Time on the Horizon

Stephen Few, Perceptual Edge

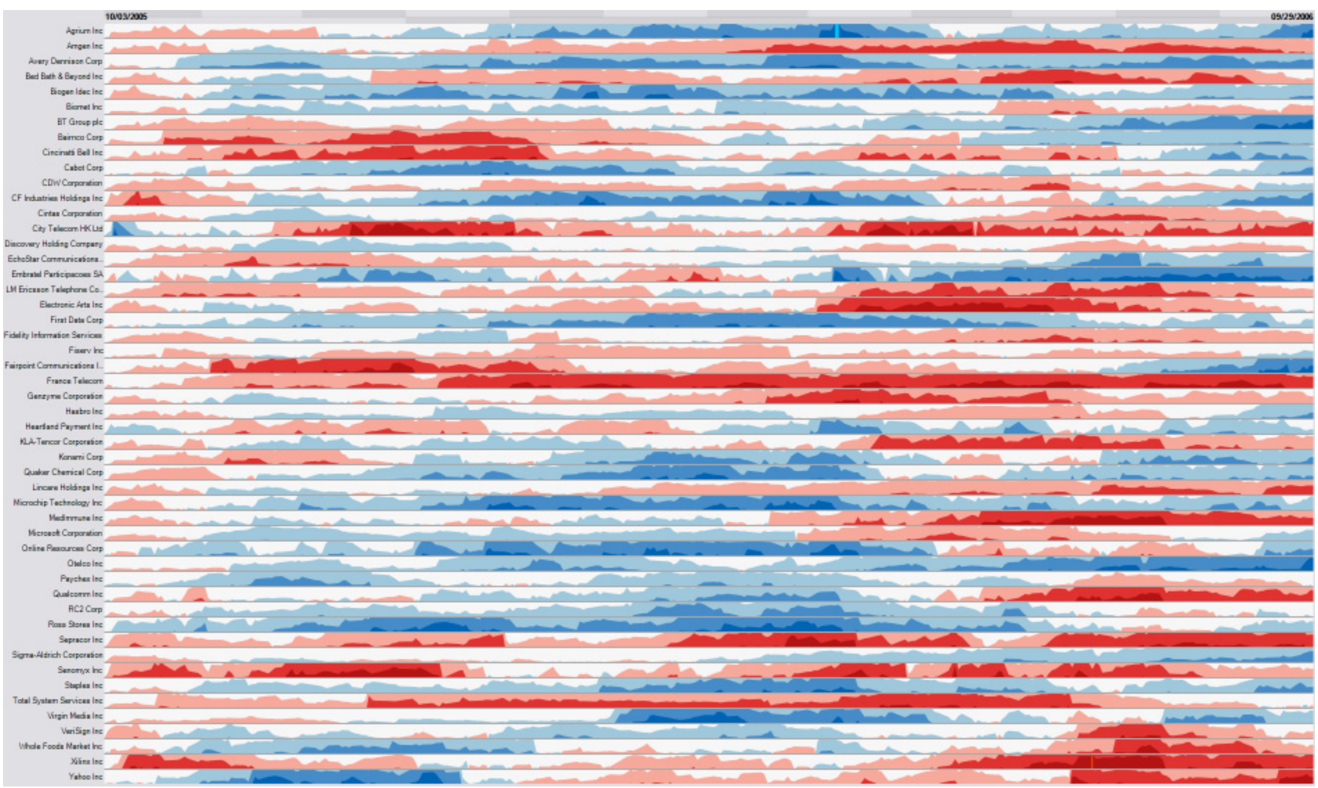
Visual Business Intelligence Newsletter

June/July 2008

To say that information visualization is still a young and evolving field is a considerable understatement. We are still only scratching the surface of its potential. As the science on which it is built continues to yield new insights and an increasing number of talented and innovative people apply these insights in search of new solutions, we can expect eye-opening progress in the next few years. In this article, I want to describe an example of how this process works. To do so, I'm turning to the work of a relatively small but gifted business intelligence software vendor named Panopticon.

