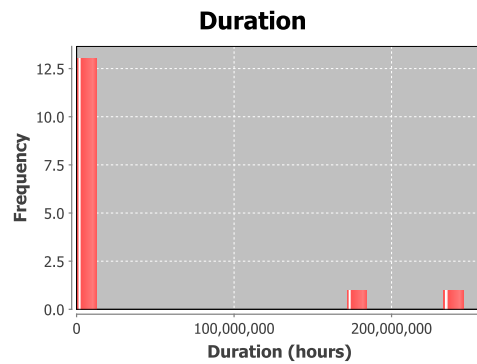


Figure .429: Histogram of the duration, time from first to last pen stroke for all students for homework one problem six for the 2012 course offering.

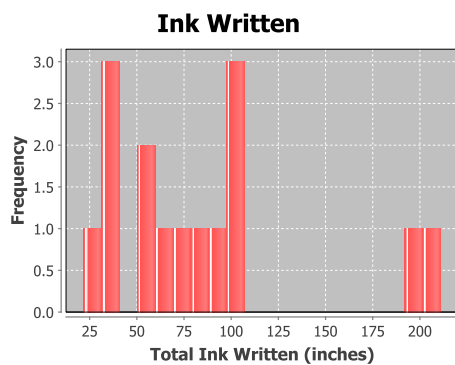


Figure .430: Histogram of the amount of ink, in inches, written, for each student on homework one problem six for the 2012 course offering.

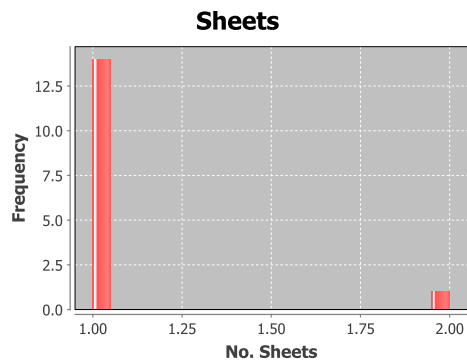


Figure .431: Histogram of the number of pages written on by each student to solve homework one problem six for the 2012 course offering.

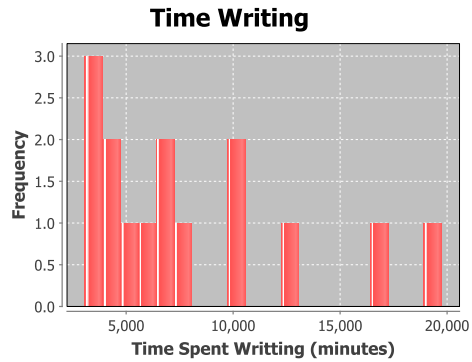


Figure .432: Histogram of the amount of time spent writing by each student on homework one problem six for the 2012 course offering.

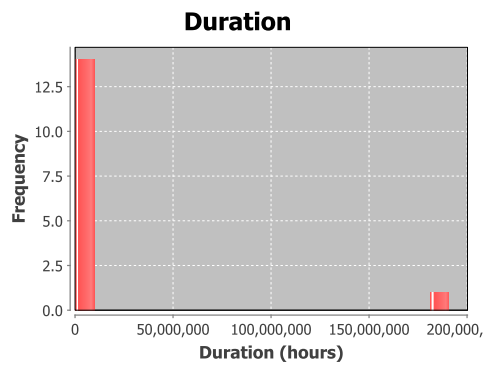


Figure .433: Histogram of the duration, time from first to last pen stroke for all students for homework one problem seven for the 2012 course offering.

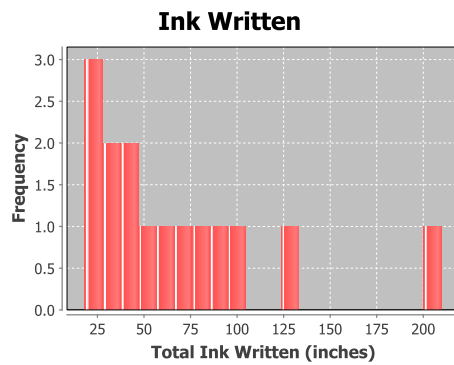


Figure .434: Histogram of the amount of ink, in inches, written, for each student on homework one problem seven for the 2012 course offering.

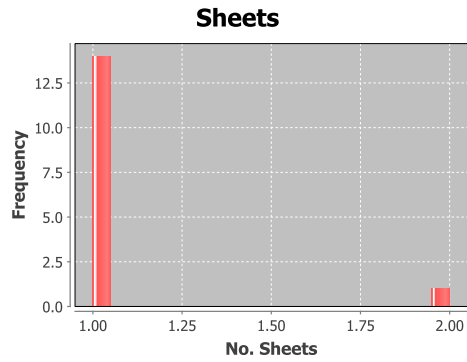


Figure .435: Histogram of the number of pages written on by each student to solve homework seven problem one for the 2012 course offering.

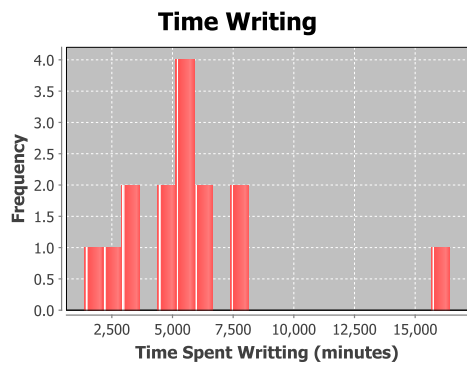


Figure .436: Histogram of the amount of time spent writing by each student on homework one problem seven for the 2012 course offering.

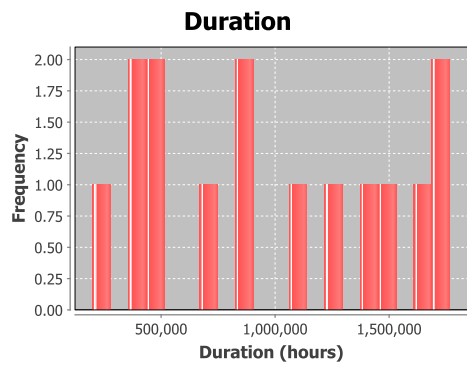


Figure .437: Histogram of the duration, time from first to last pen stroke for all students for homework one problem eight for the 2012 course offering.

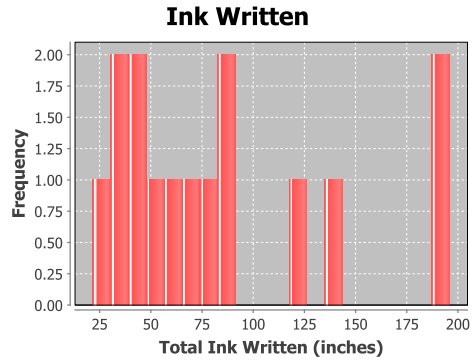


Figure .438: Histogram of the amount of ink, in inches, written, for each student on homework one problem eight for the 2012 course offering.

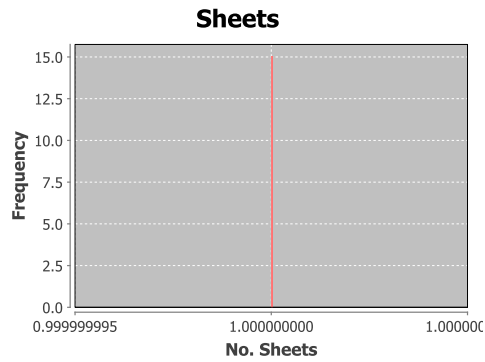


Figure .439: Histogram of the number of pages written on by each student to solve homework one problem eight for the 2012 course offering.

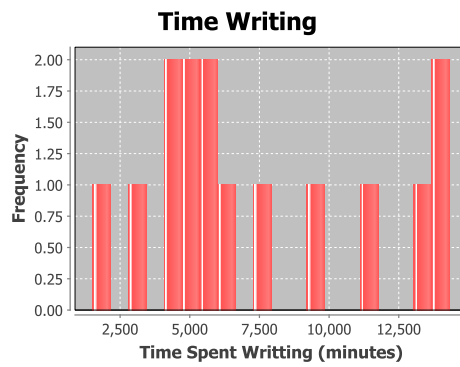


Figure .440: Histogram of the amount of time spent writing by each student on homework one problem eight for the 2012 course offering.

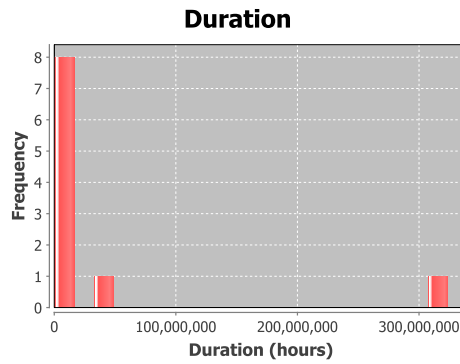


Figure .441: Histogram of the duration, time from first to last pen stroke for all students for homework two problem one for the 2012 course offering.

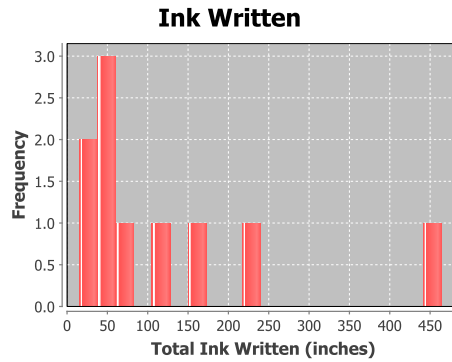


Figure .442: Histogram of the amount of ink, in inches, written, for each student on homework two problem one for the 2012 course offering.

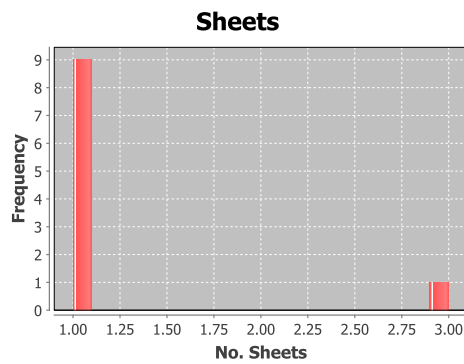


Figure .443: Histogram of the number of pages written on by each student to solve homework two problem one for the 2012 course offering.

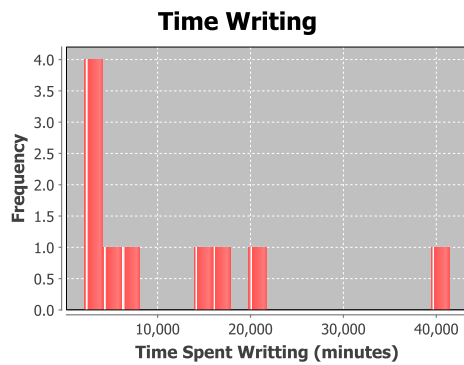


Figure .444: Histogram of the amount of time spent writing by each student on homework two problem one for the 2012 course offering.

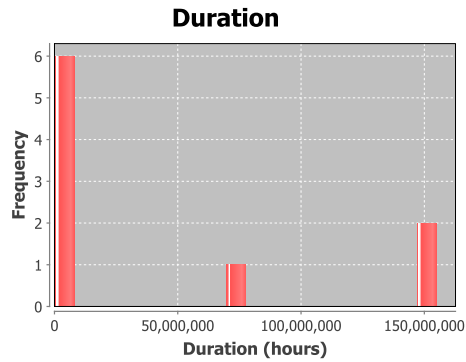


Figure .445: Histogram of the duration, time from first to last pen stroke for all students for homework two problem two for the 2012 course offering.

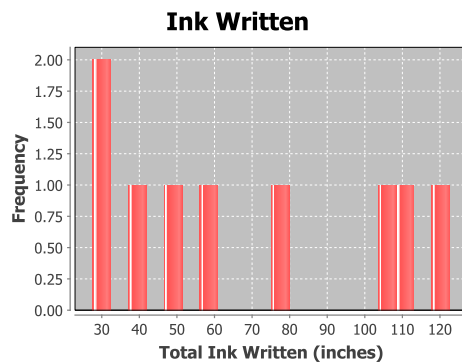


Figure .446: Histogram of the amount of ink, in inches, written, for each student on homework two problem two for the 2012 course offering.

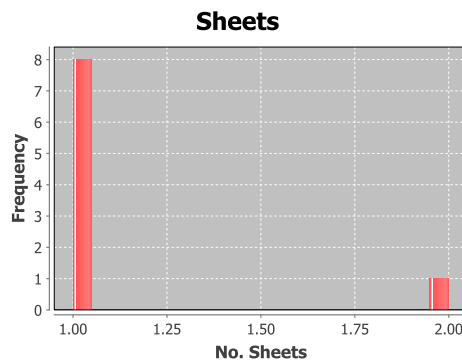


Figure .447: Histogram of the number of pages written on by each student to solve homework two problem two for the 2012 course offering.

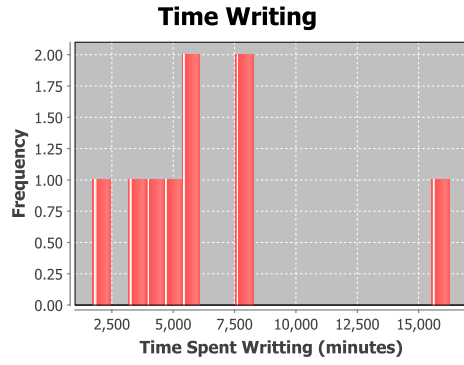


Figure .448: Histogram of the amount of time spent writing by each student on homework two problem two for the 2012 course offering.

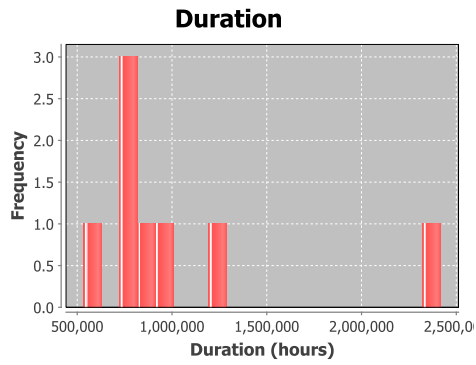


Figure .449: Histogram of the duration, time from first to last pen stroke for all students for homework two problem three for the 2012 course offering.

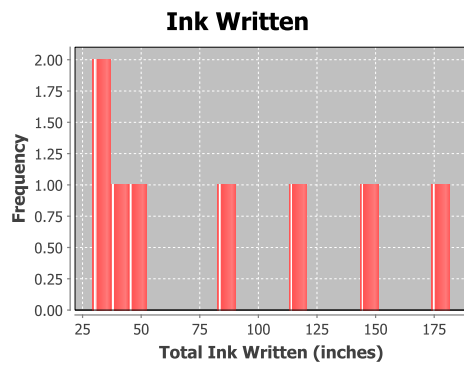


Figure .450: Histogram of the amount of ink, in inches, written, for each student on homework two problem three for the 2012 course offering.

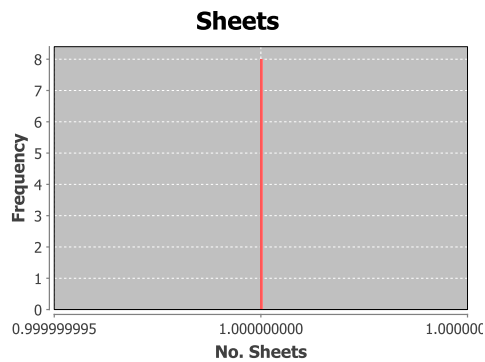


Figure .451: Histogram of the number of pages written on by each student to solve homework two problem three for the 2012 course offering.

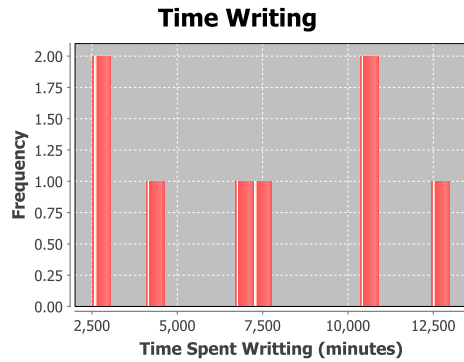


Figure .452: Histogram of the amount of time spent writing by each student on homework two problem three for the 2012 course offering.

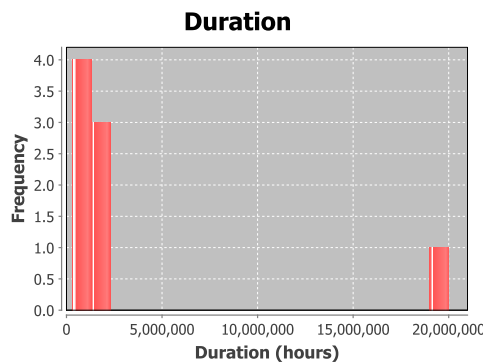


Figure .453: Histogram of the duration, time from first to last pen stroke for all students for homework two problem four for the 2012 course offering.

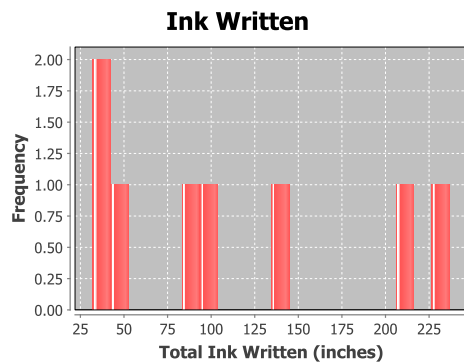


Figure .454: Histogram of the amount of ink, in inches, written, for each student on homework two problem four for the 2012 course offering.

