



Enhancing Visual Analysis by Linking Multiple Views of Data

AUTHOR:

Chris Stolte
VP, Engineering and Co-founder
Tableau Software

Ellie Fields
Director, Marketing
Tableau Software

DATE:

May 27, 2009

How Brushing and Linking Surface Hidden Relationships in Dashboards

Understanding your data is rarely as simple as knowing a single number. You often need to answer multiple questions at once, and at the same time see trends and context. Evaluating product performance brings up questions about volume, price, discounts, and locations. Evaluating system performance brings up questions about response time, utilization, and load. Whatever your data, it can be difficult to find the real story while focusing on a single answer at a time.

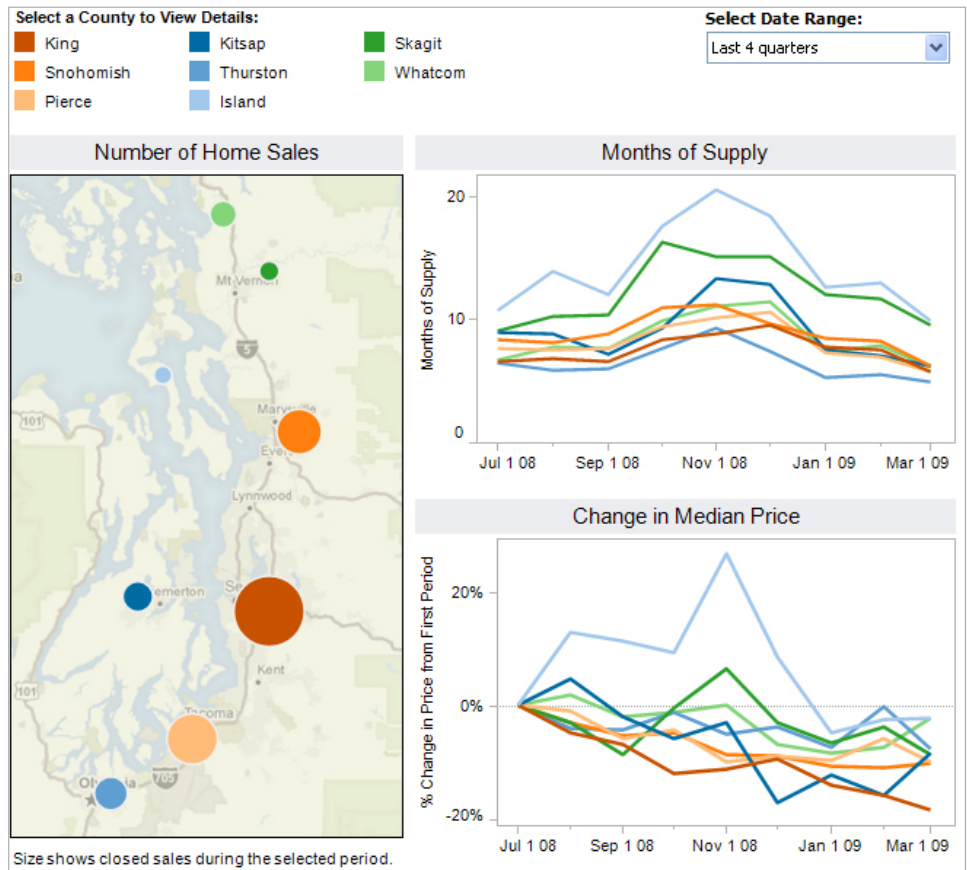
Linking multiple views can help you make sense of data. In academic research, this is called “brushing and linking.” At Tableau, we often refer to it as “highlighting and visual filtering.” We’ll use those terms interchangeably in this paper.

The concept of brushing and linking is to tie together different views while you explore your data interactively. Essentially, when you’re working with data in one view, other views adjust to highlight or filter related data. This makes it easier to answer question such as, what else was happening during a peak in traffic? Or, how important are California sales to two different store chains?

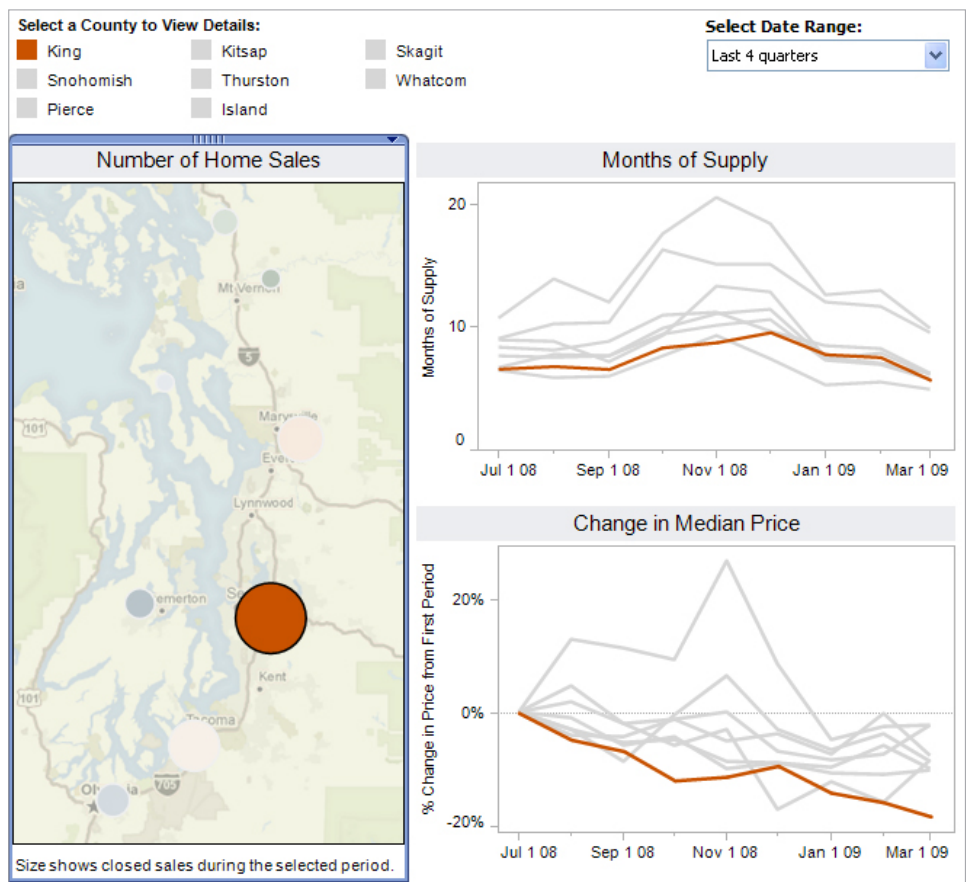
In this paper we’ll discuss several ways to link perspectives. These include highlighting by entity, highlighting by date & time, highlighting by category, highlighting to show detail, and visual filtering.