This is not a marketing piece for Panopticon, so I'll say little about the company except that until recently its products exclusively featured a particular visualization called a treemap. For information about treemaps, I invite you to read an article that their inventor, Ben Shneiderman of the University of Maryland, wrote for my newsletter back in April 2006 titled "Discovering BI Using Treemaps." The folks at Panopticon applied the potential of treemaps in several innovative and practical ways, and are now complementing their products with the addition of several traditional graphical displays (for example, bar and line graphs), including a few new variations on these themes. One of these variations is called a horizon graph.

Last month, I spent two days with Panopticon's development team in Stockholm, reviewing what they do and making suggestions, which is something that I do from time to time for several software vendors. It was during this visit to Stockholm that I looked closely for the first time at the horizon graph. It was not love at first sight. In fact, I spent over an hour stating one objection after another, only to have each politely countered with answers that I eventually found compelling. To give you a sense of the encounter, I'll show you an example of a horizon graph, much as I saw it in Stockholm for the first time.



A bit overwhelming at first, isn't it? Now let me take you on a journey that is similar to the one I was taken on by the folks at Panopticon, and we'll see if you come to agree that something meaningful can be seen on the horizon.

I'll begin with a quick explanation of the information that appears in the example above. It displays just under a year's worth of daily changes in the prices (October 3, 2005 through September 29, 2006) of 50 stocks. The stocks are arranged by row, and time is displayed horizontally, from left to right. Rises and falls in stock prices relative to the price on October 3, 2005, the first day of the period, have been encoded in two ways: 2-D position—lines that move up and down as they proceed from left to right—and hue—increases in blue and decreases in red. Each of the graphs—one per stock—shares the same quantitative scale, so an increase or decrease of a particular distance for one stock equals the same degree of change as an increase or decrease of the same distance for another stock. In other words, the 50 graphs are an example of what Edward Tufte calls "small multiples"—a series of small graphs that are laid out within eye span for the purpose of comparison, which differ from one another only in that each represents a different set of values, in this case daily prices of individual stocks. By hovering with your mouse over any point along the timeline for one of the stocks, you can view the details regarding the selected stock at that point in time (stock name, price, and percentage change from the previous day) in a small window, as illustrated in the example.

This is the initial explanation that I received when I first examined Panopticon's horizon graph. No more than a moment passed before I began expressing concerns, such as "Can a person really discern anything from this dense display that couldn't be displayed more clearly and meaningfully?" I was doubtful, but unlike some development teams to whom I've posed such questions, the team at Panopticon had put a great deal of expert thought into the design of this visualization.

When Ben Shneiderman created the treemap, he was looking for a way to present a huge amount of quantitative information in a small amount of space (that is, a single screen). The results of this effort were bound to look a bit overwhelming to someone upon first sight, but because Shneiderman is an information visualization aficionado, people who take a few minutes to learn how it works can find the treemap quite useful. Because Panopticon has focused primarily on treemaps, they're familiar with the challenge of displaying a great deal of information in little space, so when they tackled the problem of simultaneously displaying a large number of time series, they had some useful experience to guide their effort.

Almost every good invention is developed in response to a particular problem. The horizon graph was developed in response to a need shared by many organizations to examine how a large number of items (stocks, product sales, employee satisfaction, and so on) changed through time, and to do so in a way that allows them to

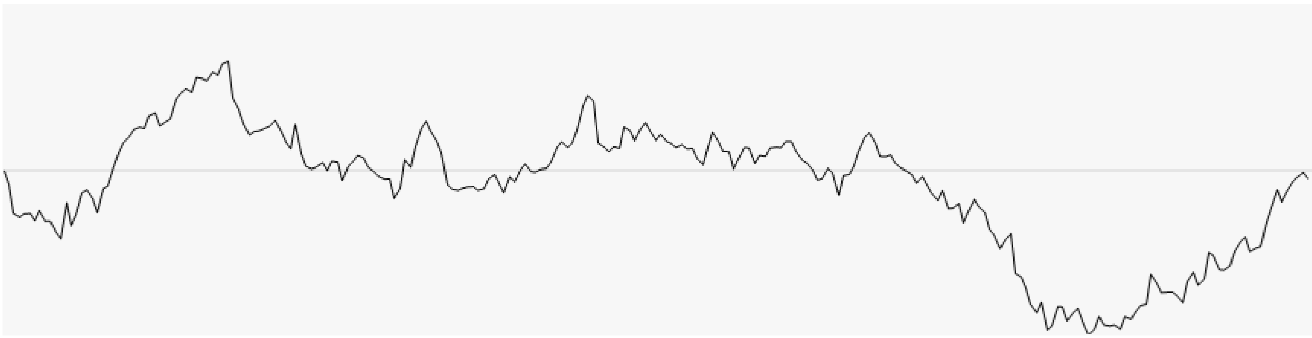
1. spot extraordinary behaviors and predominant patterns,
2. view each of the items independently from the others when they wish,
3. make comparisons between the items, and
4. view changes that occurred with enough precision to determine if further examination is required.

