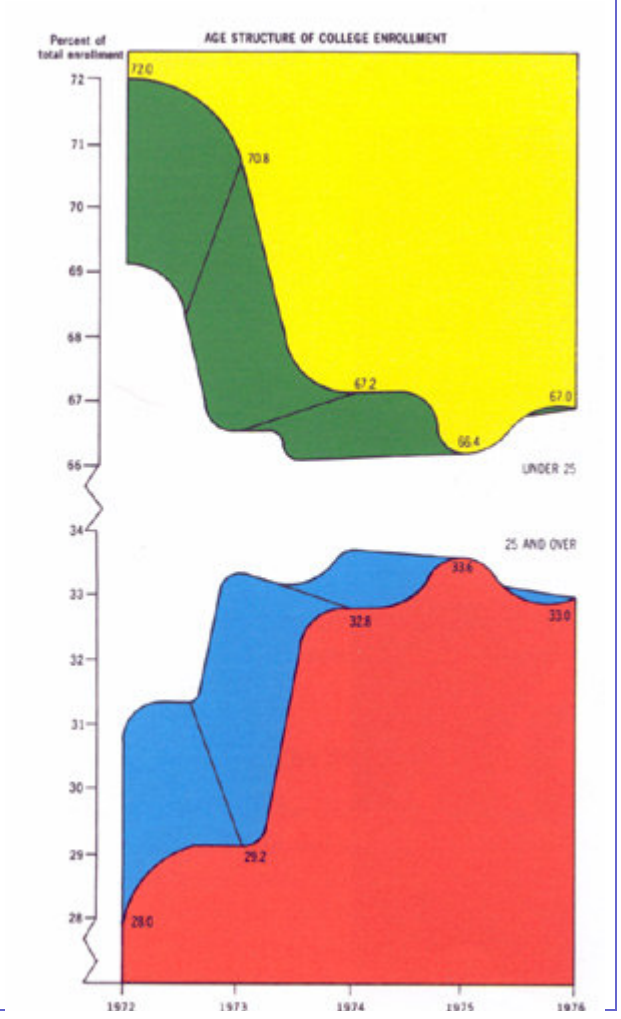
Center for Spatially Integrated Social Science

Edward Tufte: The Visual Display of Quantitative Information, 1983

By John Corbett

Background

It is said that a picture is worth a thousand words. If so, then an informational graphic is worth a thousand data entries if not more. Be they histograms, trend lines, scatterplots, pie charts, or choropleth maps, informational graphics allow us to discern what would otherwise be unmanageable loads of data indecipherable to the average person.



Because of their grounding in facts and statistics, informational graphics have a certain authority of truth and finality. For this reason, they are used by businesses, governments, newspapers, advocacy groups, and scientists as a way of presenting cogent and seemingly immutable arguments for their various causes. Unfortunately, another parallel aspect of their appeal is their ability to be manipulated. After all, a

given graphic is only as truthful—or useful—as the information from which it is composed. It was this problem that had traditionally concerned most critics of informational graphics.

Edward Tufte's groundbreaking 1983 book, *The Visual Display of Quantitative Information*, changed all of that. Tufte, a Professor of Statistics, Computer Science, and Political Science from Yale University, argued that informational graphics were not so much compromised by faulty information as they were by poor and inadvertently misleading presentation. *The Visual Display of Quantitative Information*, along with Tufte's subsequent books and his own passionate crusade for graphical correctness, brought about a revolution in the way that informational graphics were viewed and created. Although he wrote the book before the advent of the Internet or modern graphical software, Tufte's words of wisdom are every bit as relevant today as they were nearly two decades ago.

Innovation

Tufte had two major agendas with *The Visual Display of Quantitative Information*. One was to identify many of the mistakes and abuses common to informational graphics and to finger the main culprits. The second was to go beyond common errors to develop a general theory of data graphics that could be used to explore new ways to increase their efficiency and effectiveness.

Tufte spared little wrath in critiquing the way in which many informational graphics were conceived and published. First and foremost, he deplored the widespread notion that graphics were only an unsophisticated substitute for "boring" statistics. "Many believe," Tufte observed, "that graphical displays should divert and entertain those in the audience who find the words in the text too difficult." This contempt for both the information and the audience had the effect of "[blaming] the victims rather than the perpetrators," and led to what he called "graphic mediocrity."

How did this graphic mediocrity manifest itself? Tufte pointed to the fact that many of those who designed graphics were trained in the fine arts and were not familiar with data analysis. This led to graphic innovations that tended to severely distort or conceal essential information within a graphic. In one notorious example, he showed a graphic from the New York Times showing the increase in the price of a barrel of oil from 1973 to 1979. The data showed the price rising from $2.41 per barrel to $13.34, or an increase of 454%. The graphic used proportional, three-dimensional representations of barrels instead of a histogram or trendline. Although the 1979 barrel was indeed 454% taller than the original 1973 barrel, its diameter had increased by the same amount. When looking at the graphic, one sees that the 1979 barrel could contain more than 270 barrels from 1973! Other instances included trendlines that lacked contextual information and spending charts that failed to take into account

inflationary and population changes.

