perceptual edge

Bad Graphs: The Stealth Virus

Stephen Few
January 2005

I'm on a mission! No, I didn't hear the voice of God, but I have heard the voices of countless people crying out in frustration for the data they desperately need to do their jobs, presented to them clearly, simply, efficiently and meaningfully. The current state of data visualization—the use of graphs to display information for the purposes of sense-making and communication—would be amusing if it weren't so tragic. Most business graphs are poorly designed, many to the point of misinformation.

In the world of business, the damage done by bad graphics has reached epidemic proportions, but so far only a few voices are being raised in warning. Bad graphics act as a virus that stealthily destroys information with little notice. Sadly, most business intelligence (BI) software vendors aren't helping matters. They are busy throwing Botox parties, encouraging the popularity of cosmetic dazzle rather than fitness programs for healthy business communication.

I can promote good data visualization practices by writing articles and columns, writing books, consulting and teaching, but only you can work in the trenches at your own places of business to raise awareness and apply effective practices where it really counts. To communicate or not to communicate? When presenting data, that is the question—the only question.

The graph in Figure 1 brings a bit of what I'm saying to light. I found this graph shown in Figure 1 on the Web site of Visual Mining, Inc. In some ways, it is quite attractive, with lovely pastel colors and alluring 3-D. Ultimately, however, there is only one question to ask when judging its worth: How well does it communicate? The answer is: Poorly. To be fair, at least one thing was done right. Lines were used to visually encode the movement of profits through time, rather than bars or any other means. Otherwise, this graph exhibits a series of bad design choices:

- It was rendered in 3-D, even though there are only two dimensions of data (time in months and profits in dollars). Varying the positions of the four lines depth-wise along the Z-axis and rendering the lines themselves as 2-D objects (making them look like ribbons) makes it much more difficult to interpret the values by aligning them with the quantitative scale.
- Values for the month of May are missing for clamps and wrenches (and probably hammers also). Perhaps there is a good reason for this, but if so, it should have been stated.
- The quantitative scale doesn't extend far enough in the negative direction, indicated by the fact that hammer profits appear to drop below -$200,000 where the line gets cut off at the bottom of the graph between March and April.

- Using a legend rather than labeling the four lines directly forces readers to constantly shift back and forth between the data and the legend, unnecessarily stressing the limits of short-term memory and adding unnecessary effort.
- The fact that cents are included in the numbers on the quantitative scale suggests a level of numeric precision the graph can't support. Showing "$.00" for each of the numbers provides meaningless information.
- Orienting the X-axis heading "Profit Per Part ($K)" vertically makes it difficult to read, unnecessarily adding to the time needed to read the graph.
- The alternating fill shades of gray along the side and back walls of the graph serve no useful purpose, and the fact that those on the side don't match with those on the back simply creates confusion.

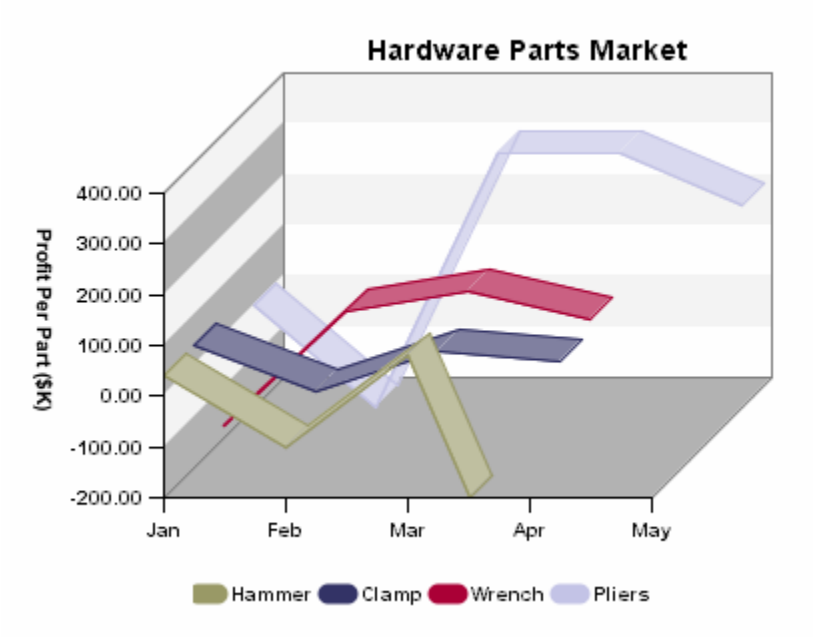

Figure 1: Visual Mining Example

The graph in Figure 1 is a fine example of catering to the desire for dazzle, not in a way that is harmless, but in a way that severely undermines the graph's ability to communicate. In this case, a simple line graph could have communicated the message nicely.

Let's look at another example from the Web site of SAS (see Figure 2). It isn't clear what message this collection of four pie charts is meant to emphasize. If it is to show the trend of each energy source's contribution to the whole across time, it fails miserably. If it is to compare the contribution of gas (the exploded slice) to the other energy sources in any given year, there are much clearer ways to accomplish this. Not only was the selection of pie charts inappropriate for this information, several aspects of the layout undermine the effectiveness of this display, including:

- The arrangement of the years into two rows, when a single row of years from left to right would provide a more intuitive representation of time.
- The excessive use of text, which slows perception of the values, overlaps in confusing ways between the various pie charts and just plain looks messy.