Highlighting by Entity: What else is happening to this?

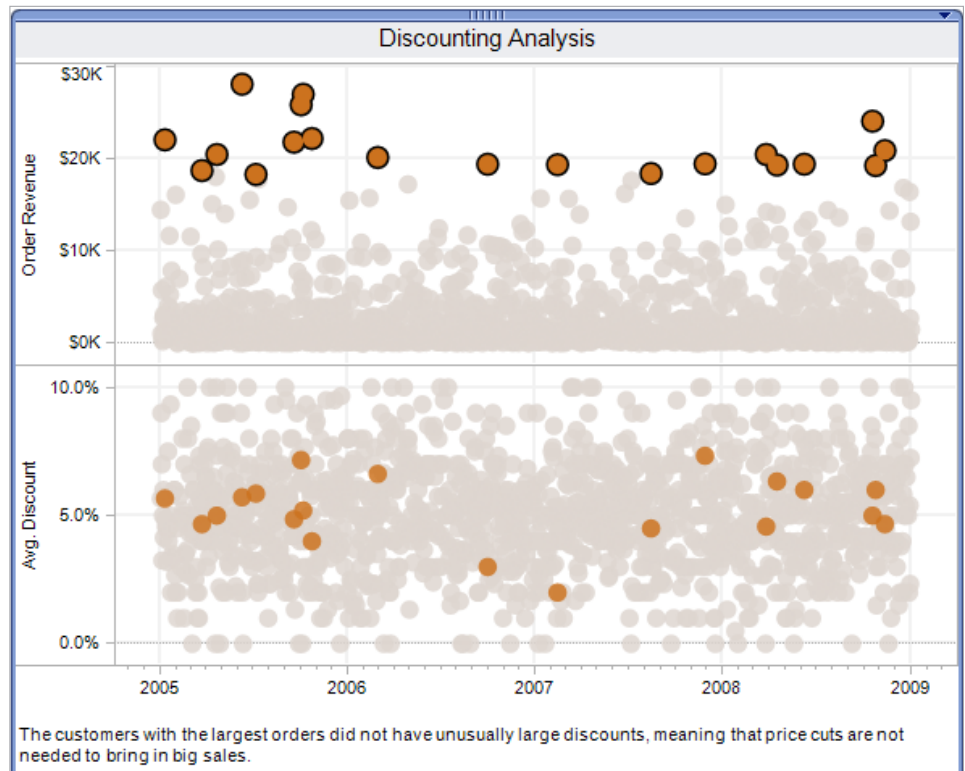
In many cases, you're showing different views of the same entity and you want to know, "what else is happening with this one?" An example using real estate data can make this clear. In the following dashboard, the entities are the counties. Each county is explicitly represented in every view and the difference in the views is that different attributes (location and time) and measures (price, supply) are represented in each.



Perhaps you want to answer the question, “How is the King County market changing?” When you click on that county in the map, data for that county is highlighted in each view. Now it’s much easier to understand the trends of supply and price in King County. As importantly, another simple click on Snohomish makes it easy and fast to see differences with respect to King County trends.



In another example of entity brushing, we're looking at sales orders including discounts and revenue amounts. When you select orders in the view on top, the same orders are highlighted in the view on the bottom, so you can easily see both the revenue amount and the discount for these orders.

