

DSI: Data Visualization

Assignment 4

Garrett Bullivant

April 3, 2023

Dataset link:

<https://open.toronto.ca/dataset/apartment-building-evaluation/>

Data visualization 1 was created using R and RStudio. The intended audience for visual 1 are Toronto residents/prospective residents who are interested in renting or purchasing apartments in Toronto. My data visualization is supposed to convey that the quality of Toronto apartment buildings varies based on the year that the building was built - specifically that building quality appears to increase with newer apartment builds. Furthermore, I've split the data according to the property type. Splitting the data according to property type shows that private apartments particularly seem to improve with newer apartment builds. I considered aesthetic principles in this visualization – I wanted my plot to be visually appealing (Healy, 2018). I was able to accomplish this by keeping the plot relatively simple – I used a minimalistic theme on R. I also used the jitter argument on R to make the data points more visible. I also tried to be substantive in this visualization – I tried to accurately display the data (Healy, 2018). I did this by plotting all the data points in a scatter plot and not manipulating the data points or axes. Finally, I tried to consider perceptual design principles – I wanted to clearly convey the message of my plot (Healy, 2018). Specifically, I added linear trend lines to my plot that helped highlight the changes occurring in apartment quality over time. I was able to ensure that this data visualization is reproducible by using a programmable language – specifically R (Zogheib, 2023a). I have tried to make this visualization accessible by using the 'Virdis' package in R to make the plot accessible to those with colourblindness (Zogheib, 2023b). Furthermore, I have created an alt text for this plot that uses level 2 and 3 descriptive content to summarize the plot (Zogheib, 2023b). This addition makes the plot accessible to those who are unable to see the image. Finally, I have made sure the font used on this plot is sans serif and the appropriate size to ensure that it is most legible (Zogheib, 2023b). The individuals who might be impacted by this visualization include those who are interested in renting or buying apartments in Toronto and the current tenants of the apartments being analyzed. There were several different features included in this dataset, but I decided to only include the year built, apartment score, and property type in this visualization. I believe that the 3 variables I chose were some of the most intuitive features to include and demonstrate how apartment builds vary according to the year they were built. 'Underwater labour' that contributed to this final data visualization is the work done by the unnamed bylaw enforcement officers who conduct the building evaluations. Furthermore, the staff at each of the apartment buildings who comply with these inspections are contributing to this visualization without any recognition.

Data visualization 2 was created using GraphPad Prism 9, which is a commercial scientific 2D graphing and statistics software. The intended audience for visual 2 are Toronto residents/prospective residents who are interested in renting or purchasing apartments in Toronto. My data visualization is supposed to demonstrate that the quality

of Toronto apartment buildings can vary based on the Toronto ward that they exist in. The message I'm trying to convey using this bar plot is that some wards tend to have a higher apartment score than others. I considered aesthetic principles in my visualization – I wanted my plot to be visually appealing (Healy, 2018). I was able to accomplish this by keeping the plot relatively simple. I decided not to include all the data points and instead used standard deviation to make the visualization less busy. Furthermore, I kept the plot aesthetic by simply using 2-D visualizations (Healy, 2018). I also tried to be substantive in this visualization – I tried to accurately display the data (Healy, 2018). By including the standard deviation, we can see that there is a lot of variability in this visualization. By being forthcoming with this data variability our audience will know that there is less certainty about the differences in apartment quality between Toronto wards. Finally, I tried to consider perceptual design principles – I wanted to clearly convey the message of my plot (Healy, 2018). Specifically, I decided to order the Toronto wards in decreasing apartment building score – this makes it easiest to identify the wards with the highest and lowest apartment building scores. Since I did not produce this visualization programmatically it is not reproducible, which makes our data visualization less trustworthy and decreases the persuasiveness (Zogheib, 2023b). Specifically, according to an idea called provenance rhetoric – the inclusion of a data source signals “transparency and trustworthiness” to the audience (Zogheib, 2023b). I have tried to make this visualization accessible by using the ‘Viridis’ colour palette in Prism to make the plot accessible to those with colourblindness (Zogheib, 2023b). Furthermore, I have created an alt text for this plot that uses level 2 and 3 descriptive content to summarize the plot (Zogheib, 2023b). This addition makes the plot accessible to those who are unable to see the image. Finally, I have made sure the font used on this plot is sans serif and the appropriate size to ensure that it is most legible (Zogheib, 2023b). The individuals who might be impacted by this visualization include those who are interested in renting or buying apartments in Toronto and the current tenants of the apartments being analyzed. There were several different features included in this dataset, but I decided to focus on the Toronto wards and apartment building score for this visualization. I believe that the 2 variables chosen were some of the most intuitive features to include and demonstrate how apartment builds vary according to the ward they exist in. ‘Underwater labour’ that contributed to this final data visualization is the work done by the unnamed bylaw enforcement officers who conduct the building evaluations. Furthermore, the staff at each of the apartment buildings who comply with these inspections are contributing to this visualization without any recognition.

References

Healy, K. (2018). *Data Visualization: A Practical Introduction*. Princeton University Press.

Zogheib, C. (2023a). *First Steps: Reproducible Data Visualization*. [PDF].

https://github.com/UofT-DSI/07-visualization/blob/main/lessons/DSI_DataViz_2.2_Reproducible.pdf

Zogheib, C. (2023b). *Visualization with Purpose: Accessible Data Visualization*. [PDF].

https://github.com/UofT-DSI/07-visualization/blob/main/lessons/DSI_DataViz_5_AccessibleDataVisualization.pdf