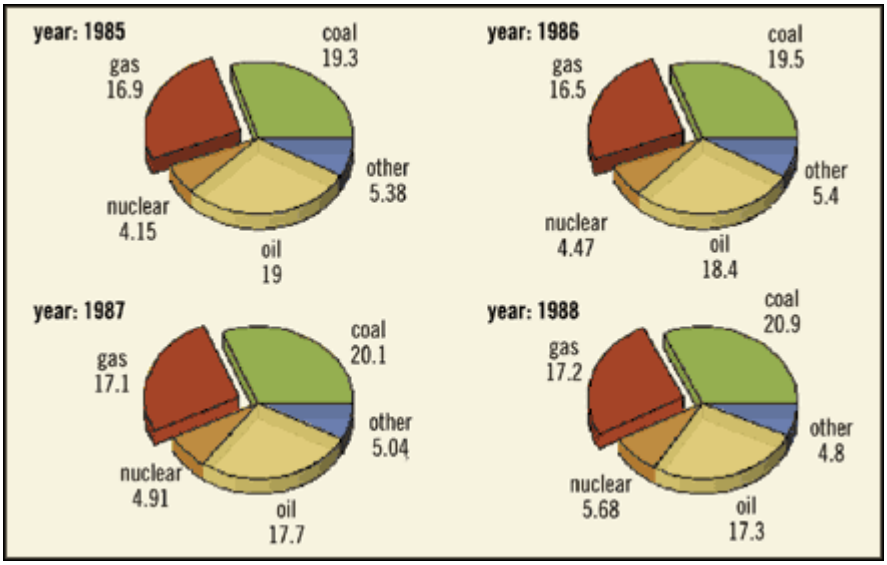

Figure 2: SAS Example

An alternative solution is shown in Figure 3. This graph makes several things immediately clear that you couldn't easily see in the previous example. You can now easily discern the trend of each energy source's contribution to the whole as it moves through time. The two green lines representing coal and nuclear both rose, while the red and orange lines both declined and gas stayed about the same. It is also easy to compare the relative contributions of each energy source in any given year, both by the vertical positions of the five data points as well as the positions of the labels (coal, oil, and so on). The emphasis on gas, shown in the original graph by the exploding slices, has been preserved here by boldfacing its labels and encoding its line as dark black. Assuming that the degree to which each energy source's contribution to the whole has changed from 1985 through 1988 is important, this has been communicated as text to the far right—otherwise, the dramatic rise of nuclear energy would have gone unnoticed.

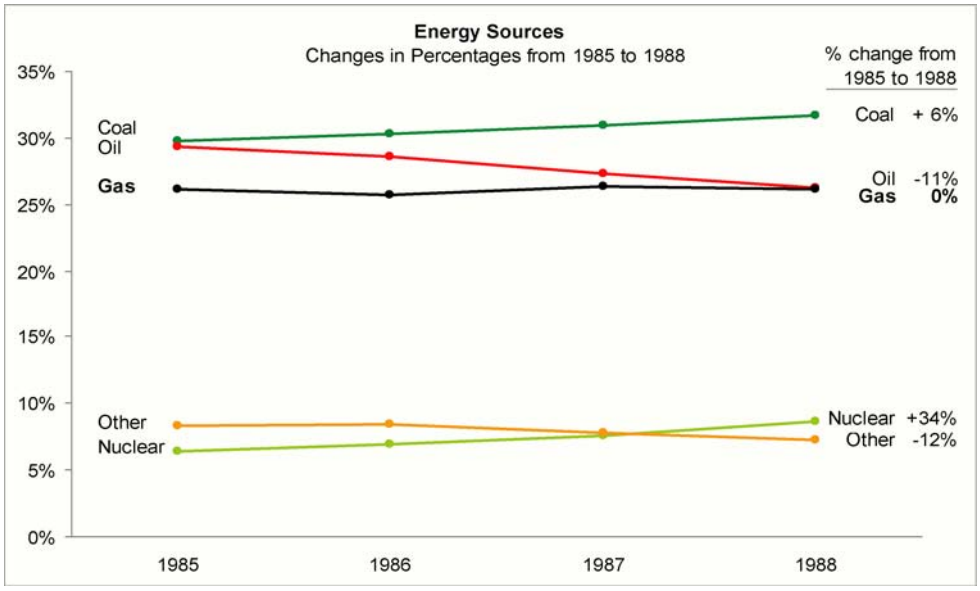

Figure 3: Alternative Solution

Don't let your BI investment lose its value in the final stage of the information life cycle - the use of that information to gain important insights and communicate them to decision-makers. Poorly designed dashboards, tables and graphs strip information of its usefulness. Analytic software assists people who already know what they are doing; it doesn't think for them. The expectation that you can put BI software on someone's desktop and they will intuitively know how to use it effectively is a marketing myth. Business intelligence requires professional skills that don't come naturally, they must be learned.

In future columns, I will focus on practical, accessible and instructional data visualization information that you can apply immediately to your work. I want to help you learn to produce what Joseph Berkson once described as an "interocular traumatic impact"—a conclusion that hits people between the eyes.

(This article was originally published in *DM Review*.)

About the Author

Stephen Few has worked for over 20 years as an IT innovator, consultant, and teacher. Today, as Principal of the consultancy Perceptual Edge, Stephen focuses on data visualization for analyzing and communicating quantitative business information. He provides training and consulting services, writes the monthly *Visual Business Intelligence Newsletter*, speaks frequently at conferences, and teaches in the MBA program at the University of California, Berkeley. He is the author of two books: *Show Me the Numbers: Designing Tables and Graphs to Enlighten* and *Information Dashboard Design: The Effective Visual Communication of Data*. You can learn more about Stephen's work and access an entire library of articles at www.perceptualedge.com. Between articles, you can read Stephen's thoughts on the industry in his blog.