In the stocks example, a single line graph with 50 lines would support some of the requirements to a limited degree, but most requirements, such as the ability to view each item in isolation from the others, would not be met. The data visualization challenge that Panopticon faced was formidable.

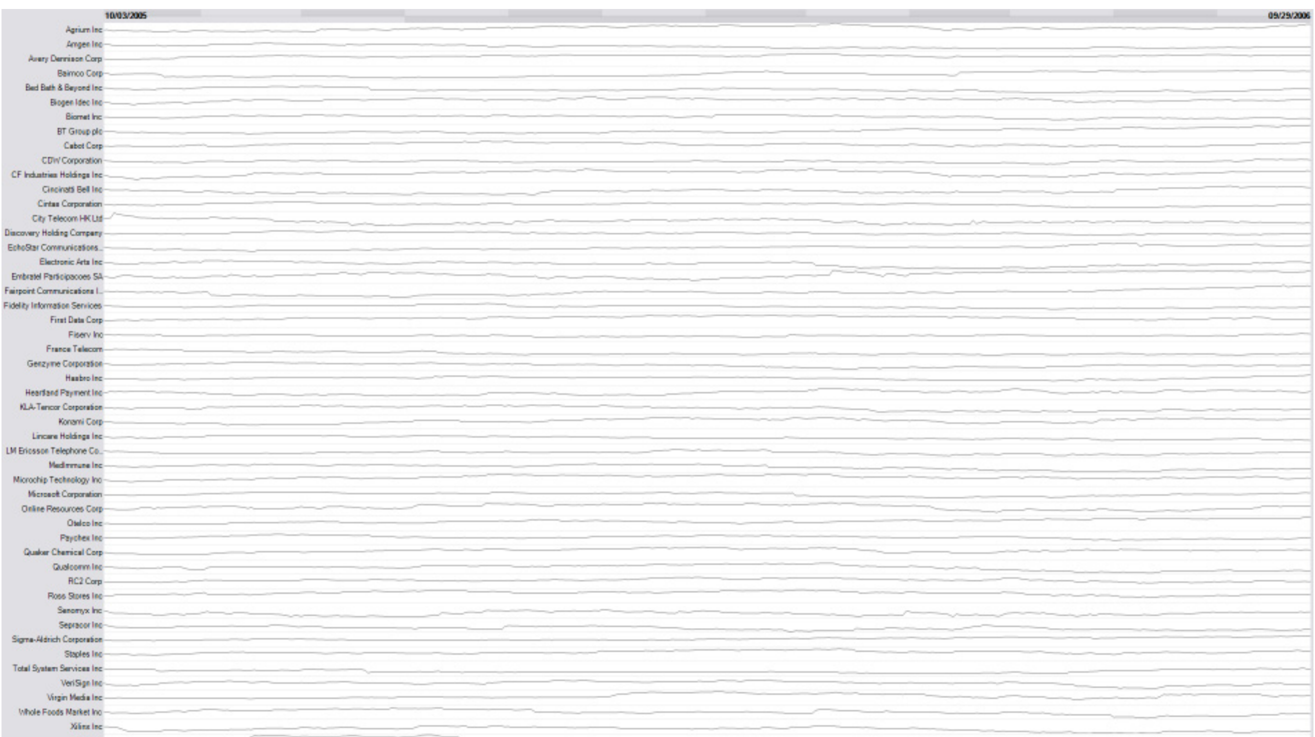
Let's walk through a simplified version of the journey that Hannes Reijner, the designer who was primarily responsible for the creation of the horizon graph, must have taken while struggling to develop a visualization that could satisfy the requirements listed above.

