GEORGETOWN UNIVERSITY
PUBLIC POLICY 646
DATA VISUALIZATION
ASSIGNMENT 3: TWO ORIGINAL VISUALIZATIONS
Prof. Wesley Joe

For this assignment, please create two original data visualizations from your choice of data.[1] You are free to choose the data and kinds of graphs. At least one of the graphs should include categorical data. You do not need to use the same data set for each of the two graphs. Please create both of the graphs using ggplot2 in R, and refine each graph in Adobe Illustrator.

For each visualization, please begin by writing a brief paragraph that explains the purpose of the graph. (This does not need to be more than a few sentences.) Do you want viewers to explore the data to answer their own questions? Or do you want the visualization to support a particular point that you want to make, perhaps to reinforce something that you explain in accompanying written narrative? Do you want to emphasize trends, particular values, or both?

Then, create your visualization. **At least one** of these visualizations should encode the values of at least three variables. For example, remember that creating "small multiples" using values of a categorical variable for each multiple constitutes encoding of a variable. You can choose the types of visualization, such as a matrix of bivariate plots, a heat map, a scatterplot that plots points in more than one color or uses text to encode identity information of each plot point, a scatterplot that encodes points as both text and colors, a matrix grouped density plots or grouped histograms, or some other type of visualization.

To demonstrate your growing range of capabilities, each graph should consist of a separate type of plot. In other words, please avoid submitting something like two separate frequency polygon visualizations.

Your visualization should also include enough contextual information (such as titles, unit labels, axis labels, and so on) for your audience to do what you intend for them to do with the visualization. In the interest of creating cleaner, more efficient graphs, try to minimize the presence of "non-data ink" in your graphs.

Also, each visualization should include annotation. In other words, help your audience zero in on your key takeaway point(s) by including at least an annotative sentence title and a sentence

[1] To be clear, an "original visualization" is *not* a replication of a graph that was created by someone else, such the authors of a World Bank report.

or two of subordinate annotation. Finally, each visualization should be polished and aesthetically pleasing to a professional audience. When creating the annotation, assume that your audience will have access *only* to the visualization. In other words, assume that your audience will *not* have access to the "statement of purpose" paragraph that you write for me. (That paragraph helps me to evaluate how well your graph fulfills your intentions for it.) Hence the data source credit, for example, should appear in the graph.

Finally, a word about sharing code and appropriating code examples from the Internet. Learning to write code from scratch is an important learning objective of this course. You'll meet this objective only by writing your own code. Consequently, for assignments in this class, students cannot obtain code from any current or former McCourt student. You can, however, use online resources, such as StackOverflow or other StackExchange sites, RStudio Community, and other the like. If you use code from such sites, even as a model, please explicitly cite the source that you used. Please include the citation in the "statement of purpose" associated with the visualization for which you used the site. Any shared or recycled code that is used, even as a model, without attribution will be considered a violation of the academic honesty guidelines.

Please turn in a single PDF file that contains the following pages:

- a brief statement of the purpose of the first original visualization
- the first original visualization
- your R code for the first original visualization
- a brief statement of the purpose of the second original visualization
- the second original visualization
- your R code for the second original visualization

As before, PDFs that have pages out of this requested order will forfeit 0.2 points on a 4-point scale. Thanks for your understanding! ☺

Please submit the PDF document using Gradescope.