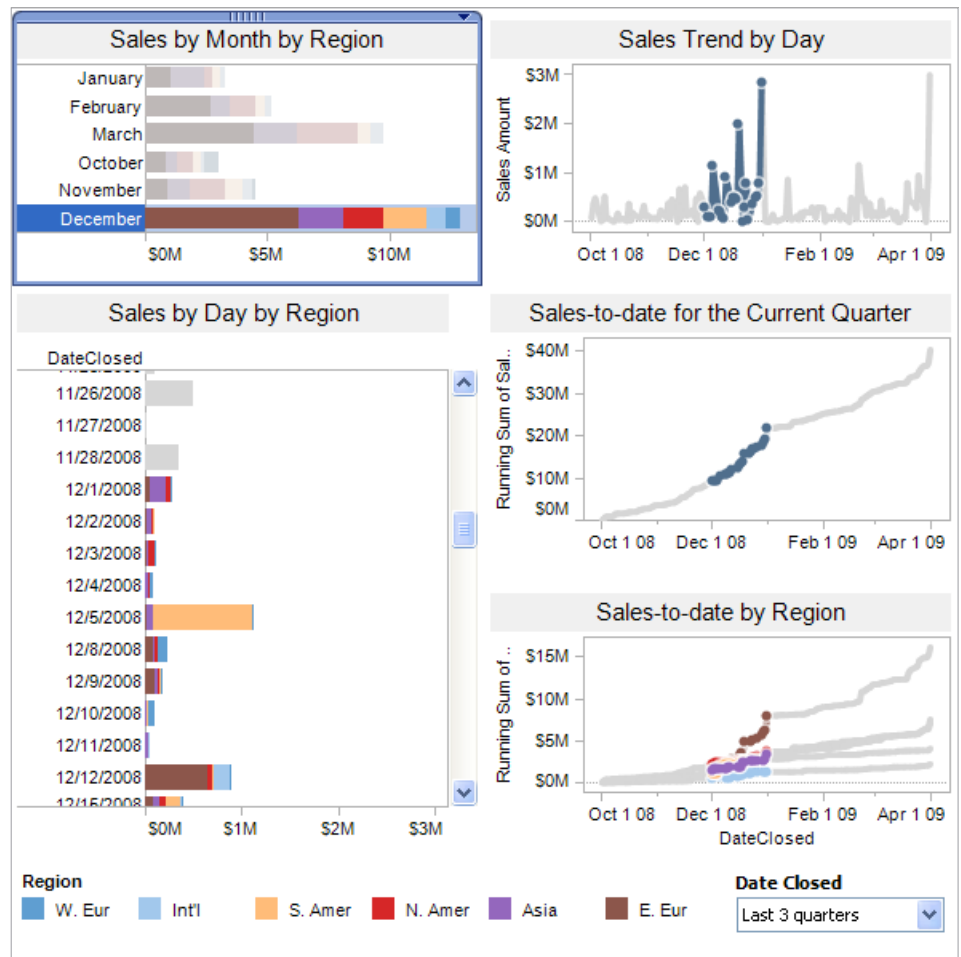


In this case, highlighting shows you that the most valuable orders did not have large discounts, so discounting may not be important to winning big sales.

Highlighting by Date & Time: What happened this month?

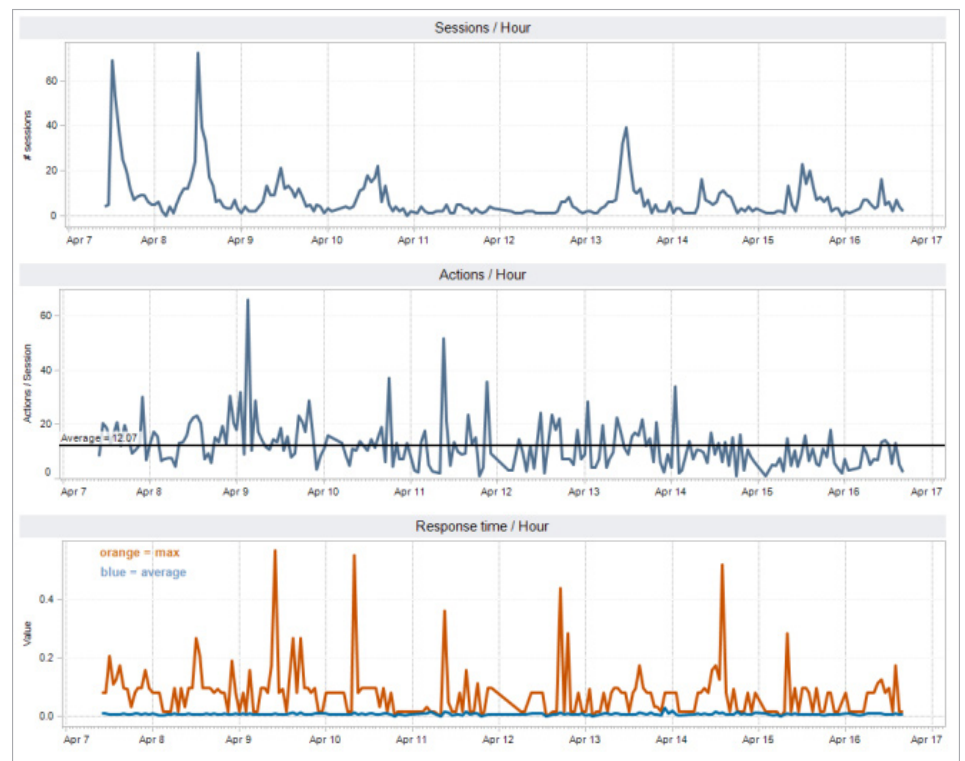
Often you're trying to find out what happened at a given time. This is true in many monitoring or event data sets, where time is common across multiple views. Date & time highlighting helps you answer questions like, what happened in that month? Or in that moment?

A good example is sales data. The dashboard below shows sales by day, by region, and sales to date. You may want to know, “How did the trend of sales change in December?” By highlighting December in the top right graph you can see in other views that December sales were much higher than October and November sales. You can also see which regions contributed the most in December (Europe and Asia) and which days were particularly strong. The sales-to-date graph shows that a larger proportion of the fourth quarter’s sales happened in the last month of that quarter. In Tableau, you can turn on date & time highlighting with one click and select December with another click to see these trends.

