GEORGETOWN UNIVERSITY
PUBLIC POLICY 646
DATA VISUALIZATION
ASSIGNMENT 4: TWO ORIGINAL VISUALIZATIONS
Prof. Wesley Joe

For this assignment, please create two original data visualizations from your choice of data.[1] You do not need to use the same data set for each of the graphs, but you can. Please create both of the graphs using R and refine each graph in Adobe Illustrator.

For each visualization, please begin by writing a brief "statement of purpose" 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 specific associations? Patterns? Trends? Exceptions?

Then, create your visualizations. Use a different idiom (such as a heat map, line plot, etc.) for each visualization. Remember that stronger visualizations provide higher degrees of cognitive amplification. Multivariate visualizations often help readers more than do univariate or bivariate graphs. For purposes of assessment, though, note that "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.

Your visualization should include enough contextual information (such as titles, legends, unit labels, axis labels, and so on) for your audience to easily 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 or two of subordinate annotation. Finally, each visualization should be polished and aesthetically pleasing

[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.

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.

Please use Adobe Illustrator to add the annotation, indicate the source of the data, and otherwise refine each of your visualizations. Enforce principles of visual hierarchy throughout the visualizations. For each text element (for example the title, subhead, annotation, axis and tic mark labels, data source credit, etc.), the size/scale and style (for example plain, bold, italic) of the type should reflect the element's relative importance in a visual hierarchy. Also, try incorporating some of the refinement recommendations in chapter 5 of Berinato's *Good Charts*.

Please turn in a single PDF file that contains six 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

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.

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! ☺