

Displaying Missing Values and Incomplete Periods in Time Series

Stephen Few, Perceptual Edge
Visual Business Intelligence Newsletter
January/February/March 2015

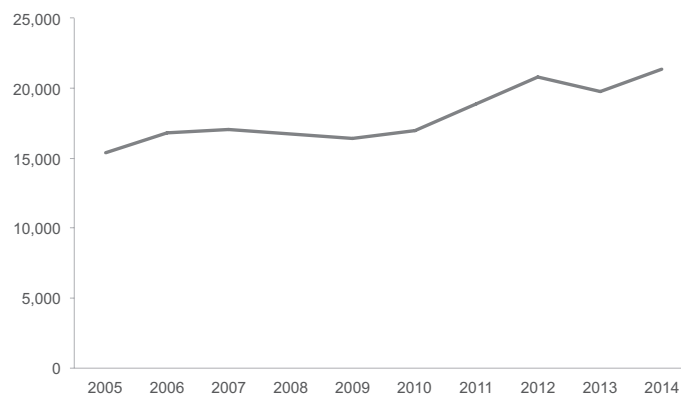
Two of the challenges that we sometimes face when displaying time series data are:

- Determining how to handle values that are missing in the series
- Determining how to display a current period that is incomplete

Both of these display challenges have simple solutions, but software products often fail to support them. Let's take a look at the ways these situations are sometimes mishandled and how they can be handled effectively.

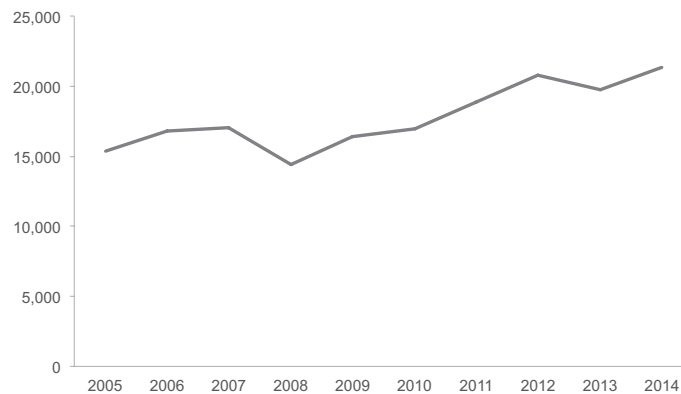
Missing Values

Something is wrong with the following graph, but you have no way of knowing what.



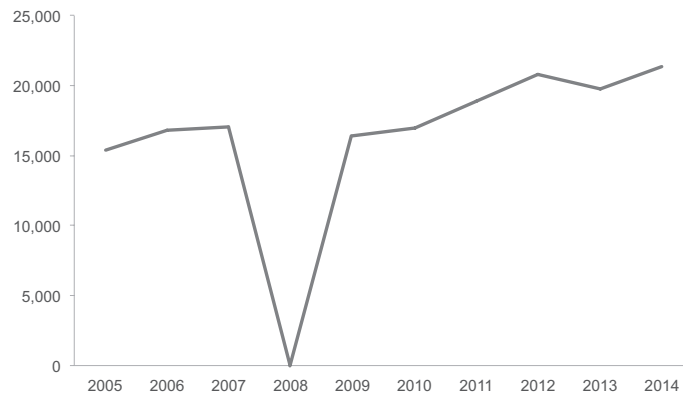
The value for the year 2008 is missing. Rather than revealing this omission, the software merely ignored it and directly connected the 2007 and 2009 values. Whenever values are missing, this fact should be obvious.

Here's the same data set, but this time the value that was missing for 2008 has been included:



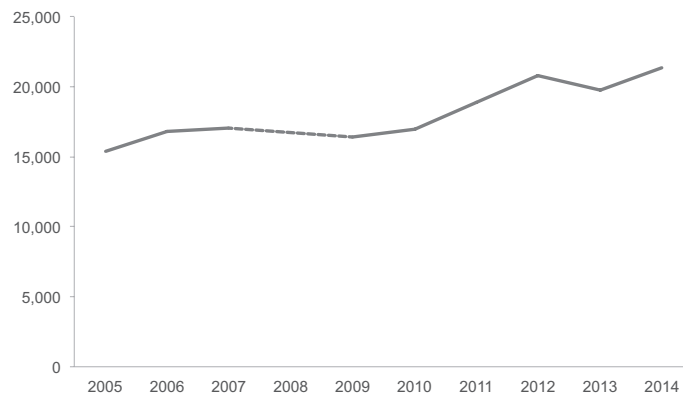
Now that the missing value is visible, it's easy to see how much the first graph deceived us by directly connecting the values for 2007 and 2009.

