Data Visualization Critique & Practice

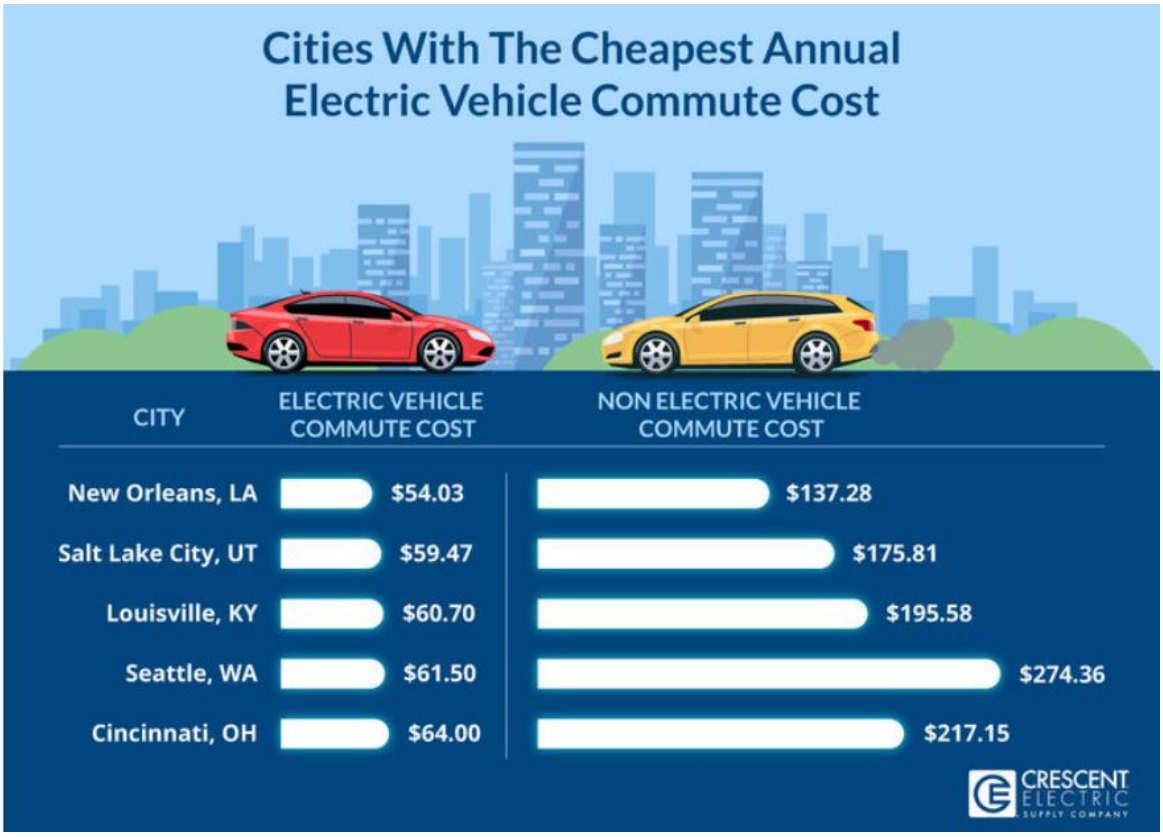
Jennifer Smith

February 24, 2022

Assignment: find an example of a poor data visualization, explain why it breaks best practices, and then fix it with a new visualization.

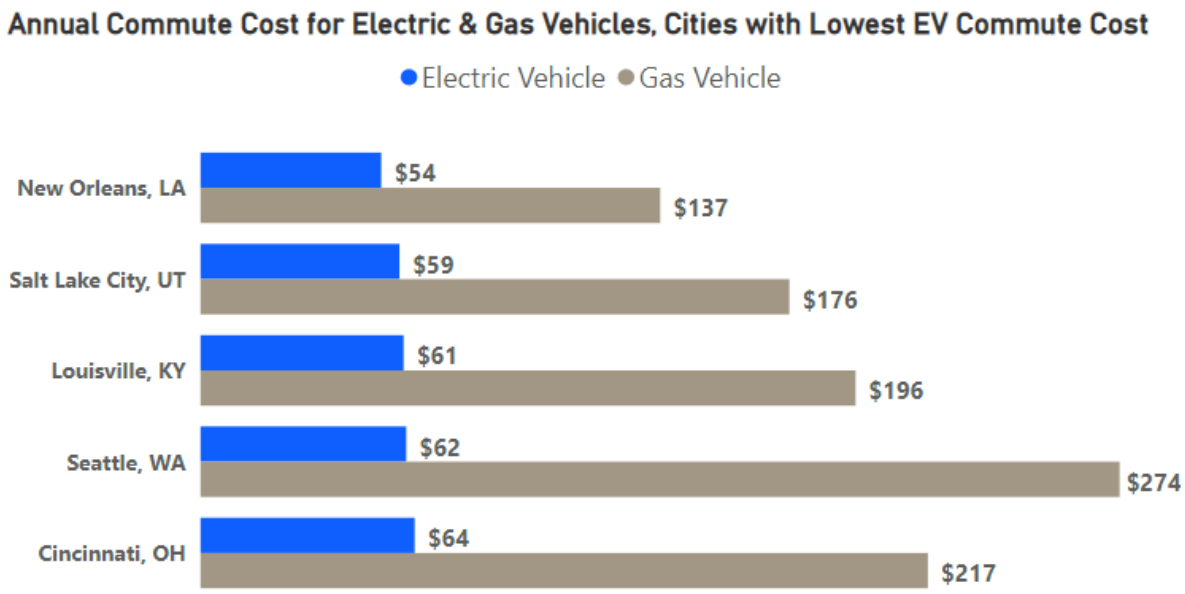
Before (poor viz example)

