Rubric for HUDM 5026 Projects

This is an exploratory data analysis project with summary statistics and graphical displays. That said, you should have some motivating research questions as you begin your exploration of the data.

Paper details:

- Write in R Markdown
- Include references in APA style in a section at the end
- Upload your work as a knitted .pdf or .html document

Introduction & literature review (20 points)

- 1-2 pages, most of which should be literature review
- There are two types of papers you would want to review here: (1) papers which have used your dataset, and (2) papers which looked at the same variables to answer questions similar yours. If you use a data set from PLOS ONE, you will already have one paper to discuss.
- When you review research, focus on two things: (1) what did they find? and (2) how does it relate to and inform my motivating questions?
- If you are working with a data set from a particular paper, try to replicate some of the fundamental findings from the paper.

Methods and Sample (30 points)

- 1-2 pages, go more in depth in this paper than you would if you were writing a research paper for another class
- Describe your sample (20 points), who are they? **You must present a table of descriptive statistics**, or you will lose points. Suppose you are working with six variables for this project. In that case, I expect to see descriptive information about these variables in the methods part so that the scale and measurement of these variables is understood. Generally, you want to report at least the mean and SD of numeric variables (I also think the min and max are very useful, as they help find data errors / give confidence that there are not issues with your variables). And for categorical variables, give frequencies and percentages (Descriptive data here should focus on the raw categorical variable, not the dummy variable)
- Describe the graphical and statistical methods you will use for this project (10 points). For example, if you use boxplots, you should note that the boxplot provides a summary of the first, second, and third quartiles and the min and the max and outliers etc. and explain how to interpret the plot. Also describe the cleaning process. How did you handle missing data? How did you organize data and what did you do to preprocess data to get it ready for analysis?

Findings (30 points)

- 1-2 pages
- Here is where you will show off your visualizations and statistical summaries
- Focus your reporting and interpretation on only the things that are relevant to your research question!
- The most successful projects will incorporate multiple and varied aspects of the coding techniques we cover in class.

Discussion (10 points)

- 1-2 pages
- Relate your findings back to your literature review, what did you find that is useful for the scientific community.
- How do your analyses inform your motivating research questions?
- All research papers have weaknesses, what are yours? You can think of weaknesses of regression, or maybe of your data set, like was an important variable missing?

Organization, grammar, and flow (10 points)

- Make sure your paper is free of grammatical errors and is organized well
- Using hyperlinks and table of contents in Markdown is a good idea.