

Data Visualization: Rules for Encoding Values in Graph

Stephen Few

January 17, 2006

When data is communicated graphically, just like verbal communication using language, certain rules of syntax and semantics apply. If you disobey the rules, you run the risk of being misunderstood. The rules of graphical communication are rarely arbitrary, but are usually based on an understanding of visual perception—how we see and the ways in which information can be visually encoded for easy and accurate decoding by our audience.

Most graphs that are used to present quantitative business data are two-dimensional with two axes (one horizontal, called the X axis, and one vertical, called the Y axis), and use one or more of three particular objects to encode values: *points*, *lines* and *bars*. The choice of which one or more of these three objects to use in a graph should never be arbitrary, and need never be, because the rules are simple to understand and follow. Why bother? It all boils down to *communication*. If your data is worth reporting, it is worth reporting well.

Points and Their Uses

Points are the data-encoding objects with the simplest possible shape. They pinpoint a specific location on a graph in a way that no other object can, due to the fact that they have negligible height and width. In the context of a 2-D, XY axis graph, points encode values by their location in association with the scale along each axis. In Figure 1, the left-most data point encodes a value of two on the Y axis and is associated with the categorical label "A" on the X axis.

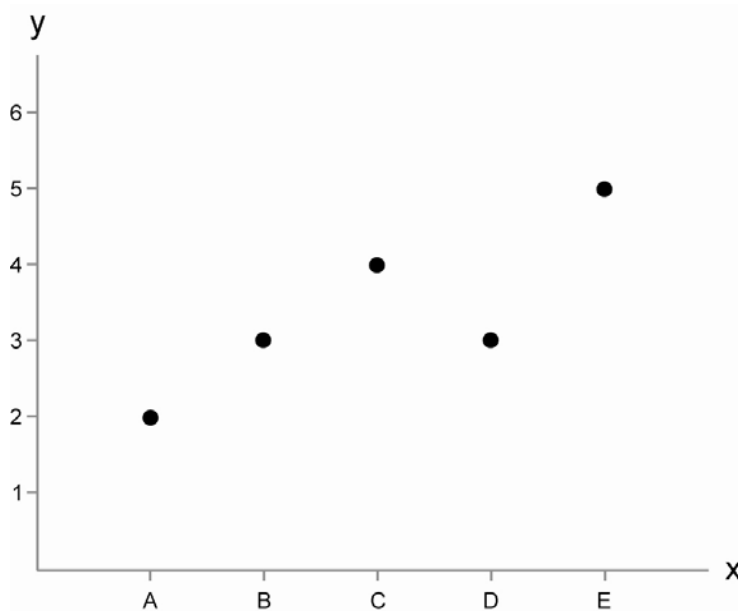


Figure 1: Points encode values based on their location in association with scales along the axes.

Points can take any simple shape, including dots, squares, triangles, diamonds, x's, plus signs, and dashes. When the points in a graph only need to encode a single set of values, and therefore require only a single shape, I prefer to use dots, because they are the simplest visually.

There are only a couple of circumstances when points by themselves (that is, without lines to connect them) are the most effective option. Figure 2 illustrates the primary circumstance. Take a look at it and see if you can determine why points alone in this example are superior to lines or bars.