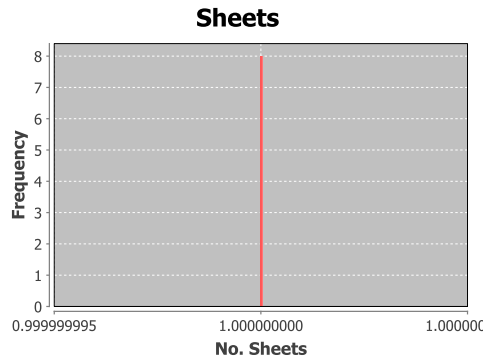


Figure .455: Histogram of the number of pages written on by each student to solve homework two problem four for the 2012 course offering.

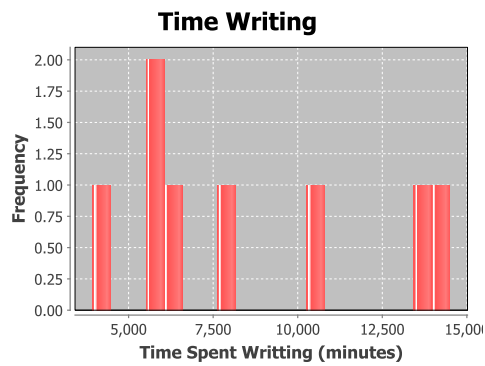


Figure .456: Histogram of the amount of time spent writing by each student on homework two problem four for the 2012 course offering.

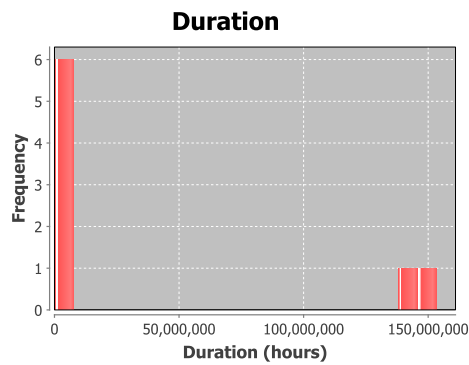


Figure .457: Histogram of the duration, time from first to last pen stroke for all students for homework two problem five for the 2012 course offering.

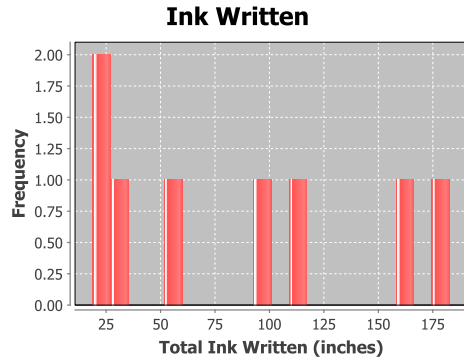


Figure .458: Histogram of the amount of ink, in inches, written, for each student on homework two problem five for the 2012 course offering.

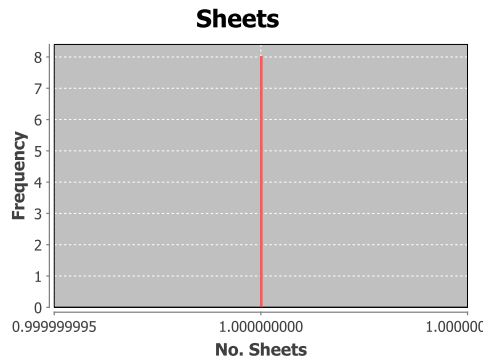


Figure .459: Histogram of the number of pages written on by each student to solve homework two problem five for the 2012 course offering.

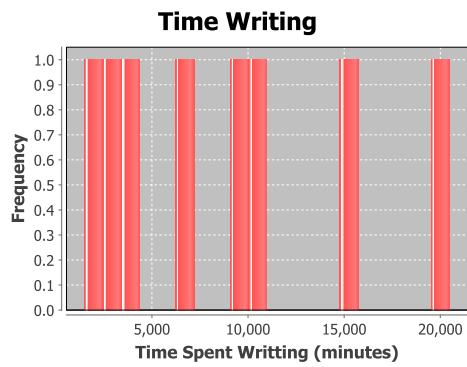


Figure .460: Histogram of the amount of time spent writing by each student on homework two problem five for the 2012 course offering.

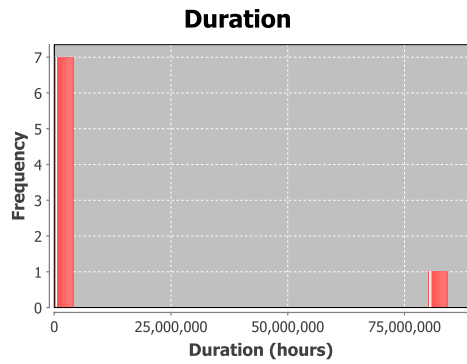


Figure .461: Histogram of the duration, time from first to last pen stroke for all students for homework two problem six for the 2012 course offering.

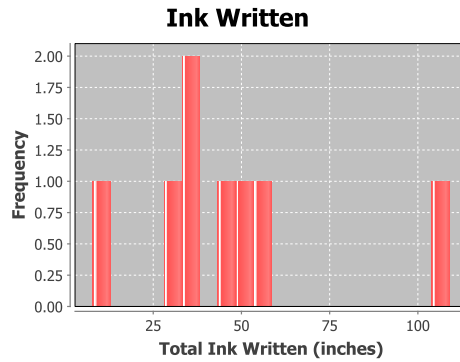


Figure .462: Histogram of the amount of ink, in inches, written, for each student on homework two problem six for the 2012 course offering.

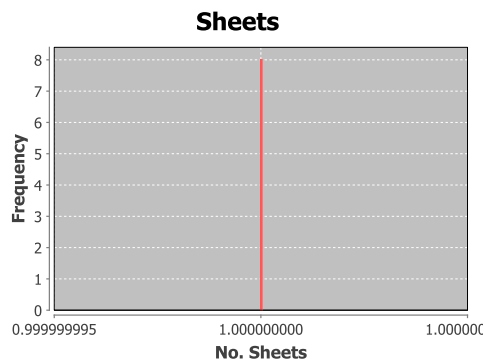


Figure .463: Histogram of the number of pages written on by each student to solve homework two problem six for the 2012 course offering.

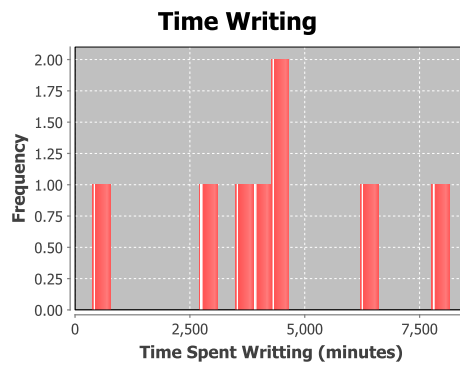


Figure .464: Histogram of the amount of time spent writing by each student on homework two problem six for the 2012 course offering.

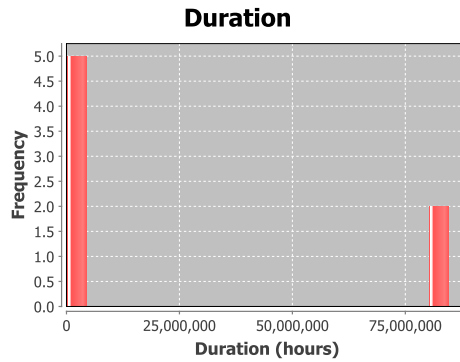


Figure .465: Histogram of the duration, time from first to last pen stroke for all students for homework two problem seven for the 2012 course offering.

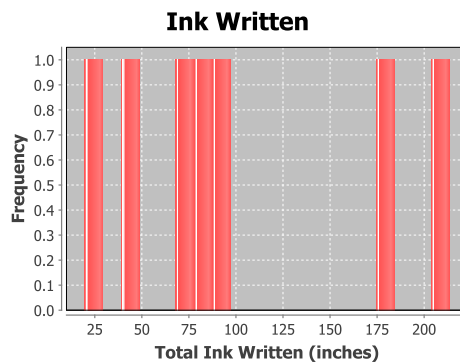


Figure .466: Histogram of the amount of ink, in inches, written, for each student on homework two problem seven for the 2012 course offering.

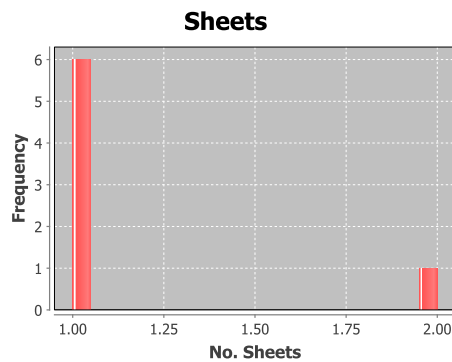


Figure .467: Histogram of the number of pages written on by each student to solve homework two problem seven for the 2012 course offering.

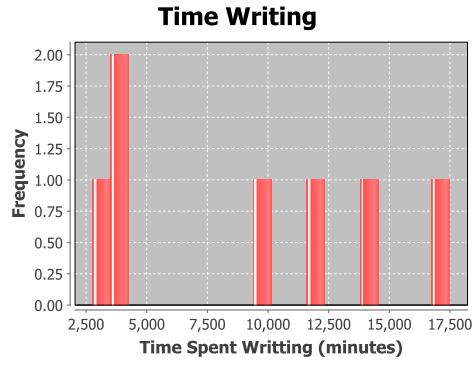


Figure .468: Histogram of the amount of time spent writing by each student on homework two problem seven for the 2012 course offering.

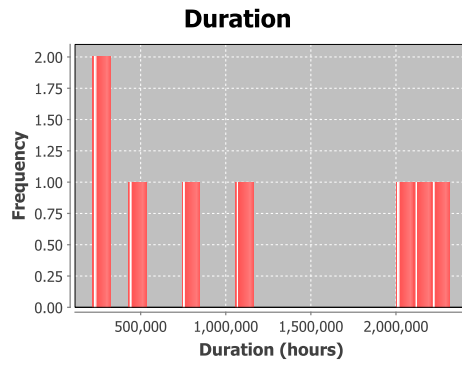


Figure .469: Histogram of the duration, time from first to last pen stroke for all students for homework two problem eight for the 2012 course offering.

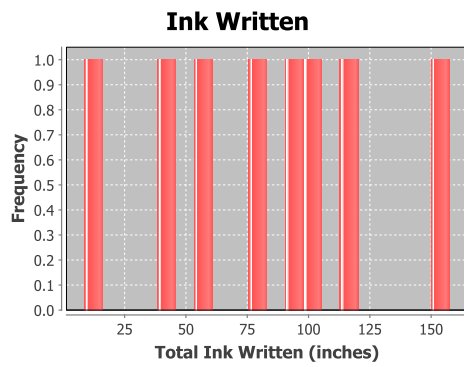


Figure .470: Histogram of the amount of ink, in inches, written, for each student on homework two problem eight for the 2012 course offering.

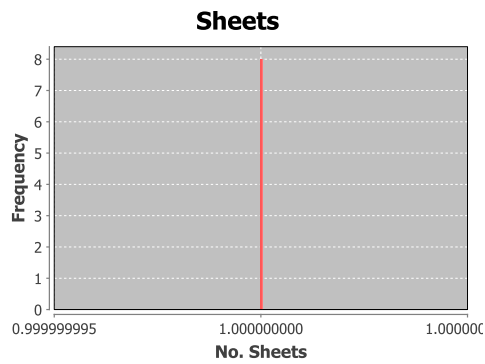


Figure .471: Histogram of the number of pages written on by each student to solve homework two problem eight for the 2012 course offering.

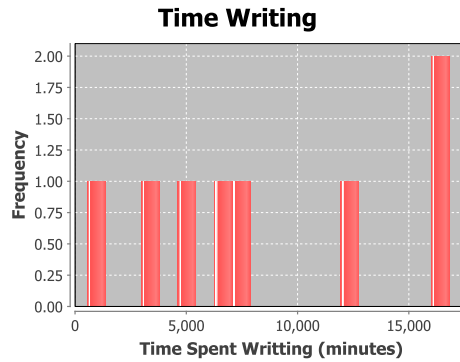


Figure .472: Histogram of the amount of time spent writing by each student on homework two problem eight for the 2012 course offering.

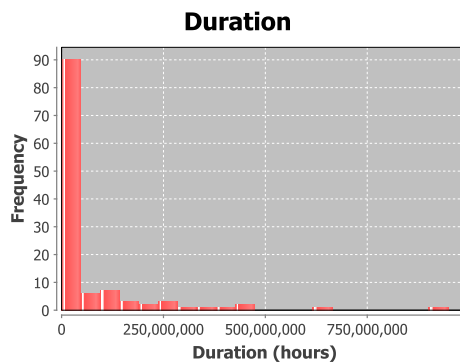


Figure .473: Histogram of the duration, time from first to last pen stroke for all students for homework three problem one for the 2012 course offering.

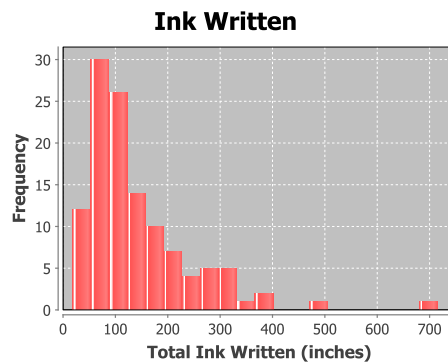


Figure .474: Histogram of the amount of ink, in inches, written, for each student on homework three problem one for the 2012 course offering.

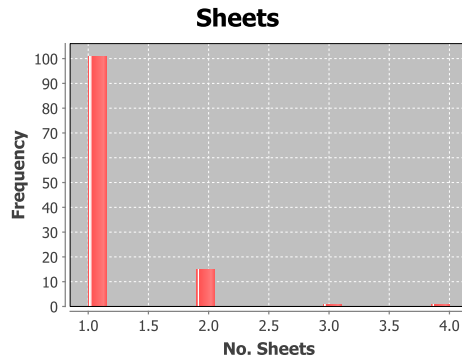


Figure .475: Histogram of the number of pages written on by each student to solve homework three problem one for the 2012 course offering.

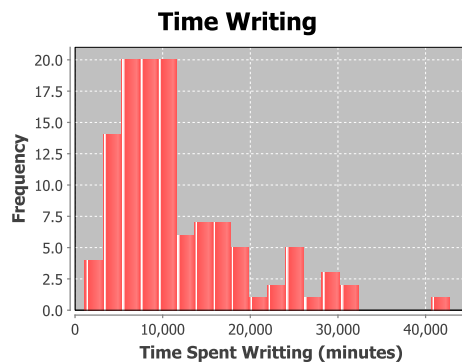


Figure .476: Histogram of the amount of time spent writing by each student on homework three problem one for the 2012 course offering.

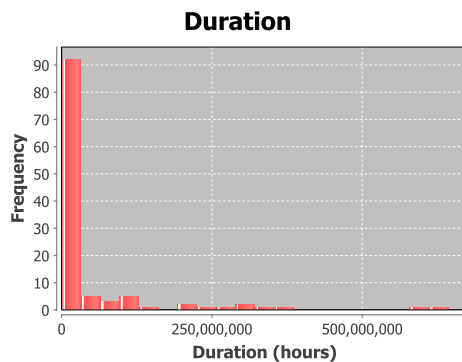


Figure .477: Histogram of the duration, time from first to last pen stroke for all students for homework three problem two for the 2012 course offering.

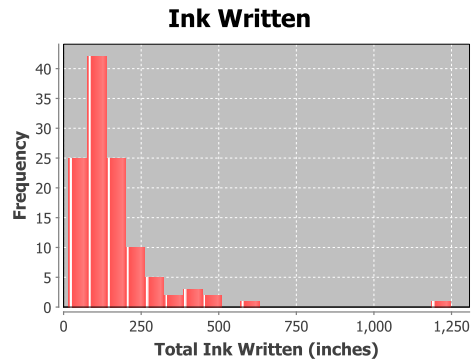


Figure .478: Histogram of the amount of ink, in inches, written, for each student on homework three problem two for the 2012 course offering.

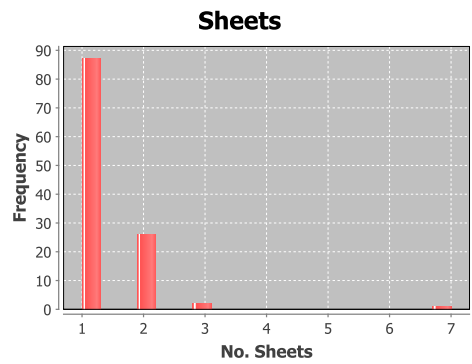


Figure .479: Histogram of the number of pages written on by each student to solve homework three problem two for the 2012 course offering.

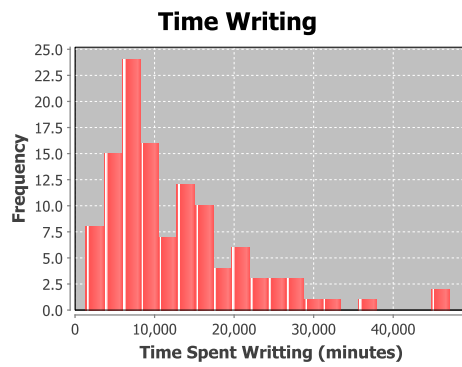


Figure .480: Histogram of the amount of time spent writing by each student on homework three problem two for the 2012 course offering.