(from https://blog.cesco.com/blog/electric-cars-vs-gas-cars)



After (re-created with best practices)

(created in Power BI)



To find a poor data visualization, I used Google image search and searched for various topics I find interesting and found that topics around clean energy and electric vehicles are ripe with busy, hard-to-read visualizations. I found several poor visualizations regarding the cost of annual commutes for gas vehicles and electric vehicles on Crescent Electric Supply Company's blog and selected one (on the previous page) that compares the annual commute cost for gas and electric vehicles in the US cities with the lowest annual EV commute cost.

This image breaks several data visualization best practices, the most obvious and egregious being the data-ink ratio. In this idea introduced by Edward Tufte, Stephen Few explains that to create an effective visualization, we should aim to maximize this ratio by reducing the non-data ink on the "page" and enhancing the data ink (Few, 2012). Crescent Electric's visualization breaks this rule to an extreme degree with its colorful background featuring unnecessary graphics and ornamentation, and with the data itself filling less than half the visual area.

Somewhat related to the first rule, another rule this visualization violates is to use soft colors except when highlighting important data (Few, 2005). This design features bright colors just about everywhere *except* the data, which is shown entirely in white. The bars do stand out from the background, but they're all the same color so we don't know what to focus on.

Finally, this visualization unnecessarily features two separate bar charts, which makes it more difficult for us to compare the cost of gas vs electric vehicles in the same city. Humans are good at comparing lengths, but our brains require the objects to be aligned (Velez, 2020).

To fix Crescent Electric's visualization, my original idea was to produce something like a dumbbell dot plot because I thought it could strengthen the story by drawing attention to the difference between the cost of operating a gas vs an electric vehicle. Unfortunately, I couldn't figure out how to create that type of visualization in Power BI, and my Google searches revealed that there may be custom add-ins that could create a dumbbell chart, but I couldn't find a way to create one with Power BI as is.

My second choice was to solve the issues of the original visualization by creating a simple clustered bar chart which, unlike a dumbbell chart, was easy in Power BI. By aligning all the bars on the same left axis, it becomes easier to compare the length (Velez, 2020) and therefore the difference in cost in each city.

Following the instructions of Stephen Few in *Show Me the Numbers*, I aimed to increase the data-ink ratio by eliminating as much non-data ink as possible (2012). I didn't feel the visualization needed axis titles and felt I could also leave out the labels for the x-axis because the data is labeled directly in the chart. I deleted the gridlines, kept the background white, and did not include a border. I also rounded the data to the nearest dollar to make comparisons even easier. I had hoped to eliminate the need for a legend by labeling the series (Electric Vehicle, Gas Vehicle) inside the bars, but I couldn't find a way to do so. I again turned to Google, but once again came up empty-handed, so I begrudgingly had to let the legend stay.

Finally, I followed Few's advice to use muted colors except when highlighting data (Few, 2005). I changed the bars for the gas vehicles to a medium gray and highlighted the electric vehicle bars in a bright blue to draw attention to the significantly lower costs. I think the result is a clean visualization that is much easier to read than the original.

Resources

Few, S. (2012). *Show me the numbers: Designing tables and graphs to enlighten.* Burlingame, Calif: Analytics Press.

Few, S. (2005, November). Uses and Misuses of Color. Retrieved March 29, 2019, from **http://www.perceptualedge.com/articles/dmreview/color.pdf**

Velez, A. (2020, February 19). What is a Bar Chart? Storytelling with data. Retrieved February 16, 2022, from **https://www.storytellingwithdata.com/blog/2020/2/19/what-is-a-bar-chart**