1. Begin with an effective visualization of a single time series

Reijner knew that nothing displays how a single set of quantitative values change through time better than a line graph, which uses position horizontally (from left to right) to encode time and vertically (up and down) to encode quantity. So, this is how he began. Here's an example of a line that encodes the changing price of a single stock over the course of roughly one year:



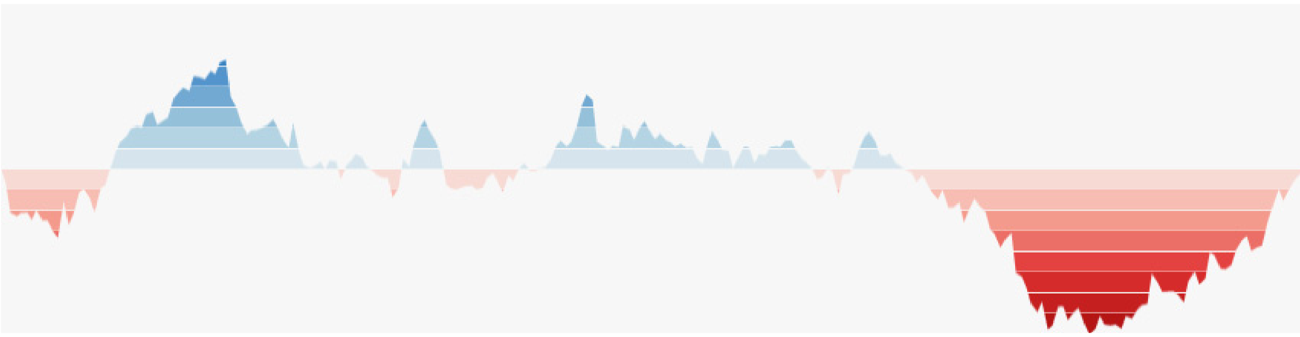
In this example, the line begins at the midpoint on the quantitative scale, which represents the price of the stock on the first day of the period. When the line moves above the horizontal reference line in the middle, it indicates that the price was greater than it was on the first day; when it dips below the reference line, as it does in this example most noticeably during the end of the period, we're seeing decreases in the stock's price relative to the first day. So far, so good. Now, how could we display many lines on the screen simultaneously, such as for 50 stocks, in a perceivable and meaningful way? As you can see in the following example, 50 line graphs arranged in rows isn't ideal.



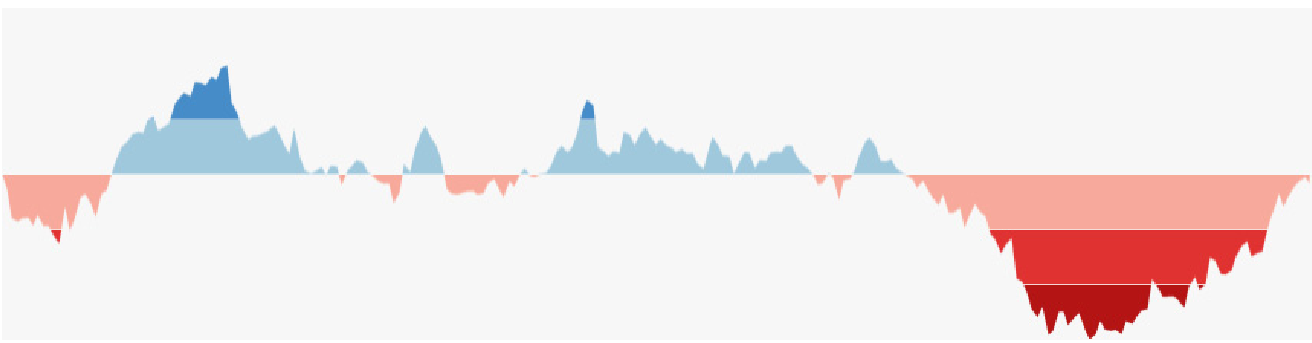
The height of each graph is no longer sufficient to display variations in value—the lines are too flat to see the patterns of change clearly. Also, with a screen full of 50 conventional line graphs, our eyes are no longer being drawn to extraordinary behavior in individual graphs or predominant patterns overall. The screen exhibits a blanket of sameness, out of which we can't spot interesting patterns without looking closely, one graph at a time. Something more must be done to draw our eyes to meaningful patterns and exceptions.