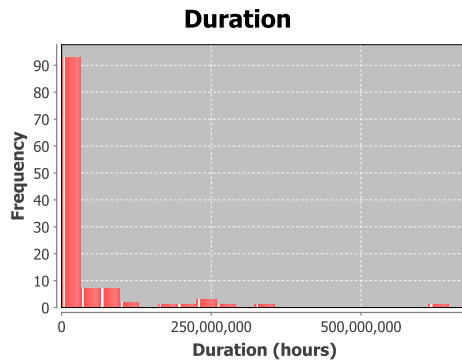


Figure .481: Histogram of the duration, time from first to last pen stroke for all students for homework three problem three for the 2012 course offering.

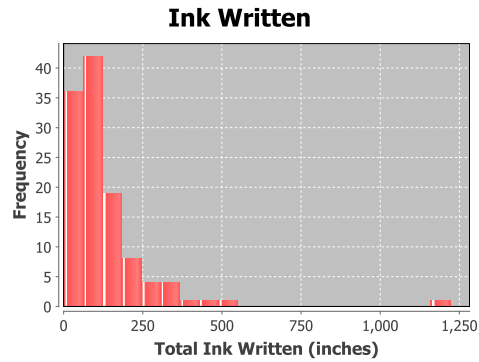


Figure .482: Histogram of the amount of ink, in inches, written, for each student on homework three problem three for the 2012 course offering.

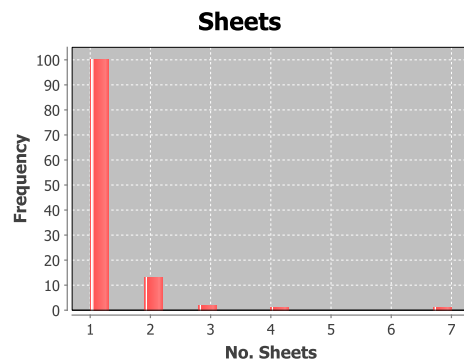


Figure .483: Histogram of the number of pages written on by each student to solve homework three problem three for the 2012 course offering.

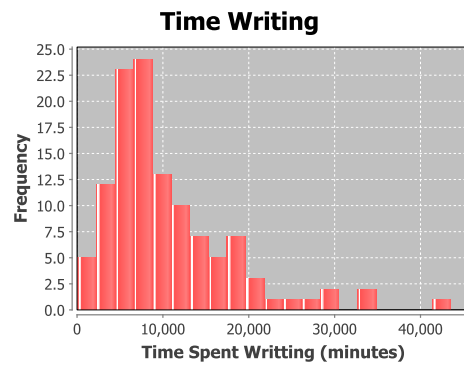


Figure .484: Histogram of the amount of time spent writing by each student on homework three problem three for the 2012 course offering.

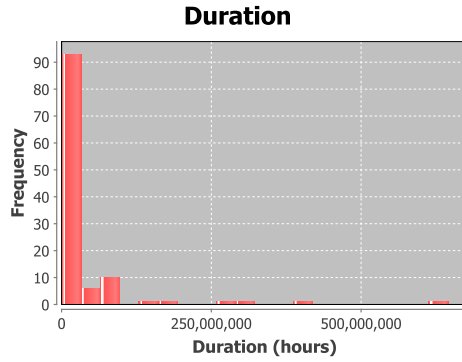


Figure .485: Histogram of the duration, time from first to last pen stroke for all students for homework three problem four for the 2012 course offering.

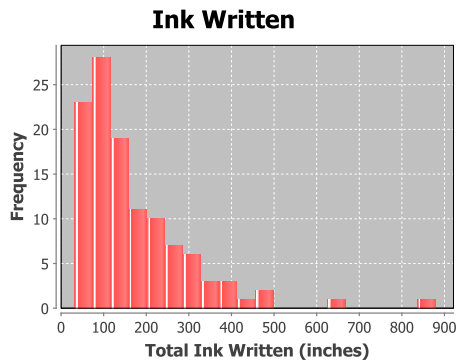


Figure .486: Histogram of the amount of ink, in inches, written, for each student on homework three problem four for the 2012 course offering.

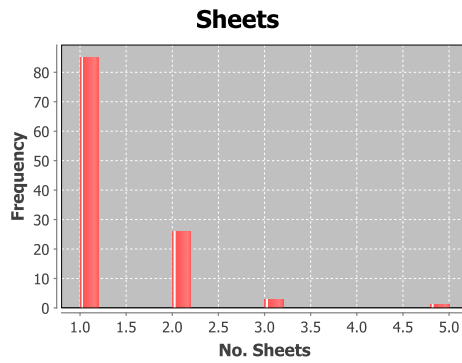


Figure .487: Histogram of the number of pages written on by each student to solve homework three problem four for the 2012 course offering.

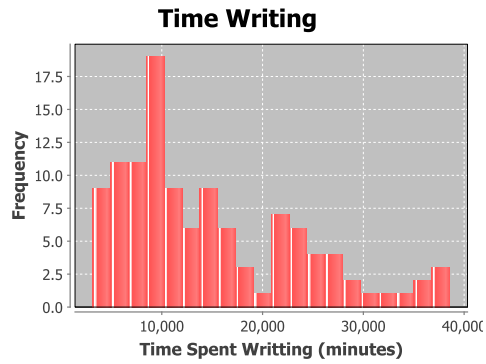


Figure .488: Histogram of the amount of time spent writing by each student on homework three problem four for the 2012 course offering.

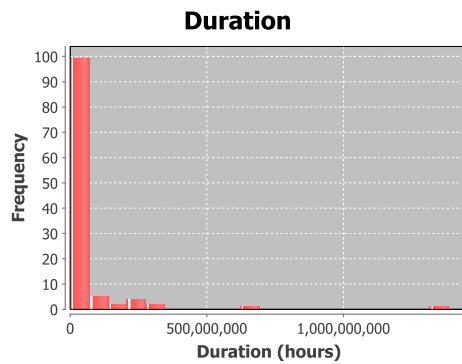


Figure .489: Histogram of the duration, time from first to last pen stroke for all students for homework three problem five for the 2012 course offering.

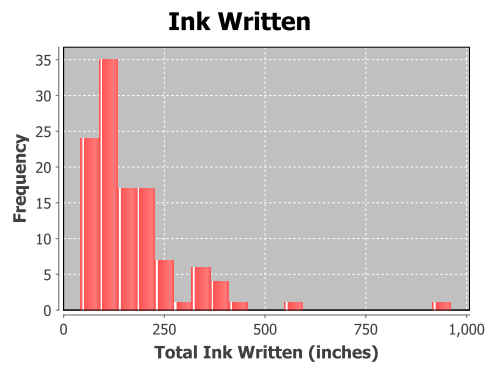


Figure .490: Histogram of the amount of ink, in inches, written, for each student on homework three problem five for the 2012 course offering.

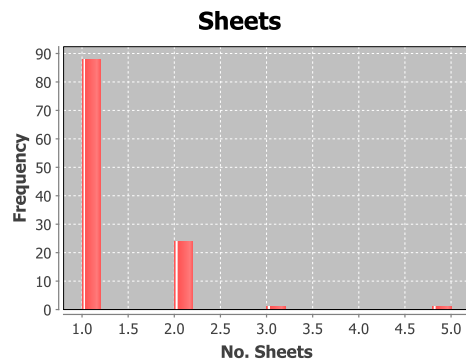


Figure .491: Histogram of the number of pages written on by each student to solve homework three problem five for the 2012 course offering.

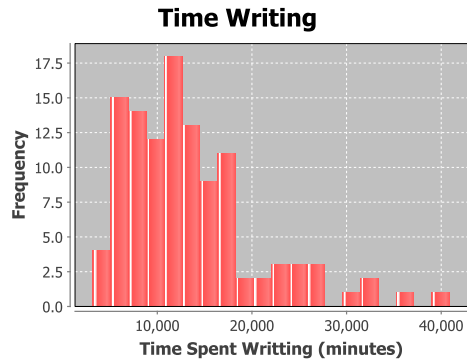


Figure .492: Histogram of the amount of time spent writing by each student on homework three problem five for the 2012 course offering.

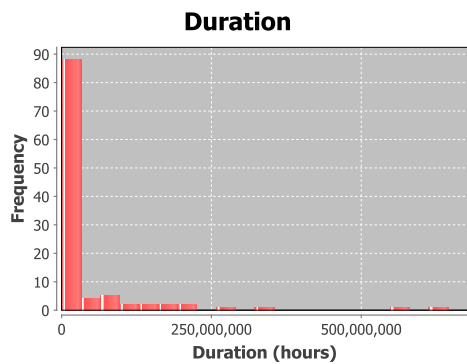


Figure .493: Histogram of the duration, time from first to last pen stroke for all students for homework three problem six for the 2012 course offering.

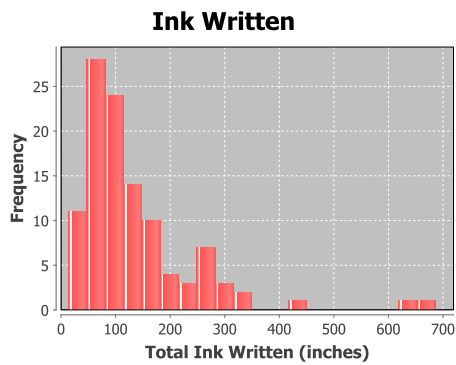


Figure .494: Histogram of the amount of ink, in inches, written, for each student on homework three problem six for the 2012 course offering.

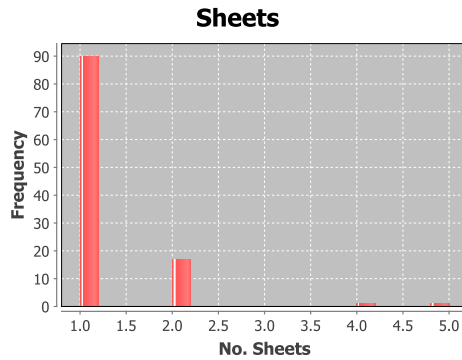


Figure .495: Histogram of the number of pages written on by each student to solve homework three problem six for the 2012 course offering.

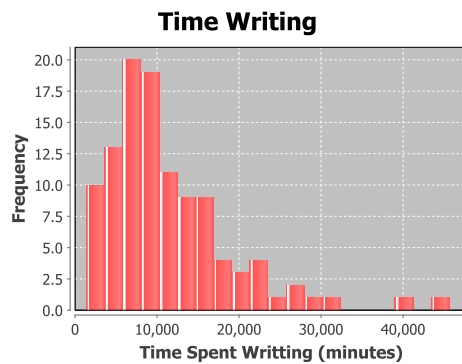


Figure .496: Histogram of the amount of time spent writing by each student on homework three problem six for the 2012 course offering.

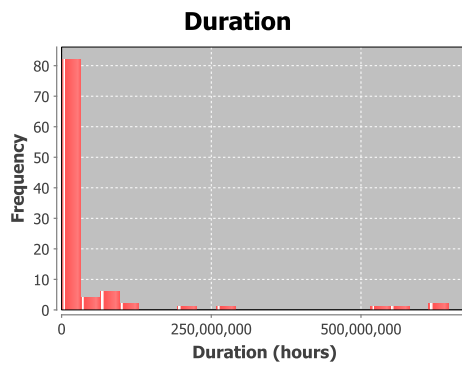


Figure .497: Histogram of the duration, time from first to last pen stroke for all students for homework three problem seven for the 2012 course offering.

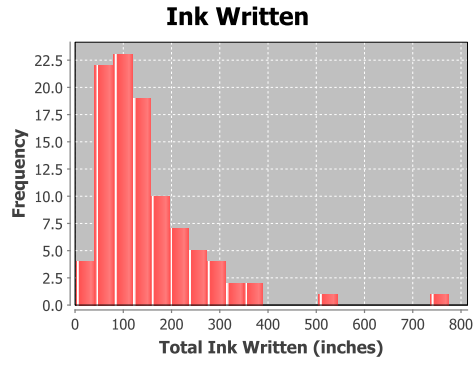


Figure .498: Histogram of the amount of ink, in inches, written, for each student on homework three problem seven for the 2012 course offering.

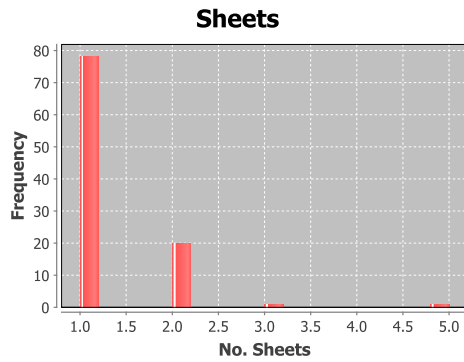


Figure .499: Histogram of the number of pages written on by each student to solve homework three problem seven for the 2012 course offering.

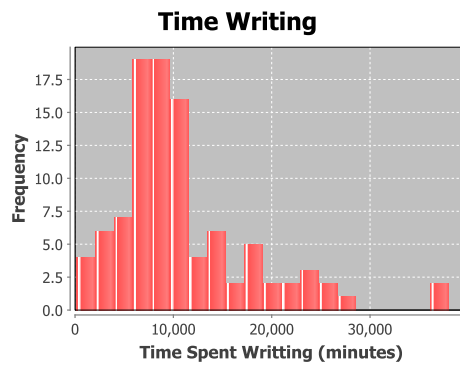


Figure .500: Histogram of the amount of time spent writing by each student on homework three problem seven for the 2012 course offering.

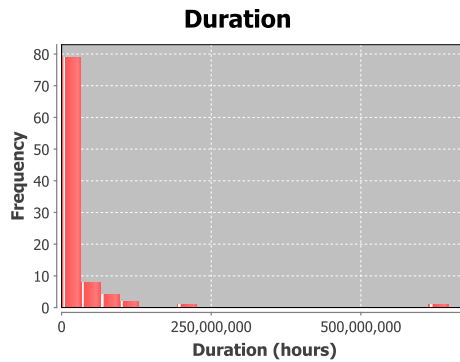


Figure .501: Histogram of the duration, time from first to last pen stroke for all students for homework three problem eight for the 2012 course offering.

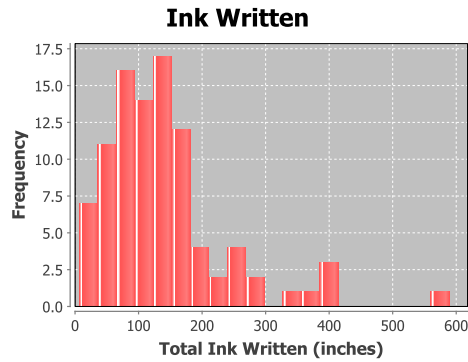


Figure .502: Histogram of the amount of ink, in inches, written, for each student on homework three problem eight for the 2012 course offering.

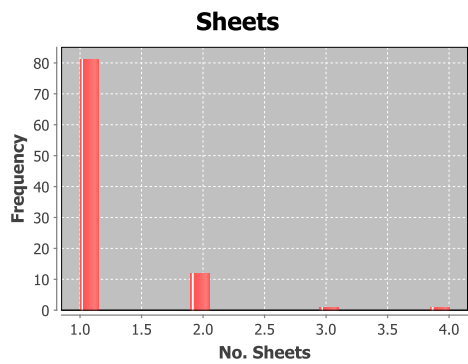


Figure .503: Histogram of the number of pages written on by each student to solve homework three problem eight for the 2012 course offering.

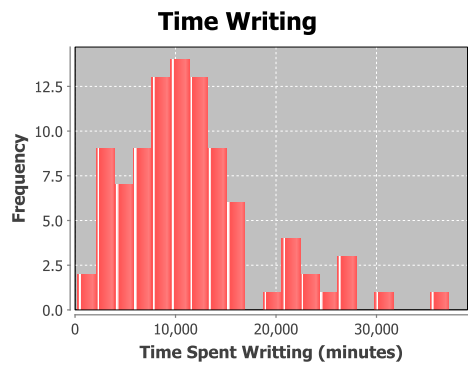


Figure .504: Histogram of the amount of time spent writing by each student on homework three problem eight for the 2012 course offering.

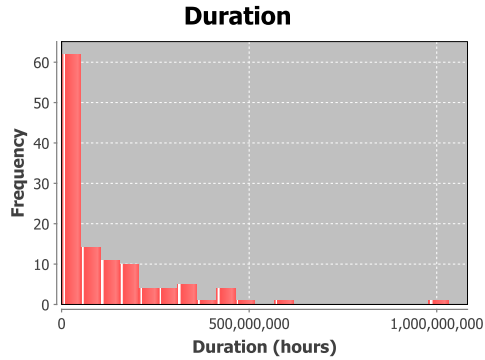


Figure .505: Histogram of the duration, time from first to last pen stroke for all students for homework four problem one for the 2012 course offering.

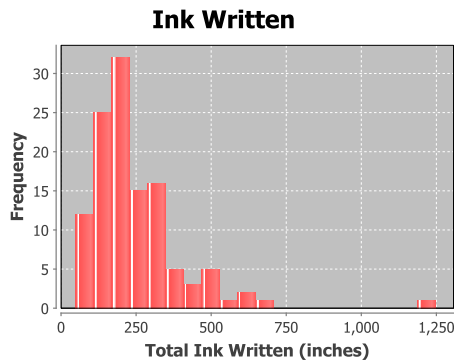


Figure .506: Histogram of the amount of ink, in inches, written, for each student on homework four problem one for the 2012 course offering.

