

Quantitative Displays for Combining Time-Series and Part-to-Whole Relationships

Stephen Few, Perceptual Edge
Visual Business Intelligence Newsletter
 January, February, and March 2011

Graphical displays of quantitative information always feature particular relationships among the values. Two of the most common are “time-series” and “part-to-whole” relationships. In the former case we feature how values change through time, and in the latter how values that combine to form a whole compare to one another and the whole (for example, a regional breakdown of total sales). Determining the type of graph to use and its design should be based primarily on the nature of the one or more relationships that it features. For example, when we want to show how values change through time (a time-series relationship) in a way that makes the trends and other patterns easy to see, a line graph is often the best solution. In this article, I will propose graphical solutions for those occasions when we must feature a combination of time-series and part-to-whole relationships—that is, how the relationships among the parts of some whole change through time.

Traditional Graph Solutions

Combined time-series and part-to-whole relationships are typically displayed using one of the following graphical solutions:

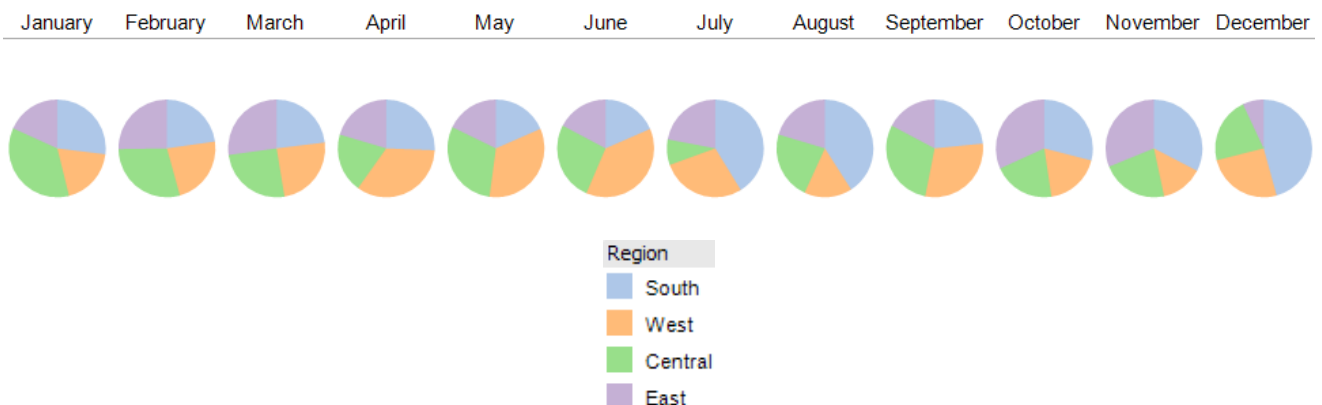
- A series of pie charts
- A series of stacked bars
- A stacked area chart

Let’s take a look at each and evaluate their strengths and weaknesses. In each case, we’ll use the following series of twelve monthly sales values (the time-series relationship) divided across four geographical regions (the part-to-whole relationship).

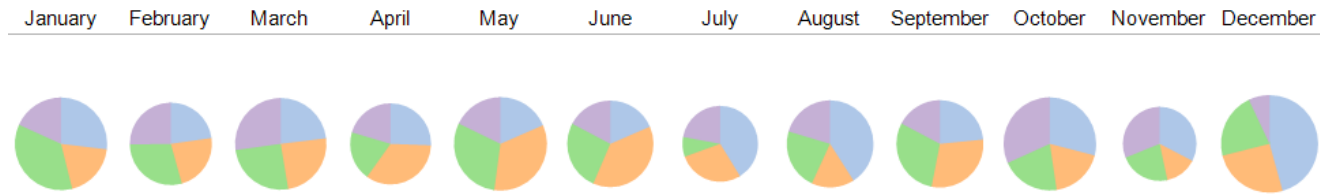
Region	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
East	64,162	70,172	93,657	55,056	63,631	53,040	51,974	63,272	54,110	112,284	69,697	27,437
Central	125,392	80,851	87,645	52,996	107,159	80,349	19,907	70,431	92,568	72,961	49,881	85,084
West	67,364	63,742	83,856	92,412	120,284	116,618	66,692	49,671	92,920	65,971	31,516	99,000
South	94,572	63,234	79,491	68,963	65,868	56,659	97,101	126,879	73,240	102,589	73,044	177,943
Total	\$351,490	\$277,998	\$344,649	\$269,425	\$356,942	\$306,666	\$235,674	\$310,254	\$312,839	\$353,805	\$224,139	\$389,465