Here's the same data with the value for 2008 missing again, this time displayed in another way that software products often do automatically:



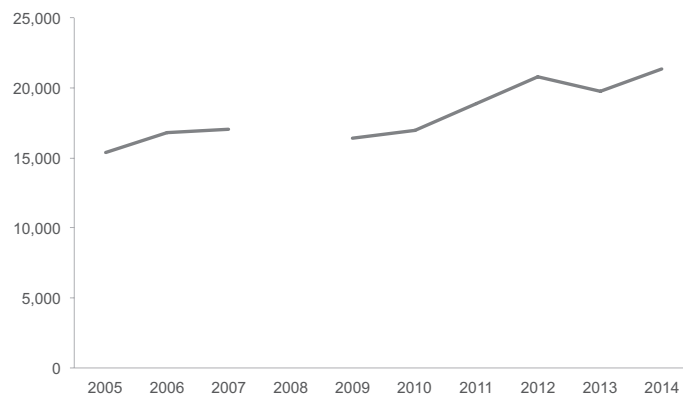
Rather than telling us that the value for 2008 is missing, this graph suggests that its value is zero. A value of zero is quite different from one that's missing, and that difference should be obvious.

Here's another way that I've seen missing time-series values displayed:



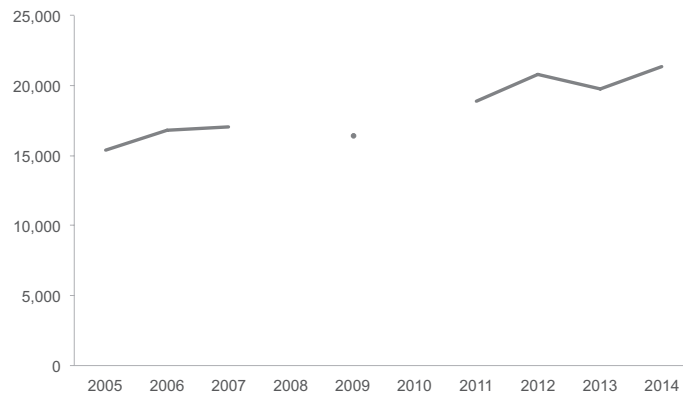
With this graph, we can see that something different is going on between the years 2007 and 2009, but similar to the first graph that we looked at, the dashed line suggests a pattern of change between 2007 and 2009 that is quite different from the actual pattern that we would see if the value for 2008 weren't missing.

The solution to this situation is simple. An effective way to show that a value is missing is to display the data in the following way:



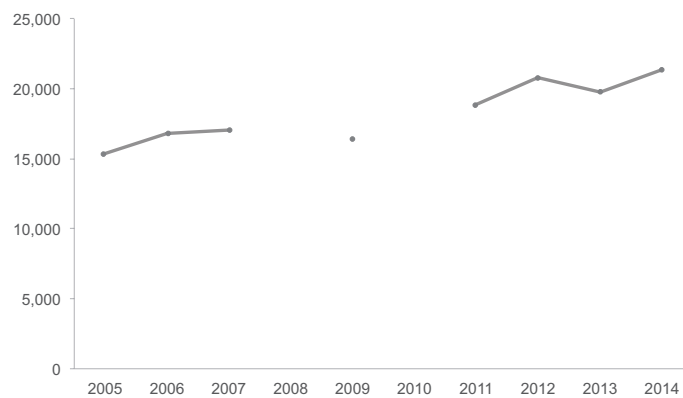
It is now obvious that a value is missing and nothing has been displayed to suggest a pattern of change between 2007 and 2009 that isn't known. Problem solved. Fortunately, this is the way that Excel handles the situation automatically, but many other products do not.

What if a single value that is present in the time series is surrounded by missing values? For example, imagine that 2008 is missing, 2009 is present, but 2010 is also missing. How can this be shown? Here's how I do it:



A single data point to show the one value that is surrounded by missing values does the job. Anything more than this would be misleading.

