Three Blind Men and an Elephant:
The Power of Faceted Analytical Displays

Stephen Few
Perceptual Edge

INTRODUCTION

There is an old Chinese folktale about three blind men who encountere an elephant for the first time and attempt to learn about it by touch alone. The experience of each is unique because each touches a different part of the elephant. This ancient story can teach us something important today about business intelligence (BI). Here's the tale:

> *One day, three blind men happened to meet each other and gossiped a long time about many things. Suddenly one of them recalled, "I heard that an elephant is a queer animal. Too bad we're blind and can't see it."*
>
> *"Ah, yes, truly too bad we don't have the good fortune to see the strange animal," another one sighed.*
>
> *The third one, quite annoyed, joined in and said, "See? Forget it! Just to feel it would be great."*
>
> *"Well, that's true. If only there were some way of touching the elephant, we'd be able to know," they all agreed.*
>
> *It so happened that a merchant with a herd of elephants was passing, and overheard their conversation. "You fellows, do you really want to feel an elephant? Then follow me; I will show you," he said.*
>
> *The three men were surprised and happy. Taking one another's hand, they quickly formed a line and followed while the merchant led the way. Each one began to contemplate how he would feel the animal, and tried to figure how he would form an image.*
>
> *After reaching their destination, the merchant asked them to sit on the ground to wait. In a few minutes he led the first blind man to feel the elephant. With outstretched hand, he touched first the left foreleg and then the right. After that he felt the two legs from the top to the bottom, and with a beaming face, turned to say, "So, the queer animal is just like that." Then he slowly returned to the group.*

Thereupon the second blind man was led to the rear of the elephant. He touched the tail which wagged a few times, and he exclaimed with satisfaction, "Ha! Truly a queer animal! Truly odd! I know now. I know." He hurriedly stepped aside.

The third blind man's turn came, and he touched the elephant's trunk which moved back and forth turning and twisting and he thought, "That's it! I've learned."

The three blind men thanked the merchant and went their way. Each one was secretly excited over the experience and had a lot to say, yet all walked rapidly without saying a word.

"Let's sit down and have a discussion about this queer animal," the second blind man said, breaking the silence.

"A very good idea. Very good." the other two agreed for they also had this in mind. Without waiting for anyone to be properly seated, the second one blurted out, "This queer animal is like our straw fans swinging back and forth to give us a breeze. However, it's not so big or well made. The main portion is rather wispy."

"No, no!" the first blind man shouted in disagreement. "This queer animal resembles two big trees without any branches."

"You're both wrong." the third man replied. "This queer animal is similar to a snake; it's long and round, and very strong."

How they argued! Each one insisted that he alone was correct. Of course, there was no conclusion for not one had thoroughly examined the whole elephant. How can anyone describe the whole until he has learned the total of the parts.

<div align="center">(Kuo, Louise and Kuo, Yuan-Hsi, Chinese Folk Tales, 1976, Celestial Arts: Millbrae, CA, pp. 83-85.)</div>

If I retold this story today to teach a lesson about BI, I might call it "Three blind analysts and a data warehouse." Business people struggle every day to make sense of data, stumbling blindly, touching only small parts of the information, and coming away with a narrow and fragmented understanding of what it means. Conventional BI tools make it unnecessarily difficult to explore data from multiple perspectives, so analysts tend to pursue only a limited set of predetermined questions. It is simply too time consuming to explore the data thoroughly, allowing fresh discoveries to lead them to comprehensive and free-flowing exploration. Without the ability to examine data from multiple perspectives simultaneously, many of the meaningful relationships that exist in our data will remain hidden.

We stumble blindly and understand only in part, mostly because we are disabled by ineffective tools. Tableau software offers a solution. Tableau makes it so easy to shift from one perspective to another while exploring and analyzing data that we, as analysts, are encouraged to pursue every question that arises during the process, almost as fast as we can think of them. Because we are not distracted by the mechanics of using the software or forced to go through time-consuming steps to get from one view of the data to another, we become immersed in the data and the analytical process. We are able to spend our time thinking about the data, not wrestling with the software.

In this paper, I want to focus on the insights that emerge when software enables us to view a set of data from several perspectives at the same time. I call a screen that contains multiple concurrent views of a common data set so that comparisons can be made a "faceted analytical display." Don't worry about the name, however. Whatever you choose to call it is fine. What matters is that you have the means to expand your analytical reach by viewing data in this way.

DATA VISUALIZATION AND THE HUMAN BRAIN

The human brain is amazing. If you are at all aware of the current research that focuses on the brain and how it works, you know that this field is now experiencing exponential growth and discoveries as scientists take advantage of new technologies such as fMRI (functional magnetic resonance imaging) scans to observe the brain in action. Despite how powerful computers have become, much that humans can do quite easily may never be possible for computers. The reverse is also true. Many of the things that computers do best are unapproachable by the human brain, which is why we rely on them for lightning fast calculations and are happy to let them do the procedural repetitive tasks that we find so boring.

The human brain is extraordinarily great at doing some things and surprisingly limited in other ways. For instance, on one hand our brains make us extremely good at recognizing visual patterns (something computers don't do very well at all), while on the other hand they are able to remember relatively little of what we perceive. When we think about things, trying to make sense of them, the place where information is temporarily stored to support this process is called short-term memory.

