Reproducibility in the Classroom

Mine Dogucu[1]

[1] Department of Statistics, University of California Irvine, United States, 92697; email: mdogucu@uci.edu

Xxxx. Xxx. Xxx. Xxx. YYYY. AA:1–19

https://doi.org/10.1146/((please add article doi))

Keywords

reproducibility education, undergraduate, graduate, teaching reproducibility, reproducible teaching, literate programming, version control

Abstract

Difficulties in reproducing results from scientific studies have lately been referred to as “reproducibility crisis”. Scientific practice depends heavily on scientific training. What gets taught in the classroom is often practiced in labs, fields, and data analysis. The importance of reproducibility in the classroom has gained momentum in statistics education in recent years. In this manuscript, we review the existing literature on reproducibility education. We delve into the relationship between computing tools and reproducibility through visiting historical developments in this area. We share examples for teaching reproducibility and reproducible teaching while discussing the pedagogical opportunities created by these examples as well as challenges that the instructors should be aware of. We detail the use of teaching reproducibility and reproducible teaching practices in an introductory data science course. Lastly, we provide recommendations on reproducibility education for instructors, administrators, and other members of the scientific community.

Contents

1. INTRODUCTION

Difficulties in reproducing results from scientific studies have lately been referred to as "reproducibility crisis". Despite varying definitions of reproducibility across different disciplines, one common definition of reproducibility is the process of obtaining the same results using the same input data and following the same computational steps pertaining to methods and analysis (National Academies of Sciences, Engineering, and Medicine 2019). Reproducibility is often associated with scientific publications but it is as important in industrial data science (Donoho 2017).

Although often used interchangeably, it is important to note the distinction between reproducibility and replicability. While reproducibility focuses on using the same data and following the same computational steps, replicability focuses on using the same research design and following the same set of methods to corroborate results. Even though both reproducibility and replicability are integral parts of science, in this manuscript, we mainly focus on reproducibility.

To clarify reproducibility further, we can consider Leek and Jager's (2017) summary as an example. They summarize the assessment of reproducibility of a scientific study with two questions "Are code and data both readily available?" and "Does running the provided code exactly produce the study results?"

The importance of reproducibility is prevalent in many disciplines including but not limited to psychology (Open Science Collaboration 2015), physics (Junk & Lyons 2020), economics (Chang & Li 2015), and life sciences (Miyakawa 2020). Despite a consensus on the importance of reproducibility, there is not a consensus on whether there is a "reproducibility crisis". Some scholars argue that the issues with reproducibility are not necessarily at a "crisis" level and do not distort the majority of the scientific literature (Fanelli 2018). Regardless of whether it is a "crisis" or not, reproducibility is a major principle of the scientific method.

Scientific practice depends heavily on scientific training. What gets taught in the classroom is often practiced in labs, fields, and data analysis. Considering that teaching and research practices go hand-in-hand, many have advocated teaching reproducibility methods at the undergraduate and graduate statistics and data sciences classes (Ball 2023, Karathanasis et al. 2022). Data science education that incorporates reproducibility in data analysis is considered the "statistical counterattack" to the reproducibility crisis (Peng 2015). The importance of reproducibility in the classroom has gained momentum in statistics education in recent years and *Journal of Statistics and Data Science Education* has published a special issue on the topic (Horton et al. 2022).

As reproducibility gained attention in recent years both in scientific practice and training, the focus has been on two main points: 1) conducting scientific research in a reproducible manner and 2) teaching reproducible research practices. The first point mainly concerns scientific practice and the latter concerns scientific training. These can be named as **reproducible research** and **teaching reproducibility** respectively. More recently, Dogucu and Çetinkaya-Rundel (2022) have drawn attention to a third point: **reproducible teaching**.

In a statistics program, whether at the undergraduate or graduate level, these three points can be brought together. Some examples can include the following.

- Reproducible research: students can be given opportunities to use reproducible practices in a research setting such as capstone projects or research assistant appointments.
- Teaching reproducibility: instructors can teach on reproducible workflows and students can become familiar with the toolkit that can support a reproducible workflow.
- Reproducible teaching: instructors can prepare their teaching materials using a reproducible workflow.

There are many benefits of teaching reproducibility. Training in reproducible workflows can improve the scientific value of students' work. Computational skills associated with reproducibility can prepare the students for the workforce and training in reproducibility skills can lead to outcomes in intellectual development and citizenship (Ball et al 2022). In addition, reproducibility education can contribute to equity in science. For instance, a blind data scientist who is unable to see a data visualization, can reproduce the visualization in an accessible format (e.g., audio or tactile) if a data analysis project provides open-access data and code (Seo & Dogucu 2024).

There are also many benefits to reproducible teaching. Practices associated with reproducibility can make course management easier especially for large classes. If instructors adopt the same workflow for reproducible teaching and teaching reproducibility, then they can serve as role models for their students. Modifications of teaching materials can be easier. If provided open access in the public domain, materials can also remove barriers for students and provide instructors opportunities to learn from each others' materials (Dogucu & Çetinkaya-Rundel 2022).

In this manuscript, we review the existing literature on reproducibility education. In Section 2, we delve into the relationship between computing tools and reproducibility through visiting historical developments in this area. In Sections 3 and 4 we share examples for teaching reproducibility and reproducible teaching respectively while discussing the pedagogical opportunities created by these examples as well as challenges that the instructors should be aware of. In Section 5 we share examples from our introductory data science course. Lastly, in Section 6, we provide recommendations on reproducibility education for instructors, administrators, and other members of the scientific community.

2. HISTORY OF COMPUTING TOOLS IN THE STATISTICS CLASSROOM

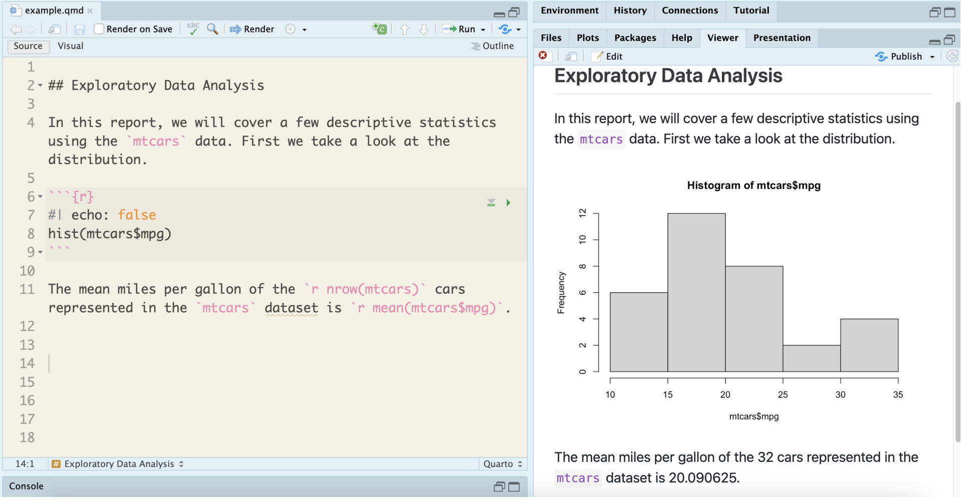
Statistics courses have historically been supported by computing tools. Carmer and Cady state the use of eight different programs that supported the teaching of statistics as early as in 1969. These programs, in today's language, would perhaps be called applications and include data generator as well as distribution plotter. In 1972, Abranovic, et al. mention that they prefer BASIC over FORTRAN in teaching and state that computers allow students to solve many statistical problems "without performing the tedious calculations by hand".