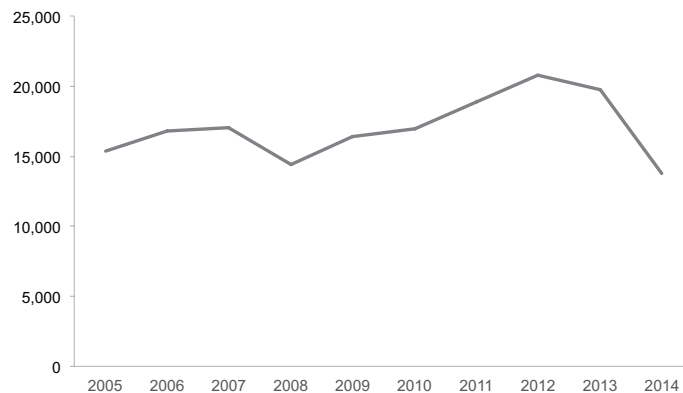
Unfortunately, if you use Excel to produce this graph, the single data point would not appear unless you enabled data points for all the values, as illustrated below.



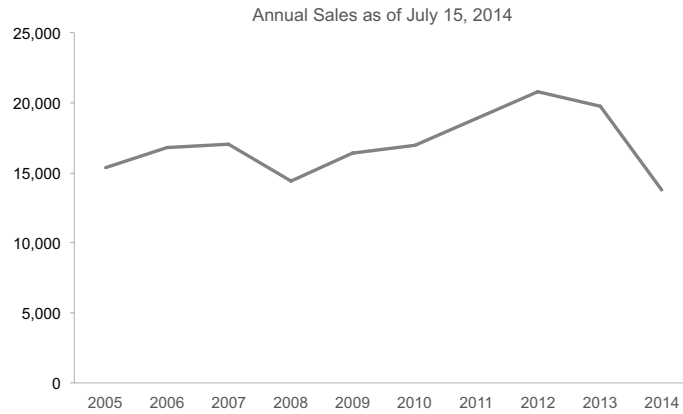
Data visualization products should handle missing time series values in the effective ways that I've illustrated as a standard practice. Unfortunately, many don't think this problem through. I suspect that many software vendors know so little about data and data visualization that they don't even recognize that this problem exists.

Incomplete Current Period

Something is wrong with the following graph, but once again, the problem isn't obvious.

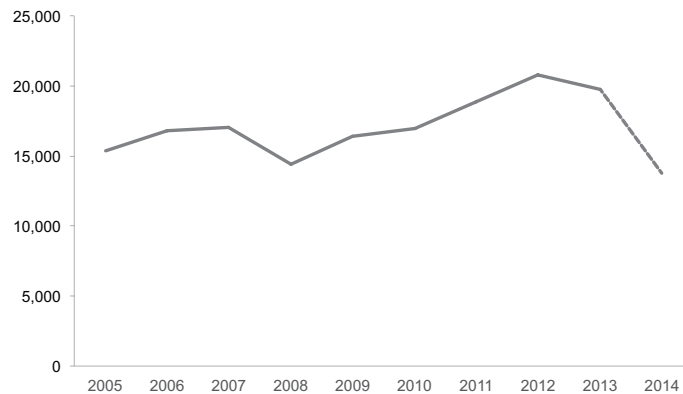


I'll add a title to the graph that will provide a useful hint about the problem.

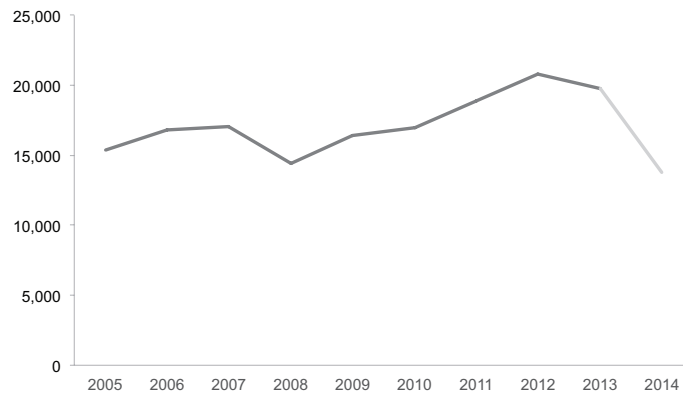


See the problem now? We're only slightly more than halfway through 2014, which is why the value for the final year is unexpectedly low. When we display a time series and the final period is not complete, we must do something to make this fact obvious and we shouldn't suggest a pattern of change at the end of the series that's misleading.

