Three steps to make your data clearer

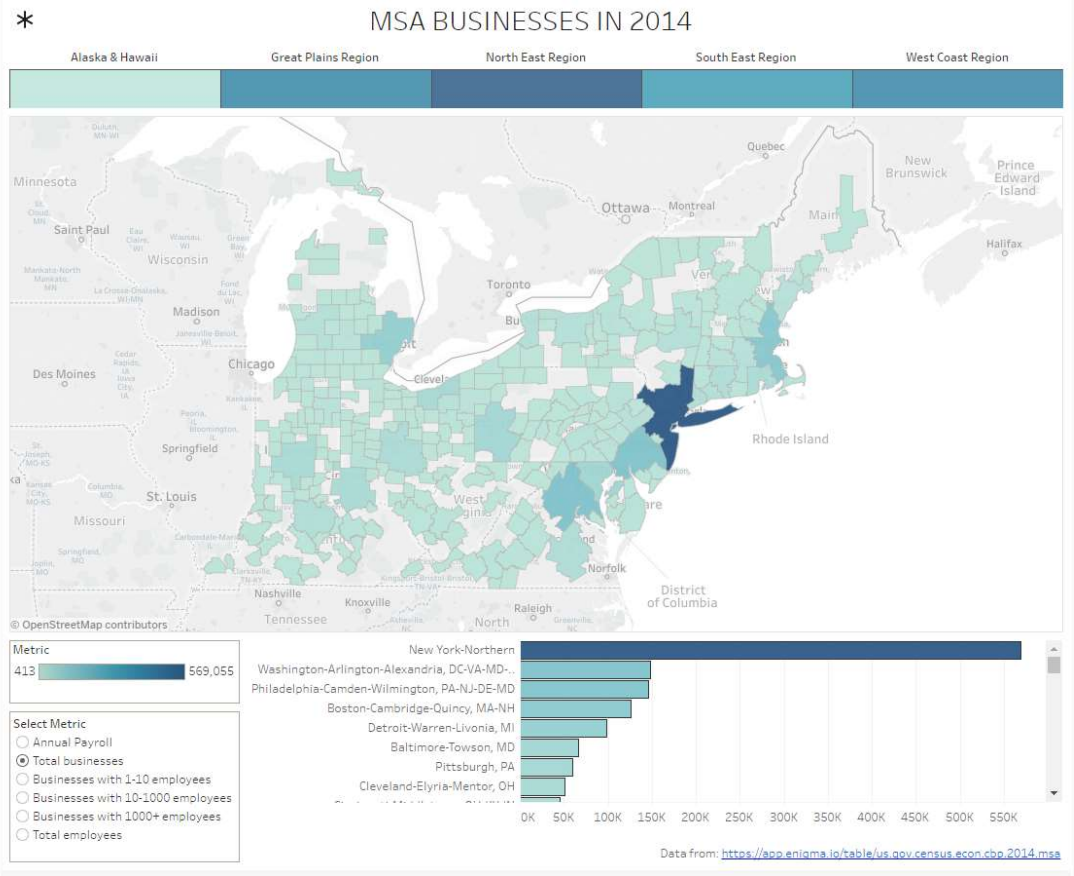
The GovInsider-Tableau guide to data analytics

GOVINSIDER

MAY 2017

Beyond the numbers, charts and graphs, good data visualisations and stories can help the average citizen to be more informed about the world they live in. After all, access to information is empowering.

This is why the United States Department of Commerce collaborated with Tableau, a business analytics software provider, and an industry partner in 2014 to visualise US economic data. With the publicly-available Tableau Public dashboard, people can now quickly answer questions such as, 'How many total businesses are in the Seattle metropolitan area?' (About 90,000) or 'How does that compare to Los Angeles?" (Nearly 250,000 more).



Dashboard of the number of Metropolitan Statistical Areas businesses in 2014.

Data and visual analytics are powerful tools for governments to educate the public on national issues, while at the same time being transparent about policies and informing policymaking. "This is a case of government taking a large data set and making it much more intelligible and useful for average people like you and I, who aren't data scientists," says Ben Jones, Director of Outreach Programmes at Tableau. Jones is currently writing a book on the do's and don'ts of visualising data with the Tableau platform.

From data to stories

Governments everywhere have been embracing the open data movement over the last half decade, but making data free to download online is not enough anymore. A very small percentage of people in any given country knows what to do with .csv or shapefiles, which are the forms that datasets typically take, Jones notes. "The best thing that [governments] can do - while also providing that data to download for people that are advanced data scientists - is create interactive data stories for people like you and me who don't have a degree in data science, but still want to know what's going on."

Here, the challenge lies in creating interactive visualisations and stories that can clearly, effectively and most of all, accurately convey the message to citizens. Jones shared with *GovInsider* some tips for government to create great data visualisations and stories.