They state the use of BASIC for many purposes including but not limited to discriminant analysis, factor analysis, and Bayesian regression.

In the late 1970s, we can also see the use of many statistical programs in teaching that have passed the test of time including Minitab, SPSS, and SAS (Thisted 1979, Tubb & Ringer 1977). The programming language S was also developed in this decade by data analysis researchers at Bell Labs which then inspired the creation of the R language in the early 1990s (Chambers 2020). The R language is now one of the most popular technologies supporting the teaching and learning of statistics.

In the 1980s, the computing world has been introduced to literate programming. Knuth introduced this paradigm shift as rather than instructing a computer what to do, we should be explaining to humans what we want the computers to do (1984). The advances in literate programming have led to the creation of notebook programming environments where the narrative and computations supporting the data analysis are weaved together. Some examples include Quarto (Allaire & Dervieux 2024), R Markdown (Xie, Allaire & Grolemund 2018), and Jupyter Notebooks (Kluvyer et al. 2016). Literate programming is a distinct kind of programming that is essential for scientific researchers and data scientists (Kery et al. 2018). Notebook environments are essential in teaching reproducibility (Baumer et al. 2014) and reproducible teaching (Dogucu & Çetinkaya-Rundel 2022).

On the left half of Figure 1, a Quarto document titled `example.qmd` is opened in RStudio. On the right hand side of the figure, one can see the rendered version, i.e., the output of this document. For instance, lines 6 through 9 of the document renders the histogram shown in the output. In addition, line 11 weaves together text along with calculations of descriptive statistics which, on the output, is shown simply as text. Having text and R code in a single document avoids copy-paste errors that students could potentially be making if they were to conduct analysis separately from writing the output.



```
1
2  ## Exploratory Data Analysis
3
4  In this report, we will cover a few descriptive statistics
   using the `mtcars` data. First we take a look at the
   distribution.
5
6  ```{r}
7  #| echo: false
8  hist(mtcars$mpg)
9  ```
10
11 The mean miles per gallon of the `r nrow(mtcars)` cars
   represented in the `mtcars` dataset is `r mean(mtcars$mpg)`.
12
13
14
15
16
17
18
```

Exploratory Data Analysis

In this report, we will cover a few descriptive statistics using the `mtcars` data. First we take a look at the distribution.

The mean miles per gallon of the 32 cars represented in the `mtcars` dataset is 20.090625.

Figure 1

Example Quarto document on the left with its output view on the right

We can see that over the last six decades, statistical computing tools have evolved and statistics educators have been adopting these tools in their teaching. Some tools have been considered to be more accessible to novice learners. Historically, statistical computing tools have been classified as 1) tools for doing statistics and 2) tools for learning statistics (Biehler 1997). In terms of reproducibility, revisiting the three dimensions in Section 1 we can also consider tools for teaching statistics as part of this classification (Dogucu & Çetinkaya-Rundel 2022).

For instance, in a hypothetical statistics classroom, students may do computations with a calculator, the instructor may prepare their teaching materials with a slideshow software, and the instructor may use R for their research projects outside of the class. In a completely different scenario, students may learn R and Quarto for data analysis, the instructor may use R and Quarto for preparing their teaching materials, and the instructor may use R and Quarto for their research projects outside of the class. Considering reproducible research, teaching reproducibility, and reproducible teaching as three separate dimensions also allows us to examine a teacher-scholar's use of tools from these three separate perspectives. Thus a teacher-scholar's reproducibility toolkit and how much their students' are exposed to this toolkit can vary across these three dimensions.

When tools for doing statistics and learning statistics are compared two points stand out. Learning statistics tools such as Tinkerplots and Fathom are point-and-click and provide an accessible environment for novice learners. However, they lack reproducibility features. On the other hand, doing statistics tools such as R when supported with a literate programming notebook, provide more reproducibility features but come with a steeper learning curve (McNamara 2015). Reproducibility methods and principles can be taught by using many scriptable software including but not limited to R, Python, MATLAB, SAS, SPSS, and Stata (Ball 2023).

In addition to tools that support data analysis, statistics classrooms also benefit from version control tools. In early 2000s, distributed version control systems were developed which allow for tracking changes in any set of computer files. One such system is the Git system which is an essential tool for researchers to keep track of different versions of research files throughout the span of the research study. Researchers, in a way, can use Git repositories as a research journal (Huskey 2023). Git repositories (a.k.a., repos) contain a set of files that are version tracked and in a typical data science project they can include data, analysis scripts, and reports.

Consider a student working on a weekly task, analyzing data. Assume that this student is assigned to work specifically using a Quarto file titled `hw1.qmd`. Without a formal training in version control, students may be inclined to name files such as `hw1-version2.qmd`, `hw1-with-figures.qmd`, `hw1-final.qmd`, or `hw1-final2.qmd`. Considering the iterative nature of data analysis, students (and data analysts) need to track different version of files. Without a strong training in version control systems this can be quite chaotic.

Table 1 shows an hypothetical history of a project that two students are working on together, not necessarily simultaneously. Rather than having different names for the file `hw1.qmd`, changes to the file (or set of files) is tracked with what is known as a commit using Git. A commit is a snapshot of a file (or set of files) at a specific time. Each commit is accompanied by a commit message written by a student (i.e., data analyst) indicating what was changed in that specific commit. Note that commit messages can be as helpful as one writes them to be. For instance the commit message "add model code" does not indicate much about the model should there be any changes to model assumptions or variables in the upcoming iterations of the model. One can always can retrieve older version of file(s) at a specific time by going to the specific commit. In addition, each commit notes the author who made the commit.

Table 1 Example of a commit history

Commit Message	Date and Time	Author
Add histogram of mpg	June 24, 2024, 00:00 UTC	Student A
Change histogram breaks to 10	June 25, 2024, 07:00 UTC	Student B
Add model code	June 25, 2024, 07:30 UTC	Student A
Switch x and y in the model	June 25, 2024, 07:36 UTC	Student A

Websites such as GitHub allow for Git repositories to be hosted online and provide additional project management features. Implementing version control with Git and GitHub has become an essential learning objective in statistics and data science courses (Beckman et al. 2021). When teaching version control with Git and GitHub, instructors need to learn two sets of new skills. First, they need to learn version control from a user perspective. In other words, Git and GitHub can be considered as tools for learning. This is what students would also learn. In addition, they need to learn GitHub from a course manager perspective (e.g., how to create students' assignments). There are tools for teaching that can support instructors in the course management process. Tools such as GitHub classroom (Fiksel et al. 2019) and ghclass R package (Rundel & Çetinkaya-Rundel 2022) can be used for managing course work. In addition, grading tools such as the gradetools R package (Ricci, Medina, & Dogucu 2022) interacts with GitHub and allows for reproducing the grading workflow from one academic term to next (Ricci, Medina, & Dogucu 2023). Use of the ghclass and gradetools package are explained further in Section 5.