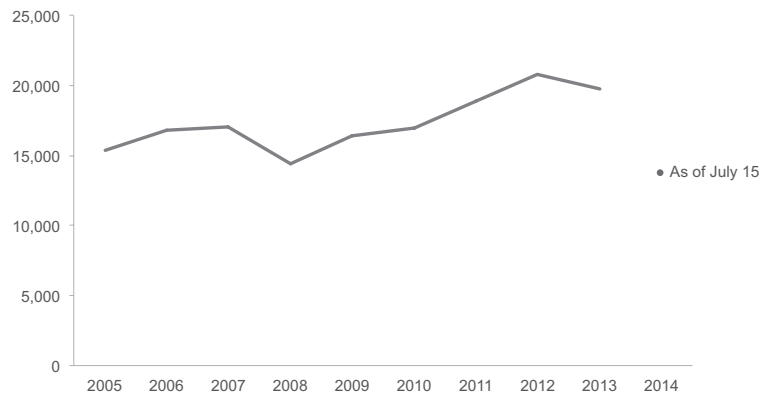
Would it work to display the final value in the series as follows?



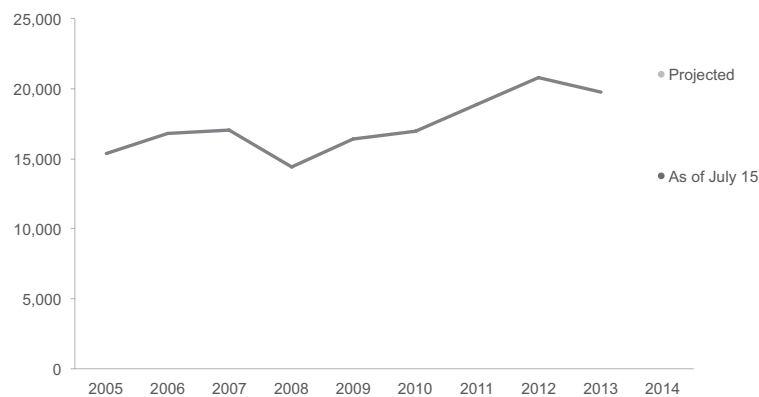
The final value certainly looks different, but even if your audience knows this means that the final period is incomplete, they might still be influenced by the misleading pattern that the downward-sloped line exhibits at the end. The following alternative would suffer from the same problem:



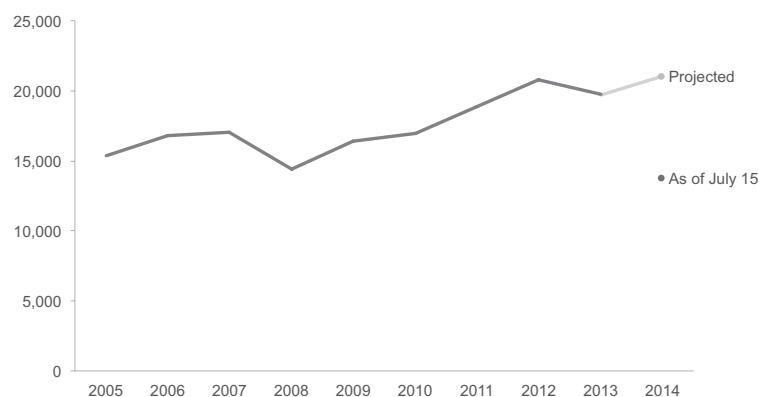
We want to make it clear that the final period isn't complete and we don't want to display a misleading pattern of change. Here's one way to handle this situation:



This supports both of our objectives and also makes the final value's date clear. What this doesn't do, which would be useful, however, is help us anticipate what the value will be when the period is complete. For this, we need a projection, but it must be one that's expertly made based on historical data and valid statistical methods. If we have a way to make a statistically informed prediction, we could add that value as well, in the following manner.



Or, because the projection was reliably calculated, we could display it as follows, using a dashed or lighter line to indicate its uncertain nature.



We would only show the line segment that connects to the projected value, however, if we wish to emphasize it more than the actual value.

These challenges aren't particularly challenging, but they are frequently handled in ineffective and misleading ways. Wouldn't it be great if the products that we rely on for visualizing data made this easy?

Discuss this Article

Share your thoughts about this article by visiting the [Displaying Missing Values and Incomplete Periods in Time Series](#) thread in our discussion forum.

About the Author

Stephen Few has worked for nearly 30 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*, Second Edition, *Information Dashboard Design: Displaying Data for at-a-Glance Monitoring*, Second Edition, 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](#).