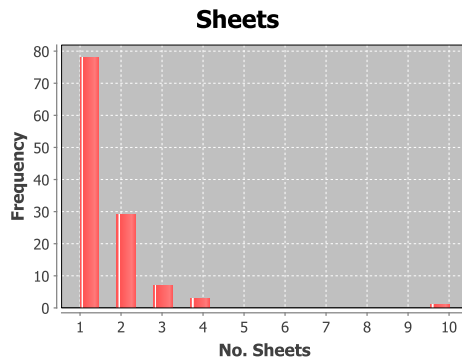


Figure .507: Histogram of the number of pages written on by each student to solve homework four problem one for the 2012 course offering.

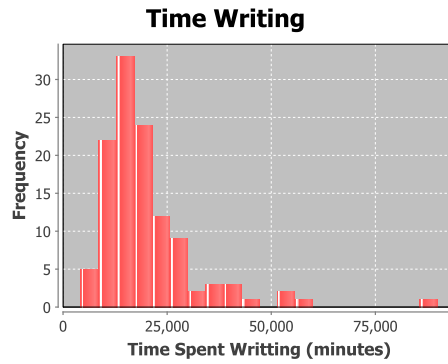


Figure .508: Histogram of the amount of time spent writing by each student on homework four problem one for the 2012 course offering.

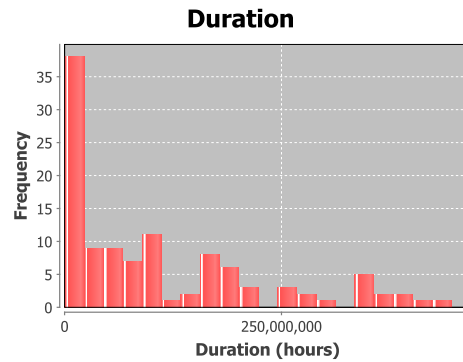


Figure .509: Histogram of the duration, time from first to last pen stroke for all students for homework four problem two for the 2012 course offering.

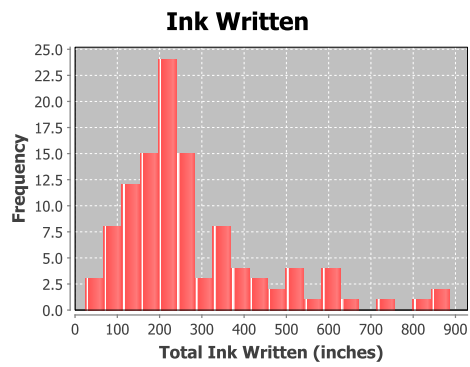


Figure .510: Histogram of the amount of ink, in inches, written, for each student on homework four problem two for the 2012 course offering.

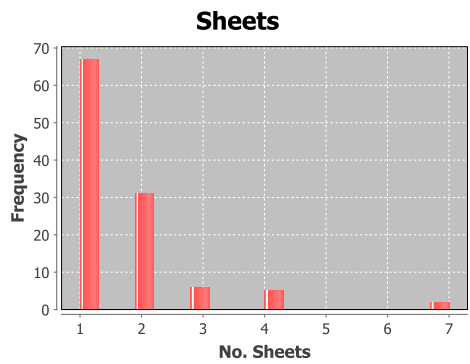


Figure .511: Histogram of the number of pages written on by each student to solve homework four problem two for the 2012 course offering.

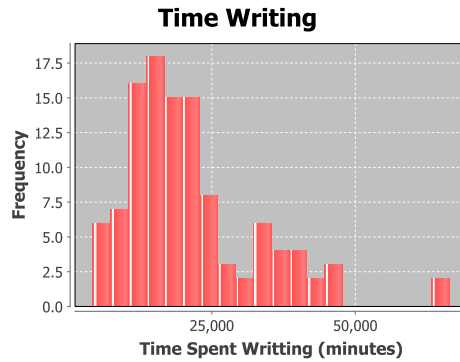


Figure .512: Histogram of the amount of time spent writing by each student on homework four problem two for the 2012 course offering.

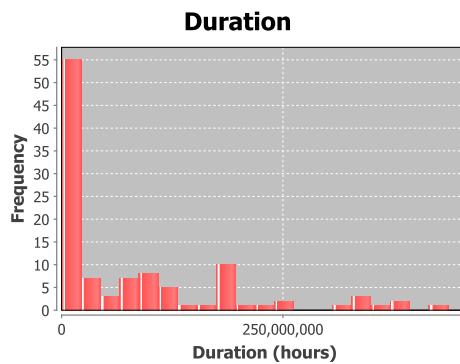


Figure .513: Histogram of the duration, time from first to last pen stroke for all students for homework four problem three for the 2012 course offering.

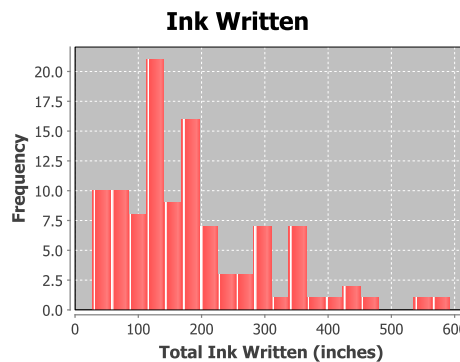


Figure .514: Histogram of the amount of ink, in inches, written, for each student on homework four problem three for the 2012 course offering.

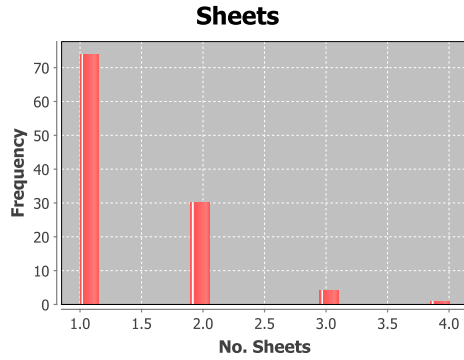


Figure .515: Histogram of the number of pages written on by each student to solve homework four problem three for the 2012 course offering.

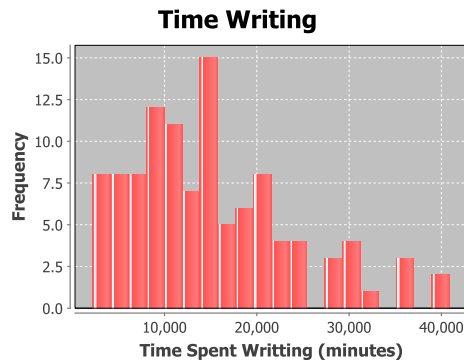


Figure .516: Histogram of the amount of time spent writing by each student on homework four problem three for the 2012 course offering.

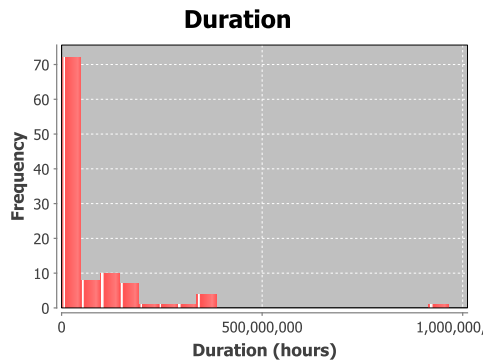


Figure .517: Histogram of the duration, time from first to last pen stroke for all students for homework four problem four for the 2012 course offering.

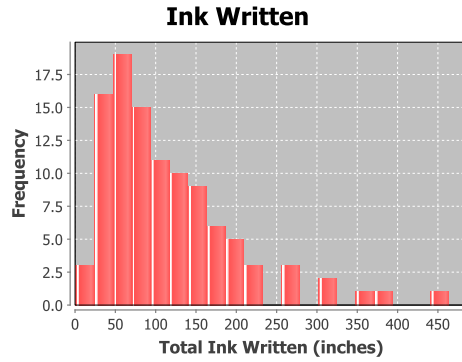


Figure .518: Histogram of the amount of ink, in inches, written, for each student on homework four problem four for the 2012 course offering.

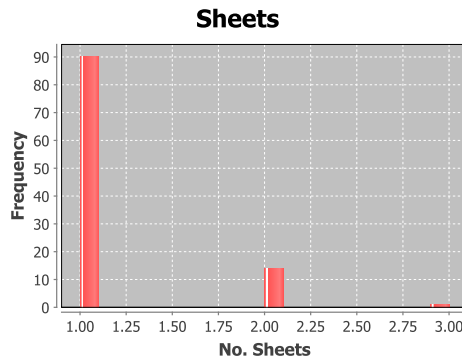


Figure .519: Histogram of the number of pages written on by each student to solve homework four problem four for the 2012 course offering.

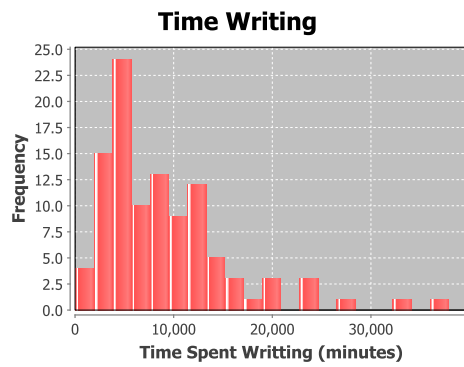


Figure .520: Histogram of the amount of time spent writing by each student on homework four problem four for the 2012 course offering.

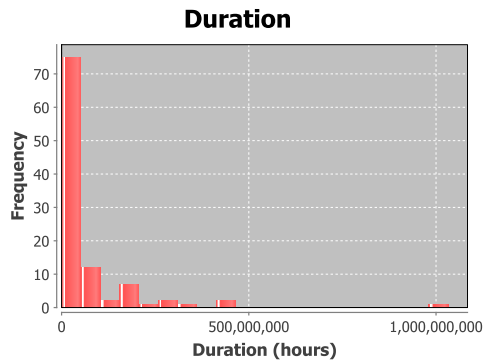


Figure .521: Histogram of the duration, time from first to last pen stroke for all students for homework four problem five for the 2012 course offering.

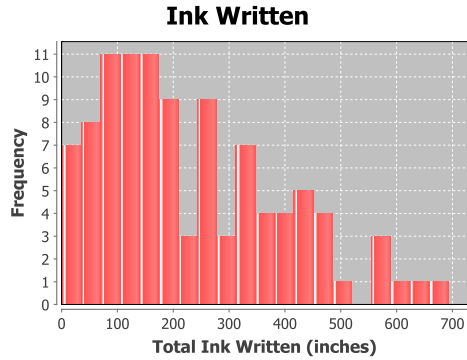


Figure .522: Histogram of the amount of ink, in inches, written, for each student on homework four problem five for the 2012 course offering.

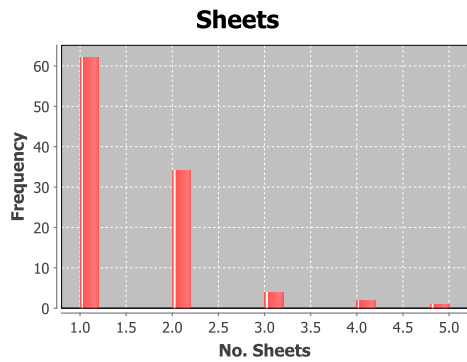


Figure .523: Histogram of the number of pages written on by each student to solve homework four problem five for the 2012 course offering.

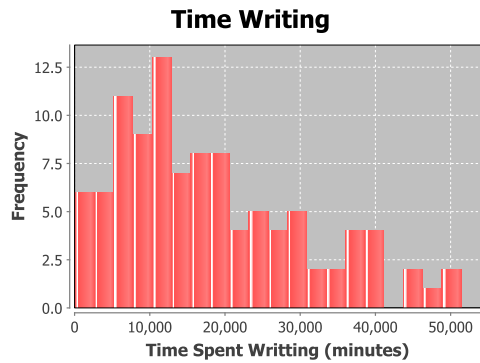


Figure .524: Histogram of the amount of time spent writing by each student on homework four problem five for the 2012 course offering.

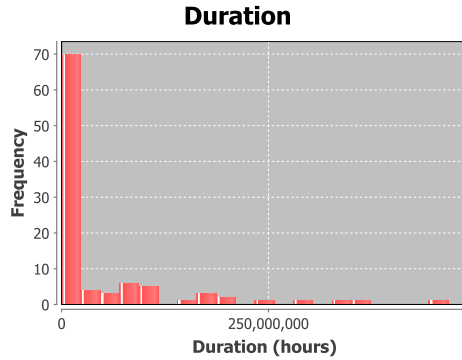


Figure .525: Histogram of the duration, time from first to last pen stroke for all students for homework four problem six for the 2012 course offering.

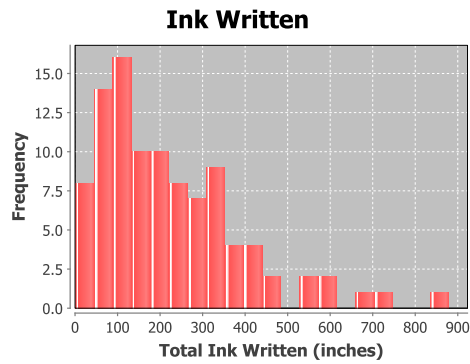


Figure .526: Histogram of the amount of ink, in inches, written, for each student on homework four problem six for the 2012 course offering.

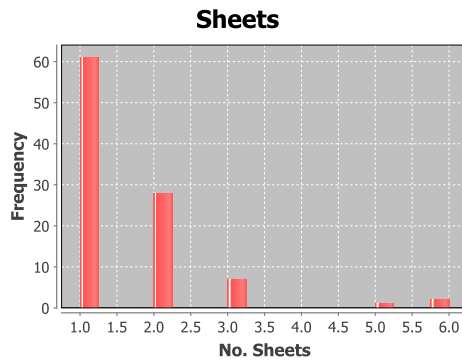


Figure .527: Histogram of the number of pages written on by each student to solve homework four problem six for the 2012 course offering.

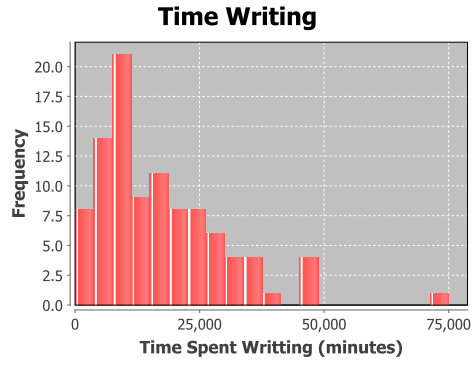


Figure .528: Histogram of the amount of time spent writing by each student on homework four problem six for the 2012 course offering.

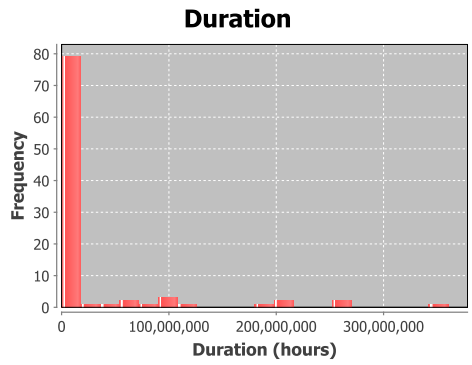


Figure .529: Histogram of the duration, time from first to last pen stroke for all students for homework four problem seven for the 2012 course offering.

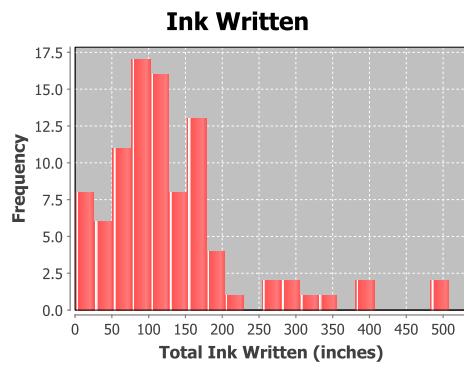


Figure .530: Histogram of the amount of ink, in inches, written, for each student on homework four problem seven for the 2012 course offering.

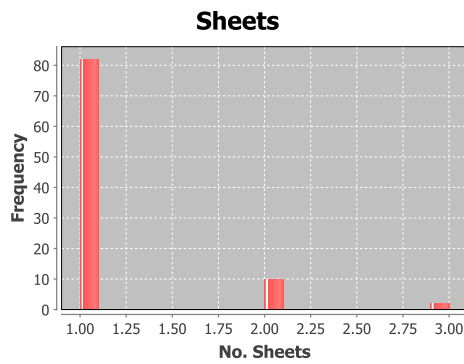


Figure .531: Histogram of the number of pages written on by each student to solve homework four problem seven for the 2012 course offering.

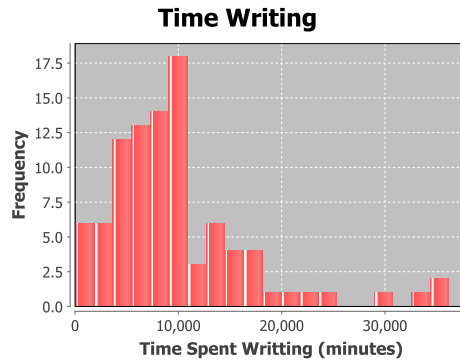


Figure .532: Histogram of the amount of time spent writing by each student on homework four problem seven for the 2012 course offering.