In addition to providing version control features, Git and GitHub also provide collaboration opportunities that can be ideal for use in team-based projects (Çetinkaya-Rundel et al. 2022, Legacy et al. 2023). This can pave the way for students to think of reproducibility of their projects by their teammates and possibly by the general public in the scientific community. In fact, many Git users learn Git out of necessity, for instance when they have to collaborate on code or want to contribute to an open source project (Milliken et al. 2021). Thus it is essential to create similar opportunities for students in the class.

In summary, teaching and learning of reproducibility in the statistics classroom is dependent on the computing tools that are available to the instructor and the students. This includes but not limited to web applications, programming languages, notebook environments, and version control tools. To spread the reproducibility culture in science, it is important to close the gap between tools for learning statistics which provide approachable ways of using a tool for novice learners and tools for doing statistics that provide more features supporting reproducibility. In addition, instructors' toolkit and training need to be supported from perspectives of tools for teaching and learning separately and reproducibility should be considered for both.

3. EXAMPLES OF TEACHING REPRODUCIBILITY

There are numerous ways students can be exposed to reproducibility methods. Some traditional methods may include lecturing and reading on the subject. Although these activities can be effective in teaching, some learners may not be able to internalize methods and practices associated with reproducibility fully. The need for computational reproducibility and the obstacles associated with it may be best understood with a project, even better a team project. Instructors can assign projects through which students learn reproducibility practices by applying them to their own data analysis. On the contrary, instructors can also assign projects through which students reproduce the work of others. Ball et al (2022) call the first type of projects where students work on their own data analysis while documenting the workflow for reproducibility as "ex ante documentation" and the second type of projects where students try to reproduce the work of others as "ex post reproduction".

Project based learning is a recommended practice for the statistics (Carver et al 2016) and data science curricula (National Academies of Sciences, Engineering, and Medicine 2018). There are many examples in introductory level statistics courses (Halvorsen 2010), data science courses (Çetinkaya-Rundel et al. 2022, Bean 2023), as well as more advanced courses that are dedicated fully to projects such as those in capstone courses (Martonosi & Williams 2016). In addition, students' theses and dissertations can also be considered projects at a larger scale. In this section, we discuss projects from a reproducibility perspective. We examine ex ante documentation and ex post reproduction projects in depth and draw comparisons.

When students work on ex ante documentation projects, to support reproducibility, an important objective is the learning of the file management as well as naming files. Figure 2 shows a directory for a hypothetical project. In terms of good practices for file management, this directory contains few key points. First of all, folders and files are named clearly indicating what could possibly be found in the file or folder as opposed to having files named as `Untitled.R`. Second of all, possible documentation is provided in the `README.md` file for the project as well for the data in the `data/README.md` file. Third of all, when needed, ordering of files are provided with enumeration. For instance, for this project, students must have first scraped data from IMDB, and then cleaned the scraped data. If the files were named `data-scraping.R` and `data-cleaning.R` then due to alphabetical ordering, the cleaning file would appear before the scraping file. This is avoided by using enumeration as 1 and 2.

In addition to file and folder organization, naming files, and documentation of data, students can also practice good coding skills in their projects. In their study of more than 2000 replication datasets with over 9000 R files, Trisovic et al (2022) found that 74% of R files failed to execute and even after doing some cleaning 56% of the files still failed to run. To prevent code execution problems, learning of code quality is an important aspect of data science projects (Pruim, Gîrjău, & Horton 2023). Projects can also teach students data citations which are also an important aspect of reproducibility (Mendez-Carbajo, & Dellachiesa 2023).

Ex post reproduction projects are also seen across multiple disciplines including psychology (Frank & Saxe 2012), artificial intelligence (Lucic et al. 2022), archaeology (Marwick et al. 2020), neuroimaging (Millman et al. 2018), and computational biology (Karathanasis et al. 2022). These projects can have varying goals. For instance, Karathanasis et al. (2022) have students reproduce figures from published articles whereas Lapane and Dube (2021) have students try to duplicate results of a previous study using the original data but applying their own analyses and interpretations.

Through ex post reproduction projects students can improve computational skills, be aware of reproducibility problems, and become familiar with publication standards. Given the diversity of methods utilized in published research, it is important for instructors to limit the types of publications that can be selected for reproduction (e.g., analysis needs to utilize logistic regression). Instructors need to be mindful of the difficulties of finding a suitable

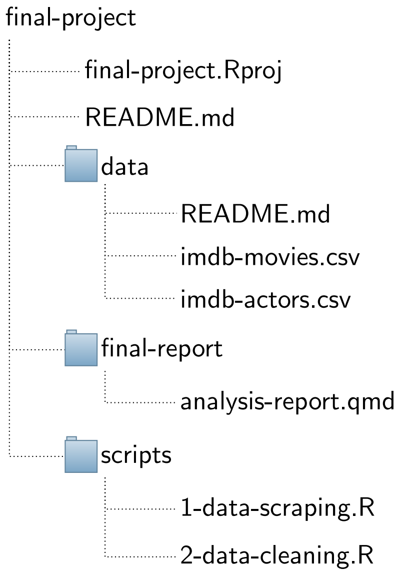


Figure 2

Example directory for a data analysis project

article to reproduce as many publications do not necessarily come with open data. Even when research products are open-access with open-data, they may still be unapproachable in practice because they are not learner-friendly (Sanchez Reyes & McTavish 2022).

Projects can also provide an opportunity for students to learn from their peers. One way of incorporating this is through peer-review. For instance, in an ex ante documentation project, students can be asked to review each other's work and reproduce each other's analysis thus creating an ex post reproduction opportunity. One such example is provided by Yu and Hu (2019).

Whether they are ex ante documentation or ex post reproduction, projects can especially be motivating if they are scientifically meaningful, in other words genuinely trying to answer a real question. This question does not necessarily have to be generated by the instructor or the student. Students can be matched with researchers (e.g., Toelch & Ostwald (2018)) or other local partners in the case of service learning (e.g., Nurse & Staiger. (2019)).

In addition to learning of reproducibility, projects provide many other pedagogical benefits. Much of statistics and data science practice is done in teams, projects allow for students to learn to collaborate from both technical perspectives (e.g., learning of collaborative tools) and social perspectives (e.g., participating as an effective member). Projects offer opportunities to work with and learn from real data. Students often find projects to be fun and engaging (Lapane & Dube 2021). Lastly, students improve their problem solving skills through project based learning.