Tufte's all-out assault on graphic mediocrity segued into a series of principles on how make effective graphics, which he called the Theory of Data Graphics. This theory had four main aspects to it: 1) Elimination of "chartjunk"; 2) Maximization of "data-ink"; 3) Multifunctioning graphical elements; and 4) High data density. Taken together, they represented a new way of thinking about how to more successfully represent data in informational graphics.

Tufte was concerned that creators of graphics were more concerned with advertising their own graphical cleverness than presenting useful data. These artifacts of unnecessary design flourish were termed "chartjunk," and included many items that had been accepted as common features of data graphics. For example, often multiple data series on a given graphic were differentiated by a bewildering array of crosshatching variations, causing visual "vibrations" that distracted the reader. Tufte also implicated the venerable grid, which added little to the data but caused distraction, and in worst-case scenarios even obscured much of the data altogether. The worst offenders went beyond mere chartjunk. Tufte derided them as "ducks", a reference to the fad of building eye-catching structures that served as visual puns, such as hamburger-shaped burger joints and duck-shaped buildings. Ducks often featured trendlines that were spruced up with flashy colors and three-dimensions, making the data look more like rollercoasters than a serious attempt to convey information.

The overuse of chartjunk tended to unnecessarily obscure the use of what Tufte dubbed "data-ink", namely the ink used in creating the elements of the chart that directly conveyed data, such as plot points or trendlines. The more chartjunk in a graphic, the smaller the proportion devoted to data-ink (what Tufte termed the "data-ink ratio"). To make a graphic effective, the data-ink ratio needed to be increased as much as was reasonable. If a grid was not necessary, remove it, or at least limit the number of gridlines to a bare minimum. "Gratuitous decoration" of graphics, such as extravagantly-curved histogram bars, should be avoided. Even the traditional histogram was under attack—Tufte believed in doing away with "redundant data-ink", where too much ink was expended on presenting a single data point. Why not replace a two-sided bar with a one-sided bar? Tufte asked. He even proposed doing away with axes that unnecessarily extend all the way to zero.

A good graphic should not only clearly and neatly present information; it should also "mobilize every graphical element, perhaps several times over, to show the data." It is this belief that is at the heart of Tufte's assertion that the 1861 map of Napoleon's Russian campaign, drawn by the French engineer Charles Joseph Minard, "may well be the best statistical graphic ever drawn." Minard's depiction of space, direction, time, quantity—even temperature—in showing the gradual

disintegration of Napoleon's army packed an enormous amount of information into a single graphic. Other maps may have made less of an impression on Tufte, but still employed novel uses of the visualization of data. One histogram from 1951 was used to show both the length and the pattern of road striping for each U.S. state. Another histogram reduced the amount of "eye-work" by using the value along the x-axis in place of the typical data plot. Still another chart used levels of gray to show the concentration of galaxies in the night sky. In contrast, Tufte held up a complex color choropleth map for criticism, noting that since the average human cannot easily grasp order out of multiple colors, it only adds extra work to the understanding of the map.

Finally, Tufte was a proponent of high data density. He defined it as the ratio of the number of entries in a data matrix to the area of the data graphic. Of course, such a ratio is often hard to quantify, but his point was clear: Don't waste a large graphic on a small amount of information. If there are only a couple data entries, a table within the text makes more sense than creating a histogram with only a handful of bars. One example of a graphic using a high density of data showed the amount of sunny versus clouded periods for each day throughout the year at a given location. Another featured a pollution map that was repeated multiple times to create a form of animation over the course of a day.

The Visual Display of Quantitative Information changed the way that people thought about the presentation of informational graphics. The discussion that began in 1983 continues to this day, partly because of Tufte's own advocacy, but in large part because of the common sense that his arguments made. To this day, web designers and publishers familiarize themselves with Tufte's work in order to fine-tune their works. Yet Edward Tufte's work is far from finished. Graphical software packages still include heavily used features that add color, gridlines, and extra visual dimensions to two-dimensional data. Worst of all, they still incorporate Tufte's greatest pet peeve of all: the pie chart.

Publications

The visual display of quantitative information. Cheshire, CT: Graphics Press, 1983.

The quantitative analysis of social problems. Reading, MA: Addison-Wesley, 1970.

Visual explanations: images and quantities, evidence and narrative. Cheshire, CT: Graphics Press, 1997.

Envisioning Information. Cheshire, CT: Graphics Press, 1997.

Data analysis for politics and policy. Englewood Cliffs, NJ: Prentice-Hall, 1974.

Related Works

Levens, A.S. *Graphics, analysis and conceptual design*. New York: John Wiley & Sons, 1968.

Harris, Robert L. *Information graphics: A comprehensive illustrated reference*. New York/London: Oxford University Press, 1999.

Wildbur, Peter, and Michael Burke. *Information graphics: Innovative solutions in contemporary design*. London: Thames and Hudson, 1998.

Wodtke, Christina. *Information architecture: Blueprints for the Web*. Indianapolis, IN: New Riders, 2002.

Stovall, James Glen. *Infographics: A journalist's guide*. Boston: Allyn & Bacon, 1997.

Links The Work of Edward Tufte