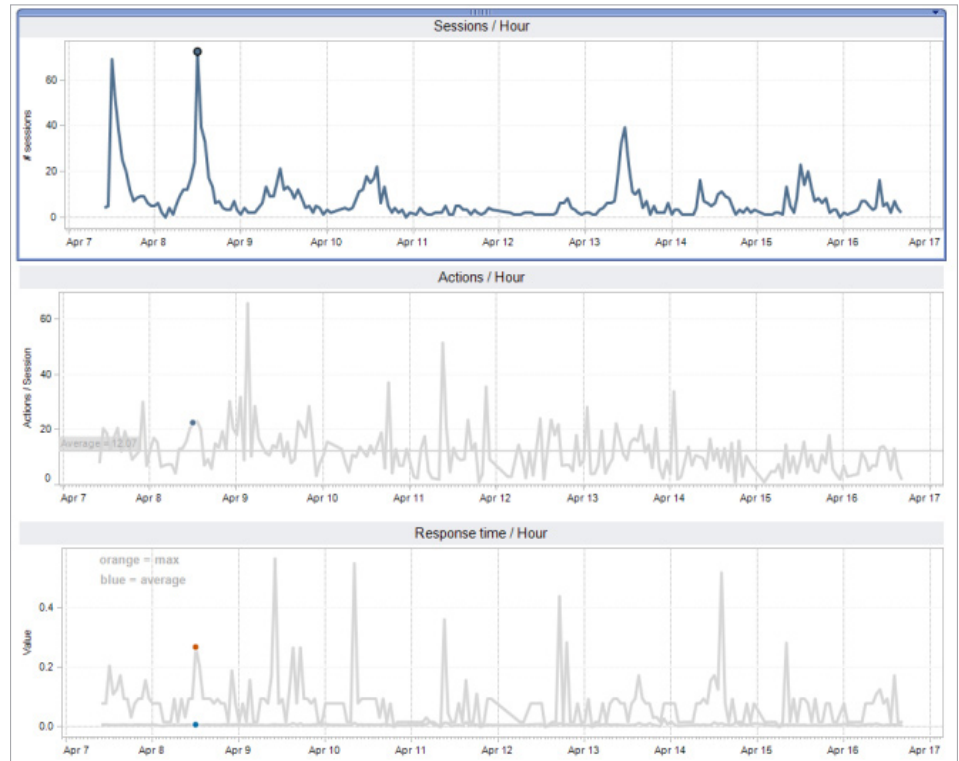


Note that you can highlight from any view to any other: you may select December 5th in the Sales by Day and December will be highlighted in the Sales by Month view. Or, as shown in this example, you can select the month of December in the Sales by Month view to see all days in December highlighted in the other views.

Performance analysis is another good example of date & time highlighting. The following graph shows a number of performance metrics for an application server by time. The interesting analytic question is “what happened during the peak on April 8?”



Highlighting by time allows that simple comparison:

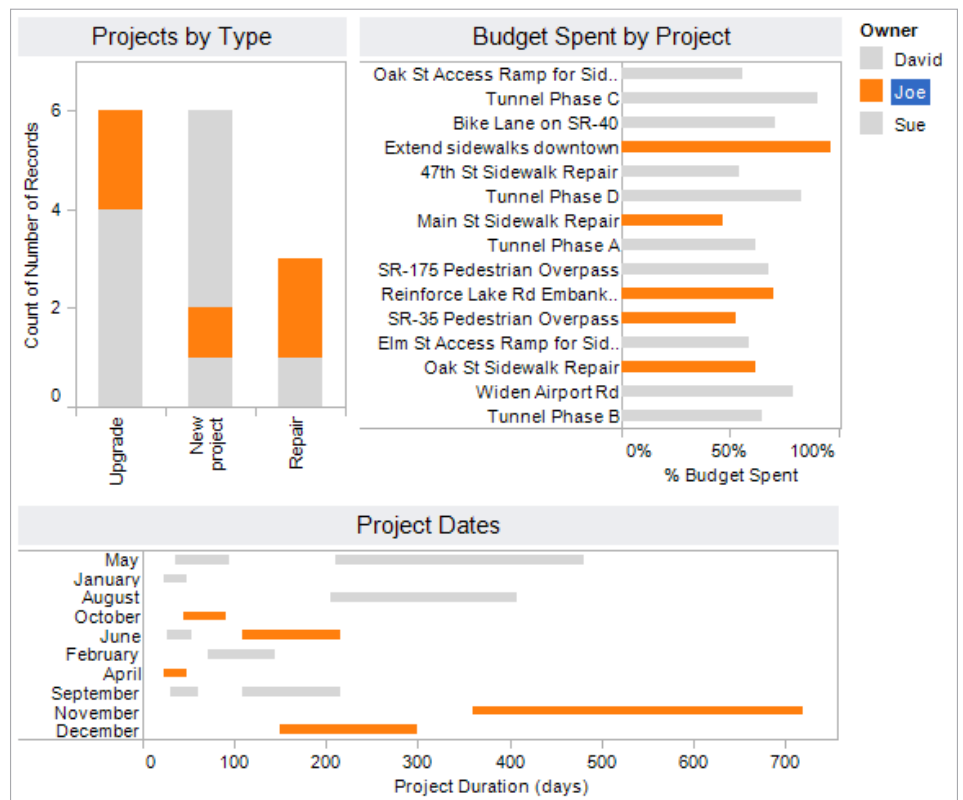


In this case, there was a sustained increase in actions requested of the server leading up to the peak, probably due to the peak in sessions. Max response time reached a local maximum but not its highest value over the time period, so the server seems to have handled the peak in sessions.