Despite many benefits, projects should not be seen as the only way to incorporate reproducibility practices and methods in the classroom. In fact, learning of new tools (e.g., version control) can be challenging for many. For the mastery of these tools students need to get “lots and lots of practice” (Ostblom, & Timbers 2022). Thus in teaching of these skills, instructors should utilize lecture, homework, and projects for teaching of reproducibility and should display consistent practices in all three (Mehta & Moore. 2022).

4. EXAMPLES OF REPRODUCIBLE TEACHING

In Section 3, we shared examples on the types of projects students work on as they learn reproducibility. In this section, we shift our focus to instructors. Instructors train the next generation of scientists and how they train shapes the scientific practice. Many instructors might also have their research agenda for which reproducibility might be relevant. However, here we solely focus on instructors' reproducibility practice in and outside the classroom for teaching purposes.

Reproducibility can serve multiple purposes from a teaching preparation perspective. An instructor may need to reproduce teaching materials from one term to the next. Similarly, instructors may share teaching materials between themselves and may want to reproduce each other's teaching materials. Whether it's one's own teaching materials or they adopted another instructor's materials, overtime, modifications can be necessary and a reproducible workflow can help make modifications easier. Thus adoption of reproducible workflow in preparation of teaching can be beneficial.

Consider an instructor teaching central limit theorem through simulations. This instructor might show perhaps three samples and their accompanying statistics in their slides before proceeding to the sampling distribution. Using the same material, another instructor might want to show five samples instead. Increasing number of samples from three to five would be an easy change if the slides were prepared with a literate programming tool where code and text for the slides coexist. However, if the instructors simulate data in one software and then copy paste the output to the slide software then this modification would be much more challenging. Similarly, if an error is found, making modifications would be a lot easier in a reproducible workflow.

An essential teaching material in statistics courses is the dataset. Use of real data is the recommended practice in teaching statistics (Carver et al 2016). One way statistics and data science educators have supported reproducible teaching is through the use of R packages in providing the datasets openly. There are many examples including but not limited to Lock5withR (Pruim 2015), openintro (Çetinkaya-Rundel et al 2022), mosaic (Pruim, Kaplan, & Horton 2024), and bayesrules (Dogucu, Johnson, & Ott 2021). An alternative method of using datasets is through reading in the data locally or through a web link. Datasets provided in packages offer simplicity of reading in the data and have more permanence than web links.

In addition to the datasets themselves, the process of acquiring the data and documentation of it is extremely crucial. A great example of importance of reproducibility in preparation of datasets for classroom use is shown by Amaliah, et al. (2022). Their work stems from the National Longitudinal Survey of Youth (NLSY79) dataset provided in the book *Applied Longitudinal Data Analysis* by Singer and Willett 2003 which has been used numerous time to showcase generalized linear models and its variations. The data comes from a sample of high school dropouts from 1979 to 1994. Amaliah, et al. (2022) update this data to include records from 1979 through to 2018. While doing this, they also document the process to allow for future updates. This method allows for future educators to refresh the dataset when needed.

Another statistics education area that has seen a reproducibility shift is textbooks. There are an increasing statistics books that are published open access and open source. Open access nature of the books allows readers (i.e., learners) to benefit from the book freely while the open source nature allows readers, including other educators, to see how the book came together, for instance, how tables and plots were generated. GitHub is one of the venues where book authors can host the source code and this also allows for readers to interact with the authors through *GitHub Issues*. For instance, readers can state any errors they might notice in the book. Examples include the books *Computational and Inferential Thinking: The Foundations of Data Science*(Adhikari, DeNero & Wagner 2022), *Beyond Multiple Linear Regression* (Roback & Legler 2021), *Modern Data Science with R* (Baumer, Kaplan & Horton 2021) and *Probability and Bayesian Modeling* (Albert & Hu 2019).

Textbook writing can be an extensive project that is not necessarily taken by all statistics educators. However, there are many other teaching materials that statistics educators produce on a regular basis such as handouts, slides, and assessments. These materials can also be produced in a reproducible manner that allows for easy modifications term-to-term or by others. Similar to textbooks, some educators also provide such materials open access and open source often compiled as a course website. Examples include courses such as Hicks' *Statistical Programming Paradigms and Workflows* (2022) and D'Agostino McGowan's *Statistical Learning* (2023).

Datasets, textbooks, handouts, slides, course websites are teaching materials that statistics educators can produce in reproducible manner. Tools such as Quarto and Jupyter Notebooks that allow for literate programming make the production of these resources easier from a technical point of view. The documentation of Quarto and Jupyter Notebooks as well as the examples provided in this section and beyond can serve as a starting point for those considering adoption of reproducible workflow in their teaching.

In summary, we can alter the aforementioned Leek and Jager's (2017) questions for reproducible research and adopt them for reproducibility in the classroom as

- Are code and data both readily available?
- Does running the provided code exactly produce the same set of teaching materials?
- Can the code and data be easily utilized to produce different set of teaching materials (e.g., book or slides or handouts)?
- Can the teaching materials be easily altered, if needed, from one term to the next or from one instructor to another?

Even though, there is not a single way to practice reproducible teaching, these questions may possibly guide instructors in their practice of reproducible teaching.

5. EXAMPLES FROM A DATA SCIENCE CLASSROOM

In Sections 3 and 4 we provided a wide range of examples on teaching reproducibility and reproducible teaching. These examples were based on the literature by multiple teacher-scholars teaching a variety of courses in a variety of scientific disciplines. In this chapter, we provide an example from a single course with the aim of clarifying how different set of tools and activities fit in together in a classroom.

Even though there is not a golden-standard workflow or set of tools to work towards reproducibility in the classroom, in this section, we provide the workflow we adapt in our undergraduate level *Introduction to Data Science* course in the Department of Statistics at University of California Irvine as an example. The course curriculum was designed with teaching reproducibility in mind and the teaching team adapts certain reproducible teaching practices. The course is taught to about 120 students in a single section with two lecture sessions in a week. The academic term consists of 10 weeks and a final exam week. The teaching team consists of an instructor who delivers the lectures and a teaching assistant (TA) who is mainly in charge of grading.

The most notable aspect of the toolkit for the course is that the tools that students learn are also the tools that the teaching team uses. This consistency allows for students to see the tools being used also by their teaching team. In addition, it allows for teaching team to get more proficient in using the toolkit that they teach. Most notably, the tools include using the R language within Quarto documents that are version controlled using Git repositories that are hosted on GitHub. We will first summarize how students use these tools and then go on to explain the use of these tools by the teaching team.

Prior to the start of the term, students are provided with instructions to do their technical setup. This includes downloading and installing R, RStudio, necessary R packages, signing up for a GitHub account and testing their installations. Installations are also provided for them on a local server. This option is crucial especially for students who may not have an easy access to a personal device or who may run into technical problems on their personal devices.

On the first day of classes, students are introduced to the toolkit. They leave the class having written a simple R code within a Quarto document. They also make changes (i.e., commits) and track their changes using a Git system that is hosted on GitHub.