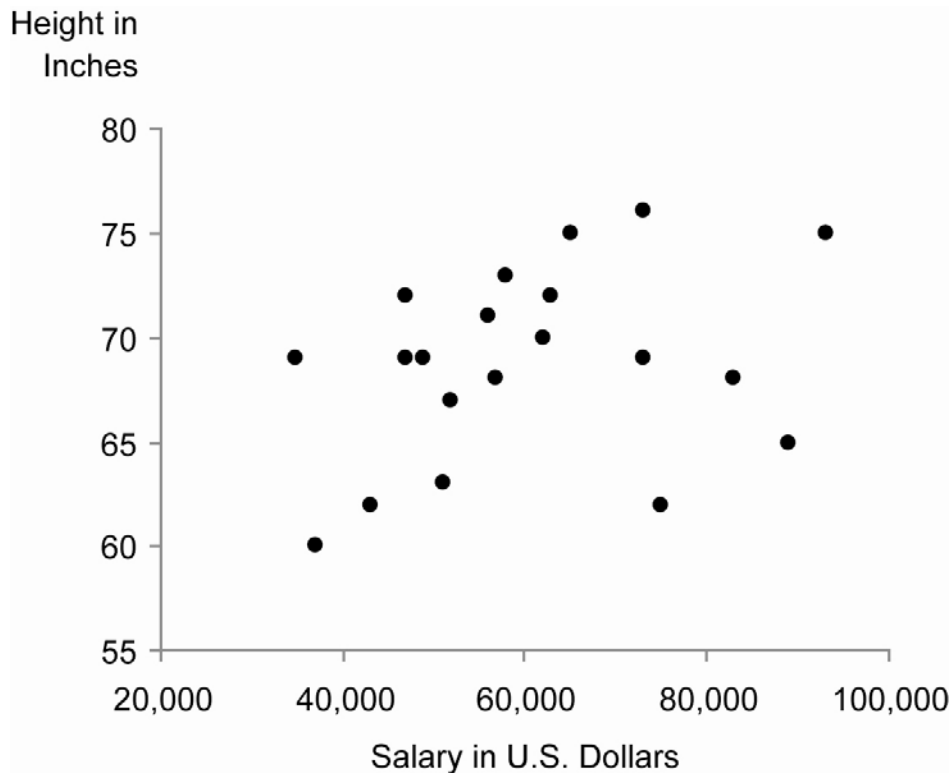


Figure 2: This graph illustrates the primary strength of points when used alone to encode data.

If the reason isn't clear, let me give you a hint. Notice that the scales on both axes are quantitative. Except in the case of this kind of graph, called a *scatter plot*, one of the axes on a graph has a categorical scale—one that labels what is being measured (for example, departments or regions) rather than quantitative values. Every point on the graph in Figure 2 encodes two quantitative values: one along the X axis (in this case a person's salary in dollars) and one along the Y axis (in this case a person's height in inches). Scatter plots include two quantitative scales because they are specifically designed to display the correlation (or lack of one) between two sets of measures, in this case whether there is a correlation between how tall someone is and the salary that person earns. Imagine using bars to encode this data. Bars extending into the graph from both axes would produce a cluttered mess. Now try to imagine using lines. Lines connecting each of these data points would produce a meaningless, meandering squiggle. Only a point can be used to simultaneously encode two quantitative values based on a horizontal and vertical location, because points alone, without height and width, occupy space in the minimum way that is needed to do this.

There is one other circumstance where points alone work better than lines or bars, but we'll come back to that a little later. Before moving on, however, I want to make it clear that points alone are

never a good option for encoding a series of values through time (that is, *time-series data*). It is too hard to follow the chronological sequence of the values when points alone are used, as you can see in Figure 3.

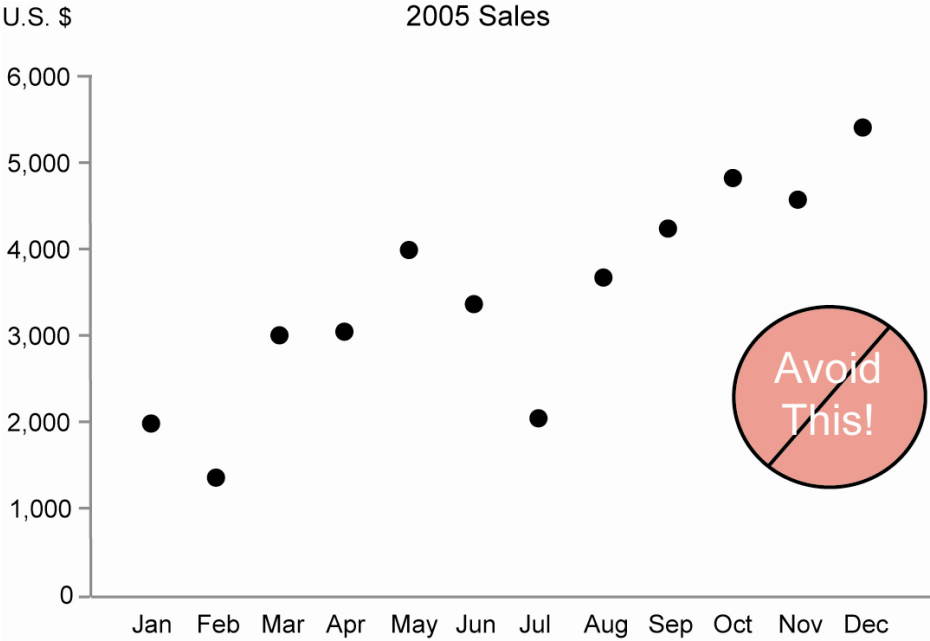


Figure 3: Points alone do not clearly display the sequence of time-series data.

