Visualizing Multidimensional Data Through Time

Stephen Few
July 2005

This is the first of three columns that will feature the winners of DM Review's 2005 data visualization competition. I want to extend my sincere appreciation to all of you who submitted solutions to the competition. I enjoyed many hours of studying and comparing your responses in an effort to select the best and learn something from them all. In this column, I am focusing on the first scenario of the competition, which involved the visual display of salary expenses along the dimensions of time, department and exempt versus non-exempt employees. Here are the requirements of the scenario as presented to participants in the competition:

This scenario involves the display of departmental salary expenses. It is used by the VP of Human Resources to compare the salary expenses of the company's eight departments as they fluctuate through time, in total and subdivided between the exempt and non-exempt employees.

As you can see, this scenario requires a visualization of the data that makes it easy to do each of the following:

1. Compare the salary expenses of the various departments to one another,
2. See how they changed through time,
3. And do so in a way that clearly presents the salaries of exempt and non-exempt employees both separately and in total (that is, exempt and non-exempt salaries combined).

Each of these requirements places particular demands on the design of the data visualization. Although they are all critical, representing the data in a way that clearly shows how data changes through time is perhaps the requirement that people struggle with the most. Of all the means available to encode values in graphs, lines are the best for highlighting the patterns and trends of values as they change through time. Take a look at the winning solution in Figure 1, which was submitted by Tableau Software, to see how effectively the lines were able to paint this picture.
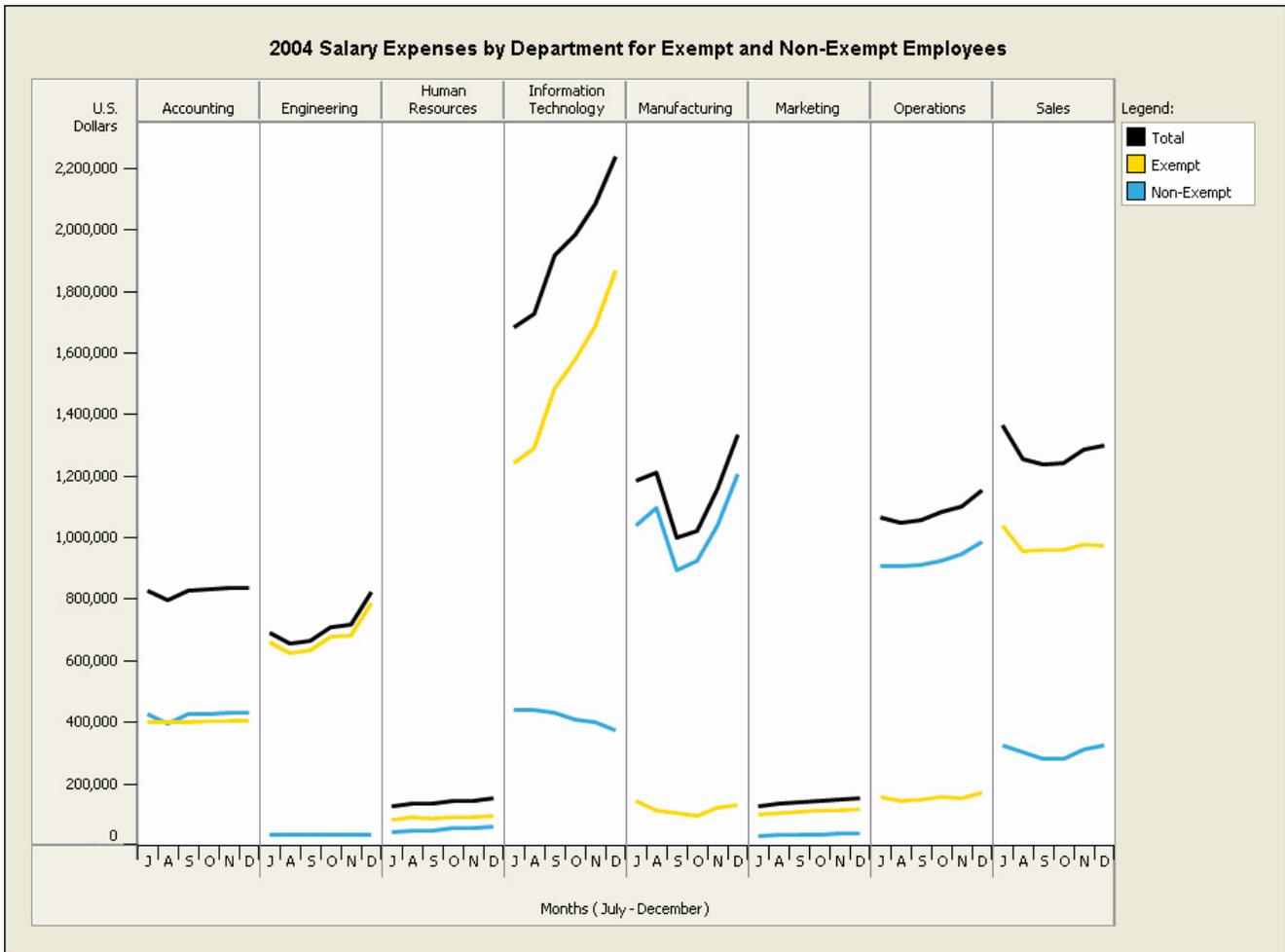
Figure 1: Winning Solution, submitted by Tableau Software