Every week, students are provided with a GitHub repository that includes lecture files as well as a homework file. A GitHub repository can be thought as a set of files and folders that is version controlled. Figure 3 shows a hypothetical repository provided to students in Week 10 covering regression topics. Note that this directory specifically shows the repository of a student with the GitHub username `mdogucu`. We differentiate students' repositories by adding their GitHub username to the end of the repository name. Students take notes during the lectures using the Quarto documents provided in the `lectures` folder, saving, committing, and pushing (i.e., uploading to GitHub) their changes throughout the lecture and later if needed. In addition to lecture notes, these files also include short learning activities to be completed during class time. In a more traditional classroom, these lecture notes can also be considered as handouts that have fill in the blank opportunities scattered throughout. Students also work on their homework in a similar fashion until the assigned deadline.

Students work on their weekly repositories starting on the first day by taking notes. They work on weekly repositories individually. Thus by the end of first class, each student leaves the class having interacted with GitHub at least once. By the fourth week of the term, students learn how to collaborate on GitHub. This is also when students are introduced to final projects and meet their teams. Thus, in addition to weekly repositories, each student also has access to the final project repository of their team on GitHub. By the end of term, each student manages a total of eleven repositories, ten for weekly work and one for a final project.

A course with about 120 students equates to a total of about 1224 repositories that the teaching team manages including 1200 weekly repositories for each students and about 24 additional repositories for the team projects (note that a team consists of about five

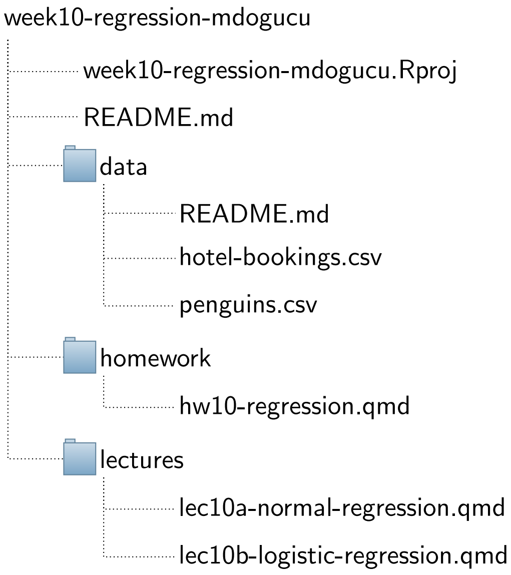

week10-regression-mdogucu
week10-regression-mdogucu.Rproj
README.md
data
README.md
hotel-bookings.csv
penguins.csv
homework
hw10-regression.qmd
lectures
lec10a-normal-regression.qmd
lec10b-logistic-regression.qmd

Figure 3

Example weekly repository for a student

students). Needless to say the course relies on automation in many aspects of course management. In order to prepare weekly repositories, the instructor prepares a single repository. Following on the example from Figure 3, the instructor prepares a repository named `week10-regression` with the accompanying lecture notes and homework assignment. This repository serves as a template to be shared with all students.

Using the class roster with students' GitHub usernames and the R package ghclass that interacts with GitHub, the instructor then replicates the template repository for each student. During the process, each student's own repository is created with their username to differentiate it from others'. These repositories are then graded by the TA using the gradetools package. Figure 4 shows each student's repository from Figure 3 downloaded on the TA's computer. Needless to say this figure only shows repositories of three students for simplicity. The gradetools package allows for opening and closing each homework file and creating feedback files based on a rubric automatically. The use of rubric allows making the grading process as fair as possible and the automation features make the grading process, especially for a course with 120 students, efficient.

In addition to repositories created for students, the instructor manages two additional repositories: one for slides, and one for the course website. The course slides and website are both prepared using Quarto and hosted on GitHub. Students also have access to these repositories and are shown the source code as soon as they start learning Quarto. Many advanced Quarto features that the instructor uses in preparing her slides are not necessarily taught in the course but many students start adopting these features as they work on their final projects.

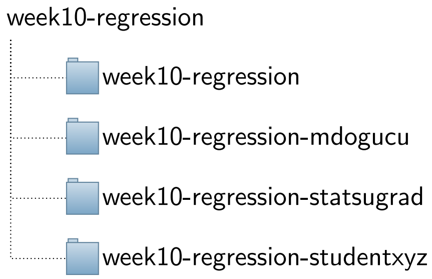

week10-regression
week10-regression
week10-regression-mdogucu
week10-regression-statsugrad
week10-regression-studentxyz

Figure 4

Example directory with template repository and repositories of three students

Between students' files and instructors' files, it is worth noting the consistency of naming files. For instance, while students take notes on `lec10a-normal-regression.qmd` document in their weekly repository, the instructor is showing output from the slides repository for the `lec10a-normal-regression.qmd` document. Students and the teaching team are aware that this is first lecture, hence the letter a, of the tenth week.

In addition to extensive exposure to the toolkit, students are also explicitly taught on reproducibility practices and evaluated on these with a full lecture dedicated to good workflow practices for reproducible data science. Much of the practice of reproducibility especially becomes important when students work with their peers on the projects. This includes but not limited to naming files in machine and human readable formats, writing code in machine and human readable formats, writing meaningful commit messages, saving session information (e.g., operating system information and versions of R packages), writing informative README files especially for documenting data acquisition. This workflows also allows for teaching team to see as needed, the commits made by different team members.

Interested readers can access the course materials at `https://www.introdata.science/` and the accompanying public GitHub repositories at `https://github.com/stats6-wi24` for the last iteration of the course in Winter 2024. Needless to say, this course is not necessarily the best or only example for teaching reproducibility and reproducible teaching but it is shared with the intent to help others who may be taking new steps towards reproducibility in the classroom.

6. RECOMMENDATIONS

The teaching of reproducibility in the classroom is an important component of reproducibility practice in science. In this section, we provide recommendations towards the goal of widening the use of reproducibility practices in the sciences.

1. Teach Reproducibility from Day 1 Introductory statistics and data science classes are often the entry point to the field for many. It is crucial to start teaching reproducible practices in these courses. Instructors can teach data analysis workflow with reproducibility in mind without providing any other option. For instance, rather than having students do computations in R and then submit word documents, instructors can have students work right away in literate programming tools where computation and text can coexist. Even though the simplicity of R scripts can be appealing for an instructor teaching at the introductory level, the benefits of literate programming for reproducibility should be considered. Many students may not find it simple to work with literate programming tools (e.g., R Markdown) initially but they do find it simple to work with by the end of the term (Baumer et al. 2014).

2. Make Reproducibility a Requirement In the classroom, one way to require reproducibility is to make it part of the grading criteria. Whether it is homework assignment or projects such as those described in Section 4, the reproducibility topics can be added to the rubric. This way there would also be a consistency between teaching and assessments. Outside of the classrooms, departments can consider assessing students on reproducibility practices as part of qualifying exams and dissertations. Lastly, teacher-scholars can also be evaluated for their reproducibility practices as well both in their teaching and research. Currently, in many institutions, reproducibility is often not part of tenure and promotion criteria (Alperin, et al. 2022) and there are only a few institutions that incentivizes such efforts (Axfors, Malički & Goodman 2024).