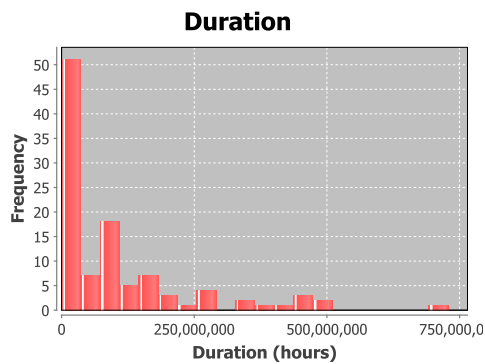


Figure .533: Histogram of the duration, time from first to last pen stroke for all students for homework five problem one for the 2012 course offering.

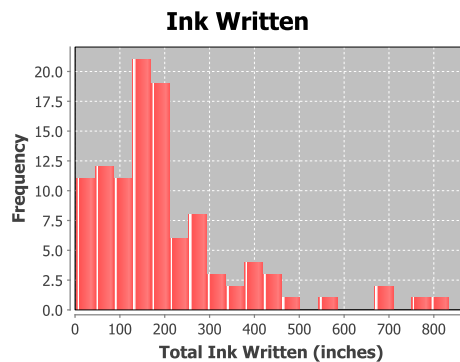


Figure .534: Histogram of the amount of ink, in inches, written, for each student on homework five problem one for the 2012 course offering.

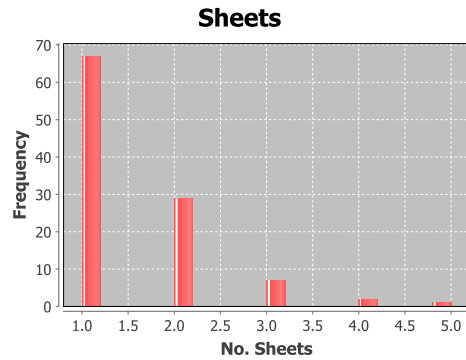


Figure .535: Histogram of the number of pages written on by each student to solve homework five problem one for the 2012 course offering.

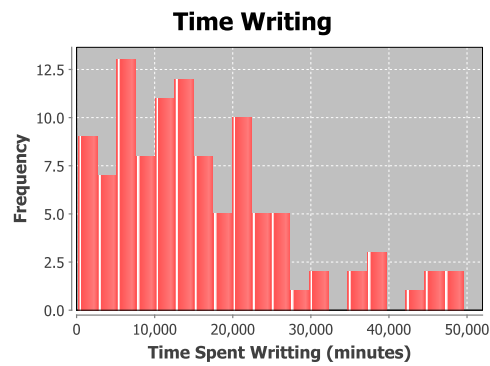


Figure .536: Histogram of the amount of time spent writing by each student on homework five problem one for the 2012 course offering.

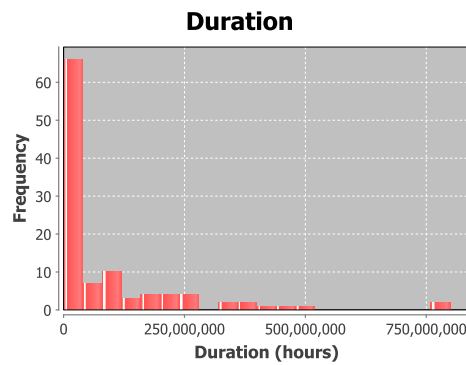


Figure .537: Histogram of the duration, time from first to last pen stroke for all students for homework five problem two for the 2012 course offering.

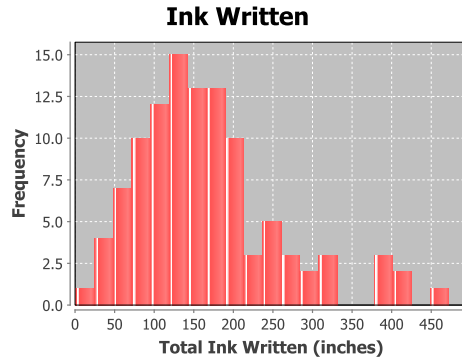


Figure .538: Histogram of the amount of ink, in inches, written, for each student on homework five problem two for the 2012 course offering.

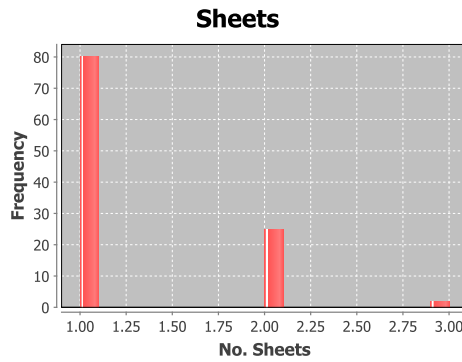


Figure .539: Histogram of the number of pages written on by each student to solve homework five problem two for the 2012 course offering.

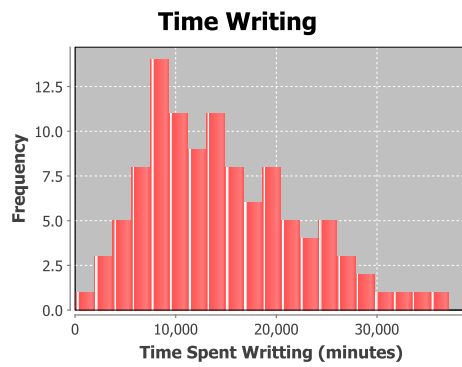


Figure .540: Histogram of the amount of time spent writing by each student on homework five problem two for the 2012 course offering.

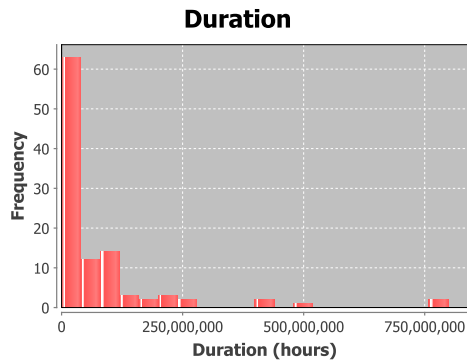


Figure .541: Histogram of the duration, time from first to last pen stroke for all students for homework five problem three for the 2012 course offering.

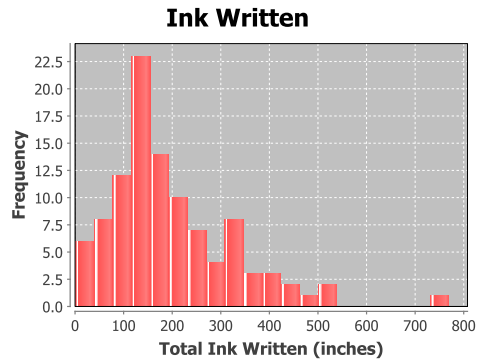


Figure .542: Histogram of the amount of ink, in inches, written, for each student on homework five problem three for the 2012 course offering.

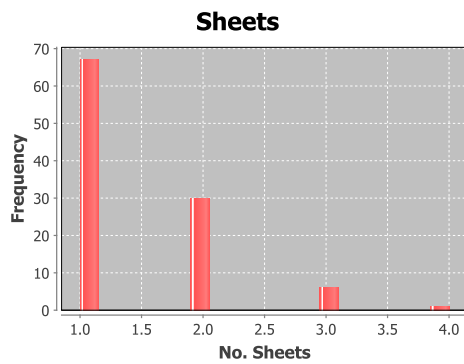


Figure .543: Histogram of the number of pages written on by each student to solve homework five problem three for the 2012 course offering.

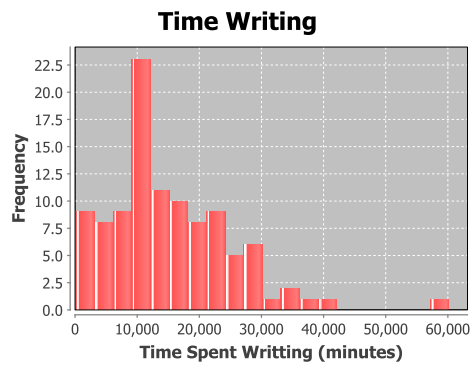


Figure .544: Histogram of the amount of time spent writing by each student on homework five problem three for the 2012 course offering.

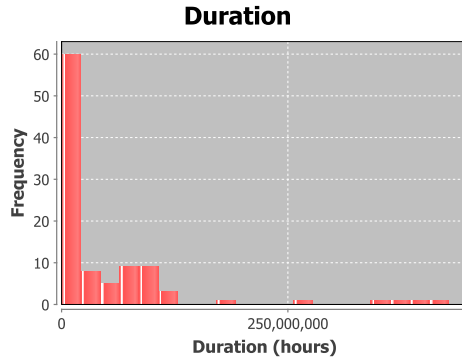


Figure .545: Histogram of the duration, time from first to last pen stroke for all students for homework five problem four for the 2012 course offering.

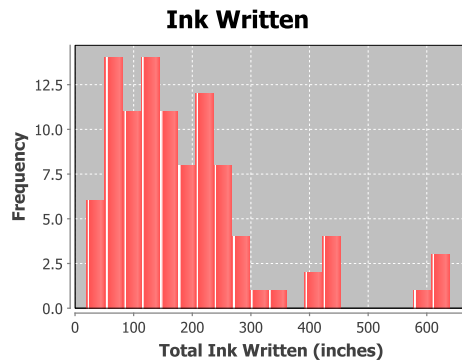


Figure .546: Histogram of the amount of ink, in inches, written, for each student on homework five problem four for the 2012 course offering.

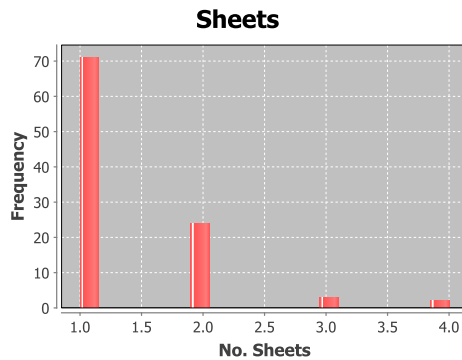


Figure .547: Histogram of the number of pages written on by each student to solve homework five problem four for the 2012 course offering.

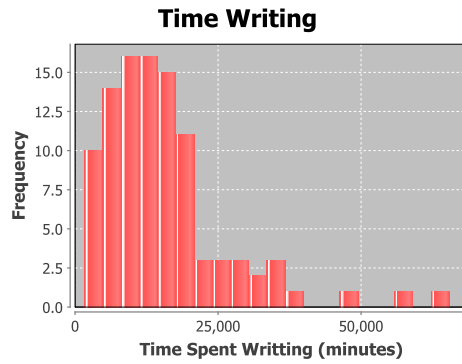


Figure .548: Histogram of the amount of time spent writing by each student on homework five problem four for the 2012 course offering.

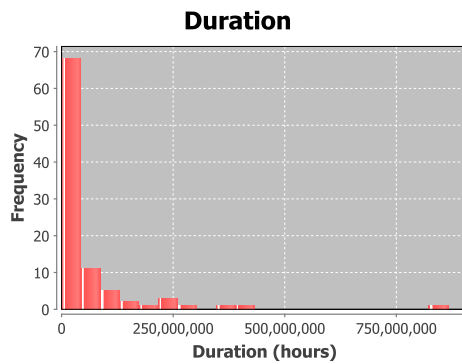


Figure .549: Histogram of the duration, time from first to last pen stroke for all students for homework five problem five for the 2012 course offering.

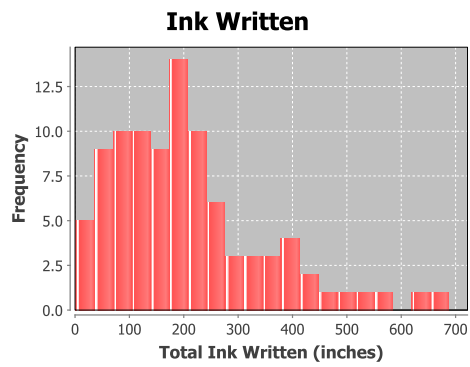


Figure .550: Histogram of the amount of ink, in inches, written, for each student on homework five problem five for the 2012 course offering.

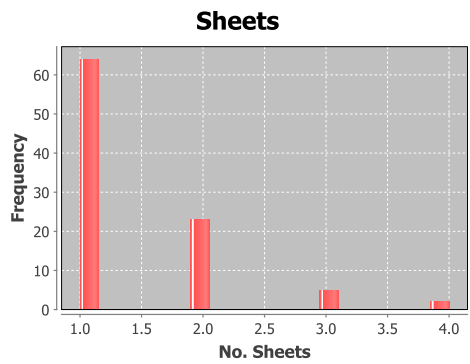


Figure .551: Histogram of the number of pages written on by each student to solve homework five problem five for the 2012 course offering.

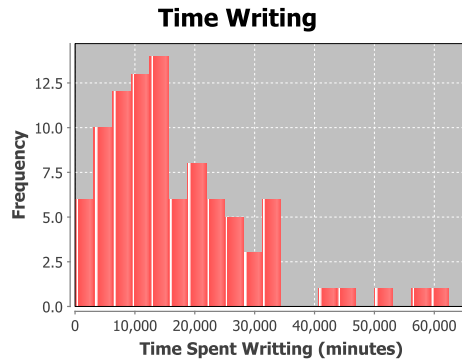


Figure .552: Histogram of the amount of time spent writing by each student on homework five problem five for the 2012 course offering.

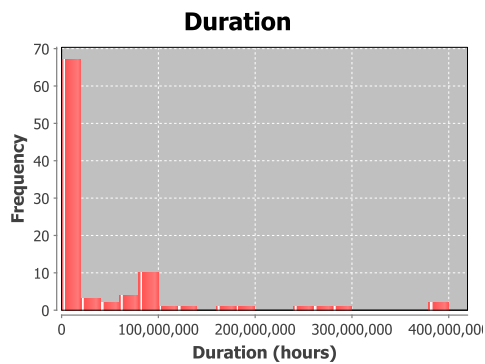


Figure .553: Histogram of the duration, time from first to last pen stroke for all students for homework five problem six for the 2012 course offering.

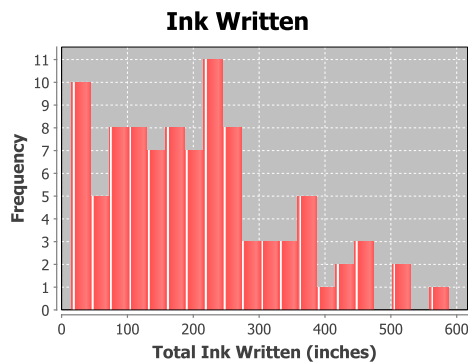


Figure .554: Histogram of the amount of ink, in inches, written, for each student on homework five problem six for the 2012 course offering.

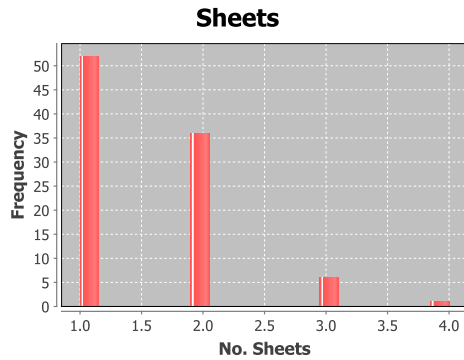


Figure .555: Histogram of the number of pages written on by each student to solve homework five problem six for the 2012 course offering.

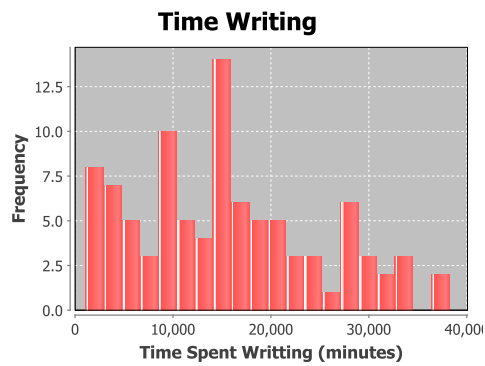


Figure .556: Histogram of the amount of time spent writing by each student on homework five problem six for the 2012 course offering.

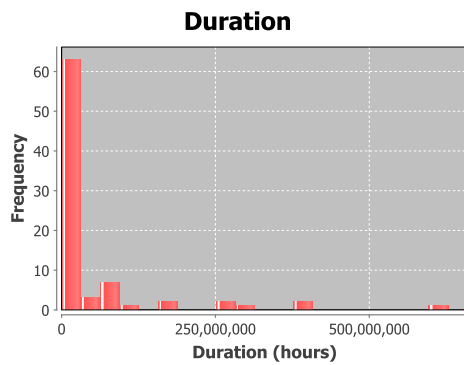


Figure .557: Histogram of the duration, time from first to last pen stroke for all students for homework five problem seven for the 2012 course offering.

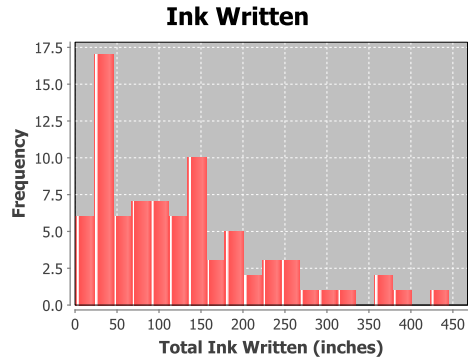


Figure .558: Histogram of the amount of ink, in inches, written, for each student on homework five problem seven for the 2012 course offering.