Short-term memory is a lot like RAM (random access memory) in a computer in that it is limited and designed for temporary storage. Compared to that hard disk drive that is built into your computer or attached to it externally, RAM seems very limited, but compared to short-term memory in the human brain, RAM seems enormous. Only around four chunks of visual information can be stored in short-term memory at any one time. Only four chunks! Information that comes in through our eyes or that is retrieved from long-term memory in the moment of thought is extremely limited in capacity. If all four storage slots are occupied, you must let something go to allow something new to come in. When you release information from short-term memory, it can take one of two possible routes on its way out: 1) it can be stored permanently in long-term memory by means of a rehearsal process that we call memorization, or 2) it can simply be forgotten.

The primary activity of data analysis is comparison. Individual facts mean nothing in isolation. Facts become meaningful when we compare them to one another. To say that quarter-to-date revenue is $1,383,593 means little until you put it into context through one of more comparisons, such as by considering it in relation to the revenue target of 1.5 million dollars or to the amount of quarter-to-date revenue that was earned by this date last year.

To compare facts, you must hold them in short-term memory simultaneously. Because we can hold so little in short-term memory at any one time, however, to do analysis effectively, we must rely on external aids to memory. This sounds like an ideal job for a computer. Even a piece of paper that you jot down notes on to keep track of information as you're analyzing data is an external memory aid that is quite powerful despite being low-tech. A computer running properly designed software, however, can augment our ability to think about information much better than paper and pen.

Good visual analysis software, such as Tableau, can help us overcome the limitations of short-term memory in several ways. The goal is to enable as many meaningful comparisons as possible. Good software can help us increase each of the following:

1. The amount of information that we can compare (that is, greater volume)
2. The range of information that we can compare (that is, broader dimensionality)
3. The different views of the information that we can compare (that is, variable perspective)

Tableau software enriches our ability to compare data in each of the three ways just mentioned:

1. Greater volume, by displaying data visually, which allows us to see and compare patterns and trends, and also allows us to chunk more data together into the limited number of storage slots in short-term memory.
2. Broader dimensionality, by displaying data in series of small graphs arranged as a visual crosstab, which allows multiple dimensions to be compared simultaneously, with less reliance on short-term memory.
3. Variable perspective, by supporting faceted analytical displays to allow multiple views of a data set to be rapidly considered and compared by quickly swapping them in and out of short-term memory.

> *Because we can hold so little in short-term memory at any one time, however, to do analysis effectively, we must rely on external aids to memory.*

Greater volume through visual encoding

Traditional BI relies mostly on tabular data displays. Tables are wonderful if you need to look up individual values, compare a single value to another, or know values precisely, but they do not display patterns and trends. This is a problem, because data analysis relies heavily on our ability to spot and make sense of patterns and trends in data. Take a look at the table in Figure 1 and compare it to the line graph in Figure 2 of the same data. Relying on the table to discern the ups and downs of sales through time and to compare the patterns of change from region to region would yield very little of the information that is immediately obvious in the graph. Visual representations give form to data, making pattern, trends, and exceptions easy to see.

	January	February	March	April	May	June	July	August	September	October	November	December
Central	21,923	21,696	21,820	22,247	21,922	22,451	23,081	22,770	22,156	21,857	21,101	22,024
East	14,367	14,294	14,297	14,900	14,617	15,720	16,024	15,536	14,718	14,668	14,414	15,024
South	8,425	8,432	8,376	8,397	8,355	8,643	8,856	9,241	8,812	8,841	8,686	8,864
West	22,149	21,869	22,115	22,506	22,191	23,122	24,038	24,190	22,671	22,489	21,906	23,026

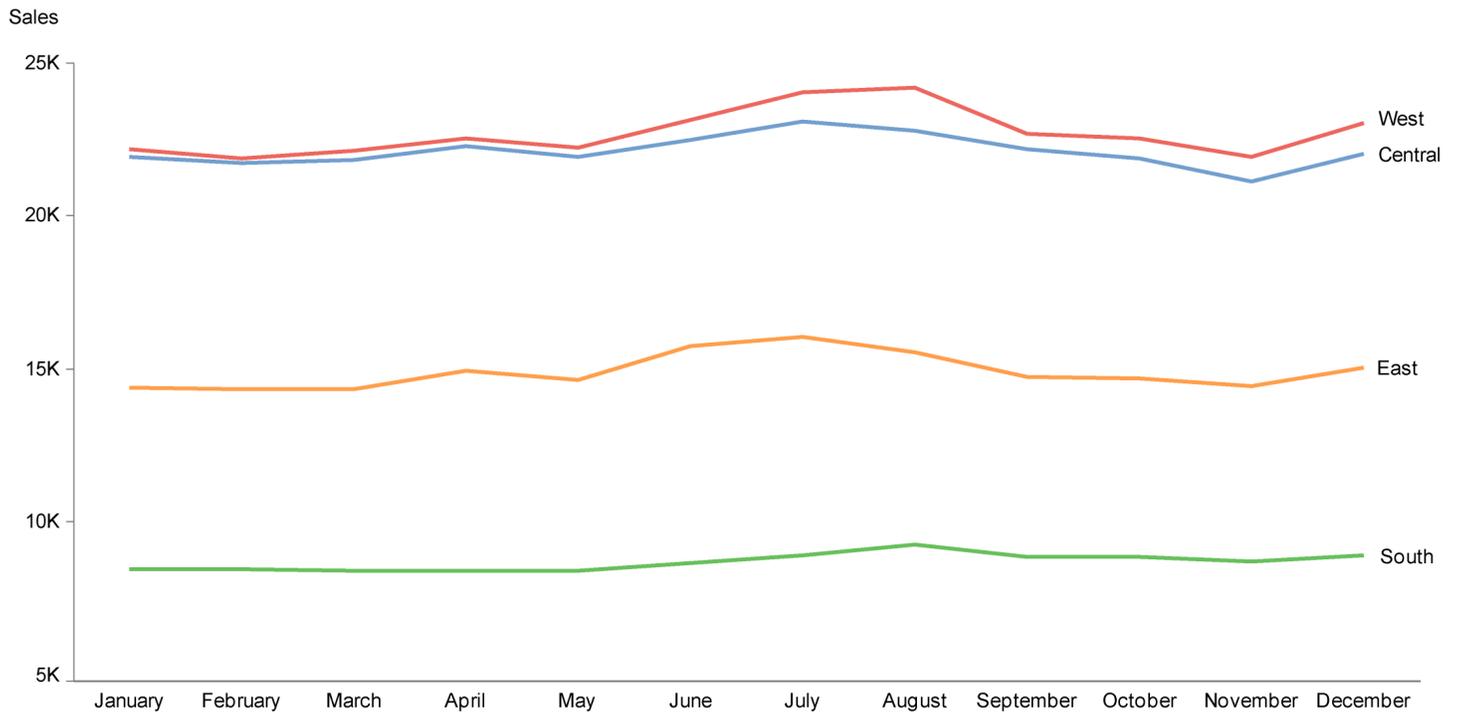**Figure 1:** A simple table of monthly sales data, grouped by region.