3. Teach Reproducibility Explicitly As instructors teach reproducible workflows, it is important for students to understand why they do certain steps the way they do. For instance, if literate programming is taught, then they should know why. For instance, instructors can show this by having students write a data analysis report. Later, they can inform students to fix an "error" in a dataset. How many steps would students take to reproduce the report with the correct version of the data? Instructors can also share examples of errors from the scientific literature and how they were discovered through reproducibility efforts. Though, instructors should be mindful when sharing these as this may also possibly lead to mistrust in science (Chopik et al 2018). For those teaching courses geared towards specific disciplines, guidelines or approaches of professional societies can be shared. For instance, if training medical students response of the National Institutes of Health to reproducibility crisis can be shared with students (as seen in Lapane & Dube (2021)).

4. Make Reproducibility a Campus-Wide Effort. Even though, reproducibility is closely linked with data analysis and statistics courses, the reproducibility education does not solely fall on the shoulders of statistics departments. There are many other stakeholders that can contribute to reproducibility education on campus and can train students, instructors, and researchers. Libraries, especially, can also be a great contributor to capacity building in this area (LaPolla et al 2022, Rethlefsen et al 2018).

5. Provide Reproducibility Training for Students, Instructors, and Researchers In this paper, we focused on reproducibility in the classroom where students are the learners. With the fast pace of change of new tools to support reproducibility as well as lack of prior training on the topic, instructors, and researchers need training as well. Lack of reproducibility in data science education research (Dogucu et al 2024) shows the urgent need for more professional development opportunities for teacher-scholars and the broader scientific community in this area. This also includes training graduate teaching assistants who teach an increasing number of classes given the enrollment increase in statistics and data science classes (Rummerfield, Ricci & Dogucu 2021). It would be difficult to train the new generation on reproducibility without ensuring the competence of the older generations. Luckily, a few learning communities provides training in this area for researchers on all career levels including but not limited to Data Carpentry (Teal et al

2015), Project TIER (Medeiros & Ball. 2017), Reproducibility for Everyone (Auer, et al. 2021), and ReproducibiliTea (Fitzgibbon et al. 2020). These communities serve those in different stages of their career and in varying disciplines.

6. Teach Reproducibility in and Outside of the Classroom. As educators, we need to consider venues outside of the classroom as part of teaching and provide as many different learning opportunities as possible for learners. For instance in recent years, datathons have become popular events for students to participate in (Anslow, et al. 2016). This also includes American Statistical Association's DataFest where students work using a single dataset over a weekend (Gould 2014). In such events it is important to emphasize reproducibility worthy of an award. Similarly datathons with specific aim of reproducing the others can be held. ReproHack is one such example (Hettne et al 2020).

To conclude, reproducibility education is the cornerstone of scientific training and thus scientific practice. Its importance cuts across training of students, instructors, and researchers. Members of the scientific community, whether students, instructors, or researchers, can benefit from and contribute to, improvements in this area. Scientific agencies and administrators should support curricular change and other efforts in this realm.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

Dogucu has been supported by NSF IIS award #2123366.

LITERATURE CITED

Abranovic W, Ageloff R, Frederick D. 1972. Time-sharing computer systems as a teaching tool. *T*he American Statistician. 26:34–38 `https://doi.org/10.1080/00031305.1972.10477322`

Adhikari A, DeNero J, Wagner D. 2022. Computational and inferential thinking: The foundations of data science.

Albert J, Hu J. 2019. Probability and Bayesian modeling. Chapman and Hall/CRC.

Allaire JJ, Dervieux C. 2024. quarto: R interface to 'Quarto' markdown publishing system. *R* package version 1.4. `hhttps://cran.r-project.org/web/packages/quarto/index.html`

Alperin JP, Schimanski LA, La M, Niles MT, McKiernan, EC. 2022. The value of data and other non-traditional scholarly outputs in academic review, promotion, and tenure in Canada and the United States. *O*pen handbook of linguistic data management MIT Press.

Amaliah D, Cook D, Tanaka E, Hyde K, Tierney N. 2022. A journey from wild to textbook data to reproducibly refresh the wages data from the National Longitudinal Survey of Youth database. *J*ournal of Statistics and Data Science Education. 30:289–303 `https://doi.org/10.1080/26939169.2022.2094300`

Anslow C, Brosz J, Maurer F, Boyes M. 2016. Datathons: An experience report of data hackathons for data science education. *P*roceedings of the 47th ACM Technical Symposium on Computing Science Education. 615–620 `https://doi.org/10.1145/2839509.2844568`

Auer S, Haeltermann NA, Weissgerber TL, Erlich JC, Susilaradeya D, Julkowska M, Gazda MA, Schwessinger B, Jadavji NM. 2021. A community-led initiative for training in reproducible research. *e*Life. 10:e64719 `https://doi.org/10.7554/eLife.64719`

Axfors C, Malički M, Goodman SN. 2024. Research rigor and reproducibility in research education:

A CTSA institutional survey. *J*ournal of Clinical and Translational Science. 8(1):e45 `https://doi.org/10.1017/cts.2024.10`

Ball R. 2023. “Yes we can!”: A practical approach to teaching reproducibility to undergraduates. *H*arvard Data Science Review. 5 `https://doi.org/10.1162/99608f92.9e002f7b`

Ball R, Medeiros N, Bussberg NW, Piekut A. 2022. “An invitation to teaching reproducible research: lessons from a symposium. *J*ournal of Statistics and Data Science Education. 30:209–218 `https://doi.org/10.1080/26939169.2022.2099489`

Baumer B, Çetinkaya-Rundel M, Bray A, Loi L, Horton N.J. 2014. R markdown: Integrating a reproducible analysis tool into introductory statistics. *T*echnology Innovations in Statistics Education. 8:1–22 `http://dx.doi.org/10.5070/T581020118`

Baumer BS, Kaplan DT, Horton, NJ 2021. Modern data science with R. Chapman and Hall/CRC. `https://mdsr-book.github.io/`

Bean BL. 2023. Teaching Reproducibility to First Year College Students: Reflections From an Introductory Data Science Course. *J*ournal on Empowering Teaching Excellence. 7 `https://doi.org/10.26077/5978-0434`

Beckman MD, Çetinkaya-Rundel M, Horton NJ, Rundel CW, Sullivan AJ, Tackett M. 2021. Implementing version control with Git and GitHub as a learning objective in statistics and data science courses. *J*ournal of Statistics and Data Science Education. 29:S132–S144 `https://doi.org/10.1080/10691898.2020.1848485`

Biehler R. 1997. Software for learning and for doing statistics. *I*nternational Statistical Review. 65:167–169 `https://doi.org/10.1111/j.1751-5823.1997.tb00399.x`

Carmer SG, Cady FB. 1969. Computerized data generation for teaching statistics. *T*he American Statistician. 23:33–35 `https://doi.org/10.1080/00031305.1969.10481895`