This columnar arrangement of line graphs (one for each department), which all share the same quantitative scale, makes it quite easy to compare the various departments' salary expenses. The fact that salary expenses for the information technology department are greatest and that those for human resources and marketing are the least each stand out quite clearly. It is also easy to compare the shape of salary expenses as they vary through time among the departments, with the information technology department exhibiting the greatest overall increase, despite the fact that non-exempt salaries actually decreased. The greatest volatility shows up in the marketing department's non-exempt salaries, which shot down in September then back up in November and December. The separate display of exempt and non-exempt salaries through the use of distinct hues (gold and blue), as well as their combined totals using a darker hue (black), allows us to easily discern the differences between these categories and how each contributes to the whole. The design of this data visualization might strike you as extremely simple in most respects, but, as you can see, this is precisely what works. Clarity of communication, not an attempt to dazzle through unnecessary visual complexity, is the clear goal of this design.

The effectiveness of this display becomes more obvious when you compare it to others that do not communicate as well. Although beautifully rendered, the graph in Figure 2 simply doesn't present the information as clearly. The vertical bars don't feature the ups and downs and overall trends of the values through time as well as lines. The 3-D design, despite its

popularity these days, is just plain difficult to interpret. Try to get a sense of the message in the data by studying this 3-D display. Even after a great deal of effort, the picture is never entirely coherent. The use of stacked bars to separately represent exempt and non-exempt salaries while simultaneously representing their combined totals does a fine job of showing the total and an adequate job of encoding exempt salaries as the lower of the two bar segments, but notice how difficult it is to interpret and compare the non-exempt salaries in the upper segment of the bars. The fact that they each begin at different levels along the quantitative scale makes this nearly impossible.
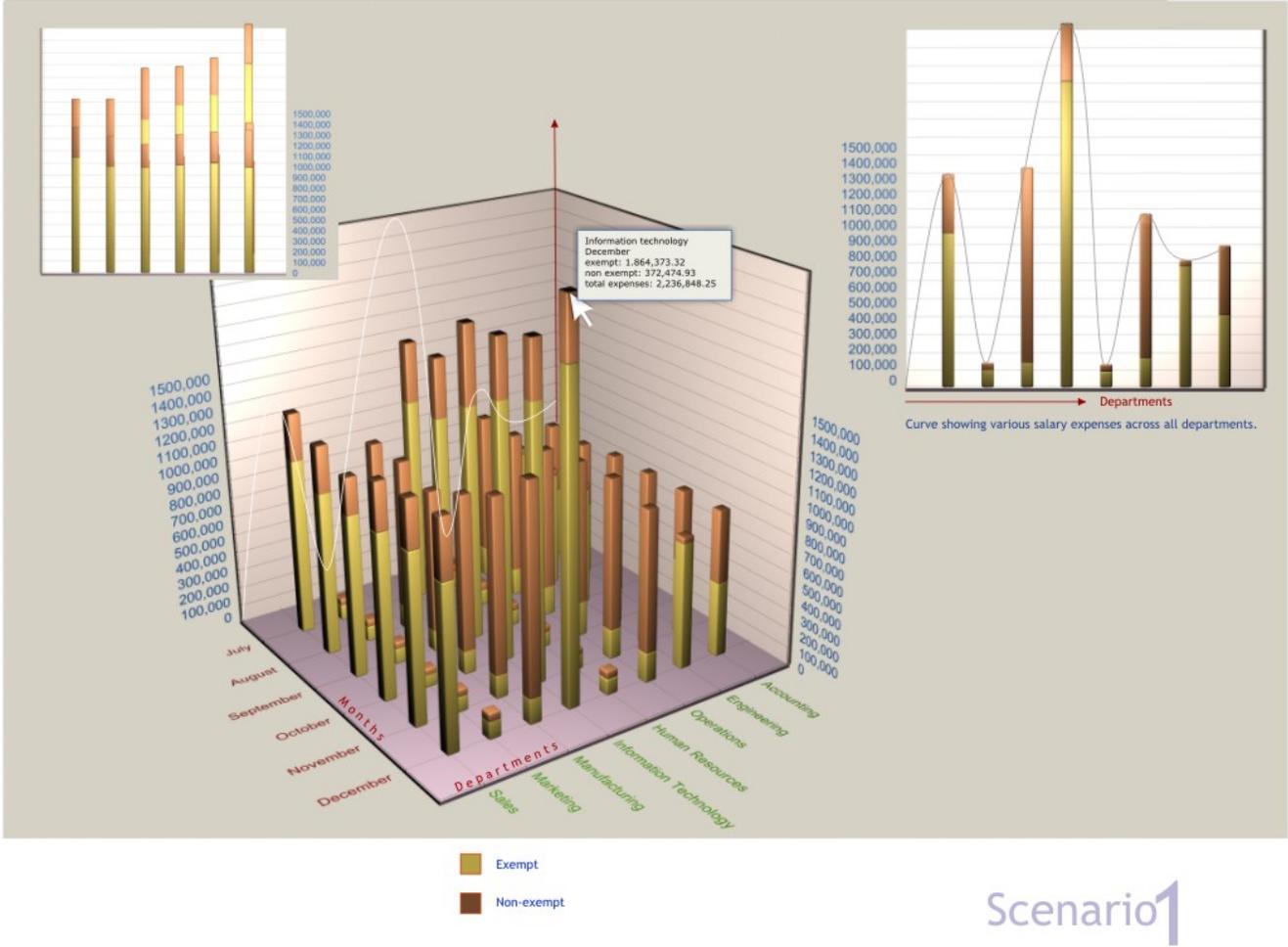