FIGURE 2: The same sales data that appears in the previous table (see Figure 1) displayed visually as a line graph.

Another advantage of properly designed graphs over tables for analytical purposes is less obvious. If you needed to remember information in the table, you could hold only about four of the values (that is, four of the monthly sales numbers) in short-term memory at any one time. By relying on the graph, however, 12 values are combined into each of the four lines to form a pattern that you might be able to hold entirely as a single chunk in short-term memory. Simply by giving the values a simple visual shape, we are able to hold much more information at one time in memory.

Broader dimensionality through visual crosstabs

Tableau extends the benefits of data visualization further by organizing many graphs on the screen at the same time in the form of a visual crosstab. The example in Figure 3 displays 24 small graphs and arranges them in the familiar crosstab fashion to present sales data across four different dimensions at once: products within product types by row, regions by column, and market size by the color of the line.

Not only does this approach make a great deal of data available to our eyes, it does so across several dimensions, thus broadening the dimensionality of the data well beyond traditional graphical displays.
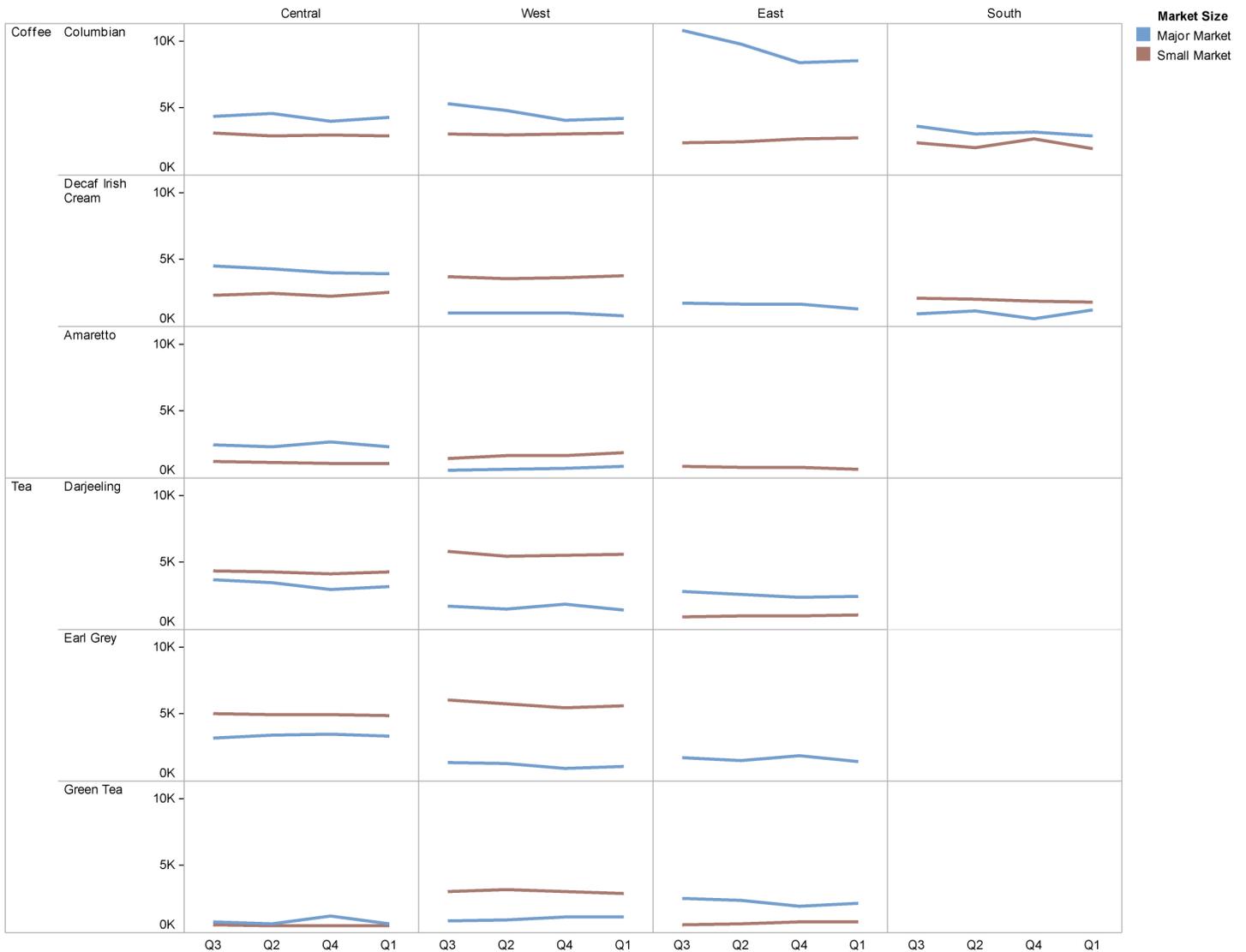
FIGURE 3: A visual crosstab.