Chopik, WJ, Bremner RH, Defever AM, Keller VN. 2018. How (and whether) to teach undergraduates about the replication crisis in psychological science. *T*eaching of Psychology. 45(2): 158-163 `https://doi.org/10.1177/0098628318762900`

Chamber JM. 2020. S, R, and data science. *P*roceedings of the ACM on Programming Languages. 4:1–17 `https://doi.org/10.1145/3386334`

Chang A, Li P. 2015. Is economics research replicable? Sixty published papers from thirteen journals say ’usually not’. *F*EDS Working Paper No. 2015-083. `https://doi.org/10.17016/FEDS.2015.083`

Çetinkaya-Rundel M, Diez D, Bray A, Kim AY, Baumer B, Ismay, C, Paterno, N, Barr, C. 2022. openintro: Data sets and supplemental functions from ’OpenIntro’ textbooks and labs. *R* package version 2.4.0. `https://cran.r-project.org/web/packages/openintro/index.html`

Çetinkaya-Rundel M, Dogucu M, Rummerfield W. 2022. The 5Ws and 1H of term projects in the introductory data science classroom. *S*tatistics Education Research Journal. 21. `https://doi.org/10.52041/serj.v21i2.37`

D’Agostino McGowan L. 2023. Statistical learning. `https://sta-363-s23.github.io/website/`

Dogucu M, Demirci S, Bendekgey H, Ricci FZ, Medina CM. 2024. Undergraduate data science education: Who has the microphone and what are they saying? *a*rXiv preprint. `https://arxiv.org/abs/2403.03387`

Dogucu M, Çetinkaya-Rundel M. 2022. Tools and recommendations for reproducible teaching. *J*ournal of Statistics and Data Science Education. 30:251–260 `https://doi.org/10.1080/26939169.2022.2138645`

Dogucu M, Johnson A, Ott M 2021. bayesrules: Datasets and supplemental functions from Bayes rules! book. *R* package version 0.0.2. `https://cran.r-project.org/web/packages/bayesrules/index.html`

Donoho D. 2017. 50 years of data science. *J*ournal Of Computational And Graphical Statistics. 26:745-766 `https://doi.org/10.1080/10618600.2017.1384734`

Fanelli D. 2018. Is science really facing a reproducibility crisis, and do we need it to? *P*roceedings of the National Academy of Sciences. 115:2628–2631 `https://doi.org/10.1073/pnas.1708272114`

Fiksel J, Jager LH, Hardin JS, Taub MA. 2019. Using GitHub classroom to teach statistics. *J*ournal of Statistics and Data Science Education. 27:110–119 `https://doi.org/10.1080/10691898.2019.1617089`

Fitzgibbon L, Brady D, Haffey A, Kurdi V, Lau J, Raw J, Roesch, E, Williams B. 2020. Brewing up a storm: developing Open Research culture through ReproducibiliTea. University of Reading `https://doi.org/10.17864/1926.92781`

Frank MC, Saxe R. 2012. Teaching replication. *P*erspectives on Psychological Science. 7:600–604 `https://doi.org/10.1177/1745691612460686`

Gould R. 2014. DataFest: Celebrating data in the data deluge. *S*ustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics. 1–4

Carver R, Everson M, Gabrosek J, Horton N, Lock R, Mocko M, Rossman A, Roswell GH, Velleman P, Witmer J, Wood B. 2016. Guidelines for assessment and instruction in statistics education (GAISE). American Statistical Association. `http://www.amstat.org/education/gaise`

Halvorsen KT. 2010. Formulating statistical questions and implementing statistics projects in an introductory applied statistics course. *P*roceedings of the 8th International Conference on Teaching Statistics.

Hettne K, Proppert R, Nab L, Rojas-Saunero LP, Gawehns D. 2020. ReprohackNL 2019: How libraries can promote research reproducibility through community engagement. *I*ASSIST Quarterly. 44:1–10 `https://doi.org/10.29173/iq977`

Hicks S. 2022. Statistical programming paradigms and workflows. `https://www.stephaniehicks.com/jhustatprogramming2022/`

Horton NJ, Alexander R, Parker M, Piekut A, Rundel C. 2022. Importance of reproducibility and responsible workflow in the data science and statistics curriculum. *J*ournal of Statistics and Data Science Education. 30:207–208 `https://doi.org/10.1080/26939169.2022.2141001`

Huskey SJ. 2023. Committing to reproducibility and explainability: Using Git as a research journal. *I*nternational Journal of Digital Humanities. 1–13 `https://doi.org/10.1007/s42803-023-00076-9`

Junk TR, Lyon L. 2020. Reproducibility and replication of experimental particle physics results. *a*rXiv preprint arXiv:2009.06864. `https://arxiv.org/abs/2009.06864`

Karathanasis N, Hwang D, Heng V, Abhimannyu R, Slogoff-Sevilla P, Buchel G, et al. 2022. Reproducibility efforts as a teaching tool: A pilot study. *P*LoS Computational Biology. 18(11):e1010615. `https://doi.org/10.1371/journal.pcbi.1010615`

Kery MB, Radensky M, Arya M, John BE, Myers BA. 2018. The story in the notebook: Exploratory data science using a literate programming tool. *P*roceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 1–11 `https://doi.org/10.1145/3173574.3173748`

Kluvyer T, Ragan-Kelley B, Pérez F, Granger B, Bussonnier M, Frederic J, Kelley K, Hamrick J, Grout J, Corlay S, Ivanov P, Avila D, Abdalla S, Willing C. 2016. Jupyter Notebooks – a publishing format for reproducible computational workflows. *P*ositioning and Power in Academic Publishing: Players, Agents and Agendas. 87-90 `https://doi.org/10.1145/3173574.3173748`

Knuth DE. 1984. Literate programming. *T*he Computer Journal. 27:97–111 `https://doi.org/10.1093/comjnl/27.2.97`

Lapane KL, Dube CE. 2021. Rigor and reproducibility training for first year medical students in research pathways. *C*linical and Translational Science. 14:1026–1036 `https://doi.org/10.1111/cts.12966`

LaPolla FWZ, Bakker CJ, Exner N, Montnech T, Surkis A, Ye H. 2022. Rigor and reproducibility instruction in academic medical libraries. *J*ournal of the Medical Library Association: JMLA. 110:281–293 `https://doi.org/10.5195/jmla.2022.1443`

Leek JT, Jager LR. 2017. Is most published research really false? *A*nnual Review of Statistics and Its Application. 4:109–122 `https://doi.org/10.1146/annurev-statistics-060116-054104`

Legacy C, Zieffler A, Baumer BS, Barr V, Horton NJ. 2023. Team-based learning, agile, skill development, collaboration, project-based learning. *F*oundations of Data Science. 5:244–265 https://doi.org/10.3934/fods.2022003

Lucic A, Bleeker M, Jullien S, Bhargav S, De Rijke M. 2022. Reproducibility as a mechanism for teaching fairness, accountability, confidentiality, and transparency in artificial intelligence. *P*roceedings of the AAAI Conference on Artificial Intelligence. 36:12792-12800 https://doi.org/10.1609/aaai.v36i11.21558