2. Use variations in color to make patterns and exceptions more visible

We need to make differences in the values more visible than up and down slopes of lines alone are able to produce in this dense display. No other visual attribute that I know can do a better job of this, under the circumstances, than variations in color. Given the fact we want the difference between increases versus decreases compared to the starting value to stand out distinctly, using different hues could do the job nicely. Differences in magnitude within either positive or negative values, however, should be encoded as differences in color intensity, which we naturally perceive quantitatively (darker is greater and lighter is less). Here's an example, which uses blue for increases and red for decreases:



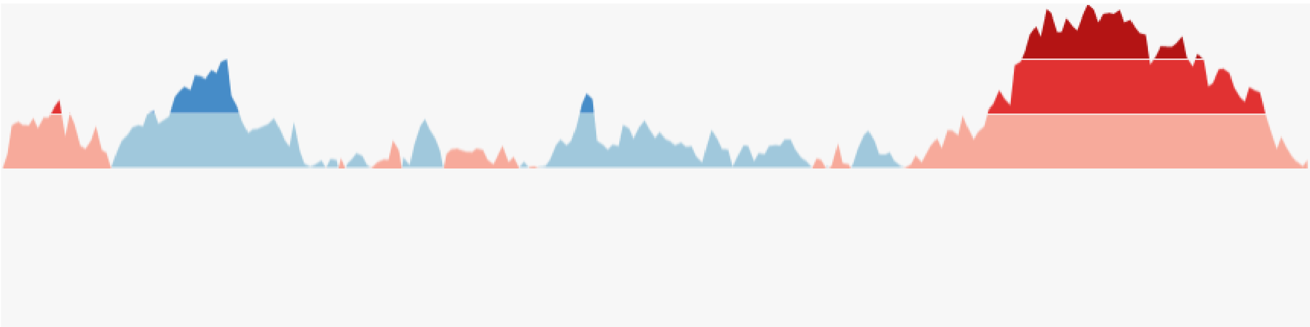
Rather than varying color intensity continuously, we're varying it in discrete steps. Notice that stock prices dipped to an extreme low towards the end of the period. At its extreme, eight discrete intensities of red have been used to encode the full range of decrease. Given the dense nature of a display consisting of 50 graphs, using discrete steps in color intensity simplifies the task of discerning differences in value. Actually, eight discrete color intensities in a 50-graph display are too much, so let's decrease the number to three.



Now it will be much easier to spot extreme values—both positive and negative—in a screen full of time-series data. Despite the improvements that we've made to the visualization so far, it would still help if the height of each graph could be increased to make it easier to see changes in values by spreading them across more vertical space. What could we change to give more vertical space to display increases and decreases in value, without decreasing the number of graphs that appear on the screen at once?

3. Display increases and decreases in value in a shared vertical space

The solution that Reijner experimented with to his eventual satisfaction involved inverting the space used to display decreases in value (the red section below the midpoint in the graph) by placing it above the horizontal reference line. Here's how it looks:

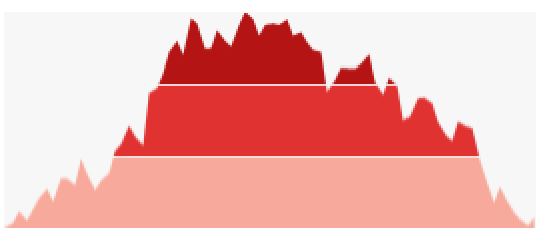


This possible solution suggested itself because, with the use of two different hues to encode increases versus decreases, placing increases above the midpoint and decreases below was no longer necessary to distinguish them. Also, because at any point in time the value could only represent either an increase (positive value in blue) or a decrease (negative value in red), there was no risk of occlusion (that is, one hiding the other) caused by both trying to occupy the same space.