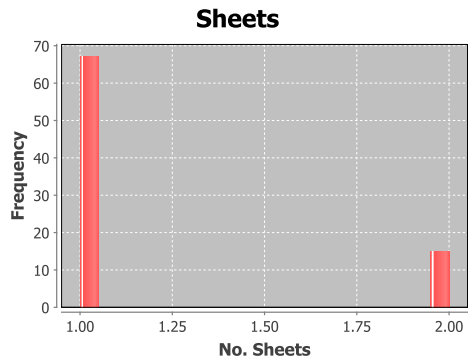


Figure .559: Histogram of the number of pages written on by each student to solve homework five problem seven for the 2012 course offering.

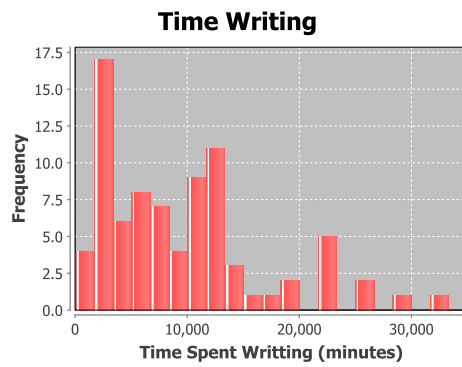
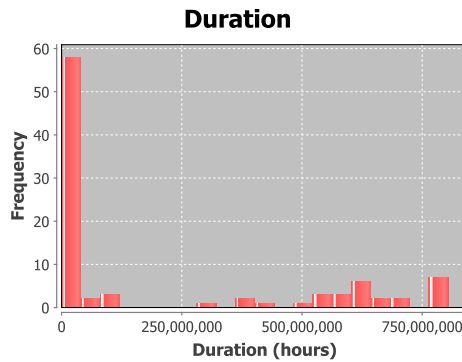


Figure .560: Histogram of the amount of time spent writing by each student on homework five problem seven for the 2012 course offering.



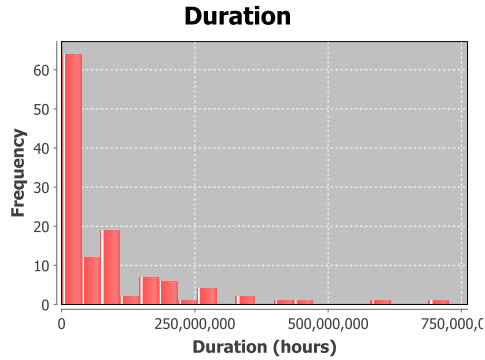


Figure .565: Histogram of the duration, time from first to last pen stroke for all students for homework six problem one for the 2012 course offering.

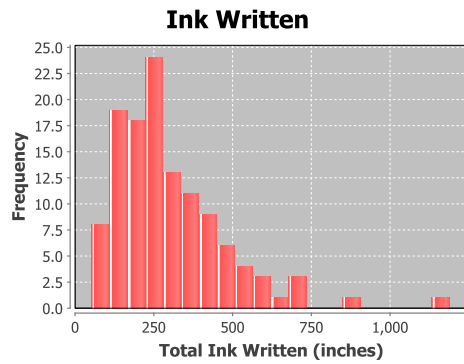


Figure .566: Histogram of the amount of ink, in inches, written, for each student on homework six problem one for the 2012 course offering.

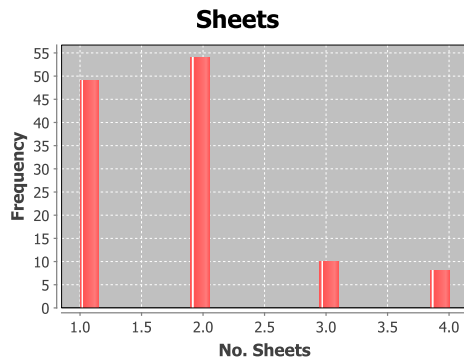


Figure .567: Histogram of the number of pages written on by each student to solve homework six problem one for the 2012 course offering.

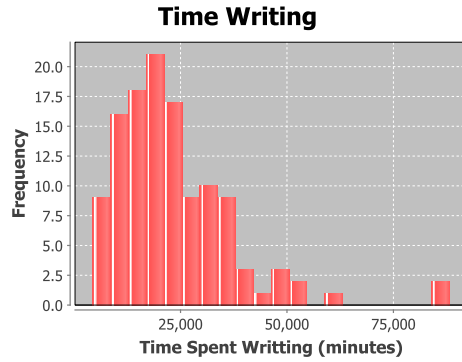


Figure .568: Histogram of the amount of time spent writing by each student on homework six problem one for the 2012 course offering.

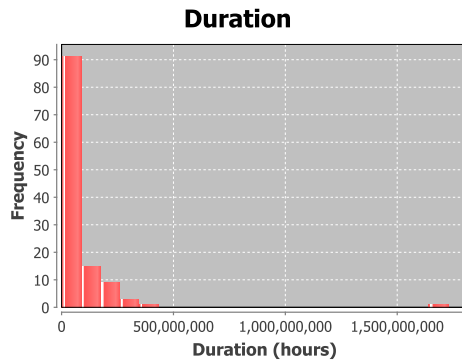


Figure .569: Histogram of the duration, time from first to last pen stroke for all students for homework six problem two for the 2012 course offering.

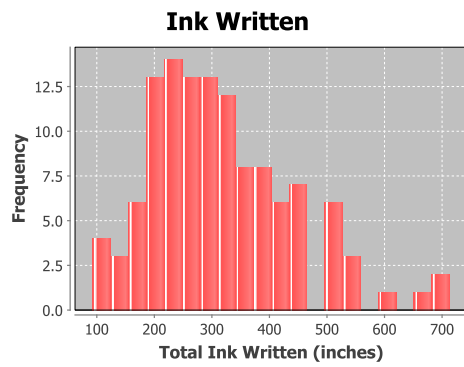


Figure .570: Histogram of the amount of ink, in inches, written, for each student on homework six problem two for the 2012 course offering.

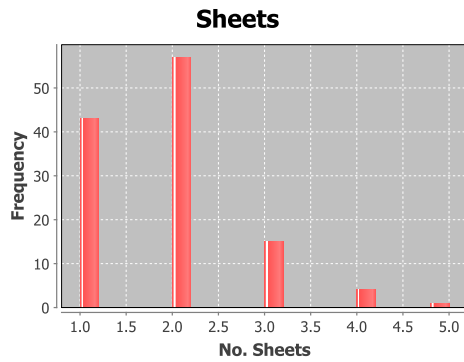


Figure .571: Histogram of the number of pages written on by each student to solve homework six problem two for the 2012 course offering.

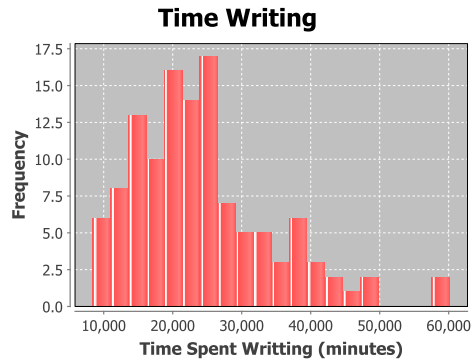


Figure .572: Histogram of the amount of time spent writing by each student on homework six problem two for the 2012 course offering.

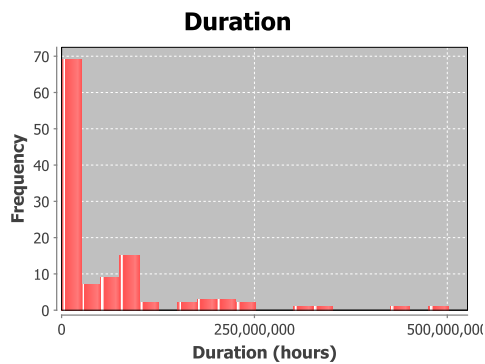


Figure .573: Histogram of the duration, time from first to last pen stroke for all students for homework six problem three for the 2012 course offering.

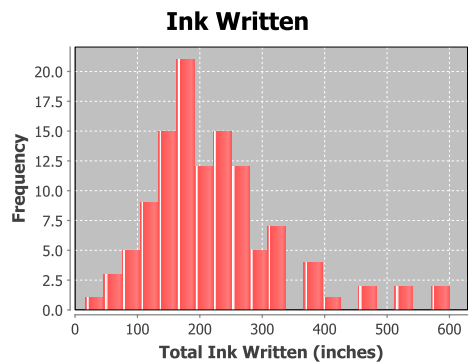


Figure .574: Histogram of the amount of ink, in inches, written, for each student on homework six problem three for the 2012 course offering.

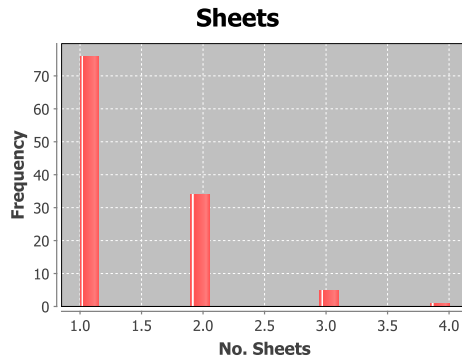


Figure .575: Histogram of the number of pages written on by each student to solve homework six problem three for the 2012 course offering.

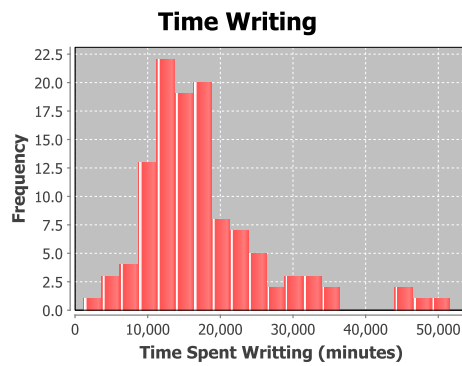


Figure .576: Histogram of the amount of time spent writing by each student on homework six problem three for the 2012 course offering.

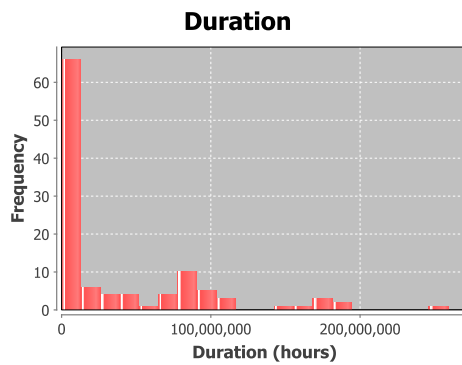


Figure .577: Histogram of the duration, time from first to last pen stroke for all students for homework six problem four for the 2012 course offering.

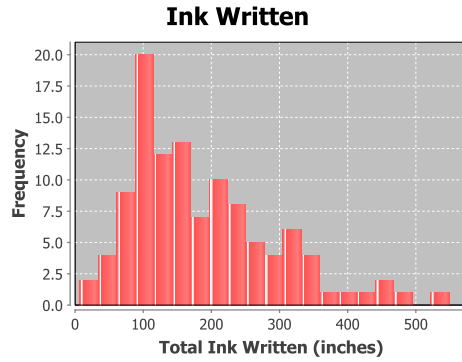


Figure .578: Histogram of the amount of ink, in inches, written, for each student on homework six problem four for the 2012 course offering.

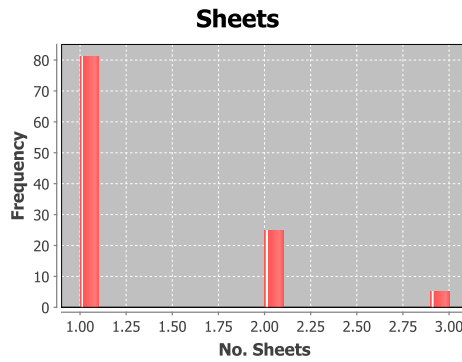


Figure .579: Histogram of the number of pages written on by each student to solve homework six problem four for the 2012 course offering.

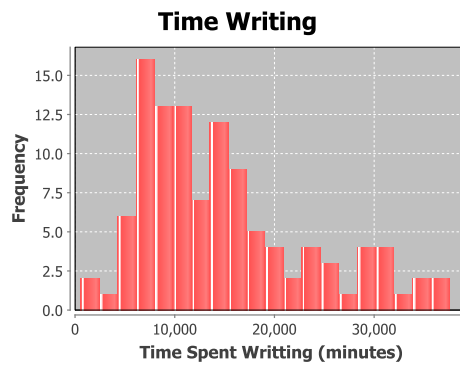


Figure .580: Histogram of the amount of time spent writing by each student on homework six problem four for the 2012 course offering.

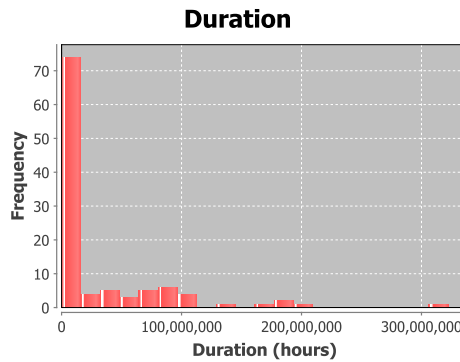


Figure .581: Histogram of the duration, time from first to last pen stroke for all students for homework six problem five for the 2012 course offering.

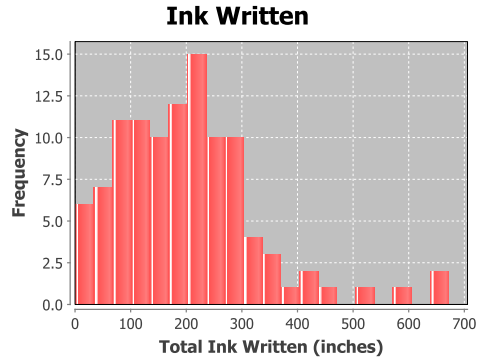


Figure .582: Histogram of the amount of ink, in inches, written, for each student on homework six problem five for the 2012 course offering.

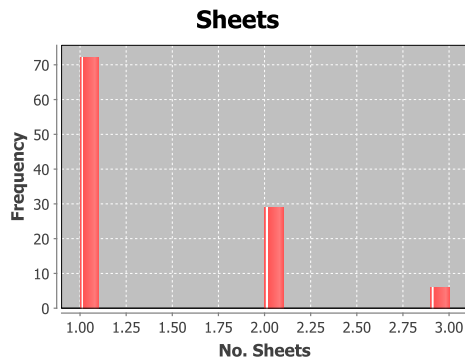


Figure .583: Histogram of the number of pages written on by each student to solve homework six problem five for the 2012 course offering.

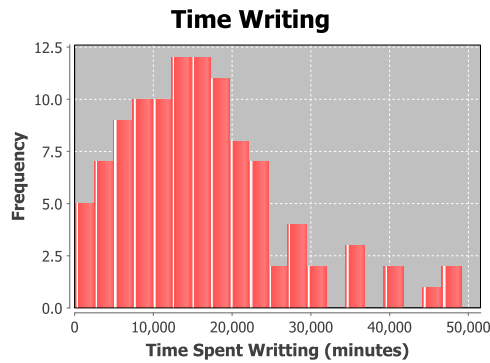


Figure .584: Histogram of the amount of time spent writing by each student on homework six problem five for the 2012 course offering.

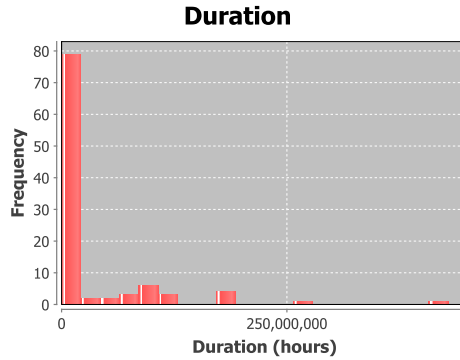


Figure .585: Histogram of the duration, time from first to last pen stroke for all students for homework six problem six for the 2012 course offering.

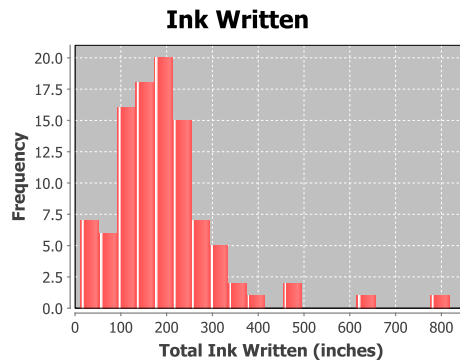


Figure .586: Histogram of the amount of ink, in inches, written, for each student on homework six problem six for the 2012 course offering.

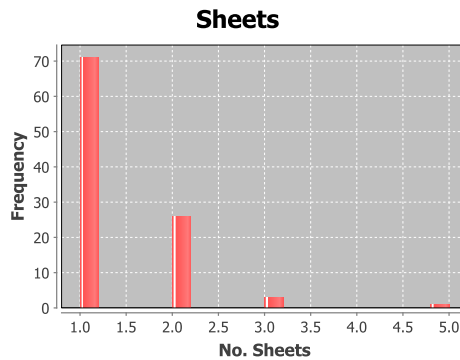


Figure .587: Histogram of the number of pages written on by each student to solve homework six problem six for the 2012 course offering.

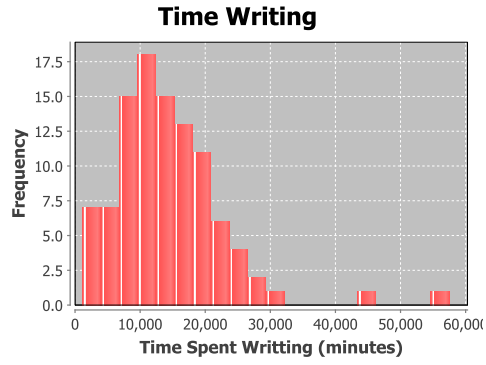


Figure .588: Histogram of the amount of time spent writing by each student on homework six problem six for the 2012 course offering.