Martonosi SE, Williams TDR. 2016. A survey of statistical capstone projects. *J*ournal of Statistics Education. 24:127–135 https://doi.org/10.1146/annurev-statistics-060116-054104

Marwick B, Wang L, Robinson R, Loiselle H. 2020. How to use replication assignments for teaching integrity in empirical archaeology. *A*dvances in Archaeological Practice. 8:78–86 https://doi.org/10.1017/aap.2019.38

McNamara A.A. 2015. Bridging the gap between tools for learning and for doing statistics. University of California Los Angeles https://www.proquest.com/dissertations-theses/bridging-gap-between-tools-learning-doing/docview/1694580439/se-2

Medeiros N, Ball RJ. 2017. Teaching integrity in empirical economics: The pedagogy of reproducible science in undergraduate education. *I*n M.K. Hensley & S. Davis-Kahl (Eds.), Undergraduate Research and the Academic Librarian: Case Studies and Best Practices. 261–268

Mehta CC, Moore RH. 2022. Third time's a charm: A tripartite approach for teaching project organization to students. *J*ournal of Statistics and Data Science Education. 30:261-265 https://doi.org/10.1080/26939169.2022.2118644

Mendez-Carbajo D, Dellachiesa A. 2023. Data Citations and Reproducibility in the Undergraduate Curriculum. *H*arvard Data Science Review. 5 https://doi.org/10.1162/99608f92.c2835391

Milliken G, Nguyen S, Steeves V. 2021. A behavioral approach to understanding the Git experience. *P*roceedings of the 54th Hawaii International Conference on System Sciences. 7239–7248 https://hdl.handle.net/10125/71493

Millman KJ, Brett M, Barnowski R, Poline JB. 2018. Teaching computational reproducibility for neuroimaging. *F*rontiers in Neuroscience. 12:1–16 https://doi.org/10.3389/fnins.2018.00727

Miyakawa, T. 2020. No raw data, no science: another possible source of the reproducibility crisis. *M*olecular brain. 26:745-766 https://doi.org/10.1186/s13041-020-0552-24

National Academies of Sciences, Engineering, and Medicine. 2019. Reproducibility and replicability in science. National Academies Press. https://nap.nationalacademies.org/catalog/25303/reproducibility-and-replicability-in-science

National Academies of Sciences, Engineering, and Medicine. 2018. Data science for undergraduates: Opportunities and options. National Academies Press. https://doi.org/10.17226/25104

Nurse AM, Staiger T. 2019. Teaching data reproducibility through service learning. *T*eaching Sociology. 47:350-357 https://doi.org/10.1177/0092055X19860577

Open Science Collaboration. 2015. Estimating the reproducibility of psychological science. *S*cience. 349:aac4716 https://doi.org/10.1126/science.aac4716

Ostblom J, Timbers T. 2022. Opinionated practices for teaching reproducibility: motivation, guided instruction and practice. *J*ournal of Statistics and Data Science Education. 30:241-250 https://doi.org/10.1080/26939169.2022.2074922

Peng R. 2015. The reproducibility crisis in science: A statistical counterattack. *S*ignificance. 12:30–32 https://doi.org/10.1080/10618600.2017.1384734

Pruim R. 2015. Lock5withR: Datasets for 'Statistics: Unlocking the power of data'. *R* package version 1.2.2. https://cran.r-project.org/web/packages/Lock5withR/index.html

Pruim R, Gîrjău MC,Horton NJ. 2023. Fostering better coding practices for data scientists. *H*arvard Data Science Review. 5. https://doi.org/10.1162/99608f92.97c9f60f

Pruim R, Kaplan DT, Horton NJ. 2024. mosaic: Project MOSAIC statistics and mathematics teaching utilities. *R* package version 1.9.1. https://cran.r-project.org/web/packages/mosaic/index.html

Rethlefsen ML, Lackey MJ, Zhao S. 2018. Building capacity to encourage research reproducibility and #MakeResearchTrue. *J*ournal of the Medical Library Association: JMLA. 106:113–119 https://doi.org/https://doi.org/10.5195%2Fjmla.2018.273

Ricci FZ, Medina CM, Dogucu M. 2022. gradetools: Tools to Assist with Providing Grades and Personalized Feedback to Students. *R* package version 0.2.0. https://github.com/federicazoe/gradetools

Ricci FZ, Medina CM, Dogucu M. 2023. Designing and implementing an automated grading workflow for providing personalized feedback to open-ended data science assignments. *a*rXiv preprint arXiv:2309.12924. https://arxiv.org/abs/2309.12924

Roback P, Legler J. 2021. Beyond multiple linear regression: applied generalized linear models and multilevel models in R. Chapman and Hall/CRC.

Rummerfield W, Ricci F, Dogucu M. 2021. Training graduate students to teach statistics and data science from a distance. *I*n Proceedings of the IASE 2021 Satellite Conference. International Association for Statistical Education.

Rundel C, Çetinkaya-Rundel M. 2022. ghclass: Tools for managing classes on GitHub. *R* package version 0.2.1. https://cran.r-project.org/web/packages/ghclass/index.html

Sanchez Reyes LL, McTavish EJ. 2022. Approachable case stu. *J*ournal of Statistics and Data Science Education. 30:304-310 https://doi.org/10.1080/26939169.2022.2099487

Seo J, Dogucu M. 2024. Data science + accessibility. *T*eaching Accessible Computing. https://bookish.press/tac/DataScience

Singer JD, Willett JB. 2003. Applied longitudinal data analysis: Modeling change and event occurrence. Oxford university press.

Teal TK, Cranston KA, Lapp H, White E, Wilson G, Ram K, Pawlik A. 2015. Data carpentry: workshops to increase data literacy for researchers. *I*nternational Journal of Digital Curation. 10:135–143 https://doi.org/10.2218/ijdc.v10i1.351

Thisted RA. 1979. Teaching statistical computing using computer packages. *T*he American Statistician. 33:27–30 https://doi.org/10.1080/00031305.1969.10481895

Toelch U, Ostwald D. 2018. Digital open science—Teaching digital tools for reproducible and transparent research. *P*LoS biology. 16:e2006022 https://doi.org/10.1371/journal.pbio.2006022

Trisovic A, Lau MK, Pasquier T, Crosas M. 2022 A large-scale study on research code quality and executions. *S*ci Data. 9:60 https://doi.org/10.1038/s41597-022-01143-6

Tubb, GW, Ringer LJ. 1977. Current use of computers in the teaching of statistics. *N*ational Bureau of Standards Special Publication 503.

Xie Y, Allaire JJ, GrolemundG. 2018. R markdown: The definitive guide. Chapman and Hall/CRC.

Yu B, Hu X. 2019 Toward training and assessing reproducible data analysis in data science education. *D*ata intelligence. 1: 381-392 https://doi.org/10.1162/dint_a_00053