Variable perspective through faceted analytical displays

To understand something, you often have to examine it from many angles and focus on many parts. Too much business data analysis involves looking only for one thing in particular. Is revenue going up? According to this line graph, the answer is "yes," end of story. Perhaps, however, you ought to look at expenses, profits, marketing campaigns, seasonality, composition

of the sales force, new product introductions, and the competition to understand the richer story that your data has to tell.

Multi-faceted data analysis like this is seldom pursued, not so much because its value isn't recognized, but because the tools that business analysts generally use discourage a thorough pursuit of the facts. It would be very difficult—nearly impossible—to build a picture of what's going on in one's mind if you were forced to examine each useful view of the data independently, separated by individual queries and the mechanics of constructing each new table or graph. This slow and fragmented process will resign you to the narrow and misleading world of the three blind men who encountered the elephant.

You'll never be able to stitch together a whole cloth from these individual swatches, partly due to the limitations of short-term memory, but also because the cumbersome and time-consuming mechanics of the process will discourage any momentum that your inquiring mind tries to build. For your eyes to be opened and your mind to become enlightened, you need tools that can display the data from several perspectives simultaneously, allowing you to shift your gaze rapidly between them to spot connections and build a rich picture of what's going on.

> " *Multi-faceted data analysis like this is seldom pursued, not so much because its value isn't recognized, but because the tools that business analysts generally use discourage a thorough pursuit of the facts.* "

FACETED ANALYTICAL DISPLAYS IN ACTION

To understand the strength of faceted analytical displays, you definitely need to see one. It would be even better if I could show you in a live demonstration the ways that you can interact with them to uncover meanings in data, but in this paper I'm stuck with static images. I'll do my best to walk you through a few typical interactions using a series of screenshots.

Here's an example that I quickly threw together using the new multi-chart functionality that Tableau calls a dashboard, but I prefer to call a faceted analytical display to distinguish a

display like this that is used for interactive data explorations and analysis from one that is used to monitor what's going on:

FIGURE 4: A faceted analytical display.

This single screen combines several views of a single data set, which each focus on a different aspect of the data. In the upper left corner you see a visual crosstab of line graphs that display monthly sales by product type (the rows) and region (the columns), sorted by total sales in descending order from top to bottom by product type and from left to right by region. To the right of that you see a comparison of actual sales (the asterisks) to the sales budget (the circles) by product, with color coding to associate individual products with product types. In the lower left corner you see sales by state, further divided into product types by color. To the right of that you see a scatterplot that examines the correlation between marketing expenses and sales, grouped by product type using color, with a separate data point for each combination of state and product type. In the lower right corner you see sales by product, once again grouped into product types by color. Finally, in the panel at the far right, two legends appear at the top and below that appear five filter controls for product type, product, state, sales, and profit. What we have here is a display that will allow us to explore a year's worth of sales data from several perspectives with little effort and no delay as we move rapidly from view to view and also filter the

data to examine particular subsets in detail.

I can discern several facts about sales using this single display, even before beginning to filter the data. Here are a few of the facts that I notice:

- Espresso sales were best overall, which I can see because the Monthly Sales graphs are sorted by product type in descending order, as are the segments of the stacked bars in the Sales by State graph.
- Despite the fact that Espresso is the leading product type, the leading product is Columbian Coffee, followed by Lemon Tea in the number two position. Espresso sales take the overall lead, however, because the third and fourth products both fall into the Espresso category.
- Despite the fact that Columbian Coffee is the best sales performer, it is one of only two products that failed to meet the sales budget. The other is Decaf Irish Cream. Both are coffee products (the blue items). Perhaps the person in charge of budgeting coffee sales doesn't handle budgeting as well as the other product type managers. Of course, there are other possible explanations, so the truth requires

deeper digging. I suppose it's possible that the high sales performance of tea products (the red items) compared to budget might also be due to poor budgeting skills.

- Sales are highest in the West region, probably due to the fact that sales in the state of California lead the nation. Although the Central region does not perform as well as the West overall, it outperforms the West in coffee sales. Coffee sales increase in the Central region in the month of July, as they also do in the East, which happens a month later in the South. A similar summer peak, however, does not occur in the West for coffee, but it does for herbal tea.