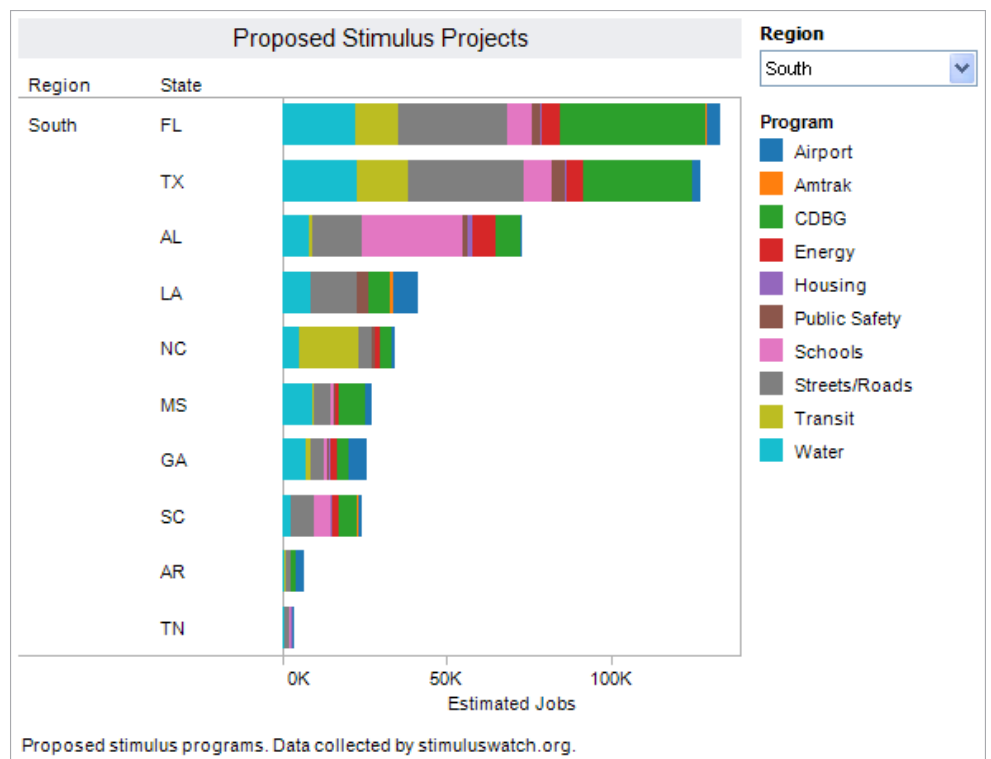
Highlighting by Category: What's happening with Joe's projects?

Another common type of highlighting is by category, or shared attributes. The shared attribute may be location (all the stores in California), type (all the programs classified as Transit) or association (all of Joe's projects.) Date & time highlighting is one specific type of highlighting by category, where time is the shared attribute among entities.

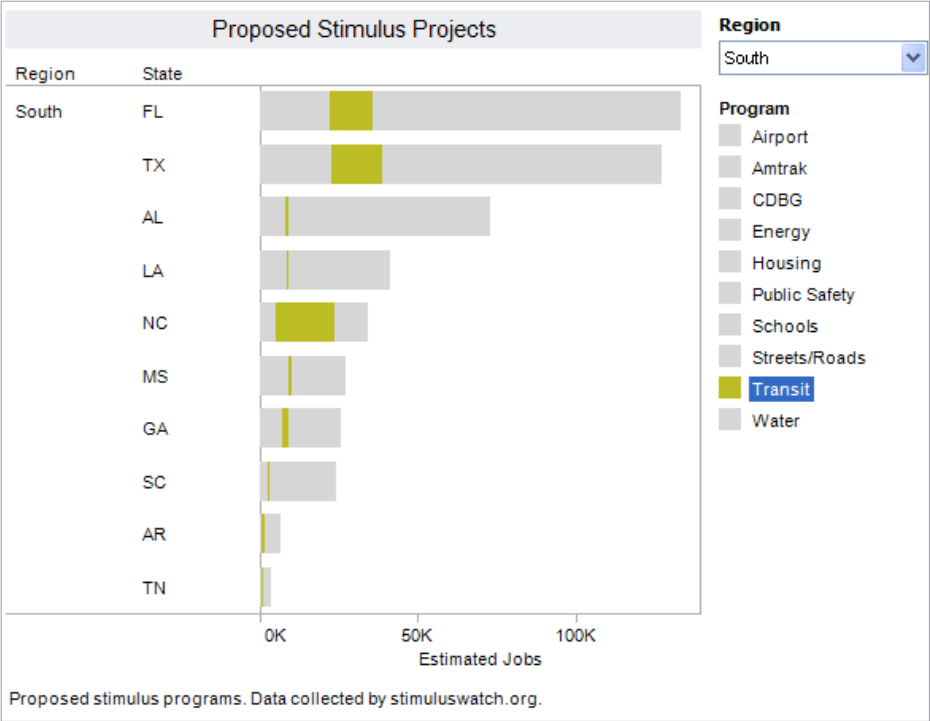
Here we're defining entities as the main subjects of the analysis, each of which has a set of characteristics, or attributes. Therefore, in the view below you can analyze projects (the entities) by attributes such as owner (Joe, David, Sue), type (upgrade, repair, new project), or start date (February, March). We're interested in what's happening with all projects whose owner is Joe:



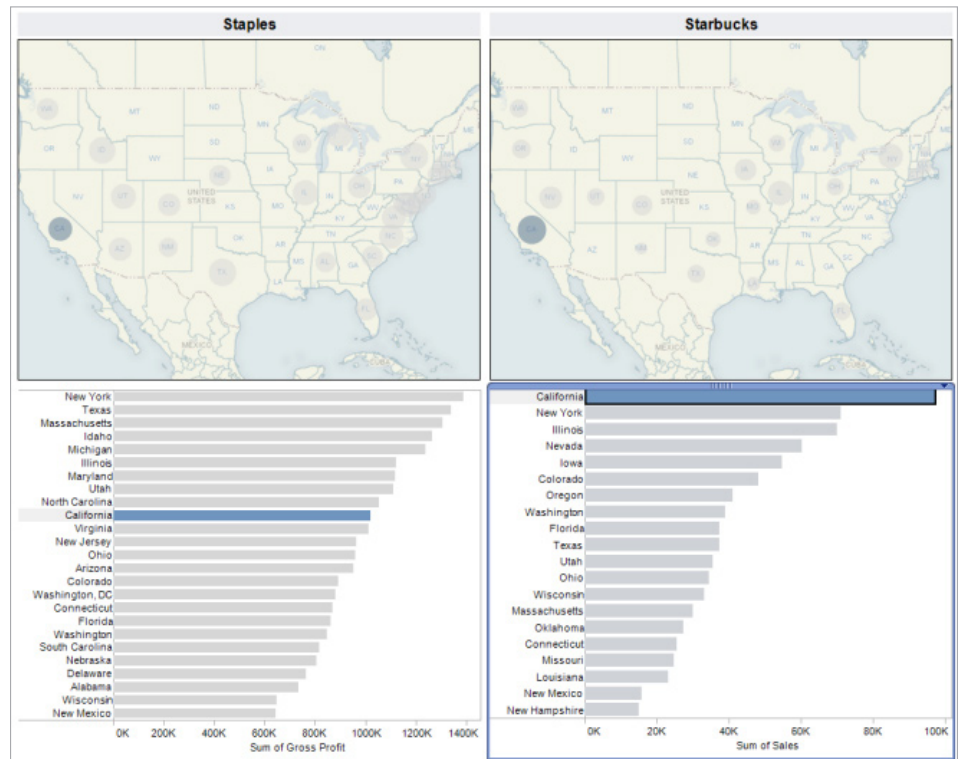
Another simple example shows proposed stimulus projects, broken down by state and program.



You may want to know where Transit project generate the most jobs. Highlighting the projects (entities) by program type (attribute) shows that in general, states with more jobs overall have more Transit jobs (Florida, Texas) but that North Carolina has a high number of Transit jobs relative to total jobs in that state. Turning on highlighting with one click yielded this result.



Another common question is, “What drove behavior in that place?” The dashboard below displays hypothetical sales for two different companies (an office supply store like Staples and a coffee retailer like Starbucks). We can use highlighting by category to focus on specific data based on location: sales in California for the two companies.



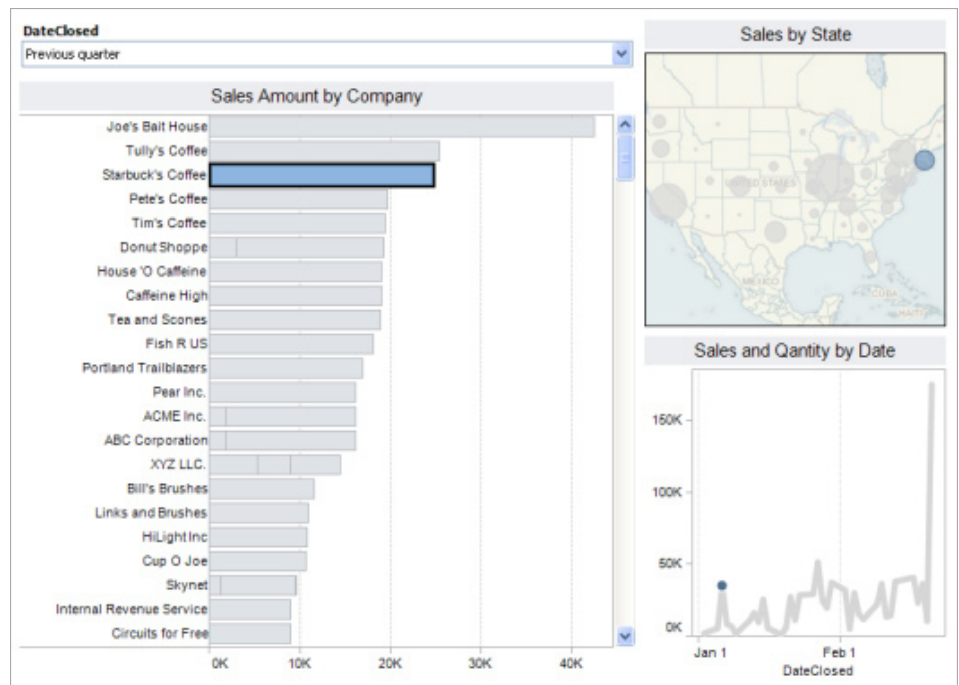
Highlighting by state shows clearly that for the office supply store California is one of many important states, but for the coffee chain California is the most important state.

It's important to note that the example above draws from two different data sources. Because it's common to have related data in different places, highlighting and visual filtering need to work across data sources. Otherwise the questions you can ask are constrained by what's in your database or spreadsheet.