Pie Chart Solutions

Here’s how this story could be told using pie charts:



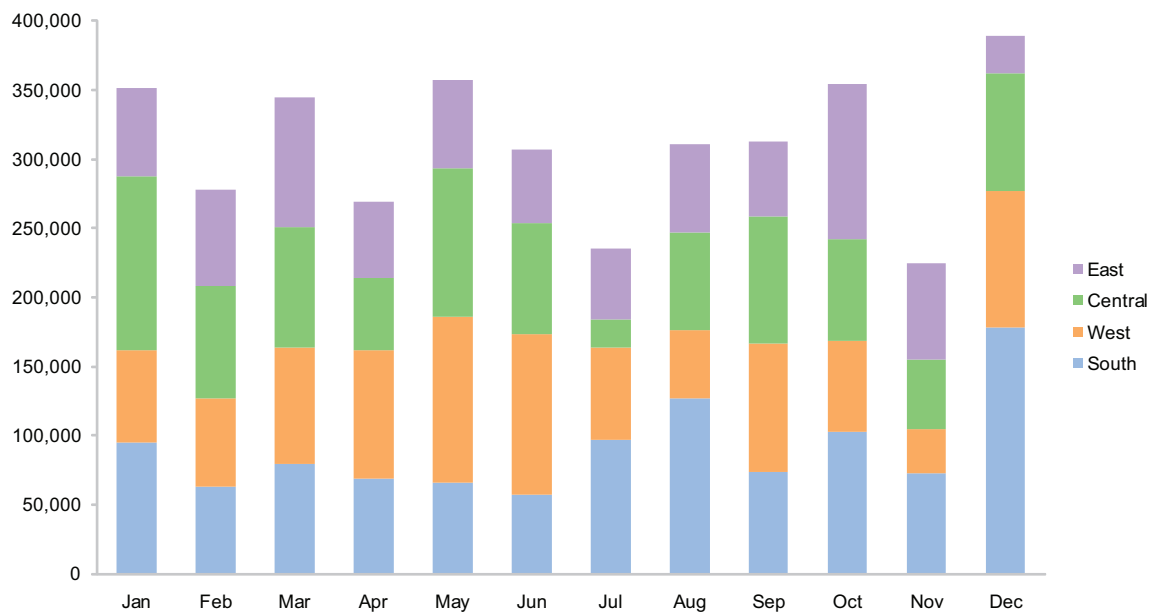
One of the first things you might have noticed is the fact that, while the slices of each pie show the regional percentages of sales, total monthly sales aren't displayed at all. One way to add this information to the display is to vary the sizes of the pies, which I've done in the following example.



If we were to list the benefits of this particular solution, the list would remain empty. This solution fails in every respect. A single pie chart already fails as a way to compare parts of the whole easily and accurately. Visual perception simply isn't tuned to compare the sizes or angles of the slices, except in a very rough sense. Multiplying the number of pie charts magnifies the failure by making it virtually impossible to see how the regional sales values changed through time. Try it for yourself. See if you can explain how regional sales went up and down through time and how the relationship between the regions changed through time. As you can see, this is not a viable solution.