- In general, there is a positive correlation between marketing expenses and sales, and no single product type stands out as being better or worse than the others in this respect. The scatterplot reveals a few outliers in the data—data points that seem far removed from the norm—and there appears to be two separate groups of data points that form linear positive correlations, revealed by the one linear series of points that appears above the trend line and the other that appears below

it. This is worthy of further investigation. The two outliers in the bottom right of the scatterplot, with low sales corresponding to high marketing expenditures, are Green Tea in Nevada and Café Mocha in New York. I was able to determine this simply by pointing to these points, which caused the details to appear as text in a popup window.

The story that this display tells without filtering is probably much richer than this, but these are the facts that my eyes noticed with little effort.

Faceted analytical displays really come alive when you begin to filter the data. Filters in Tableau can either function locally, affecting a single view (that is, a single chart), or globally, simultaneously affecting all views. The ability to filter all of the views in the same way is especially revealing. Let's look at a few simple examples. We'll start with something simple. What happens if we remove California, the top state, from the mix? When I uncheck California in the state filter, I'm able to watch the display change in appearance to this new picture of the data:
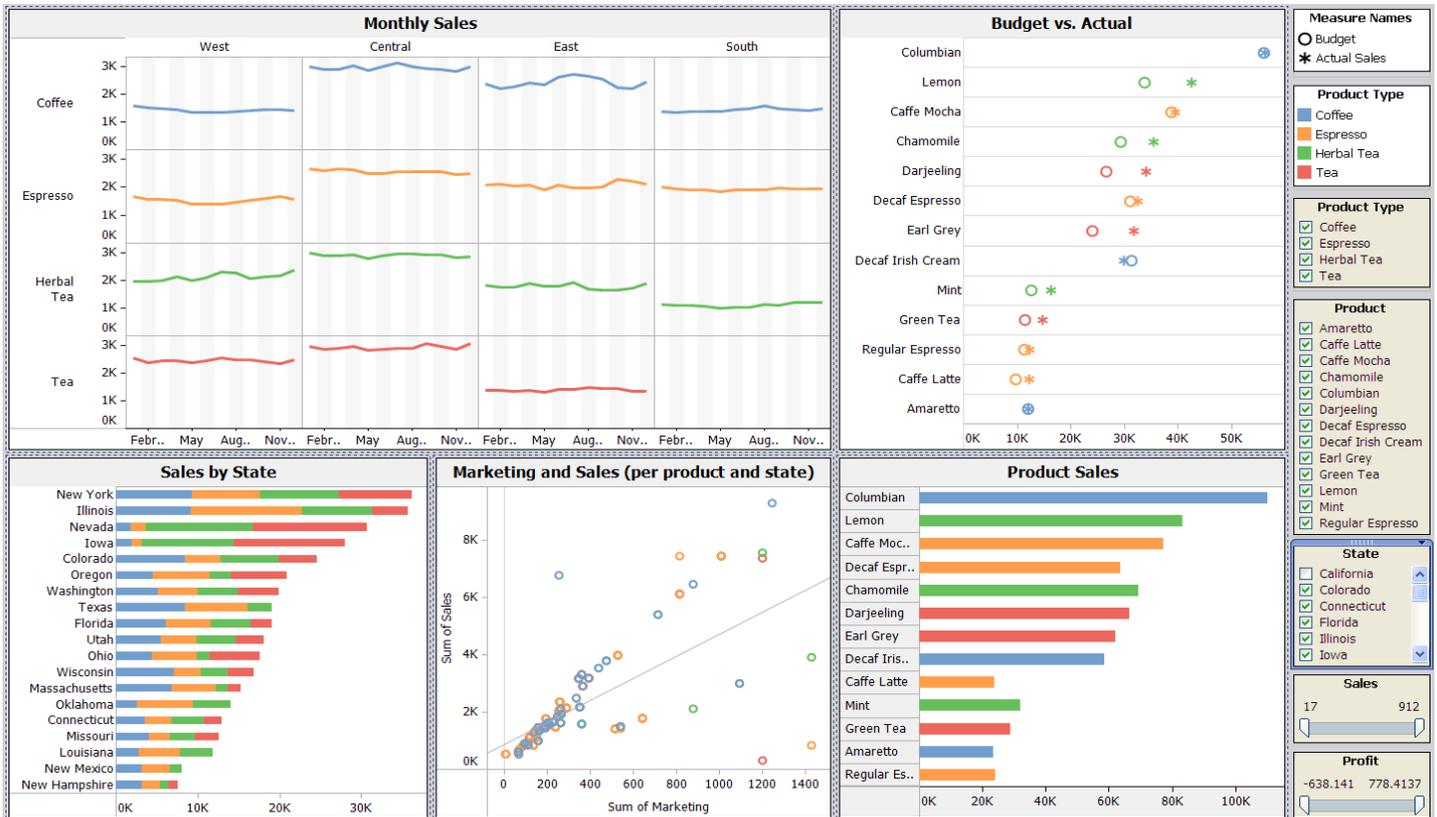

Figure 5: The state of California has been filtered out of this faceted analytical display.