Highlighting to Show Detail in Context: What's the rest of the story?

In this type of highlighting you use one view to select data, and the other views to get additional details about that data. For example, you may be trying to better understand a sale – when did it happen, how does it rank, where did it happen?

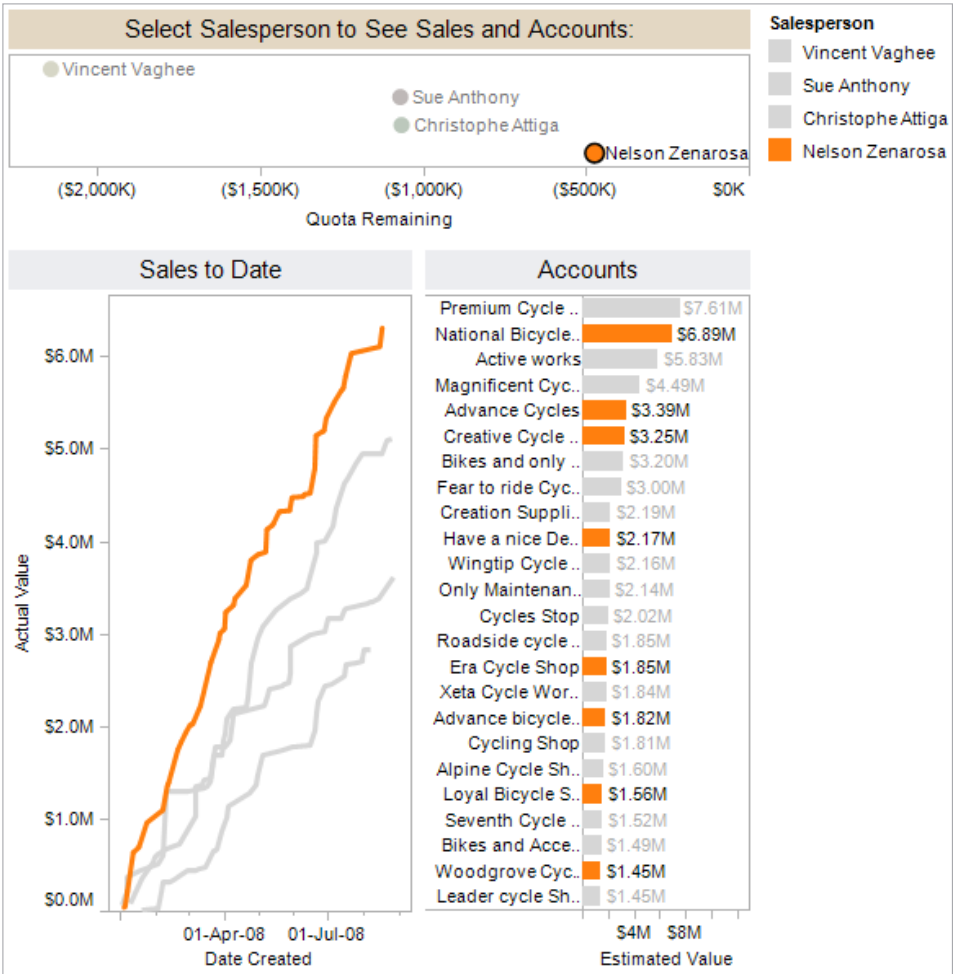
In the dashboard below, the left-most view ranks sales by amount. When you select one sale in the left view (to Starbucks's Coffee), the location of the sale is shown in the top right, and the date of the sale is shown in the bottom right. The other two views simply exist to visually communicate the attributes “when” and “where” about the selected sale. Alternately, you could select a state in the top right view and see all the customers in that state highlighted in the left-most view, with the dates of those sales highlighted in the bottom-right view.



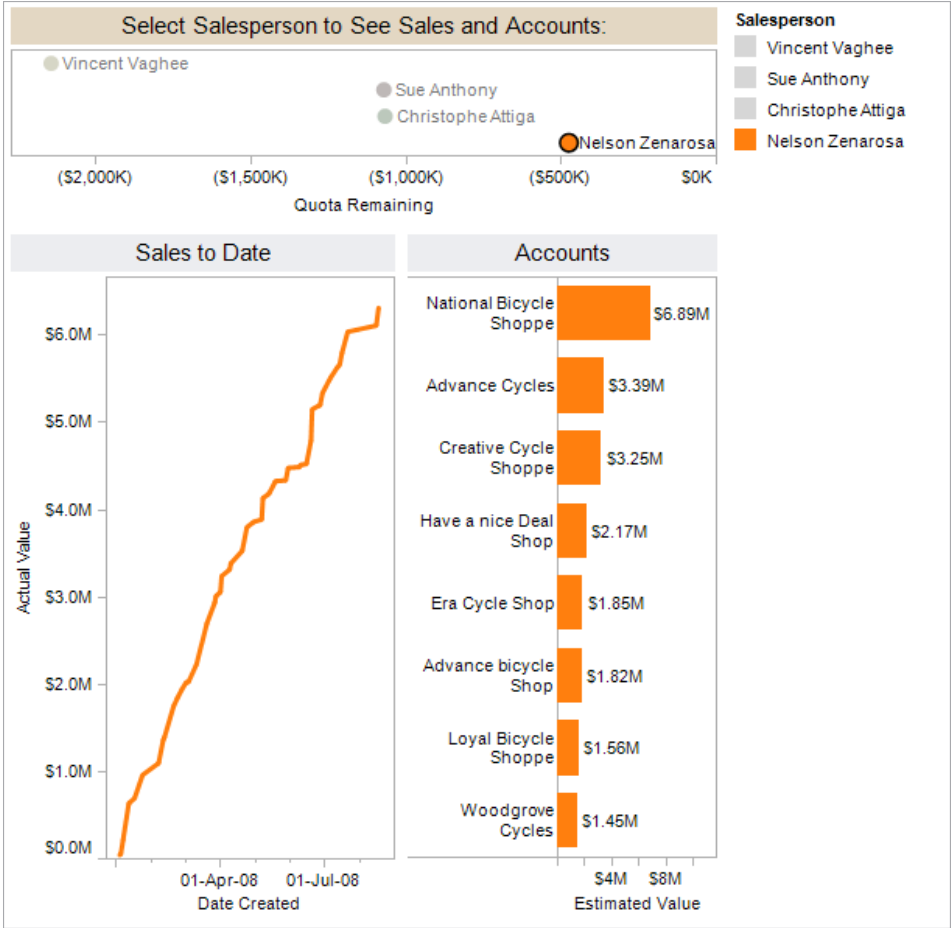
Visual Filtering: What's the next level of detail? And the next?