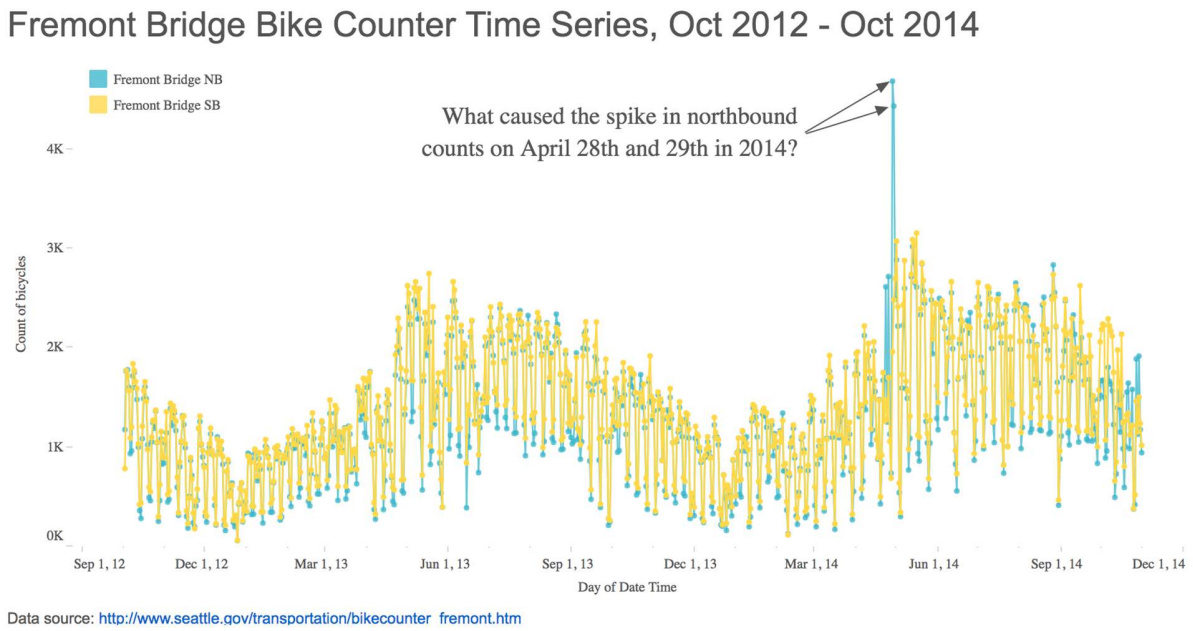
1. Data doesn't always reflect reality

Often, governments make mistakes when visualising their data. Among the most common is the pitfall of "assuming that your data perfectly reflects reality," says Jones.

In Fremont, Seattle, where Jones is based, the Department of Transportation had installed a bicycle counter on both sides of the Fremont Bridge. When Jones

visualised this data over a few years, something immediately jumped out: there were two days in April 2014 when there were almost double the number of bicycles that crossed the bridge on one side, but not the other.

Many interesting theories abounded for this bizarre reading: "People were saying, maybe it was Bike to Work Day or there was a race, and for some reason they didn't ride their bikes back on the other side of the bridge," Jones shared. "Maybe they threw their bikes in the water when they were done?" he added with a laugh.