The first thing I notice is that without California, the relative order of product sales has been rearranged. I have intentionally set the Budget vs. Actual graph to automatically re-sort if filtering affects the relative order of product sales, but the Product Sales graph is not set to automatically re-sort. This allows me to look at the Product Sales graph to easily spot when bars are no longer in order by size, which tells me that short bars that appear above longer bars represent products that have decreased in sales relative to those that previously had lower sales. For example, notice that the Decaf Espresso and Café Latte bars are both smaller than those immediately below them.

Now by looking up at the Budget vs. Actual graph, I can see that Decaf Espresso has fallen from the 4th position to 6th and that Café Latte has fallen from 9th to 12th. These products obviously sold better in California than they did overall. I can also tell by glancing at the Budget vs. Actual graph that the poor budget performance of Columbian Coffee has been corrected by removing the state of California, which must have been primarily responsible.

Now let's see if we can learn something about profits, first by filtering out all but negative profits (that is, losses). Here's the result:

Figure 6: Profits have been filtered out of this faceted analytical display.

Filtering out all but the losses has affected the overall sales picture dramatically. Notice how many differences now catch your eye. Looking at the Sales by State graph, you can easily see that the states of Missouri, Utah, and New Mexico now show higher sales than all of the other states, including the three top ranking states of

California, New York, and Nevada. Move your attention to the Marketing vs. Sales scatterplot and you will probably be surprised to find that with the exception of two outliers in the bottom right corner, most sales seem to exhibit a fairly linear positive correlation between marketing expenses and sales, so we probably can't blame the losses on

marketing expenses. Keep moving to the right, this time looking at the Product Sales graph, and you can see that Café Mocha and Decaf Irish Cream appear to have a disproportionate influence on losses, which you can further confirm by glancing up at the Budget vs. Actual graph where these products are now ranked first and second. Before moving on, notice also how little Columbian Coffee contributed to the losses, which appears dead last in rank, even though it previously ranked as #1. Lastly, notice how jagged the lines are in the

Monthly sales graphs. With the exception of Herbal Tea in the West region and Coffee in the Central and South regions, low profits appear to be related to volatile sales with lots of dramatic ups and downs through the year.

Let's look at one more example, this time the reverse of what we just examined. Here's the picture that results from filtering out all sales except those with high profits ranging from $362 to the maximum amount of $778 per customer:
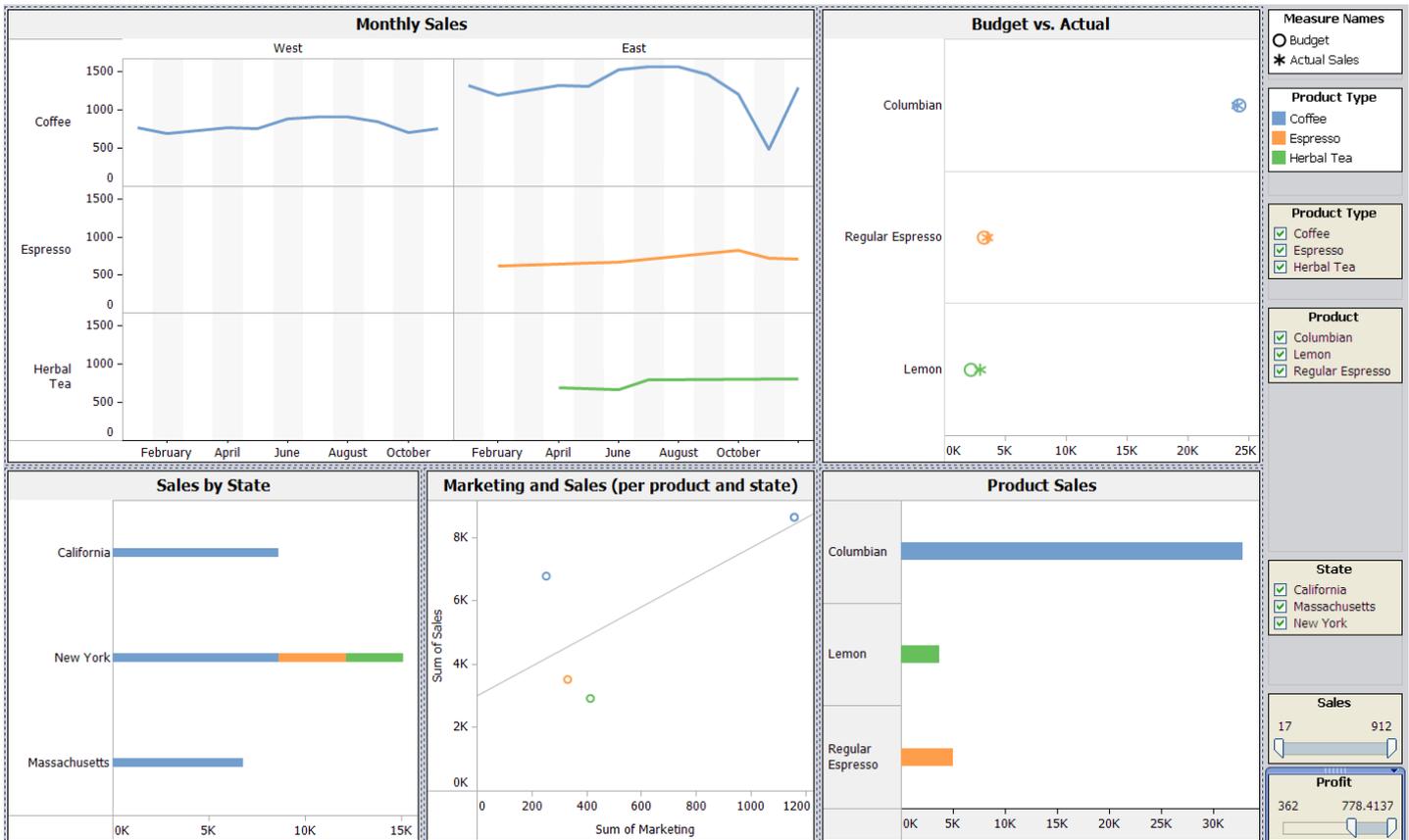

Figure 7:
All but high profits have been filtered out of this faceted analytical display.

