GEORGETOWN UNIVERSITY
PUBLIC POLICY 646
DATA VISUALIZATION
ASSIGNMENT 2
Prof. Wesley Joe

This assignment requires you to create two data visualizations. The first project is to replicate a visualization from Nathan Yau's book *Visualize This*. The second is to create an original visualization from a data set of your choosing.

A. Please create a replication of graph 6-15 in Nathan Yau's book *Visualize This*.

The .tsv extension in the file name means that this is a TAB-DELIMITED text data file. We have not yet discussed how to use the read_csv() function to import this data. So I'll just give you the code required to obtain the data:

```
library(readr)
crimeRatesByState2005 <-
  read_delim("http://datasets.flowingdata.com/crimeRatesByState2005.tsv",
                                "\t", escape_double = FALSE, trim_ws = TRUE)
View(crimeRatesByState2005)
```

You do not need to create a perfect replica of Yau's graph. Just try to make something that is reasonably similar. *You do not need to recreate the legend that appears in Yau's graph*. (He created this legend in base R using code that appears in his book.) You can just rely on a legend from ggplot. Also, you do not need to match the type font family that Yau used. You will learn how to modify type font families, sizes, and so on in the Adobe Illustrator training.

After you finish the plot, export the plot from RStudio Desktop into a PDF file. In the plots window there is a button named "Export." Click on the "Export" button, and select "Save as PDF. You can refine the graph in Illustrator. Note: You should select the "Landscape" orientation for the page, as shown next:
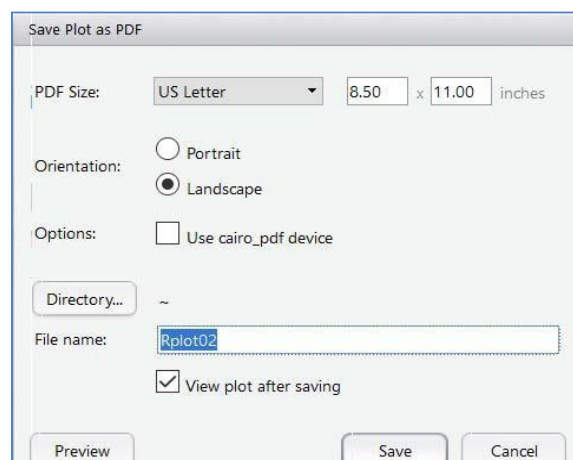
B. Using a data set that you have found, create an original data visualization. Begin by writing a <u>brief</u> paragraph that explains the purpose of your visualization. (This does not need to be more than a few sentences.) Do you want viewers to explore the data to answer their own questions? If so, what kinds of 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 the accompanying written narrative? Do you want to emphasize trends or particular values?

Then, create your visualization. Keep in mind the expressiveness and effectiveness criteria for evaluating data visualizations. You can choose the type of plot. You are also free to encode the data into visual variables (shapes, colors, etc.) of your choice. Your selection of visualization type and variables and values should advance the purpose of your visualization as described in your paragraph.

Your visualization should also include enough contextual information (such as titles, unit labels, axis labels, a data source credit, and so on) for your audience to use the visualization for the purposes you intend. You must also report the source of your data.

Finally, your visualization should be aesthetically pleasing to a professional audience.

Please turn in <u>a single PDF file</u> that contains the following pages:

- Page 1: your replication of Yau's visualization.
- Page 2: your brief statement of the purpose of your original visualization.
- Page 3: your original visualization.
- Page 4: your R code for your original visualization

<u>Please make sure that the pages are in the specified order. PDFs that are out of the required order will lose 0.2 points on a 0-4 point grading scale</u>

Please submit the PDF file to Gradescope.