GEORGETOWN UNIVERSITY
PUBLIC POLICY 646: DATA VISUALIZATION
PORFOLIO PROPOSAL/PLANNING DOCUMENT
Spring 2023
Prof. Wesley Joe

Your first assignment is to prepare a proposal/planning document for your data visualizations portfolio. Your final portfolio is the culmination of your work for this course. The portfolio will consist of revised original visualizations (that is, not the replications) that you submitted for your interim visualization assignments or your Data Story assignment for this class.

The final portfolio should contain between 10 and 12 original, polished visualizations that address the public policy issue (e.g. renewable energy, K-12 education, criminal justice) discussed in your portfolio proposal. Each visualization and the data story should be refined and complete. The portfolio must include at least 8 original visualizations created in R. The entire portfolio, except for any interactive Tableau visualization(s), should be submitted as a single PDF document. Your portfolio must include one interactive visualization created in Tableau.

Your proposal must:

1. Describe your main topic. This should be a public policy topic that will serve as the focus of your portfolio. The topic should be broad enough to permit a wide variety of types of visualizations (e.g., time series, proportions, relationships between categorical and quantitative variables, sets of quantitative variables) but not so broad as to result in an incoherent final product.

2. Identify at least three easily accessible data sources. (If possible, please include URL links to each source.)

Choose data come from reputable sources. (The aphorism "garbage in, garbage out" applies to data visualization.) Examples of such sources include, but are not limited to, government agencies; major, credible nongovernmental organizations (such as the World Bank, the Pew Research Center); universities or university-based research centers; and so on. Try to find at least 60-70 variables. Data sets with at least a few hundred cases provide more flexibility. But some smaller data sets (e.g., for regional maps) can work well for a few assignments.

Once you find promising data sources, find *usable* data sets. Usable data sets have proportionally few missing values. Additionally, choose survey data sets that require the use of sample weights only if you are familiar with working with such weights. If you decide to work with survey data, try to find at least one or two additional data sources that do not require the use of sampling weights. Your complete set of data should include at least some categorical and quantitative variables in both cross-section and time series. You should also have some data that are suitable for geospatial visualizations that include a geographic unit that Tableau recognizes. A list of such units is available here. You might also want to try to find some data that shows movement of cases from one status or location to another. For example, migration of residents from one state to another state. The structure of such a data table would resemble that of the US Census Bureau's "State-to-State Migration Flows" tables, available here.

Finally, avoid collecting original data or relying on data that has been "promised" to you. These approaches almost invariably lead to *major* regrets!

3. A preliminary plan for your portfolio contents. Provide a series of initial expectations of the visualizations your portfolio will include and the kinds of insights that one could reasonably expect the visualizations to show. For example, what kinds of relationships would a visualization explore between variables or between entities (such as human beings or nations)?

Your proposal should not specify the idioms (e.g., scatterplot, small multiples of bar charts, a chord diagram) of each visualization because we will learn about new types throughout the semester.

Please keep your proposal brief; I recommend a length of between 300 and 500 words. Please submit your proposal as a PDF file to Gradescope, a cloud-based grading workflow management platform that we'll use this semester.