By slowing moving the left end of the profit filter slider control (see bottom right corner of Figure 6) to the right, gradually removing low profits and narrowing the data down to higher and higher profits, I noticed that as soon as I passed the threshold of $361 dollars in profit, the number of products dropped from seven to the three that

now remain, eliminating the Tea product type, and the number of states dropped from six to three, eliminating the Central and South regions. Looking at the Monthly Sales graphs, we can see that high profits are associated with Coffee sales throughout the year in the West, specifically Columbian Coffee.

> *The real proof that faceted analytical displays are powerful is demonstrated when you try them out on your own, using your own data. Every new fact that you discover could be one that translates into big benefits to your business.*

Each of the examples of using faceted analytical displays above involves the use of global filters, which allow you to remove the same subset data from all views in the display. As an alternative to this approach to filtering, let me show you something that you can do with a faceted analytical display that involves the use of filters that don't affect all views.
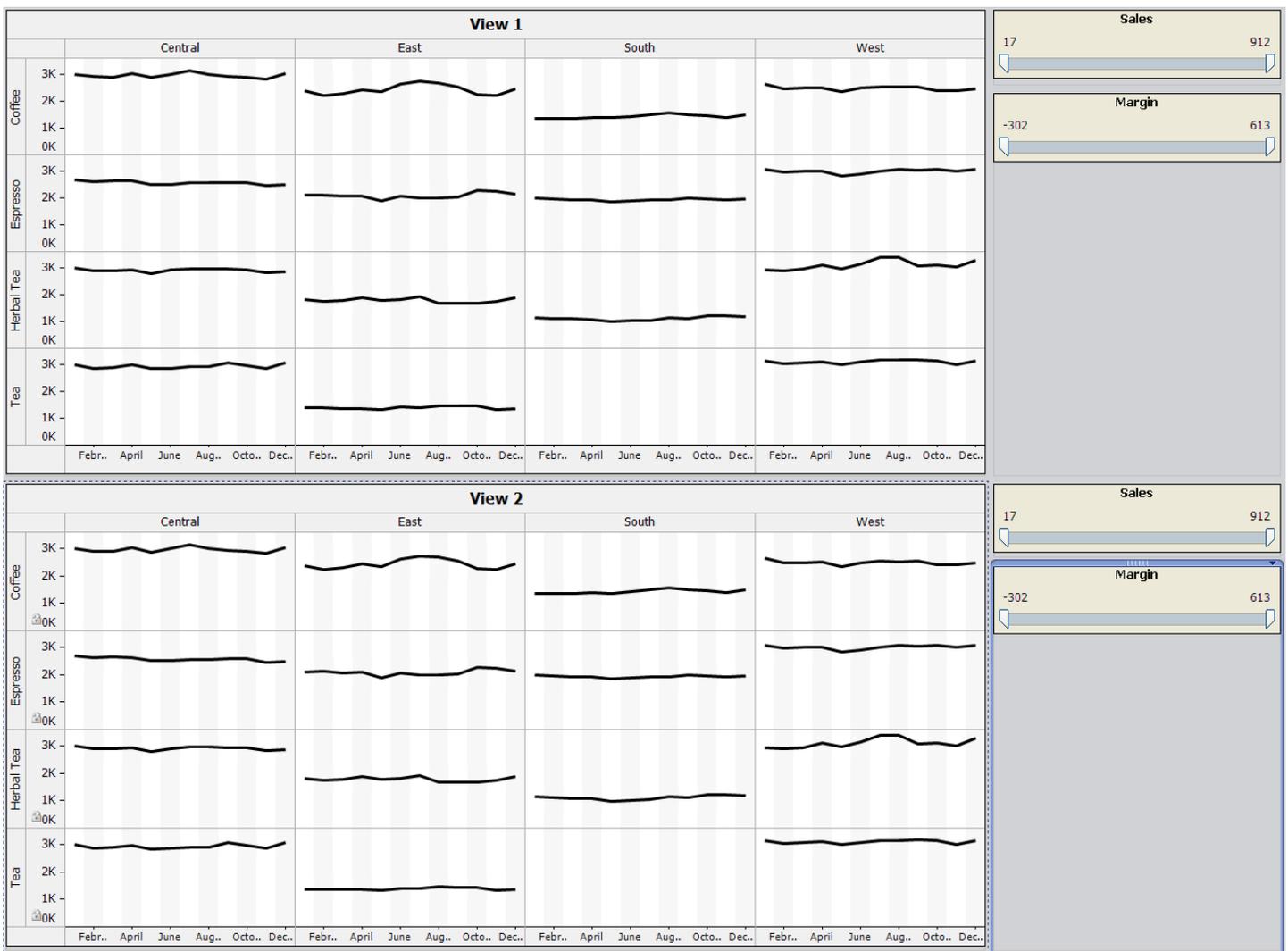
FIGURE 8:

This display shows the same view of the data twice, but allows separate filters to be applied to each.

