Dual-Scaled Axes in Graphs

Are They Ever the Best Solution?

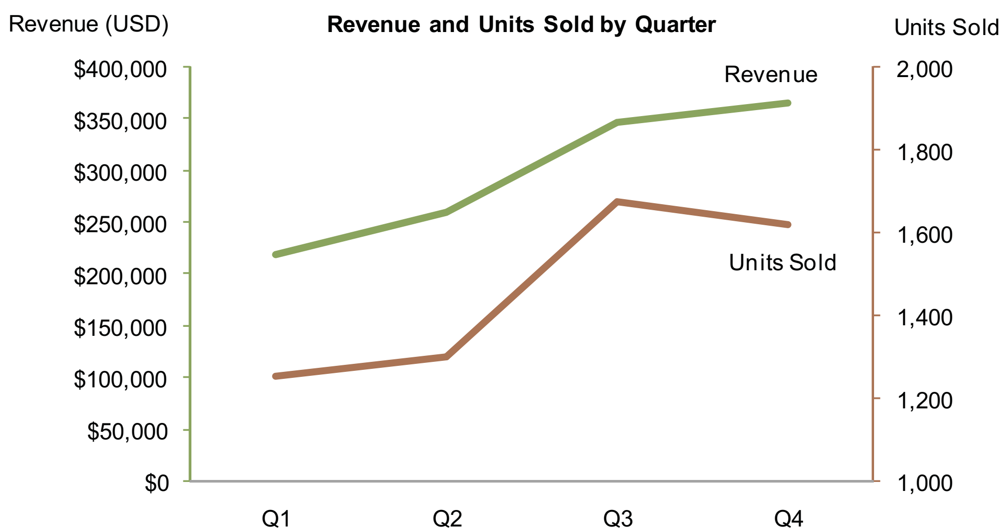
Stephen Few, Perceptual Edge

Visual Business Intelligence Newsletter

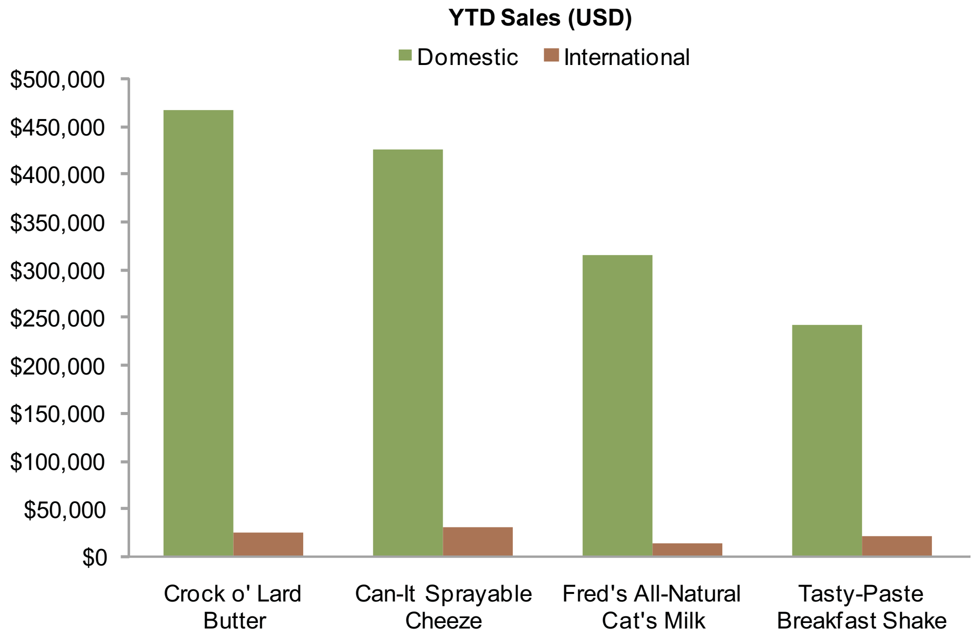
March 2008

In 2004, when I wrote the book *Show Me the Numbers*, and even more recently when I wrote *Information Dashboard Design*, I considered graphs with two quantitative scales on a single axis (either X or Y) a viable option. *Show Me the Numbers* originally included a solution to a graph design problem that displayed two quantitative scales on the Y-axis—one that measured quarterly revenue in U.S. dollars and another that measured the number of units sold for the same four quarters. In *Information Dashboard Design*, my sample sales dashboard to this day still includes a graph that once again combines revenue dollars and units sold. I have always been hesitant about using graphs with dual-scaled axes, but primarily because of concern that some people might find them confusing, not because I thought they suffered from more fundamental problems that undermine their effectiveness. Several months ago, however, my hesitation began to grow, and I decided to examine them more closely. Today, I can't think of a single case