When points are connected by lines, however, the sequence of time-series data becomes easy to follow, illustrated in Figure 4.



Figure 4: By adding lines to connect the points, the sequence of time-series data can be clarified.

Lines and Their Uses

Lines can be thought of as points extended through space from one value to the next to connect them. Similar to points, lines encode values by means of their location in relation to the scales along the axes; unlike points, only the ends of each segment of the line mark the locations of values.

The strength of lines is more obvious than points or bars. Looking at the graph in Figure 4 above, it is easy to see that lines, by connecting each value to the next, show the overall shape of the values in a way that points or bars cannot. Figure 5 shows the same data as Figure 4, this time using only lines.

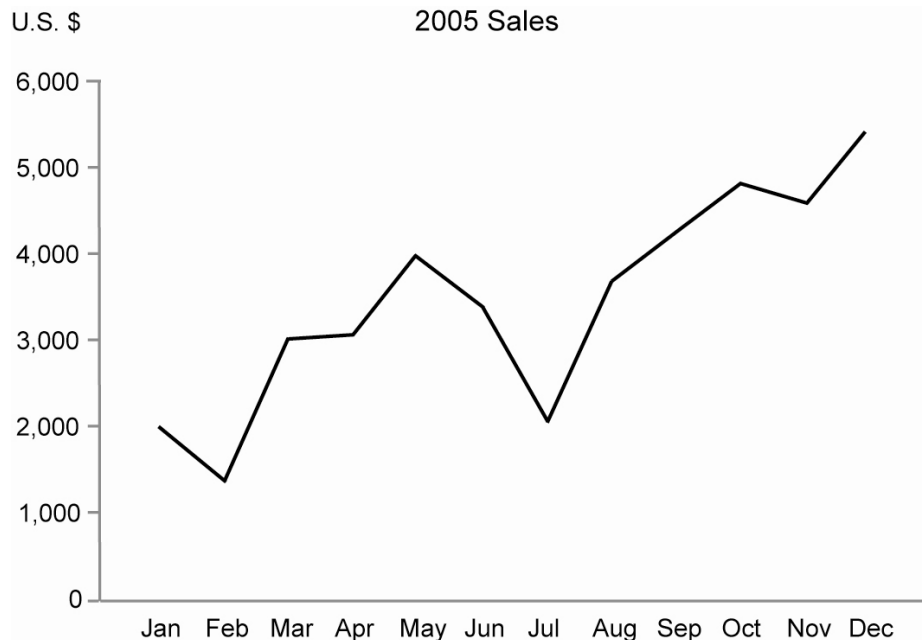


Figure 5: Lines alone show the overall shape of values as they move up and down through time.

As you can see, lines alone display the overall shape of data as it changes through time more clearly than any other encoding method. Nothing distracts from the pure movement of the data up and down from one value to the next. Lines excel in their ability to show trends and patterns of change.

When you wish to emphasize the overall shape of time-series data but must also place additional focus on the individual values, the combination of points and lines works well. This is especially useful when you display multiple data sets across a time series, each as a separate line, and want to make it easy to compare individual values at particular points in time to one another. Including points in Figure 6 makes it a little easier to compare the three values for the month of May, for example, than it would be if the data were encoded using lines alone.