In all the highlighting examples above, we've discussed presenting related data while keeping all data in the view. Visual filtering lets you focus on only the data related to your selection. In visual filtering you typically select one view as the master and make selections in that view to determine which data is shown in the other views.

The main difference between highlighting and visual filtering is that with highlighting, all the data stays in the view to provide context. With visual filtering, you drill down to specific related data. For example, here's sales data with accounts by salesperson highlighted across the views:

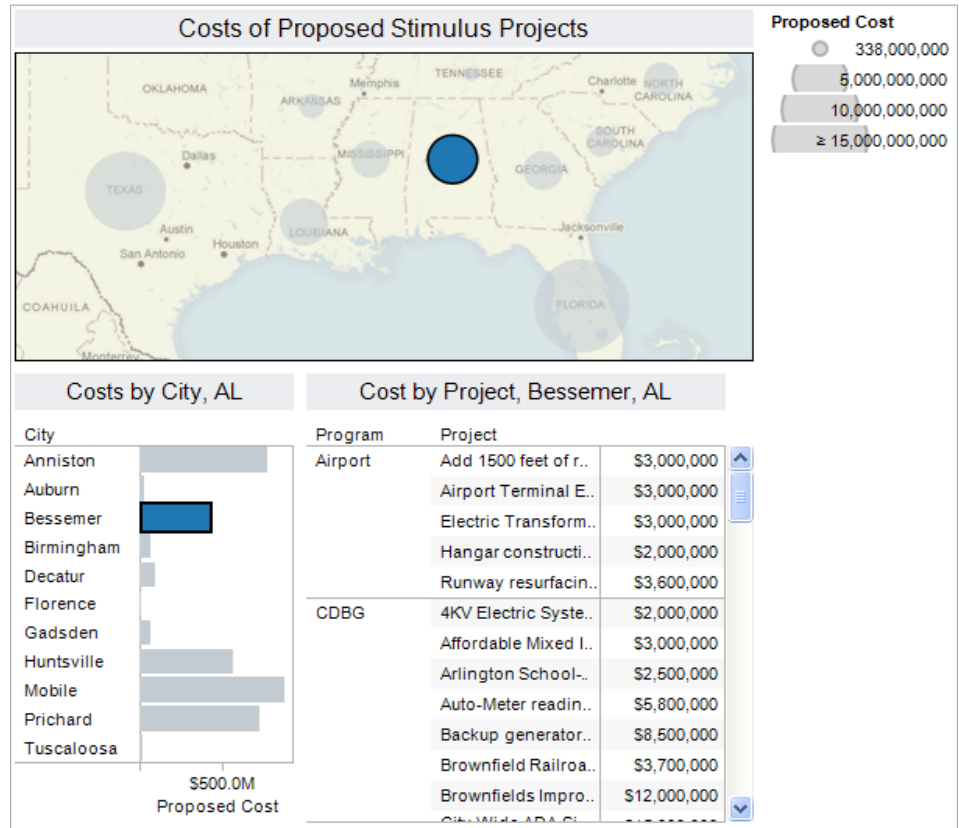


Here's the same view, now filtered by salesperson with the top view used as the master view. Note the difference in the bottom two views, *Sales to Date* and *Accounts*:



With visual filtering we don't see the context of all the other salespeople so it's easier to focus on Nelson's sales and accounts. Discovering unique patterns across data for specific data points is a matter of a few clicks.

Visual filtering can be extremely useful when you need to browse large data sets. One application of filtering is to cascade into many levels of detail. Below is a dashboard of proposed Stimulus projects. When you select Alabama in the top view, the bottom-left view updates to show cities in Alabama and their proposed projects. When you select a city in Alabama (in this case, Bessemer) in the bottom left, the view on the bottom right updates to show you detail about all the proposed projects in Bessemer. With Tableau, you can even define one more level of detail by defining a link out to the project's web page on stimuluswatch.org.



Consider how versatile this makes your views: you can work with thousands of rows of stimulus data and still be able to filter quickly from state to city to project to project detail, all in a few clicks.

Conclusion

Highlighting and visual filtering are simple but powerful tools that help you explore and understand related data in dashboards. While you're asking questions, seeing related data can give you answers in many dimensions: what, when, how, who, where or what type.

With Tableau, you can turn on highlighting or visual filtering with one click. You don't need programming expertise, you only need questions.