when there isn't a better solution than a graph with a dual-scaled axis. In this article, I'll explain the reasoning—step by step—that has led me to this position. I don't consider this article the final word on the subject. Rather, it is nothing more than my thoughts to date. My mind remains open.

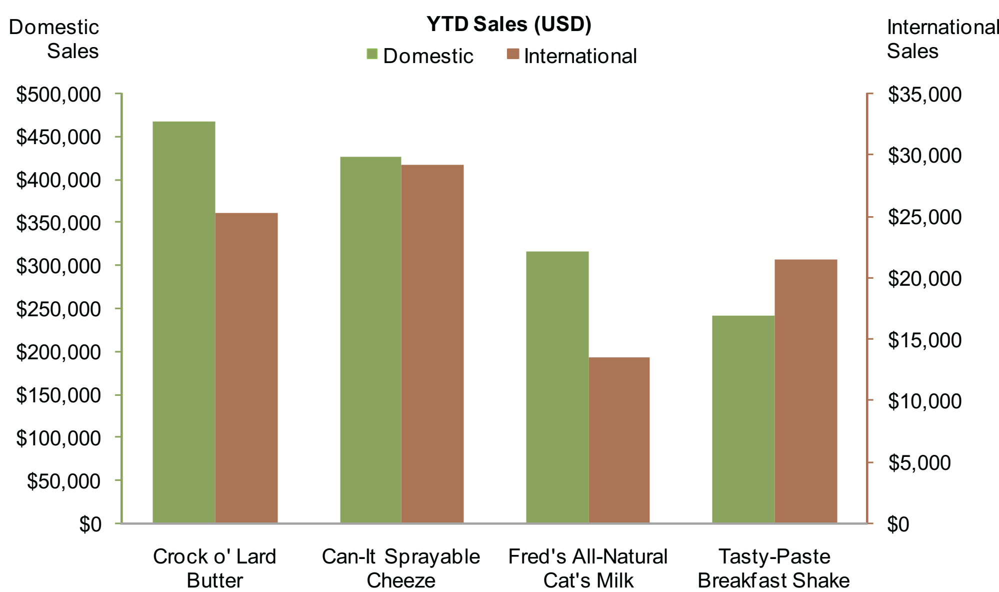
Let's start by clarifying why we might ever want to include two quantitative scales on a single axis. Sometimes it's useful to compare values that have different units of measure, because the things they measure are related in some meaningful way. This was my intention when I designed the dual-scaled axis graphs in my books. What better way could there be to compare these related but unlike sets of values than by displaying them in the same graph?—or so I reasoned.



Would there ever be a reason to use a dual-scaled axis, except to combine two different units of measure in a single graph to support meaningful comparisons? People sometimes used them when sets of values that are being compared share the same unit of measure, such as dollars, but differ significantly in scale. For instance, you might want to compare high domestic sales to relatively low international sales, as illustrated below.