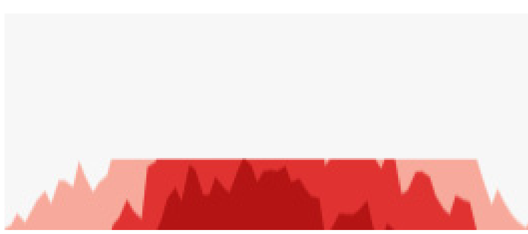
The downside of this solution, of course, is that the intuitive encoding of increases as up and decreases as down has been sacrificed. The benefit of this sacrifice is the ability to spread the data across more vertical space, which was accomplished because the bottom half of the graph is now empty and no longer needed. Is this sacrifice worth the cost? For awhile when it was first shown to me, I resisted this aspect of the design, unwilling to give up the obvious benefits of up and down, above and below, as a powerful means of distinguishing increases and decreases. In time, however, my resolve gave way as I recognized the value of the space that was gained. This space is so valuable, in fact, we should ask, as Reijner did, if there is a way to free up even more.

4. Collapse the color bands to display the values in less vertical space

Imagine what it would look like if we took the three distinct bands of red (light red, medium red, and dark red) shown below and collapsed them such that they all sat on the same baseline, with the bands of greater intensity, representing greater values, in front of the less intense.

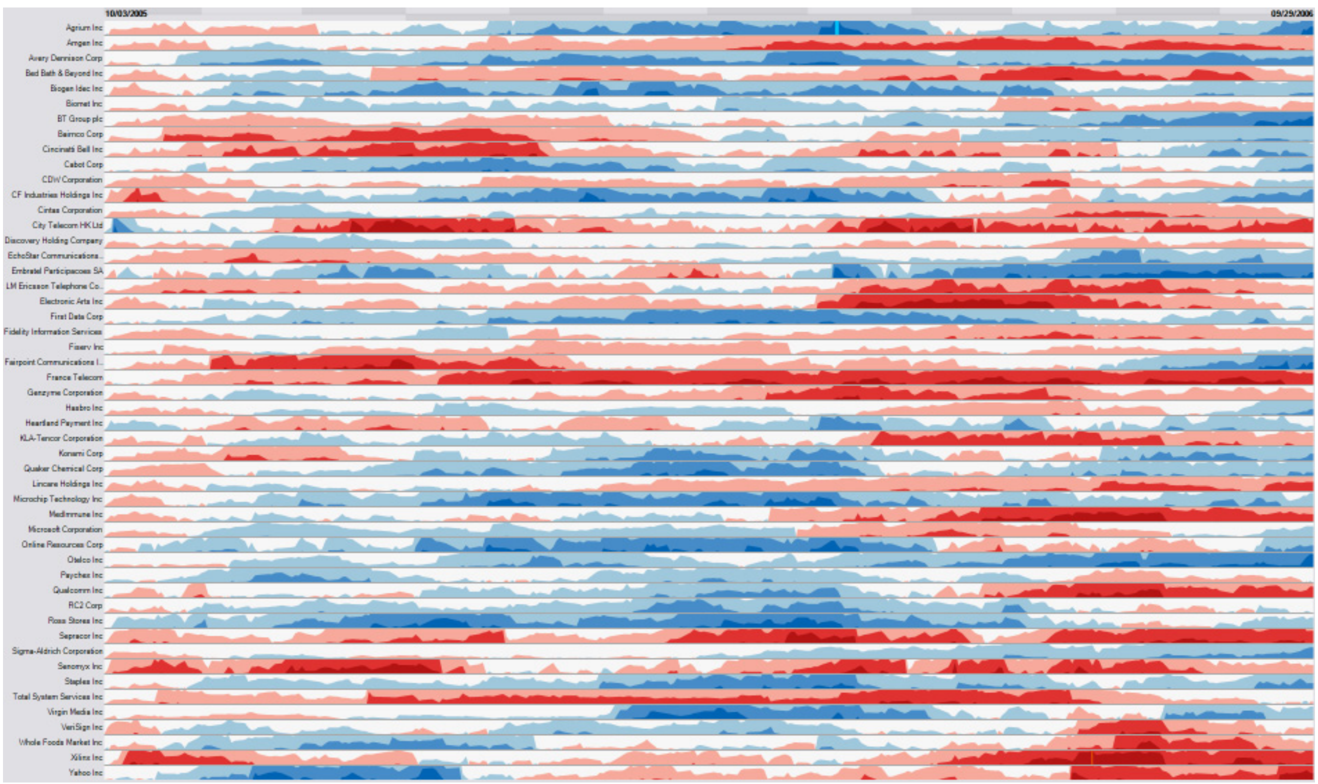


I suspect that Reijner used his imagination in just such a manner, and then tested the notion to see how it would look, as you see here:

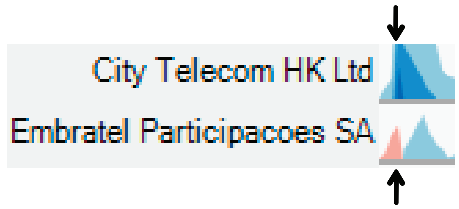


What have we gained, what have we lost, and what is the net effect? We have obviously freed up more vertical space, which we could use to either heighten the display of these values or to display more rows of graphs on the screen. Does this benefit outweigh the cost? What we've lost is the ease of seeing a set of time-series values as a continuous flow of up and down contours. Now, thresholds between one band of color and the next exhibit breaks in the flow, which require us to imagine the next greater band of color repositioned above the first to picture the continuous flow of change. This is a significant loss, but let's not be too quick to dismiss it as worthwhile. When scanning a screen full of time-series data to spot extremes and predominant patterns, we can still use color differences to do the job. In fact, when scanning this much information at once, it would be difficult to perceive anything more nuanced than what we can currently see anyway. Once an extreme value or interesting pattern catches our eyes and we wish to discern those patterns and their magnitudes more precisely, the collapsed overlapping contours of color can be used to do this. Clearly, the collapsed contours require extra effort to see and compare patterns and magnitudes of change, but does the space that we've gained justify the additional perceptual load? This is a concession that I steadfastly resisted during my meeting with Panopticon, but once again I eventually gave in. I believe that when (and only when) you must see and compare a great deal of time-series data, the extra cost in perceptual effort and loss of pattern clarity is outweighed by the benefit of seeing data.

Let's look at the original horizon graph with all 50 stocks once again to see if it now makes more sense and, most important, does the job.



Notice how easy it is to spot sections of extremely high or low values, because they stand out as darker patches. Notice the opposite as well—how easy it is to spot sections of values that vary relatively little from a stock's initial price (both increases in blue and decreases in red) because they stand out against a background of white. Can you see predominant patterns across the entire collection of 50 stocks, such as periods of time when the majority were priced higher or lower than they were in the beginning? Notice that in the middle of this period of time only a few of the stocks were valued far below their starting value? This section features mostly blues of varying intensities and light reds, with relatively few exceptions of dark red. Now, isolating the small section at the beginning of the time series for only two of the stocks as shown below, can you compare their magnitudes without excessive effort? I'm not asking you to determine the actual values, which can't be done without the quantitative scale, but merely to compare the magnitude of the City Telecom HK Ltd's stock price at its height (the darkest blue) to Embratel Participacoes SA's stock price at the same point in time.



Can you see that the increase in City Telecom HK Ltd's price is roughly five times the amount that Embratel Participacoes SA's stock decreased? The first stock's increase spanned the quantitative scale 2 ½ times (light blue spanned the scale entirely, medium blue also spanned it entirely, and dark blue spans it halfway), while the second stock's decrease (the light red peak) spanned only half the scale. With a little practice, it is perhaps no harder to compare the magnitudes of these values using these collapsed scales than it would be using two regular line graphs positioned similarly, one above the other.

The horizon graph appears to achieve its objectives. Like all innovations, I'm sure it can be improved. During my meeting at Panopticon's headquarters in Stockholm, I suggested that it might be useful to provide the means to select a few of the time series with a few clicks of the mouse and then view only them in a single, conventionally-designed line graph. This would support simpler, more precise comparisons when closer examination is needed. Despite its potential for improvement, the horizon graph succeeds because its designer, Hannes Reijner, kept the rules of visual perception in mind as he worked his way, one design step at a time, testing each as he proceeded, towards the goal. This is a good example of how effective innovations in the field of information visualization are developed.

The information visualization research community produces many innovations each year, which I'm always excited to discover in the research literature or during visits to research labs. I get a special thrill, however, when I find an effective innovation that has been incorporated into commercial software, where it must reside to reach a broad audience. Only a few commercial software vendors are producing effective visualization solutions today. When I run across an example like the horizon graph, it gives me joy to make it known and praise its worth. It gives me greater joy by far than what I get when bemoaning the poor designs of most business intelligence vendors, which is rarely fun at all.

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.