In this example, notice that the same exact visual crosstab appears twice: above and below. Both present monthly sales by product type and region. The sales and profit margin filters in the top right affect only the top graph and those below affect only the bottom graph. A display of this type allows me to compare different subsets of sales by filtering each of the graphs differently. For instance, look at what happens when I leave the upper view alone, without applying any filters, but in the lower view filter out all but sales with profit margins of $100 and greater:
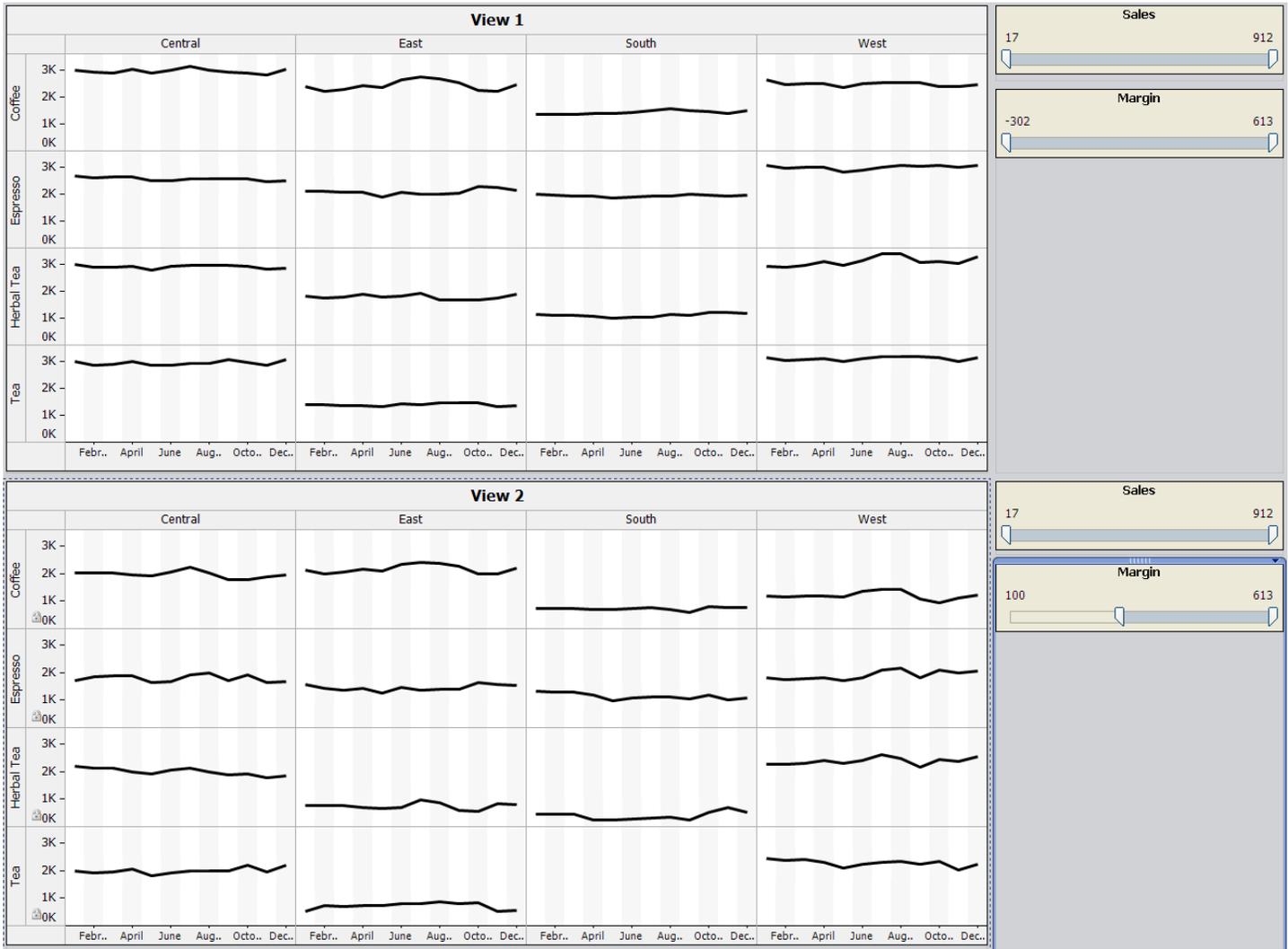

Figure 9:
Profit margins less than 100 have been filtered out of the lower visual crosstab (View 2).

By displaying these two versions of the data simultaneously within eye span, you can make subtle comparisons with ease. For instance, you don't need to look beyond the first graph in the top row of each view to see that the elimination of sales with low profit margins has caused the peak in coffee sales in the central region that occurred in April to disappear completely.

What I hope is becoming obvious with these examples is that faceted analytical displays, with the multiple perspectives that they simultaneously reveal, can lead to insights that would be difficult to unearth otherwise. The real proof that faceted analytical displays are powerful is demonstrated when you try them out on your own, using your own data. Every new fact that you discover could be one that translates into big benefits to your business

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 consulting and training 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.*

ACKNOWLEDGEMENT