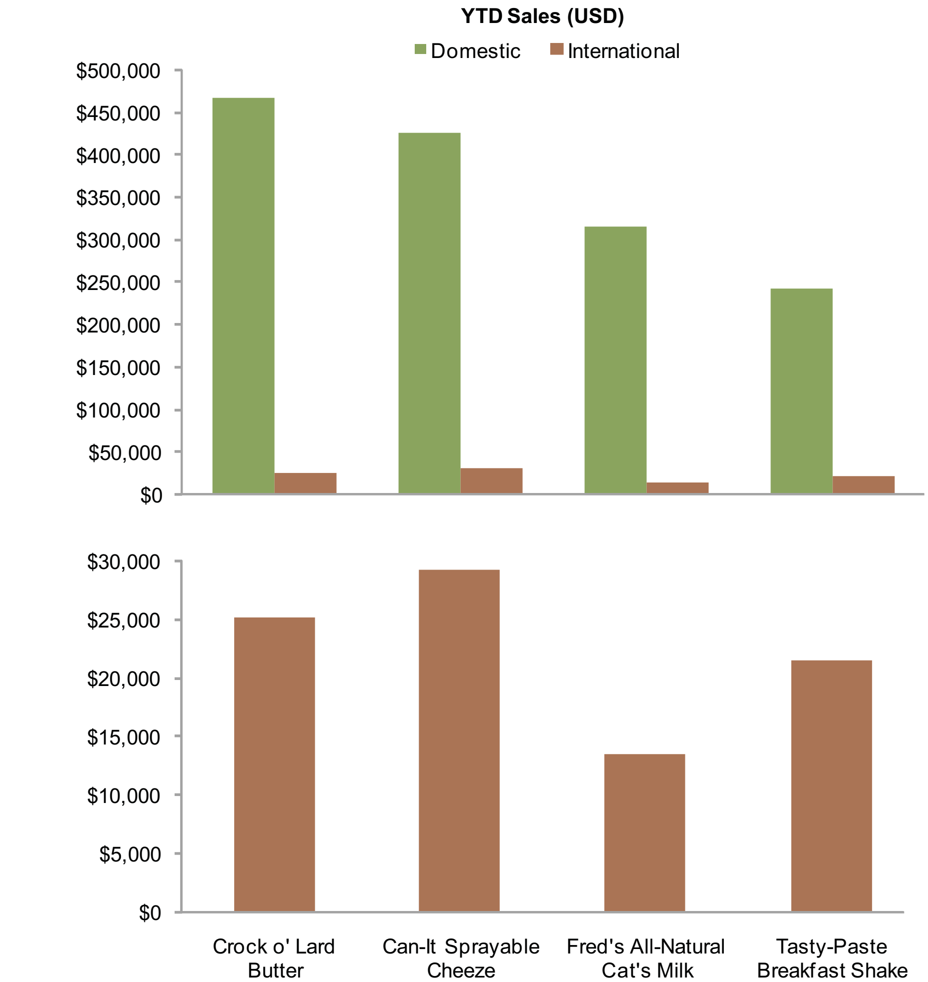
In this example, because international sales are much lower than domestic sales, the bars representing them are too short to clearly display differences between them. I have seen graphs that address this problem by assigning a quantitative scale on the left axis to one set of values and another on the right axis to the other, independently scaled to make full use of the graph's height for both, as shown below.



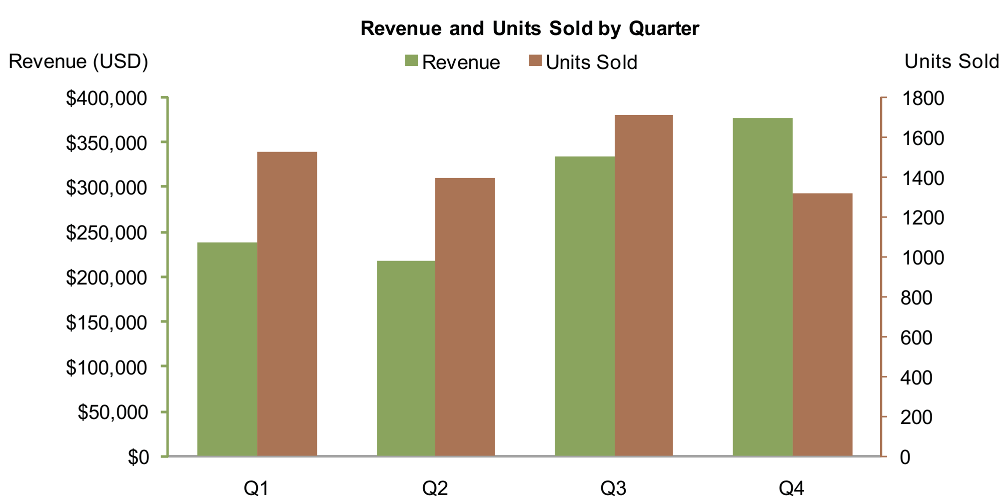
Now the bars that encode international sales are easier to read and compare to one another, but in solving the one problem, we have introduced one that is more serious: the relative heights of the bars can no longer be used to compare domestic and international sales. By independently scaling domestic and international sales in a single graph, we have encouraged people to compare their magnitudes, but this is completely meaningless and inaccurate. This observation led to my first conclusion:

> A graph should only include a dual-scaled axis (assuming for the moment that one might be useful on occasions) when needed to compare data sets that have different units of measure.

So, how can we solve the problem posed above? How can we handle data sets that differ significantly in scale? In this case, two graphs could be used: one that compares domestic and international sales in a single graph with a single scale, to show how different they are, and another to display international sales by itself, scaled in a manner that makes it easy to see differences between the products, as shown below.



Let's continue to assess the effectiveness of dual-scaled axes. Take a look at the following example, which associates the scale on the left axis with product revenue and the one on the right axis with the number of products sold:



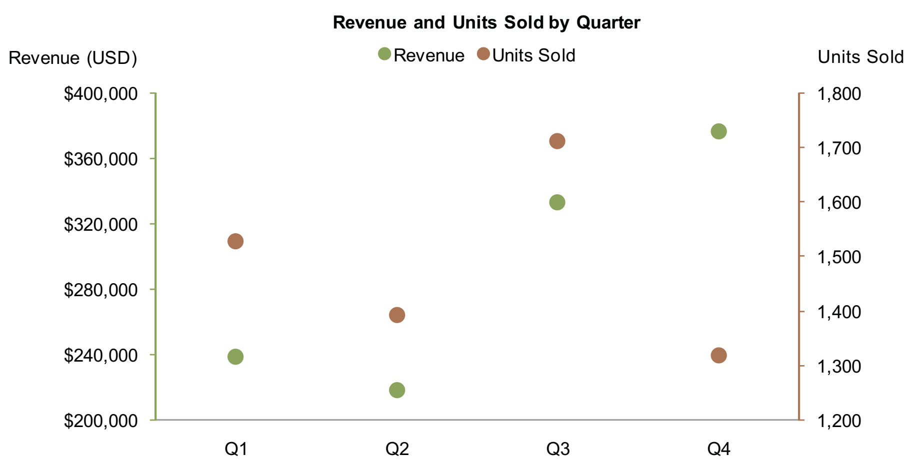
Bar graphs have two primary strengths:

1. The visual weight and independence of each bar focuses attention on individual values, rather than the overall pattern formed by the entire set of values.
2. Our ability to compare the lengths of bars to one another, which visual perception accurately supports, makes bar graphs exceptionally effective for comparing the magnitudes of individual values to one another.

In other words, by using bars to encode the values in the bar graph above, we have encouraged magnitude comparisons, but because the two sets of bars have completely independent units of measure and scales, magnitude comparisons between them are completely meaningless. This observation led to my next conclusion:

> Magnitude comparisons between values with different units of measure and scales are not appropriate, and should therefore be discouraged. Because bar graphs are designed for magnitude comparisons, a graph with a dual-scaled axis should never exclusively encode values as bars.

If bars are excluded, what other means of encoding values in graphs could we use with a dual-scaled axis, which would discourage magnitude comparisons? In the example below, the bars have been replaced with data points in the form of a dot plot. Have magnitude comparisons been discouraged, and if so, has some other useful comparison been supported?