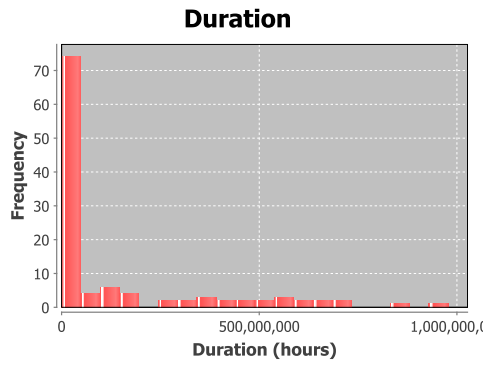


Figure .589: Histogram of the duration, time from first to last pen stroke for all students for homework seven problem one for the 2012 course offering.

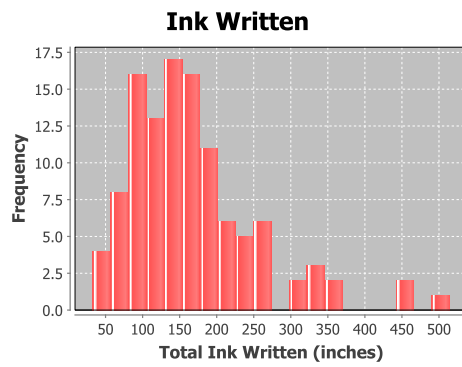


Figure .590: Histogram of the amount of ink, in inches, written, for each student on homework seven problem one for the 2012 course offering.

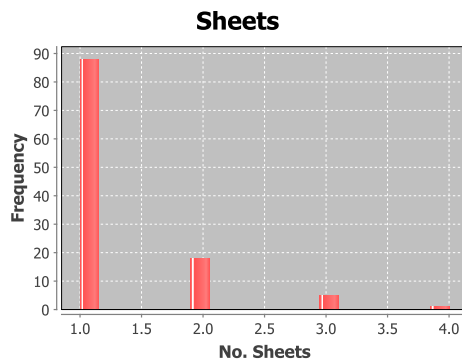


Figure .591: Histogram of the number of pages written on by each student to solve homework seven problem one for the 2012 course offering.

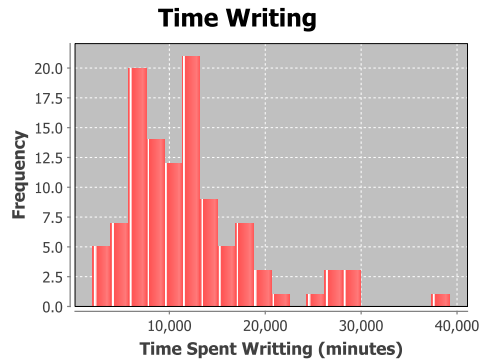


Figure .592: Histogram of the amount of time spent writing by each student on homework seven problem one for the 2012 course offering.

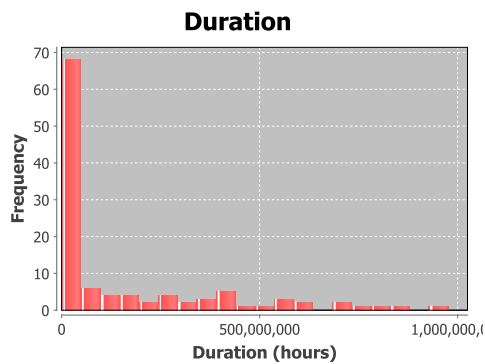


Figure .593: Histogram of the duration, time from first to last pen stroke for all students for homework seven problem two for the 2012 course offering.

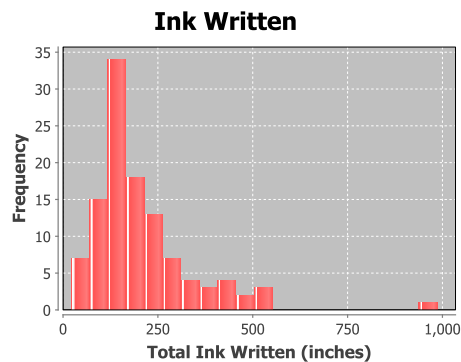


Figure .594: Histogram of the amount of ink, in inches, written, for each student on homework seven problem two for the 2012 course offering.

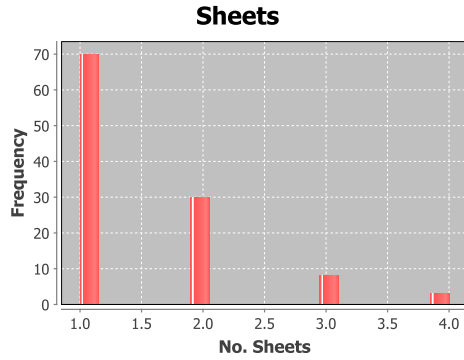


Figure .595: Histogram of the number of pages written on by each student to solve homework seven problem two for the 2012 course offering.

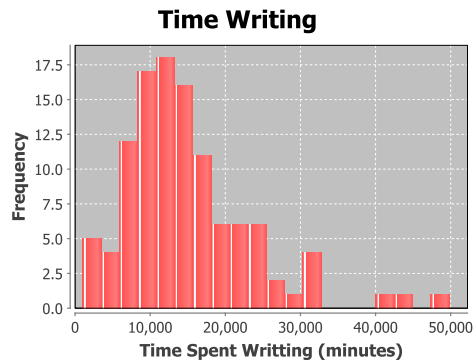


Figure .596: Histogram of the amount of time spent writing by each student on homework seven problem two for the 2012 course offering.

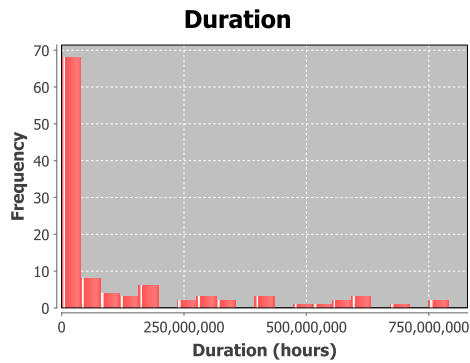


Figure .597: Histogram of the duration, time from first to last pen stroke for all students for homework seven problem three for the 2012 course offering.

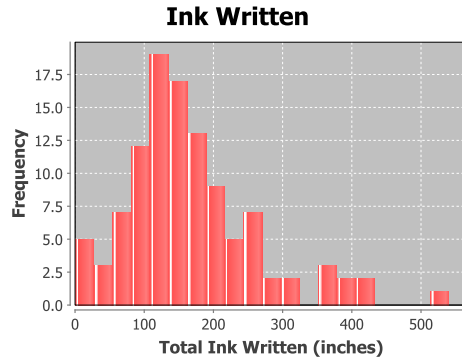


Figure .598: Histogram of the amount of ink, in inches, written, for each student on homework seven problem three for the 2012 course offering.

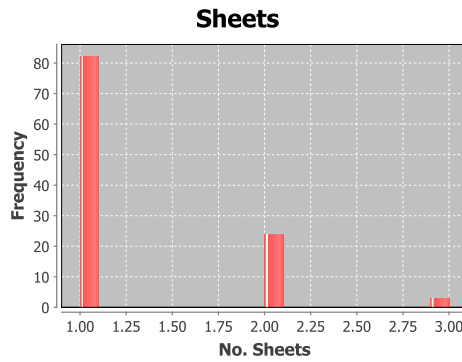


Figure .599: Histogram of the number of pages written on by each student to solve homework seven problem three for the 2012 course offering.

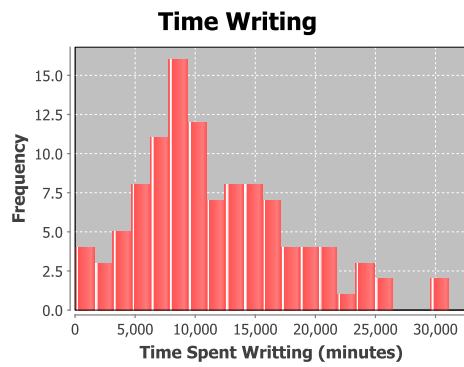


Figure .600: Histogram of the amount of time spent writing by each student on homework seven problem three for the 2012 course offering.

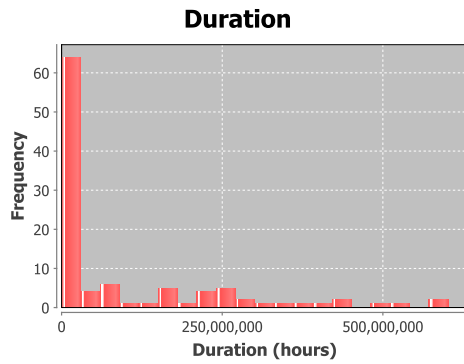


Figure .601: Histogram of the duration, time from first to last pen stroke for all students for homework seven problem four for the 2012 course offering.

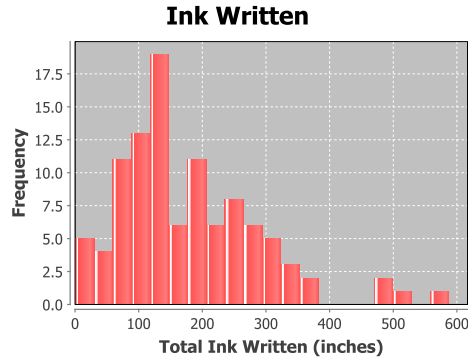


Figure .602: Histogram of the amount of ink, in inches, written, for each student on homework seven problem four for the 2012 course offering.

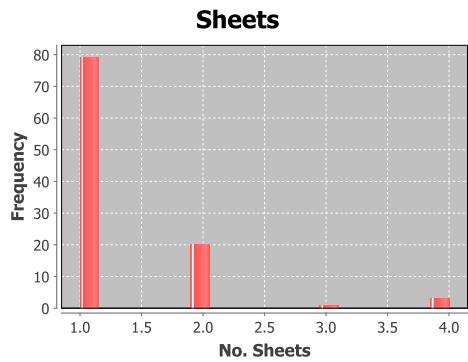


Figure .603: Histogram of the number of pages written on by each student to solve homework seven problem four for the 2012 course offering.

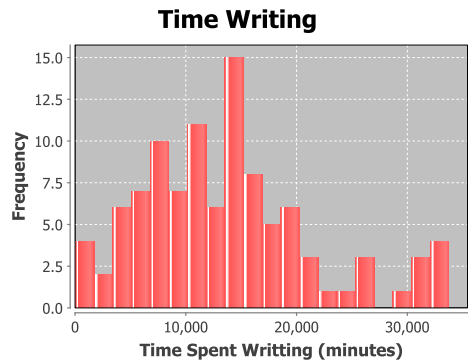


Figure .604: Histogram of the amount of time spent writing by each student on homework seven problem four for the 2012 course offering.

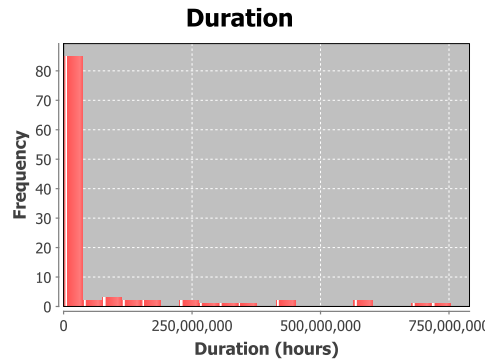


Figure .605: Histogram of the duration, time from first to last pen stroke for all students for homework seven problem five for the 2012 course offering.

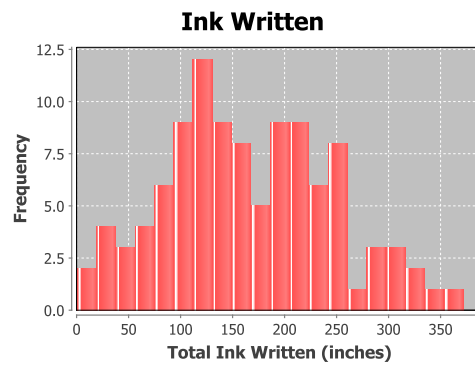


Figure .606: Histogram of the amount of ink, in inches, written, for each student on homework seven problem five for the 2012 course offering.

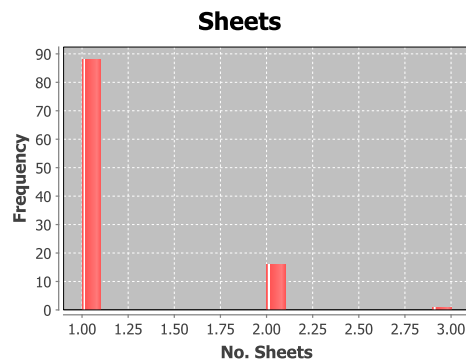


Figure .607: Histogram of the number of pages written on by each student to solve homework seven problem five for the 2012 course offering.

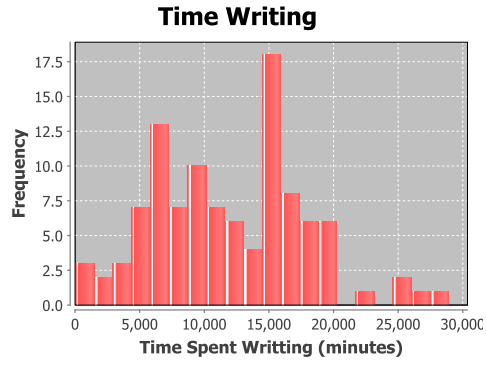


Figure .608: Histogram of the amount of time spent writing by each student on homework seven problem five for the 2012 course offering.

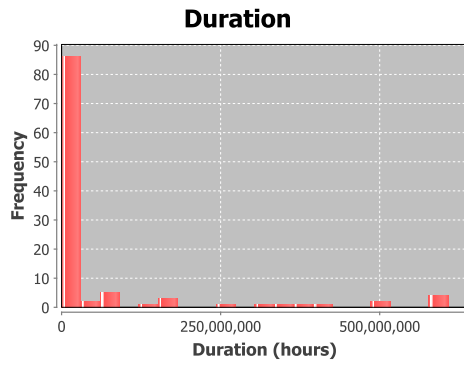


Figure .609: Histogram of the duration, time from first to last pen stroke for all students for homework seven problem six for the 2012 course offering.

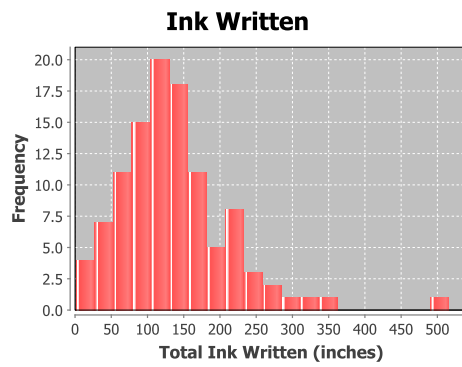


Figure .610: Histogram of the amount of ink, in inches, written, for each student on homework seven problem six for the 2012 course offering.

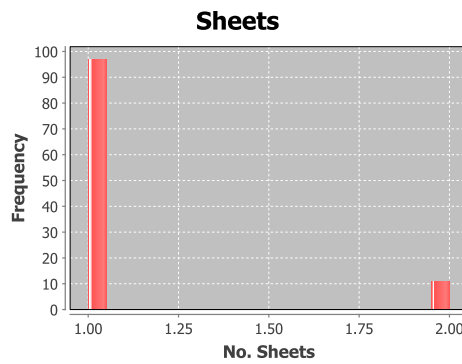


Figure .611: Histogram of the number of pages written on by each student to solve homework seven problem six for the 2012 course offering.

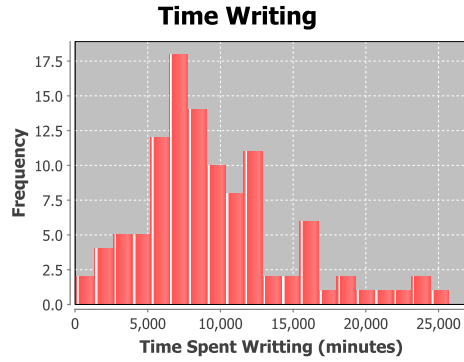


Figure .612: Histogram of the amount of time spent writing by each student on homework seven problem six for the 2012 course offering.

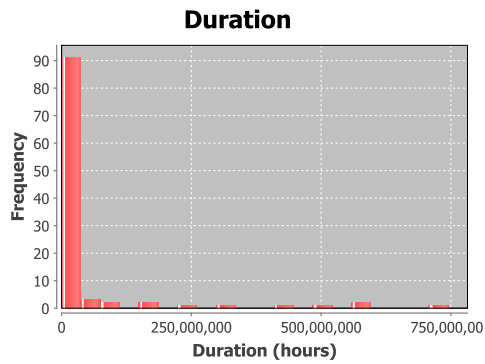


Figure .613: Histogram of the duration, time from first to last pen stroke for all students for homework seven problem one for the 2012 course offering.

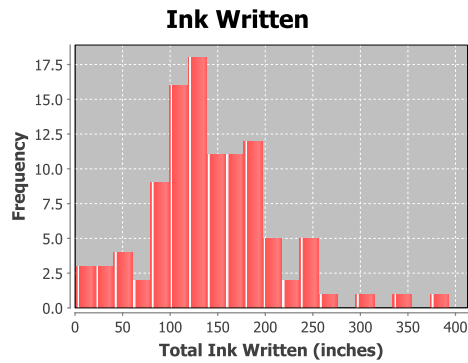


Figure .614: Histogram of the amount of ink, in inches, written, for each student on homework seven problem seven for the 2012 course offering.