Figure 2: Example Solution

The approach shown in Figure 3 strays even further from one that fits the way that visual perception works. Arranging individual pie charts from left to right, one for each month, to present how the values changed through time (in this case encoded as a range of colors from white through blue, green, brown and eventually black) simply doesn't work. We don't intuitively perceive a spectrum of hues quantitatively. It looks interesting, but fails to communicate.
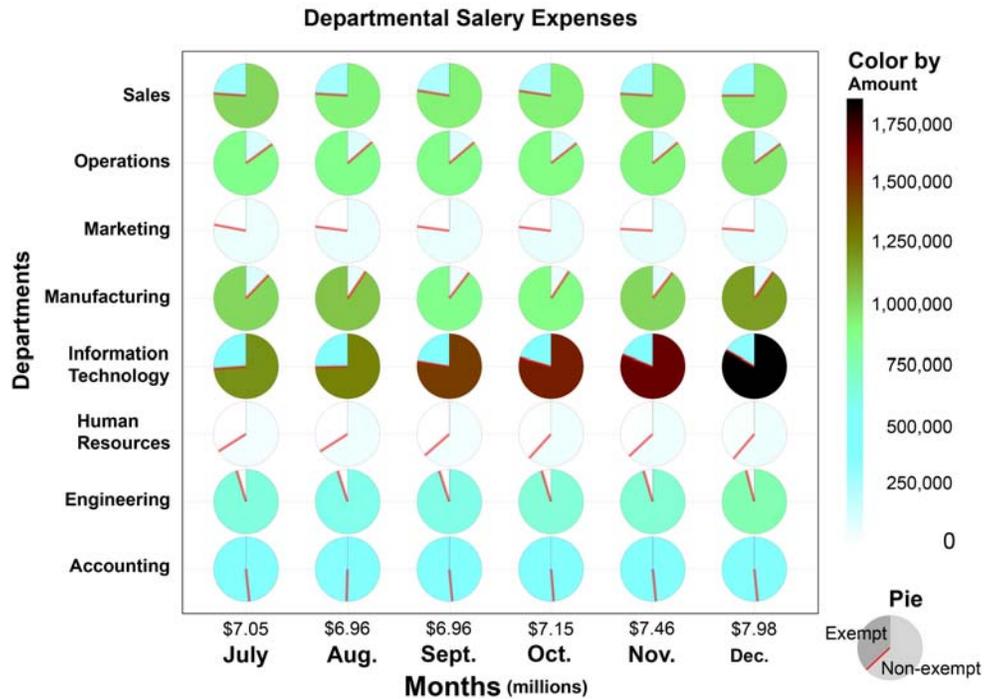
Figure 3: Example Solution