Unlike a bar graph, which encodes absolute values as bar lengths along a scale that should always starts at zero, dot plots don't require a zero-based scale. When the scale in a dot plot begins at some quantity greater than zero, the distance of a dot from the base of the graph does not encode its value. To compare values in to one another, we must decode the value of each by referencing the scale, and then subtract one from the other to get the difference. We cannot rely on the relative positions of the dots alone to discern the difference. As such, when dot plots forego a zero-based scale, they encourage people to focus primarily on (1) greater and lesser differences between values in a graph, rather than the actual values of those differences, and (2) basic patterns formed by entire series of dots. Of these two strengths, only the second is appropriate in a graph with a dual-scaled axis, because we want to discourage all forms of magnitude comparisons between values with different scales, however imprecise the visual representation of those differences might be.

When series of values in a dot plot can be arranged in any order (a *nominal scale*), such as those associated with departments, products, or regions, the patterns that they form are meaningless. If you change the order of the items, the pattern will change. Even if they are sequenced by value to form an ordinal scale (for example, top selling product, second best selling product, and so on) from either high to low or vice versa, the resulting pattern still features magnitude differences from one value to the next, which we must discourage in graphs with dual-scaled axes. What patterns are left? When would a comparison of the pattern formed by one set of values (for example, sales revenue) to the pattern formed by another set of values (for example, units sold) be meaningful, other than those that encourage magnitude comparisons? The answer is, when the values are arranged along an *interval scale*, rather than a nominal scale or ordinal scale.

An interval scale is like an ordinal scale in that its items are ordered, but different in that it started out as a quantitative scale—a range of numbers—which was divided into a series of equally sized intervals and each interval was labeled. For instance, if you create a graph to display the number of orders that your company received in a month, broken down by order size ranging from orders of less than $1 to the largest totaling $100 (for example, with intervals of "Greater than $0 to $10," "Greater than $10 to $20," and so on ending with "Greater than $90 to $100"), the graph would display values along an interval scale. The most common example of an interval scale is one that features time, divided into intervals of years, quarters, months, weeks, days, or some other unit of time. We measure the passage of time quantitatively, so when we divide a span from one point in time to another into equal intervals, we create an interval scale. Unlike nominal and ordinal scales, each item along an interval scale is intimately connected to the one before and the one after. It is meaningful to examine the shape that is formed when we connect all the values along an interval scale with a line. With time series, the slopes formed by these lines represent change.

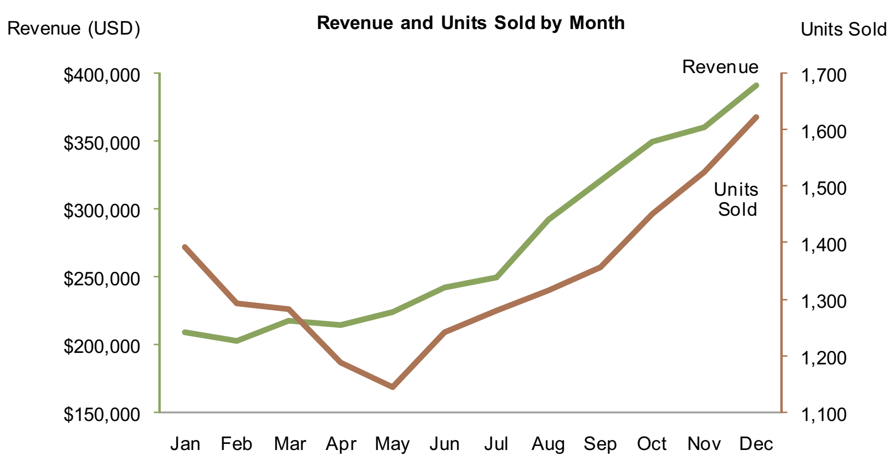
The next example is a line graph that displays sales data by month through an entire year. It is designed to help us to compare changes in revenue to changes in the number of items sold. It is meaningful to compare the shape of change in revenue to that of items sold through time. At last, we might have found an example of a dual-scaled axis that supports a meaningful comparison. When sets of values are encoded as lines, we are invited to examine the overall patterns formed by them, rather than the individual values, and to compare their overall patterns to one another, rather than their magnitudes. This observation led to my next conclusion:

> Nothing but lines should be used to display values in graphs with dual-scaled axes, because only lines focus attention on the overall pattern formed by the values and comparisons of those patterns.