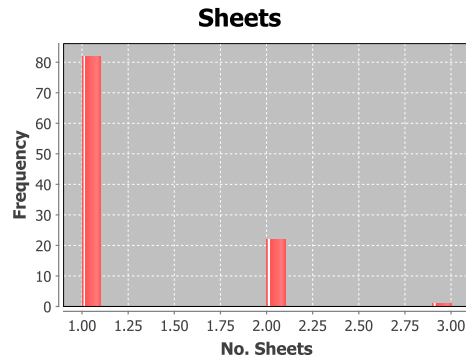


Figure .615: Histogram of the number of pages written on by each student to solve homework seven problem seven for the 2012 course offering.

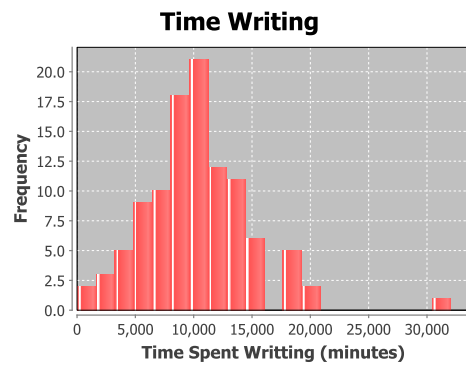


Figure .616: Histogram of the amount of time spent writing by each student on homework seven problem seven for the 2012 course offering.

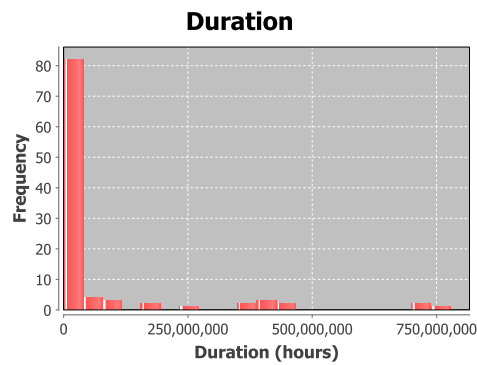


Figure .617: Histogram of the duration, time from first to last pen stroke for all students for homework eight problem one for the 2012 course offering.

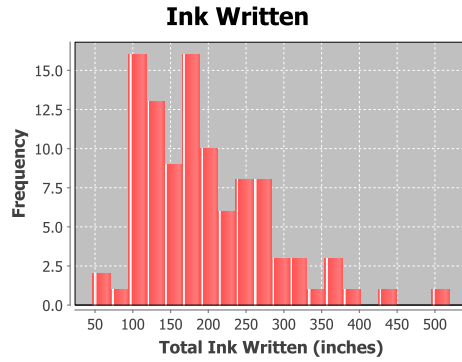


Figure .618: Histogram of the amount of ink, in inches, written, for each student on homework eight problem one for the 2012 course offering.

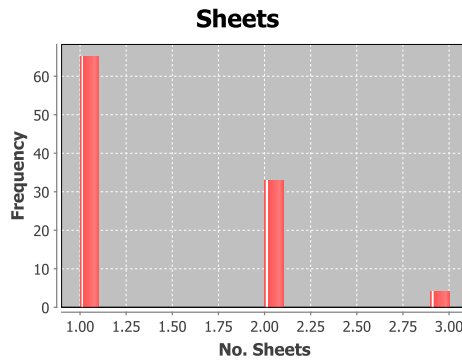


Figure .619: Histogram of the number of pages written on by each student to solve homework eight problem one for the 2012 course offering.

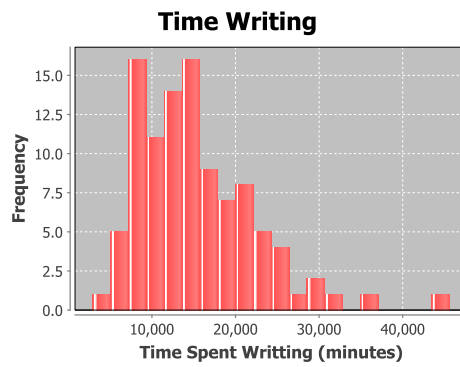


Figure .620: Histogram of the amount of time spent writing by each student on homework eight problem one for the 2012 course offering.

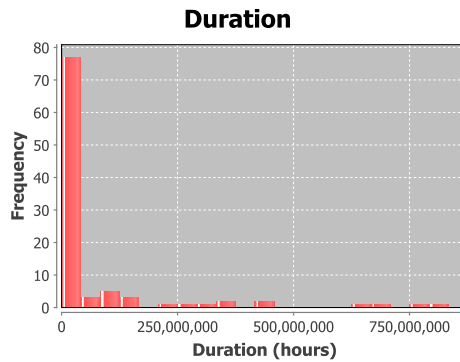


Figure .621: Histogram of the duration, time from first to last pen stroke for all students for homework eight problem two for the 2012 course offering.

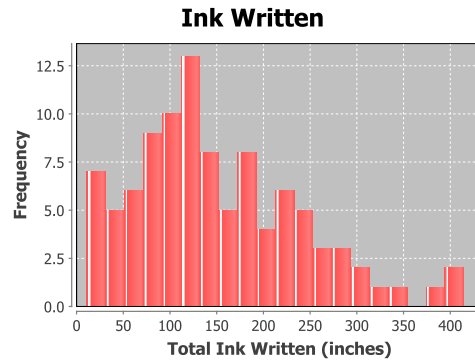


Figure .622: Histogram of the amount of ink, in inches, written, for each student on homework eight problem two for the 2012 course offering.

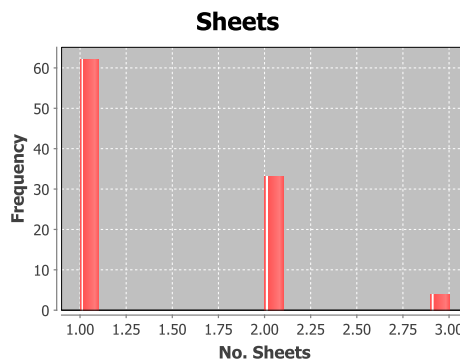


Figure .623: Histogram of the number of pages written on by each student to solve homework eight problem two for the 2012 course offering.

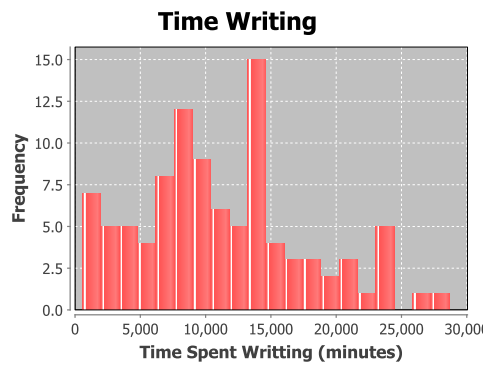


Figure .624: Histogram of the amount of time spent writing by each student on homework eight problem two for the 2012 course offering.

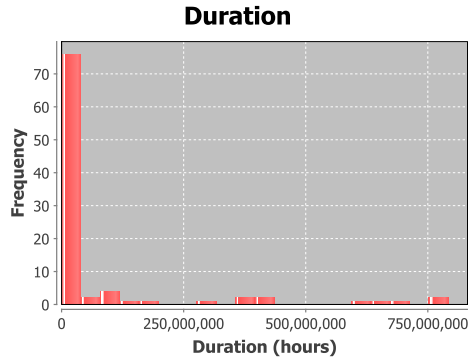


Figure .625: Histogram of the duration, time from first to last pen stroke for all students for homework eight problem three for the 2012 course offering.

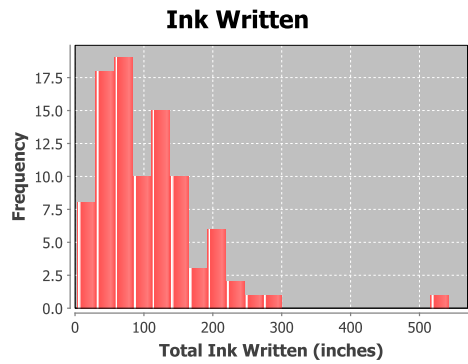


Figure .626: Histogram of the amount of ink, in inches, written, for each student on homework eight problem three for the 2012 course offering.

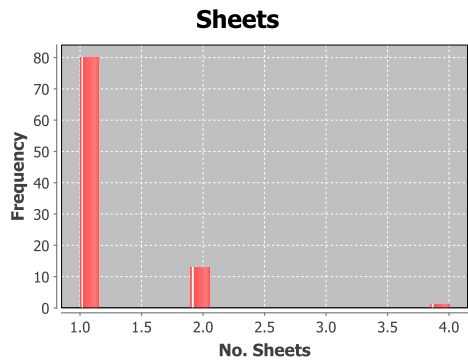


Figure .627: Histogram of the number of pages written on by each student to solve homework eight problem three for the 2012 course offering.

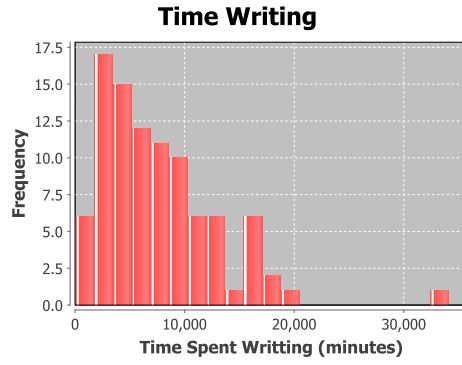


Figure .628: Histogram of the amount of time spent writing by each student on homework eight problem three for the 2012 course offering.

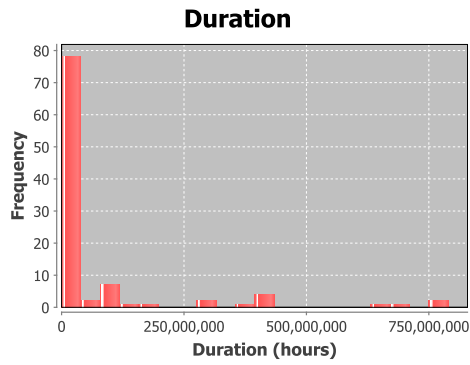


Figure .629: Histogram of the duration, time from first to last pen stroke for all students for homework eight problem four for the 2012 course offering.

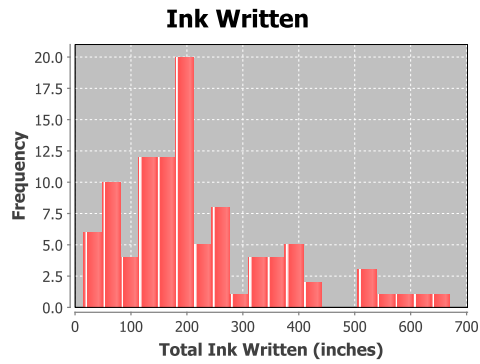


Figure .630: Histogram of the amount of ink, in inches, written, for each student on homework eight problem four for the 2012 course offering.

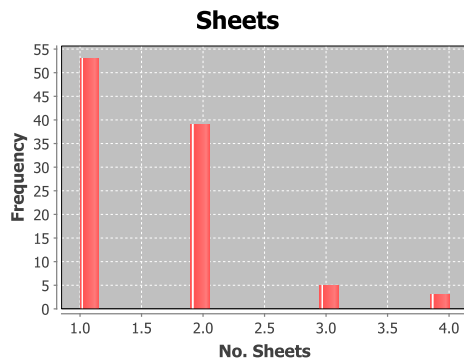


Figure .631: Histogram of the number of pages written on by each student to solve homework eight problem four for the 2012 course offering.

Algorithm 1 Transforms a candidate stroke into a template. First the stroke is resampled via RESAMPLE, it is then converted to a one-dimensional vector via C-DISTANCE. Last the vector is normalized via Z-NORMALIZE.

function RESAMPLE(*points*, n)

$I \leftarrow \text{PATH-LENGTH}(\mathit{points}) / (n - 1)$

$D \leftarrow 0$

$\mathit{newPoints} \leftarrow \mathit{points}_0$

for all point p_i for $i \geq 1$ in *points* **do**

if $D + d \geq I$ **then**

$q_x \leftarrow p_{i-1_x} + ((I - D)/d) \times (p_{i_x} - p_{i-1_x})$

$q_y \leftarrow p_{i-1_y} + ((I - D)/d) \times (p_{i_y} - p_{i-1_y})$

 APPEND(*newPoints*, q)

 INSERT(*points*, i , q)

$D \leftarrow 0$

else

$D \leftarrow D + d$

end if

end for

return *newPoints*

end function

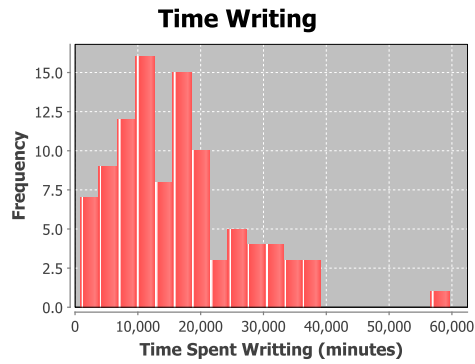


Figure .632: Histogram of the amount of time spent writing by each student on homework eight problem four for the 2012 course offering.

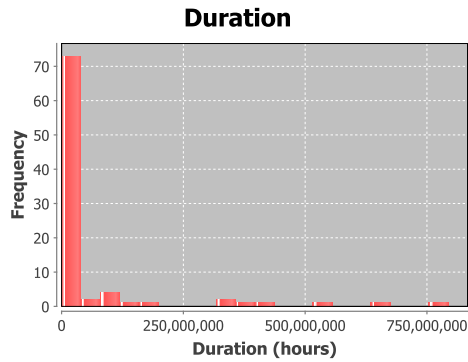


Figure .633: Histogram of the duration, time from first to last pen stroke for all students for homework eight problem five for the 2012 course offering.

Appendix - Pseudocode

The following pseudocode has been implemented in JAVA and is available at:

<https://sourceforge.net/projects/onecentrec/>

Algorithm 2 Computes the path length of a stroke, computed as the sum of Euclidean distances between each pair of consecutive points.

function PATH-LENGTH(\mathbf{A})

$d \leftarrow 0$

for i **from** 1 **to** $|\mathbf{A}|$ **step** 1 **do**

$d \leftarrow d + \text{DISTANCE}(A_{i-1}, A_i)$

end for

return d

end function

Algorithm 3 Computes the centroid of a pen stroke.

function C-DISTANCE(\mathbf{points}) ▷ Equation 10.3.3.2

$c \leftarrow \text{CENTROID}(\mathbf{points})$ ▷ Computes (μ_x, μ_y)

for all point p **in** \mathbf{points} **do**

$d \leftarrow \text{DISTANCE}(c, p)$

APPEND($\mathbf{distances}, d$)

end for

return $\mathbf{distances}$

end function

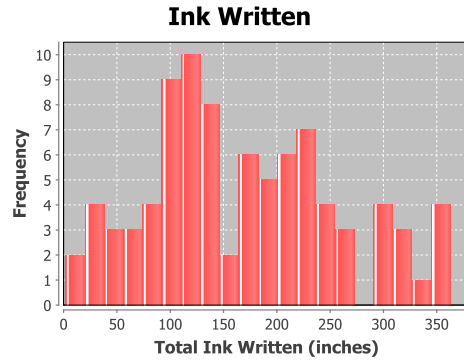


Figure .634: Histogram of the amount of ink, in inches, written, for each student on homework eight problem five for the 2012 course offering.

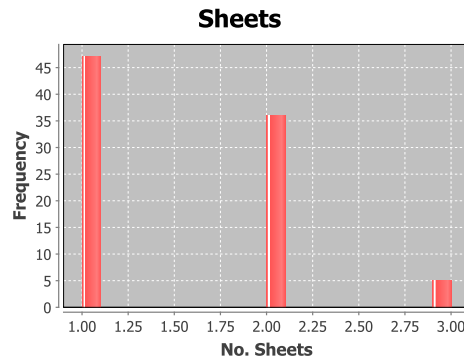


Figure .635: Histogram of the number of pages written on by each student to solve homework eight problem five for the 2012 course offering.

Algorithm 4 Z-normalizes a vector of real values by subtracting each element in the vector by the mean and then dividing by the standard deviation.

function Z-NORMALIZE(S) ▷ Equation 10.3.3.2

$\mu = \text{AVERAGE}(S)$

$\sigma = \text{STANDARD-DEVIATION}(S)$

for all d **in** S **do**

$z \leftarrow (d - \mu) / \sigma$

APPEND(z, d_z)

end for

return d_z

end function

Algorithm 5 The following functions comprise all the steps needed to recognize a candidate stroke that has been converted into a template, given a set of training templates.

function RECOGNIZE(S , $Templates$) ▷ Equation 10.3.3.3

for all template T in $Templates$ **do**

$b \leftarrow +\infty$

$d \leftarrow L2(S, T)$

if $d < b$ **then**

$b \leftarrow d$

$T^* \leftarrow T$

end if

end for

return T^*

end function

function L2(S , T) ▷ Equation 10.3.3.3

$d \leftarrow 0$

for all s_i, t_i for $i \geq 1$ in S, T **do**

$d \leftarrow d + (s_i - t_i)^2$

end for

return d

end function