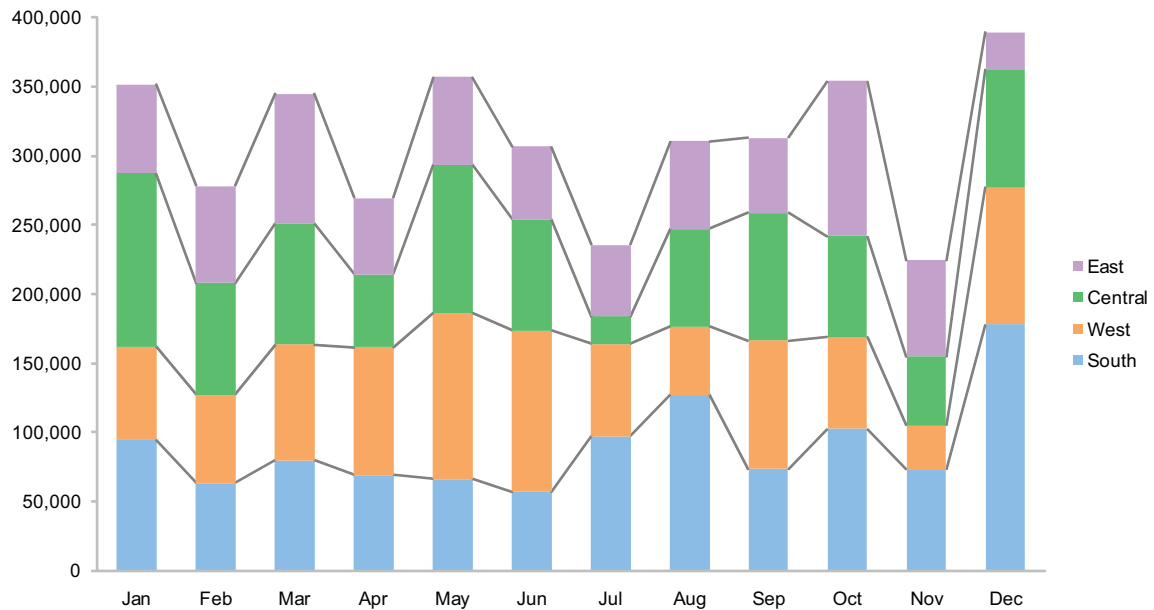
Stacked Bar Solutions

Here's a stacked bar chart that displays the same information:



Just as it's difficult to compare the slices of a pie, it is difficult to compare the segments of a stacked bar. If the segments were arranged side-by-side and all grew upwards from the same baseline, it would be easy to compare their heights, but when stacked upon one another, the task becomes hard. Plus, although it's fairly easy to see how total sales changed from month to month and how sales in the South region changed from month to month (the blue segments of the bars at the bottom, which share a common baseline), it is quite difficult to see how sales in the other regions changed. Difficult, but still a bit easier to do this using stacked bars rather than pie charts.

Bars don't do a good job of showing overall trends and other patterns of change through time. Lines do this much better. If you're tempted to believe that this problem can be solved by adding lines to connect the tops of each segment across the entire set of bars, take a look at the following example and determine for yourself if this helps.

