Lance Barto

Data Visualization

Dr. Aaron Gordon

May 4, 2019

Visualization Choices for Storytelling

First, while approaching the final project, it was a struggle to find a subject; there are so many stories to be told which can be accentuated with data visualizations. In the end I have chosen to use some data visualizations to narrate a story about the growing demand for housing in Denver as our population has grown, and how that has affected the cost of both rentals and homes for sale. This is a subject that hits close to home, as I am a Denver native who has struggled to maintain housing in the current market.

The housing market in Denver, much like most housing markets, is primarily influenced by the invisible hand of the free market. As our population has grown, so has the demand for housing, and while construction of new homes in the area has been booming; the available housing has not been able to keep up. This has driven costs of housing through the roof. For my first plot I wanted to express how much the population has grown, while also highlighting which specific areas have grown the most, and also to provide a sense of what areas have the largest contributions to the overall population make-up.

To express these ideas, I chose a stacked bar graph. While this, like all of my plots, is relatively simple I think that it does an excellent job of showing the root cause of a rise in demand for our housing. When dealing with time-series data, one typically doesn’t have a wide variety of choices to present the data. The choices that were primarily available in my mind were a stacked bar chart, a line graph, or a stacked area graph. There are many pitfalls in using stacked area graphs, and because of that I chose to steer clear of using that type of plot. The main pitfalls of stacked area charts are that they can be deceiving, the primary reason being that a user can struggle to understand where each layer is rooted. I didn’t feel as if I could effectively show each individual contribution to the overall population with a simple line chart, and because of these issues I chose a stacked bar graph. This plot allows the user to interact through the use of the plotly and ggplot2 packages designed for the R programming language. If I would have used Python and matplotlib to create this plot, I would have been able to achieve similar results. However, the overhead to would have been much greater and I don’t think the plots would have been as nice as the results achieved by this collection of tools. The source of the data used in this plot was the Metro Denver Economic Development Corporation.

For my second and third plot, I chose line charts. Each of the plots show time series data, and a line graph is a common choice to display this type of data. The nature of the story which I’ve decided to show is how things have changed over time; it was important to me to focus on pushing the narrative, even if that translated to making the decision to use relatively simple plots to further the story. Some designers choose to use plots which are more complicated, and while I appreciate the ambition of such decisions it’s not uncommon to see the narrative get lost in the process. This ‘less-is-more’ approach is what I’ve seen to be the most effective, and I made the decision to emulate that idea.

The second plot shows only a single line, but that single line shows what an incredible rise in the prices of purchased and appraised home values. Each data point represents a quarterly home price index value which was sourced from the Federal Housing Finance Agency. Once again, I have employed the plotly package to create hover over labels for each data point, allowing the user to see the exact index value and which quarter it is related to. This plot is very simple, but it tells a striking story about the incredible rise in our housing costs; a single line plot with such an extreme increase tells a striking story in my opinion, and that is why I’ve made the decision to utilize a plot which is so simple.

In the next plot, I again chose to use a simple line graph with some interactivity applied using the plotly package. This graph looks at the growth in median and mean rents, both in terms of the entire nation, the entire state, and the metro Denver area. I’ve chosen an interactive line graph here, because the nature of my data is represented as time-series, my choices were limited. Although it is simple the data does an effective job of furthering the narrative being told, and clear displays the incredible growth in housing costs, and how much faster the values are growing in Denver and Colorado when compared to the national average. The data used for this plot was curated from the United States Census Bureau website.

The final two plots were chosen to represent the growth of wages in Denver. I chose two plots to represent this data, the first being a line chart. I used a line chart for similar reasons as my other plots that were implemented previously, but also, because I wanted to show that the rates of wage increase are not nearly as steep as the increases in housing costs. I also wanted to present this data in a way that was potentially more dynamic, and attention-grabbing. To achieve this, I animated the data using a bar chart. Here, we can see small differences in the rates of growth for each region or measure of center being used. I think this interpretation gives a more straight-forward feeling of growth, but without a line graph it would be difficult to see exactly how the rates of growth differ when making a comparison to the housing visualizations.

In conclusion, I’d largely say that I’ve chosen a route of simplicity to further the narrative being told. In general, that approach aligns with my holistic ideas about data visualization; I have often found some of the most effective visualizations to progress a narrative are those which are the simplest. I have also tried to take care of many small details to create effective visualizations throughout the narrative. These include: consistency of formatting, consistency of color palettes, being sure to create interactivity when possible to allow users to understand as much of the story as possible, clear legends, clear separations in the use of red/green color combinations (and the ability to interact with the legend to isolate if there are any color perception issues), clearly labeled axes, and clear introductions to the graph with the text.

References:

1. Federal Housing Finance Agency -

<https://www.fhfa.gov/DataTools/Downloads/Pages/House-Price-Index.aspx>

1. Metro Denver Economic Development Corporation –

<http://www.metrodenver.org/do-business/demographics/population/>

1. United States Census Bureau: Housing -

<https://www.census.gov/topics/housing.html>

1. Bureau of Labor and Statistics: OES data

<https://www.bls.gov/oes/tables.htm>

1. Ward, M. (2015). *Interactive data visualization: Foundations, techniques, and applications*. Boca Raton, FL: CRC Press.
2. Tufte, E. R. (2015). *The visual display of quantitative information*. Cheshire (Connecticut): Graphics Press.