Technology Spotlight: Data Visualization

Data is exploding at an incredible rate. According to Forbes, there are ***2.5 quintillion bytes of data created each day*** at our current pace, and that pace is only accelerating. To put that into perspective, if you laid 2.5 quintillion pennies out flat, they would cover the Earth five times. This creates the significant challenge of making sense of such large datasets. We've all heard of Data Scientists, who are able to use mathematical, statistical and computational methods to perform data processing and analysis on this massive amount of data. However, an equally important "last mile" skill is creating visualizations that communicate the insights that data scientists help unearth. HII-TSD provides data visualizations to numerous clients, including the US Postal Service, where we produce custom visualizations using the D3.js JavaScript library toolset, as well as the commercial application, Qlik. Also, on our Veterans Affairs program, we've used Tableau, another popular commercial data visualization product. What we've learned is that producing good data visualizations do not require sophisticated mathematical skills, but rather a solid understanding of your data.

What makes a good data visualization?

Like good design, you know when you see it. And similar to good design, good data visualization doesn't just happen – it requires thought and planning, and knowledge of your data and the information it represents can't be replaced by purely mathematical skills. Understanding the underlying subject matter is key to creating meaningful visualizations, visualizations that tell a story. Fortunately, modern tools democratize the process of creating reports: creating visualizations in Excel, Qlik, or Tableau is not technically difficult, so subject matter experts are less beholden to report developers for new visualizations. So have fun and try some visualizations for your next report or presentation! Here's some things to keep in mind to take your visualizations to the next level:

- *What do you want to communicate?* Before creating a visualization, have a specific message or idea that you want to get across. Then when you create a graphic, go back and ask, "is that clearly communicated"?
- *Think about the medium and the audience.* Is this visualization for a presentation or for self-paced consumption? How quickly does the visualization need to communicate its message? If viewers need to absorb a message quickly, it may be necessary to use multiple graphics. What is the audience's level of familiarity with the information represented by the data? Just because a graphic many not be meaningful to a layman doesn't mean it needs to be redesigned, as long as the audience will understand it.
- *Finally, iterate.* If your first visualization doesn't clearly communicate your message, make changes – remove data not needed for the message, simplify busy graphics, or try different formats. The flexibility of the tools and the ease of generation gives you the opportunity to refine your visualizations quickly and easily.

For further reading, check out some of these articles:

https://www.forbes.com/sites/brentdykes/2016/03/31/data-storytelling-the-essential-data-science-skill-everyone-needs/#3ebd219c52ad
https://news.nationalgeographic.com/2015/09/150922-data-points-visualization-eye-candy-efficiency/
https://hbr.org/2016/06/visualizations-that-really-work