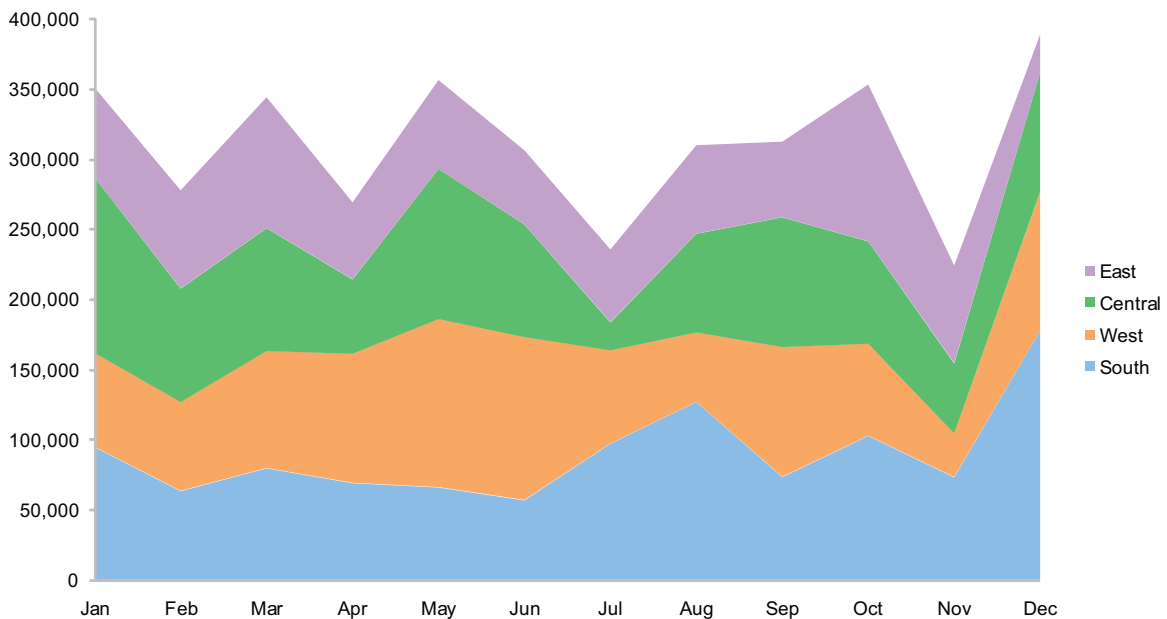


One of the immediate problems you'll notice is that this looks quite cluttered, even though I've kept the design clean and simple. Beyond this, the bigger problem is the fact that you cannot use the lines to follow the pattern of change in regional sales, except for the bottom segment (the South region in blue in this case). Other than this one exception, the lines can only be used to ever so slightly assist in determining the amount of change from one bar to the next—not by following the slopes of the lines up or down, but by examining the difference between the vertical distance from the lower to the upper line where they connect to the bar on the left versus where they connect to the bar on the right. The slight assistance that the lines provide does not overcome the problem caused by the clutter or our tendency to focus on the slopes of the lines.

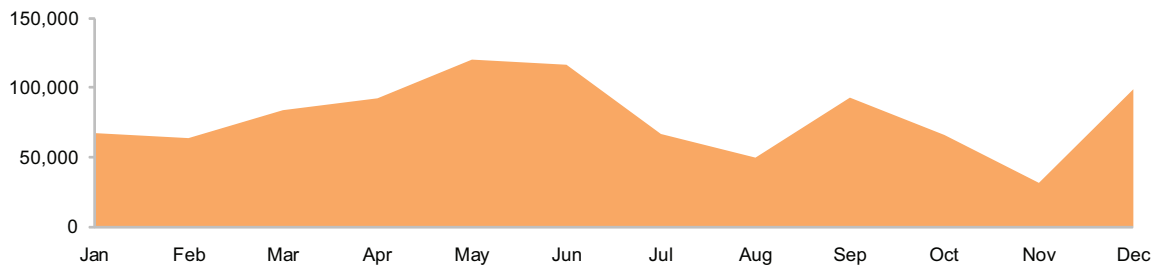
Because bars are so visually weighty and separate from one another, they do a good job of focusing our attention on one month at a time. In situations when patterns of change aren't part of the story that you're trying to tell and you instead want to focus attention on sales in particular months, one at a time, or to compare one particular month to another, a stacked bar chart will do the job. Otherwise, better solutions are available.

Stacked Area Solutions

Here's the same data in a stacked area chart:



With a stacked area chart, which we read in some respects much as we read line charts, we now have a solution that does a better job of showing patterns and trends of change through time. Unfortunately, our tendency to read this chart as we would a line graph gets us into trouble. We can follow the pattern formed along the top of the chart (the top of the purple section) to examine total monthly sales, and we can follow the pattern formed along the top of the blue section to trace monthly sales in the South, but we cannot follow the tops of the orange, green, or purple sections to follow monthly sales in the West, Central, or East regions. This is because the slopes of the lines at the tops of these areas are influenced by the slopes of the sections below them and therefore don't represent the patterns of change for those regions. To see the pattern of change for the West region, we must look only at the vertical distance between the bottom and top of the orange section above each Month label, and try to imagine those values as a line going up and down. Here's what the pattern of monthly sales for the West looks like when shown by itself:



Constructing this pattern in your head by viewing the West region in the stacked area chart simply isn't possible, unless you're an extremely rare savant with this particular ability.

Although limited in use for the reason that I've explained, stacked area charts work when you primarily want to show how the whole changes through time and only give a rough sense of how the parts compare to one another in their contributions to the whole at particular points of time. Just make sure when you use this approach that your audience knows not to follow the slopes of the lines along the tops of each section to see how a part of the whole changed through time.