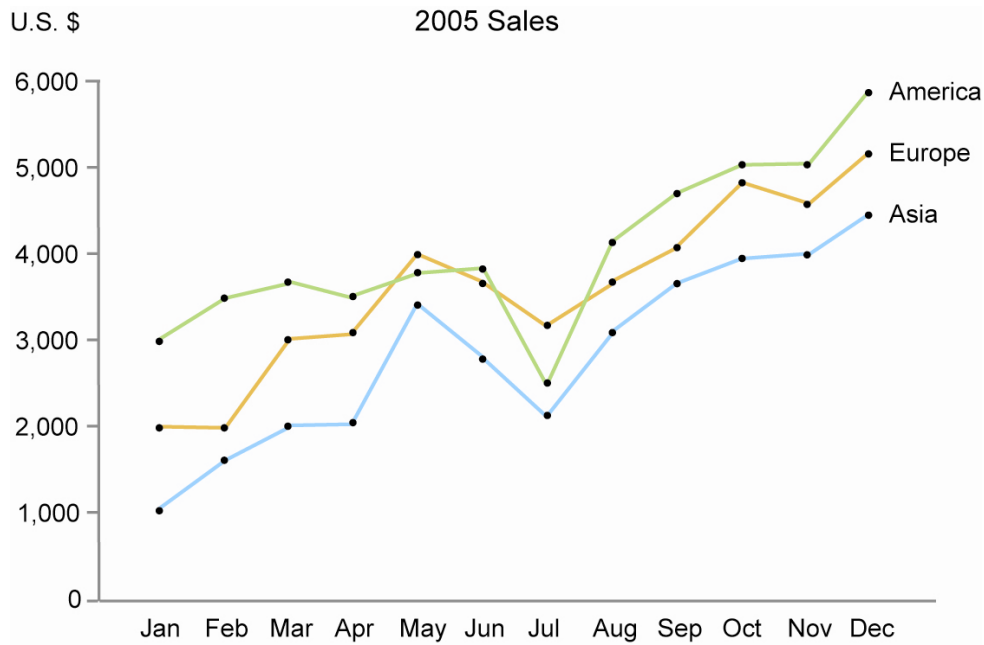


Figure 6: Points make it easier to compare multiple values at the same point in time.

One of the most important guidelines to keep in mind about lines is that they should only be used to connect values that are themselves intimately connected to one another. Changes in the amount of sales from one month, quarter, or year to the next are intimately connected to one another. On the other hand, a set of values that measure the expenses of different departments are not intimately connected; they are discrete and should be displayed in a manner that visually suggests their discreteness. Connecting discrete values with a line does not properly depict the relationship between those values, as shown in Figure 7.

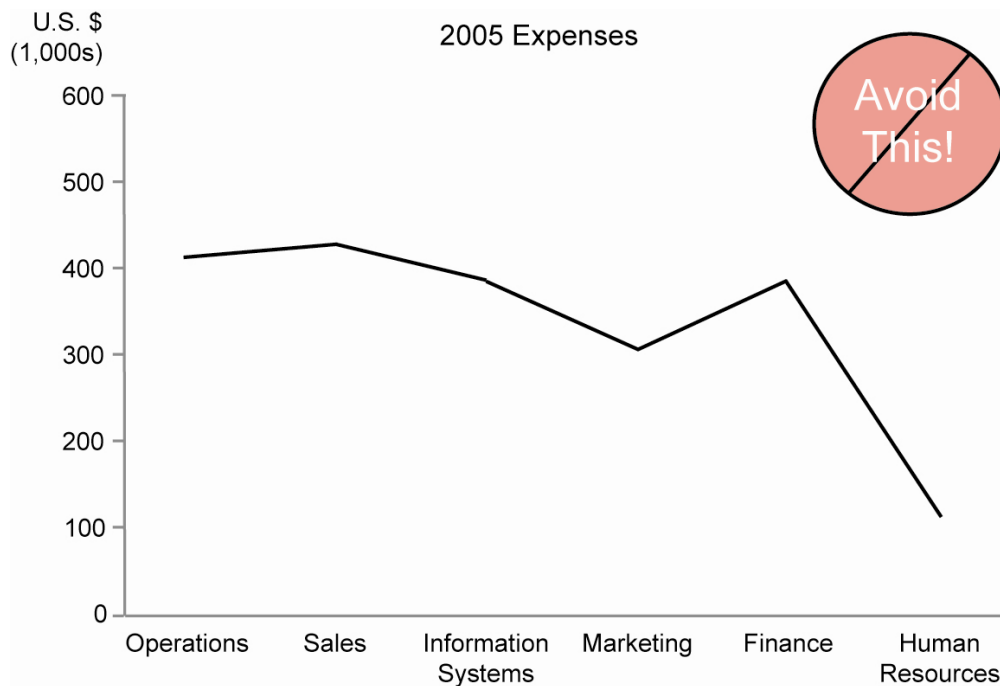


Figure 7: Using lines to connect discrete values doesn't make sense.

When lines are used properly in graphs, their slope is meaningful. For instance, with time-series data, the slope of the line from one value to the next represents the rate of change—the steeper the slope the greater the rate—but with discrete values, such as those for each of the departments above, the pattern formed by the line and the slopes from one value to the next have no meaning.

Bars and Their Uses

Bars are visually the most weighty and dominant of the three objects that we commonly use to encode data in graphs. Unlike points and lines, which encode values as location relative to a scale along an axis, bars encode quantitative values in two ways. Similar to points and lines, the location of a bar's endpoint encodes its value. Unlike them, a bar's length (or height, depending on how you look at it) also encodes its quantitative value. When bars are properly used, you can compare their values by comparing their lengths. Because bars have a great deal of visual weight and stand out so clearly as separate from one another, it is very easy to compare their lengths to determine the relative magnitudes of the values they encode. Take a look at Figure 8 to see how well this works.