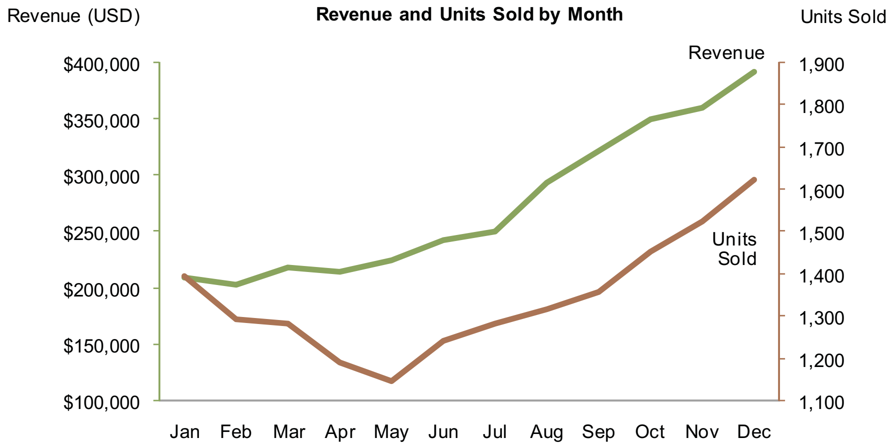
It is only appropriate to connect values with a line along an interval scale, for only in such cases are the values, from one to the next, intimately connected to one another, and only then are the slopes of the lines meaningful (for example, to encode relative degrees of change). This fact led immediately to my next conclusion:

> It is inappropriate and meaningless to use a dual-scaled axis in a graph that does not display values along an interval scale.

So, we've narrowed the use of dual-scaled axes down to line graphs and interval scales, based on the fact that they focus attention on the overall trends and patterns in the data and the comparison of those trends and patterns. We have determined that dual-scaled axes are meaningful in such cases, but we have not yet concluded that they are appropriate, for other problems might be lurking that have not yet been identified. Take a look at the next example to see if you can spot a potential problem.



Notice that a salient feature of this graph is the point where the two lines intersect. Usually, in a line graph, if two lines intersect, the fact that our eyes are powerfully drawn to their intersection is useful, because the intersection is meaningful. The intersection indicates that one set of values surpassed the other. When lines are associated with different quantitative scales, however, their intersection means nothing. The same values appear in the following example, but because the scales are slightly different, the lines now begin at the same point and never intersect.