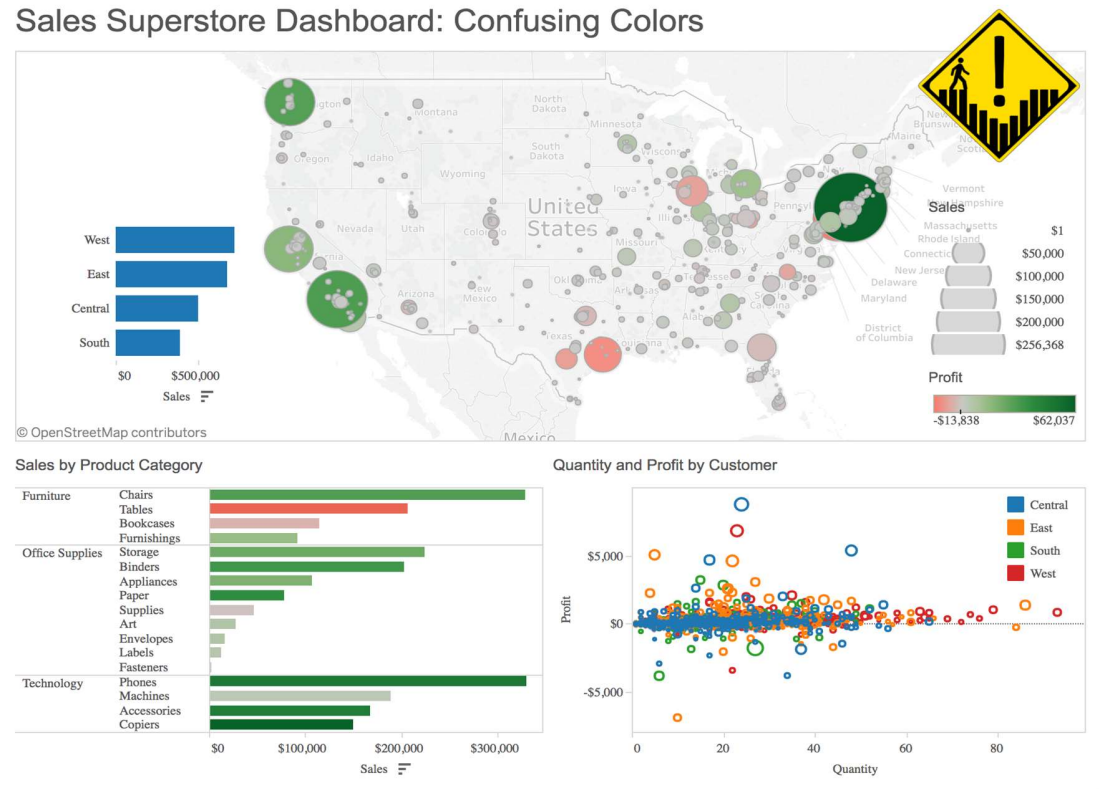
Data visualisation of readings from the Fremont Bridge bike counter from 2012 to 2014.

Jones found out later that it was none of those things: the bicycle counter had in fact malfunctioned on those days, and returned faulty readings. As someone who worked with data all the time, it was an eye-opener for him. "No one stopped to think that maybe something was wrong," he says. "Your data is never a perfect reflection of reality."

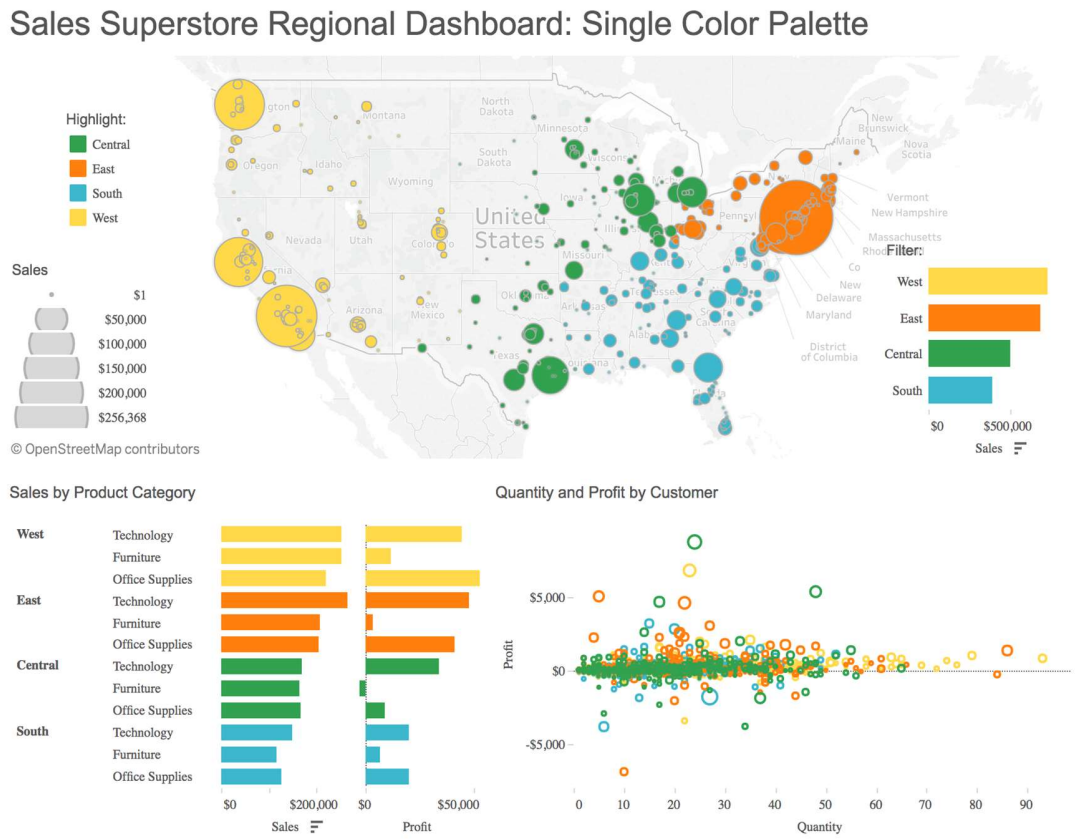
But some good comes out of working with flawed datasets, he says. It forces people and governments to ask new, different questions that they would not have asked before. "That involves you getting away from the data, picking up the phone to talk to somebody, or getting more context from the person who created the data."