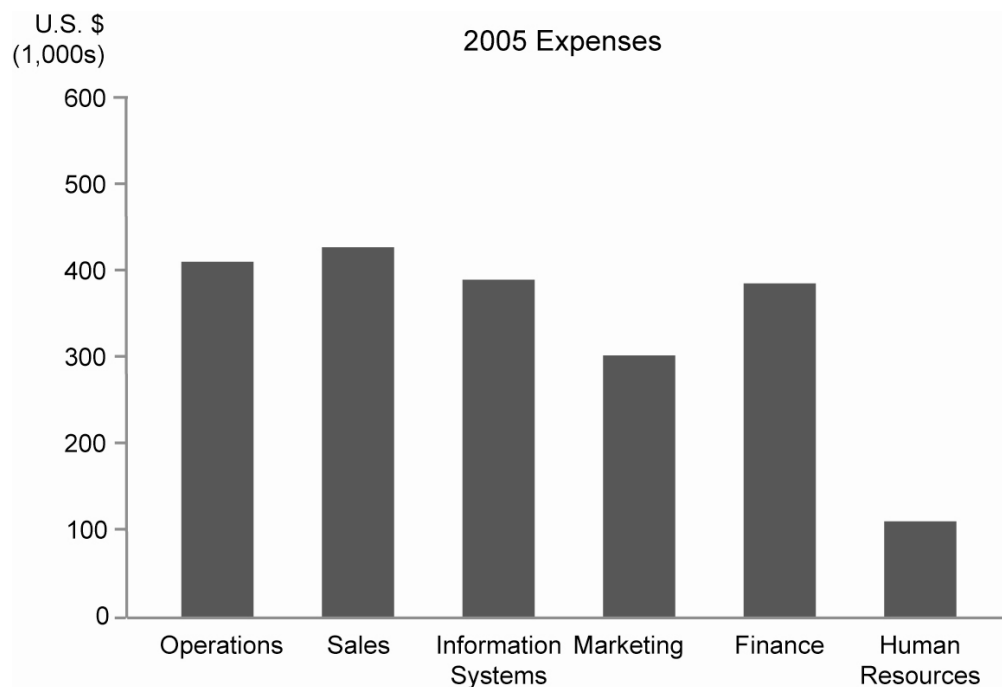


Figure 8: Bars do a great job of representing discrete values and supporting their comparisons.

When you wish to help your audience focus on individual values and compare individual values to one another, bars are ideal. This works especially well for discrete values that are not intimately connected to one another. Bars can be used, however, to encode values along a series of intimately connected values, such as a time series, when you're less concerned about showing the overall shape of the values (which a line would do better) and more concerned about helping people examine and compare individual values. Figure 9 illustrates an occasion when bars work well for time-series data, because the primary purpose is to help people compare actual to budgeted expenses at a particular point in time rather than to see the overall trend of those expenses through time.

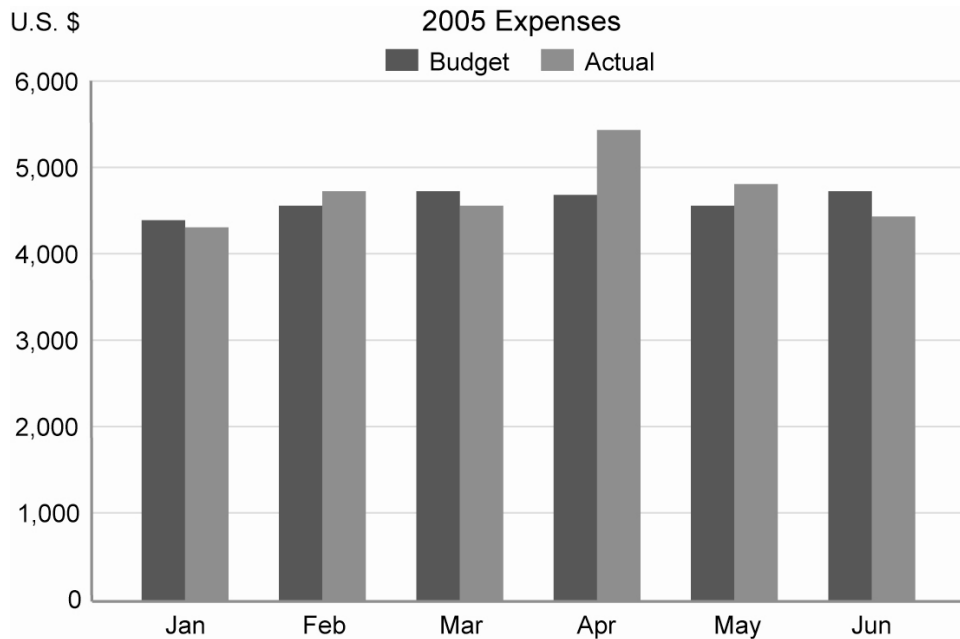


Figure 9: Bars can be used to encode time-series data when you intend primarily to support the examination and comparison of individual values.