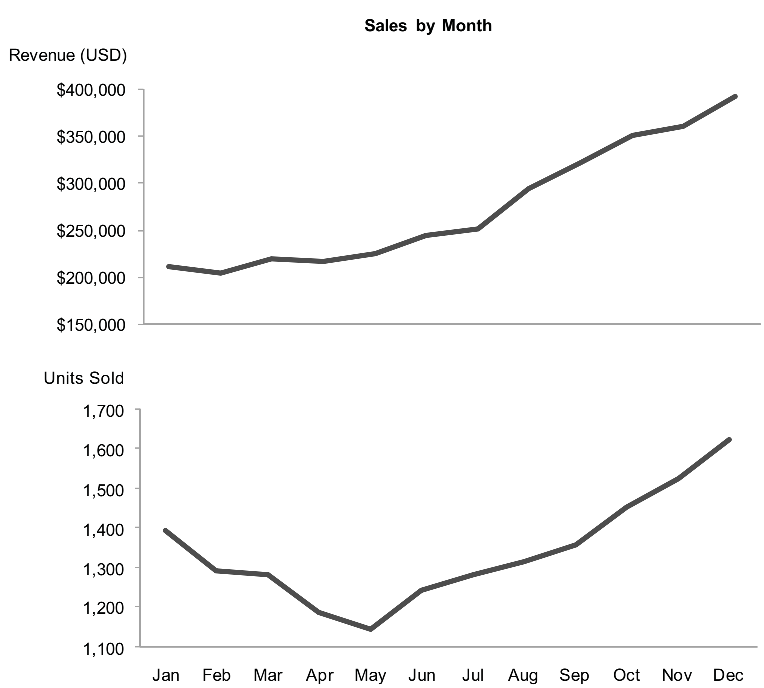


Whether the lines intersect or not, and if so where, is artibitrary. We could stamp a warning across the graph, “Beware of intersecting lines, because they don’t mean what you probably think they mean,” but why force people to resist their natural inclination, attempting to ignore as meaningless that which draws their attention so powerfully? In essence, the intersection of lines and our inclination to interpret it in terms of magnitude (one set of values exceeded another), reveals that even lines, which feature overall trends and patterns, are nevertheless subject to magnitude comparisons. By placing two lines in a single graph and associating each with a different quantitative scale, we are still encouraging people to compare their magnitudes. Thus, my final conclusion:

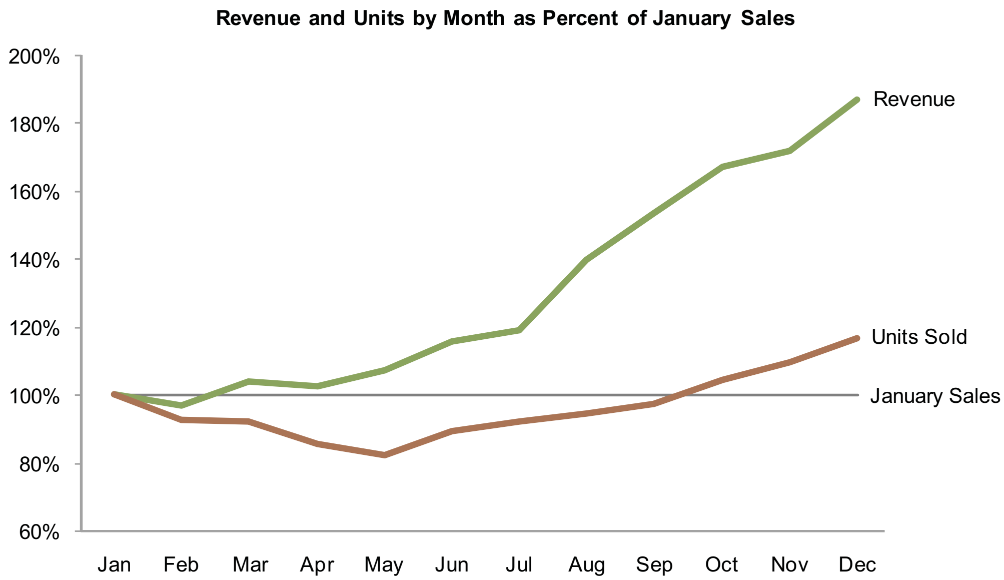
> It is inappropriate to use more than one quantitative scale on a single axis, because, to some degree, this encourages people to compare magnitudes of values between them, but this is meaningless.

All circumstances for using two quantitative scales on a single axis that come to my mind have been disqualified, because all encourage people to make meaningless comparisons.

How then do we make it easy for people to compare related sets of values when they are associated with different units of measure? Two answers come to mind. The first and most obvious is to place them in separate graphs, positioned close to one another so that the patterns in each can be compared to one another, but magnitude comparisons will be discouraged. Here’s an example, using the same revenue and units sold data that we’ve seen before:



The other less obvious solution, which works only for time series, is to convert all sets of values to a common quantitative scale by displaying percentage differences between each value and a reference (or index) value. For instance, select a particular point in time, such as the first interval that appears in the graph, and express each subsequent value as the percentage difference between it and the initial value. This is done by dividing the value at each point in time by the value for the initial point in time and then multiplying it by 100 to convert the rate to a percentage, as illustrated below.



This gives us a way to compare the data sets' patterns of change in a single graph along a common scale, so magnitude comparisons are appropriate and the intersections of lines are meaningful.

I certainly cannot conclude, once and for all, that graphs with dual-scaled axes are never useful; only that I cannot think of a situation that warrants them in light of other, better solutions. I invite you to propose viable exceptions, which I will welcome with open arms.

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.