The final example, which appears in Figure 4, fails to communicate on a number of fronts. Take the time to examine it closely and see if you can identify the characteristics of this design that undermine its ability to clearly present the data.
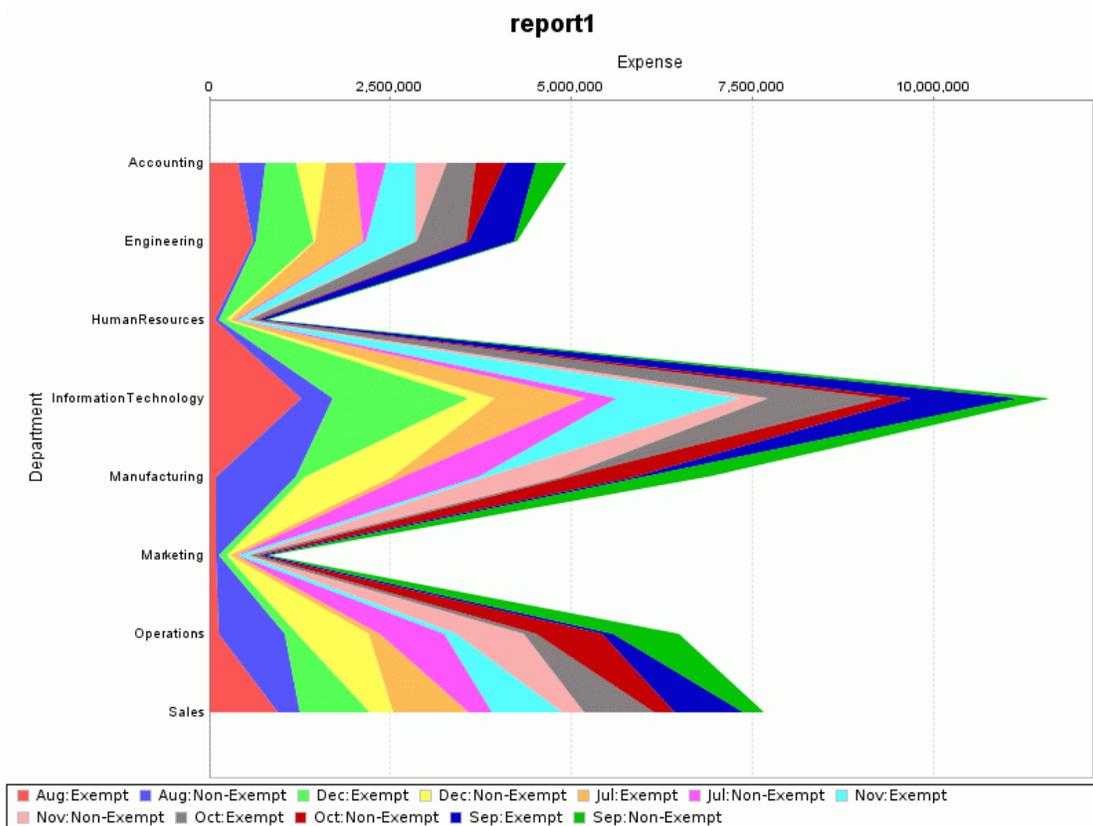

Figure 4: Example Solution

One problem that jumps out immediately to me is the way that the months and the distinction between exempt and non-exempt salaries have both been encoded through the use of different hues. This approach paints a bright picture, but not one that can be meaningfully discerned. It is easy to see which department's salary expenses were greatest for the entire six-month period and which were least, but to interpret the values for a specific department, you must compare the widths of the individual segments of color by scanning horizontally to the right, starting from a department's name, adding each pair of colored segments to get monthly totals. The patterns formed by each of the hues as its area flows and increases or decreases moving vertically from one department to the next suggests change, but the departments are distinct entities and the differences in values between them are distinct measures, not something that is changing from one instance to the next.

As you can see, there are many possible approaches to visualizing the same information, but some are clearly more effective than others. Unfortunately, many of the approaches that are common today hide important information in a confusing picture.

(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.