Figure 9 can help us learn another lesson about bars. Notice how all of the bars fall within the range of \$4,250 and \$5,500. They are also all fairly tall, and, with only one exception, fairly close in size. Sometimes when values fall within a relatively narrow range and are all far from zero, it is useful to narrow the quantitative scale such that it begins just below the lowest value and ends just above the highest value. This spreads the values over more space in the graph, enhancing their differences so they can be examined in greater detail. Look what happens, however, when we do this with bars.

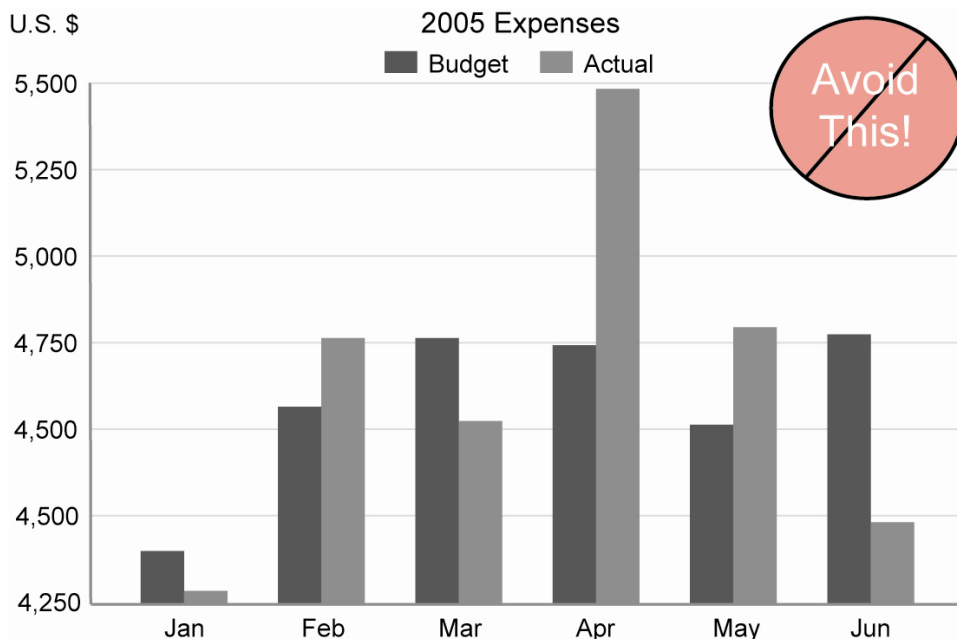


Figure 10: Bars don't work unless the quantitative scale begins at zero.

We now have a problem, because the length of each bar is supposed to represent its value such that comparisons of their lengths support accurate comparisons of their values. Based on a comparison of bar lengths, actual expenses in January appear to be about 1/15th the amount of those in February, but this is hardly the case. For bars to work properly, they can only be used with a quantitative scale that begins at zero. What do you do, then, when you wish to narrow the quantitative scale to make more fine-tuned comparisons of the values? With the data in Figure 10, we could use lines and points, because lines can be used to display time series, but what if the values are discrete and not intimately related? The answer is to take advantage of the discrete appearance of points, which can be used with a narrowed quantitative scale that does not begin at zero, because they encode values using location alone, not length as bars do. Figure 11 provides an example of a point graph that solves the problem.



Figure 11: Points can be used in exchange for bars when you wish to narrow the quantitative scale, but still focus on individual discrete values and their comparisons.

These guidelines for the effective use of points, lines and bars to display quantitative data in graphs are quite easy to understand and follow using just about any software, from *Microsoft Excel* to the most sophisticated business intelligence software available. As you can see, these guidelines aren't arbitrary, and neither is the effective communication that results when you follow them. Happy graphing!

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](http://www.perceptualedge.com) of articles at www.perceptualedge.com. Between articles, you can read Stephen's thoughts on the industry in his [blog](#).