2. Don't mislead by using confusing colours

Other pitfalls include the common one of "assuming causation when all you have is correlation", he shares. In particular, he believes that design should come into play as well: colour palettes should be kept as simple as possible, while still achieving the objective of the visualisation or data story. Often, Jones sees visualisations and data stories that make use of two colour schemes, or uses colour in a misleading way - users can't help but make connections where none existed before. "Data people tend to not have been educated in design, so we make these kinds of mistakes," he remarks.



An example of a confusing use of colour.



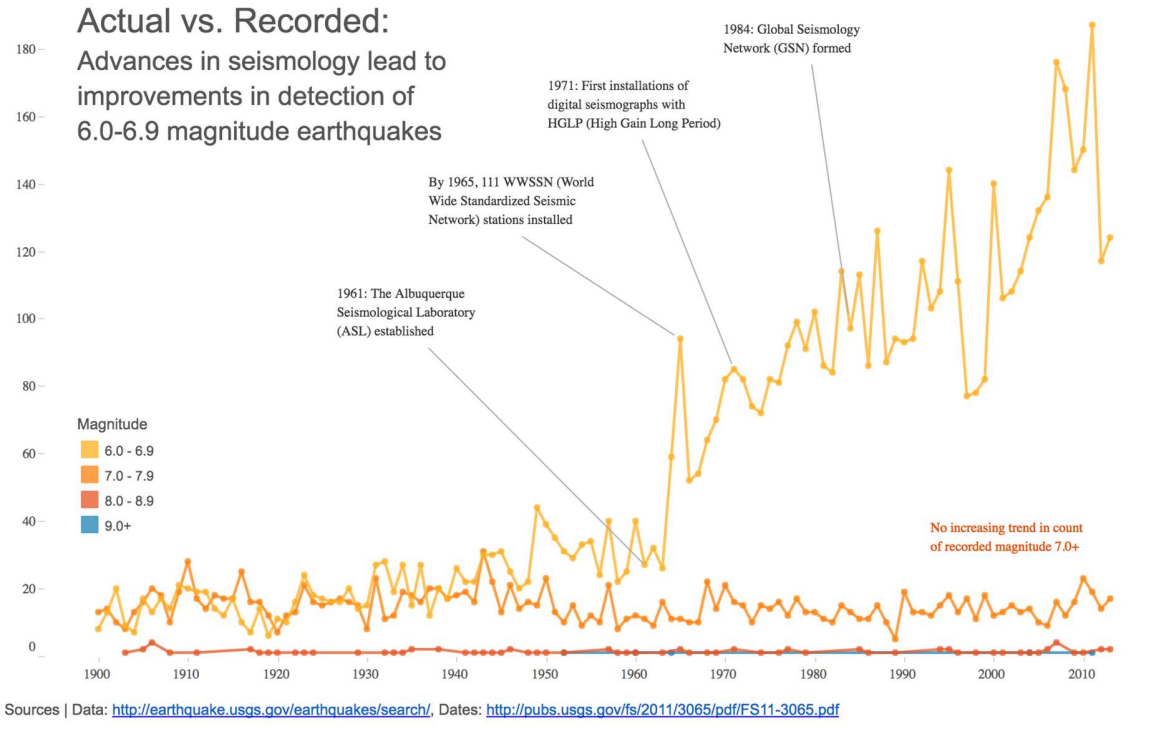
A clearer use of colour.

Ultimately, it's about learning how to communicate well with data, which will go a long way in creating crisp, clear visualisations and data stories. "Regardless of what [data scientists] went to school for, they probably should spend some time learning skills and techniques that help them communicate data better," Jones concludes.

3. Annotate trends

Annotating the data is a very powerful way to help people understand what's happening within the data, Jones continues, adding that "this is journalism best-practice as well".

After the data is visually mapped out, it helps the reader if there is a notation that calls out specific details, outliers or trends. Where there are peaks or valleys, it is worth labelling with a call-out so that viewers have a sense of their significance. However, he is quick to point out that these annotations should not go too far and specify cause and effect, but simply show trends and correlations.



This data visualisation shows how advances in seismology during the 20th century translated into a dramatic increase in the number of recorded earthquakes. It is annotated with important information for the reader's sake.

"This way, [people] are not just exposed to a timeline that they then need to make heads or tails of," Jones explains.

Now more than ever, governments have the tools to help them communicate data better to citizens and colleagues. The time is right to harness these opportunities, to kickstart meaningful conversations on a national level.