STAT 451, Fall 2019

Final Exam (Take home)

Due Date: 5pm on Wednesday, 11 December 2019

Format: Digital submission in PDF **and** Markdown formats (Markdown should have all the code and the PDF should be the output). I will be testing all of your code so make sure it works!

Everything else: You are encouraged to complete this project in groups of up to two people (2 = max). Each person should contribute about half of the work. Each member of the group needs to submit the project with both names on it.

Assignment:

1) When using color in visualizations there are different kinds of pallets that can be used.
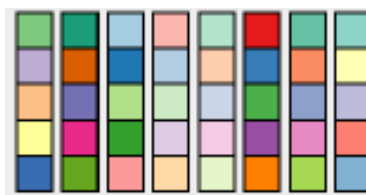
a. Describe the pallet and explain when you would use such a pallet.

b. Describe the pallet and explain when you would use such a pallet.

c. Describe the pallet and explain when you would use such a pallet.

2) Here is the link to the USGS website where the worldwide earthquake data can be downloaded. Download all earthquake data for the past 30 days in .csv format. Using R, make a map of the world with points where the earthquakes occurred. Make a bubble map using the magnitude. Thoroughly discuss your visualizations.

3) The data can be from any source of your choice. Include a description of the data and its source. You solution should be narrative that tells a story that is supported by visualizations.

Here are some suggestions about deciding on a dataset and making your visualizations.

Suggestions:

a.   Find a website that uses data visualization to communicate information to readers of the website. The data should be related to a topic of interest to you. For example, if you are interested in Economic Data you could use the Bureau of Labor Statistics (BLS) website. If you are interested in Health Data you could use the Centers for Disease Control and Prevention (CDC) website. If you are interested in Baseball data you could use the blog Visual-Baseball Project website.
b.   It is okay to try to reproduce a visualization(s) you have found, but they should not be the only visualization(s).
However, you need to give the source of the data and cite the original visualization. For example, if you use the BLS > Publications > The Economics Daily (TED) or the Spotlight on Statistics website, you might try to reproduce of the following: The Recession or Employment in political organizations or State unemployment. (Note that the BLS uses HighCharts, an open-source JavaScript software.) If you choose this path, clearly state that you will be reproducing and modifying visualizations.
c.   Make at least 4 plots using the data you selected. These plots could be for the same data, that is, using various visualizations to look at the different variables in the data. Or you could use different variables and at least 3 variables in each plot.
d.   Tell a story about your data and your visualizations. Describe the kind of data you are working with, the source, and what is being communicated in your visualizations.
e.   No screenshots.