Other, More Useful Solutions

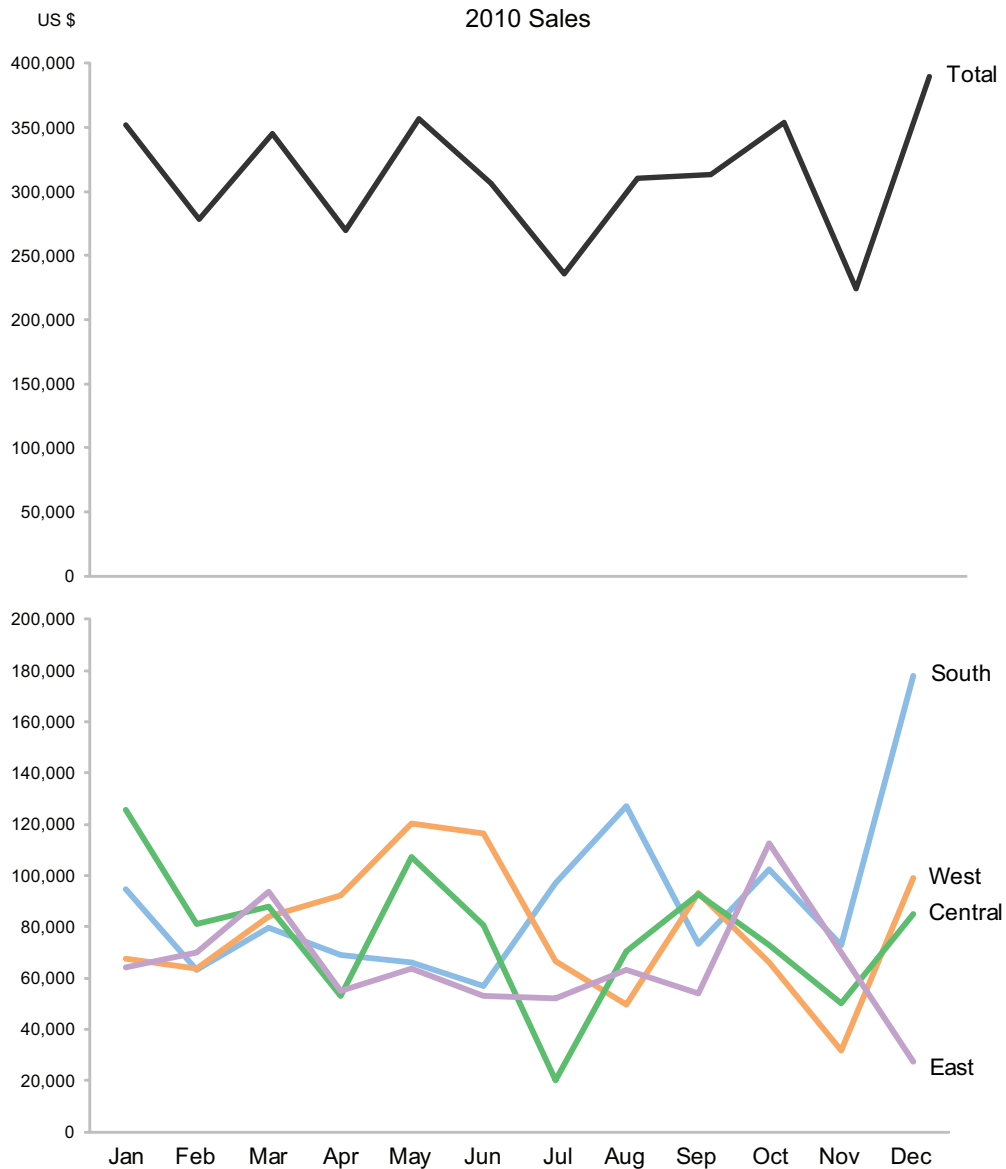
I'd like to propose a few solutions that usually work better than the traditional approaches that we've already examined. In most cases, the solution will involve more than one graph. This is because different aspects of the time-series/part-to-whole stories that we tell are best told using different graphical approaches, which cannot or should not be combined in a single graph. If you're inclined to think that a data visualization should limit itself to a single graph, it's time to leave that constraint behind. The best solutions often require multiple graphs.

The solutions that I'll propose are divided into two types:

- Solutions with few parts
- Solutions with many parts

Solutions with Few Parts

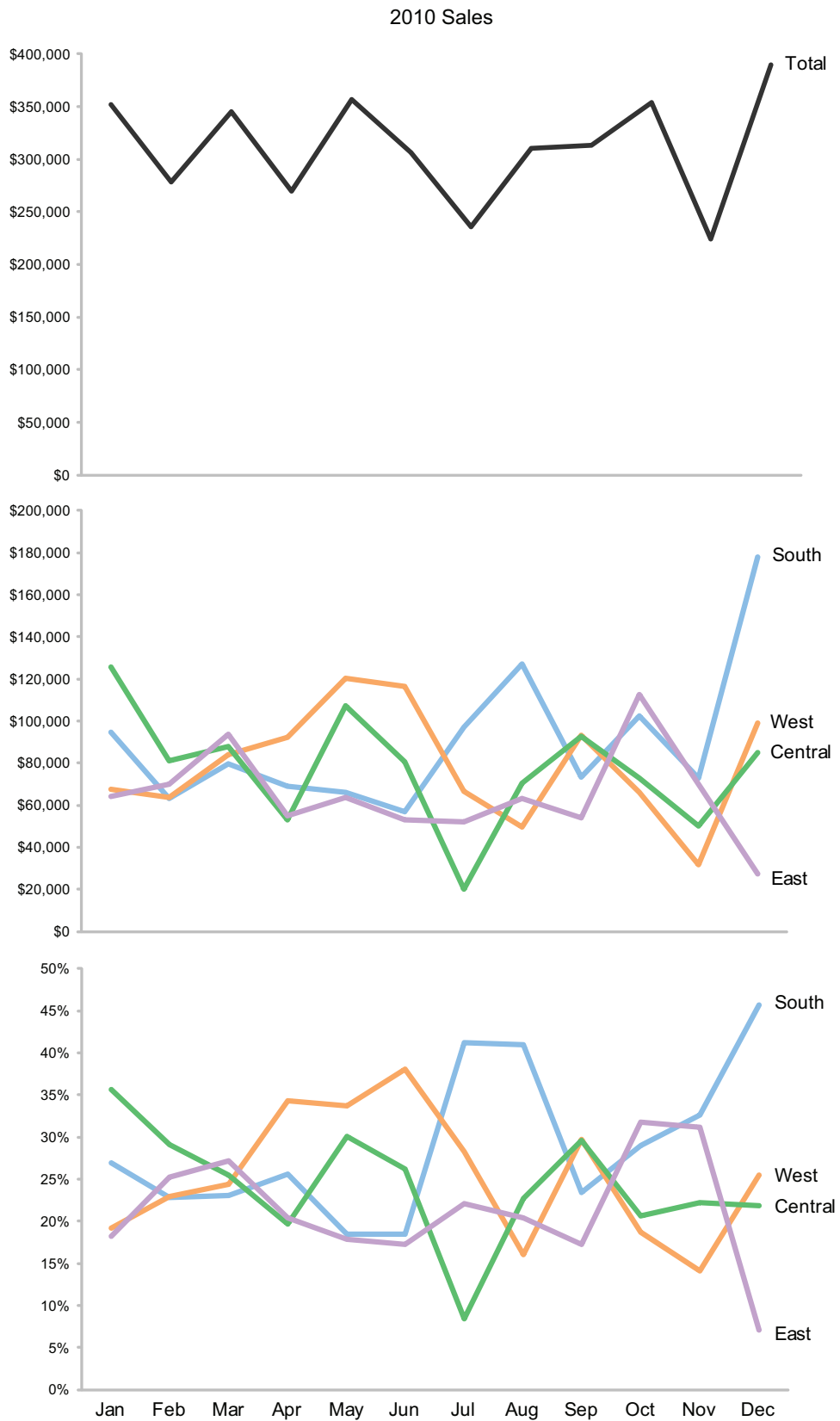
For this situation, I usually recommend the following solution:



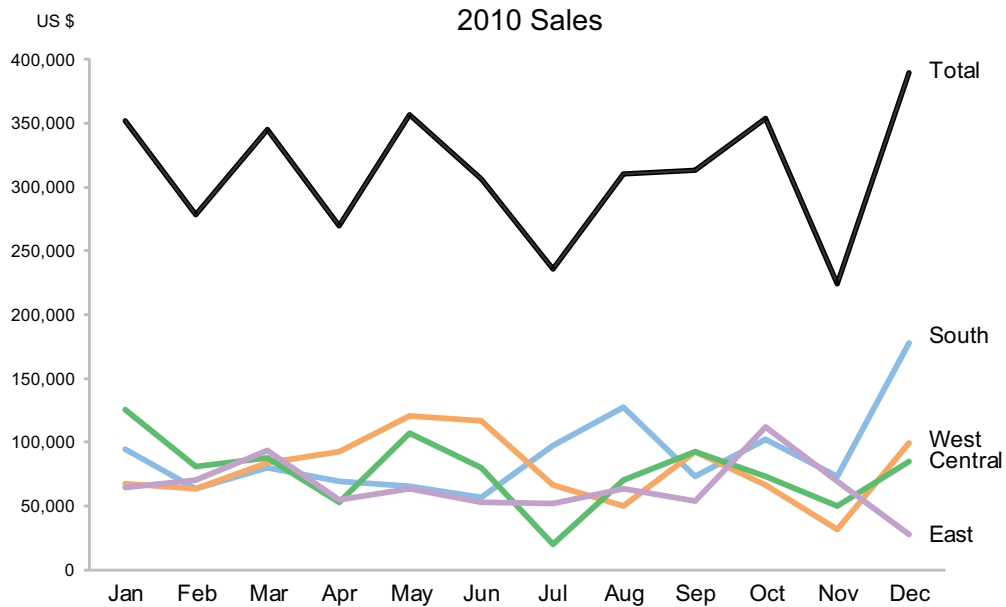
Consider the attributes of this approach:

1. The trend and patterns of change in total sales are readily available to our eyes. We can focus on total sales without being distracted by the regions.
2. The trend and patterns of change in regional sales are easy to see and compare.
3. The contributions of the parts are easy to see and compare at any point in time.
4. Because specific points in time in the two graphs are aligned, you can easily see the total sales at a particular point in time in relation to regional parts that make up that total.
5. The trend and patterns of change in total sales are easy to compare to the trend and patterns of change in any of the regions.

Although there are occasions when slight variations of this solution might work as well or better (for example, replacing the line graph for the whole with a stacked area chart), this simple solution seems to work well most of the time. A variation that's sometimes useful adds a third graph that expresses the contributions of the parts as percentages of the whole, illustrated below.



You might wonder if it would ever make sense to place the values of the whole and the parts in a single graph rather than two. Here's what this would look like using the same data:



In this particular case, because there are only four parts, values of the whole are not so much greater than values of the parts that they can't effectively be displayed in a single graph. In many cases, however, placing the whole and its parts in a single graph would force the lines representing the parts into a narrow range at the bottom of the scale, which would flatten the lines to such a degree it would be difficult to see and compare their patterns of change.

Solutions with Many Parts

As you can imagine, the solution that I proposed above would decline in effectiveness as the number of parts increased, due in part to clutter. This same solution could continue to work when there are many parts, however, if we enhanced it in one of two following ways:

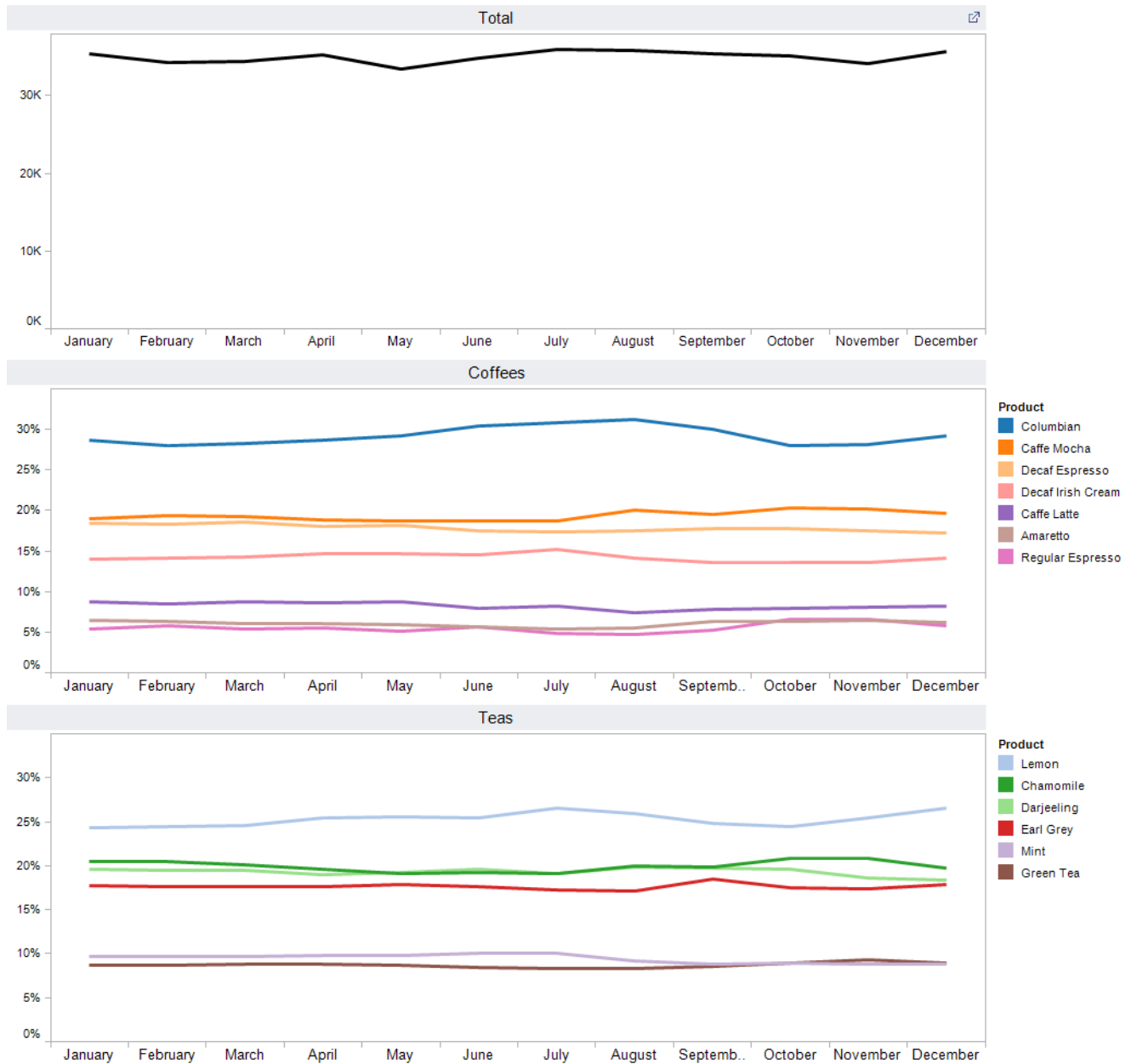
- By allowing specific parts to be highlighted interactively to feature them
- By placing each of the parts or smaller groups of parts into separate graphs

If the software that you're using to display data supports selective highlighting of lines in a line graph, even though the graph is cluttered with many lines, this problem can be reduced by highlighting particular lines when you want to focus on them, thus removing distraction caused by the others. In the following example, 13 products are displayed in the bottom graph, but only two have been highlighted, which makes it easy to compare them with little distraction from the other products:



In this example, created using Tableau Software, I highlighted two products in the lower graph—Chamomile and Green Tea—by selecting them in the legend to the right of the graph. Other interactions can be enabled as well for a richer analytical experience, such as filters using the controls that appear in the upper right.

If you're stuck with a static view with no opportunity to highlight particular lines, you can solve the clutter problem by breaking the parts into multiple graphs, as illustrated below.



When you use this approach, it's important that you group the parts in a meaningful way. In this example, it was easy, because these beverage products could be divided into two manageable groups based on the natural groupings of coffees versus teas. It isn't always this easy. If no natural way to break the parts into groups can be found based on a categorical dimension, it might make sense to break them into groups based on quantitative ranges. For example, using the products in the example above, they could be divided into two graphs—one with revenues up to the 15% range and the other for everything higher—or perhaps into three groups ranging from (1) 0 to 10%, (2) over 10% to 20%, and (3) over 20%.

For frequent graphical display requirements, such as combined time-series and part-to-whole relationships, it helps to have a few proven solutions at hand and to understand the strengths and limitations of each. Reserve the mentally taxing work of creating a new solution for those occasions when the requirements are genuinely novel.

Discuss this Article

Share your thoughts about this article by visiting the [Quantitative Displays for Combining Time-Series and Part-to-Whole Relationships](#) thread in our discussion forum.

About the Author

Stephen Few has worked for over 25 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 quarterly [Visual Business Intelligence Newsletter](#), and speaks frequently at conferences. He is the author of three books: *Show Me the Numbers: Designing Tables and Graphs to Enlighten*, *Information Dashboard Design: The Effective Visual Communication of Data*, and *Now You See It: Simple Visualization Techniques for Quantitative Analysis*. 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](#).