**Project Instructions for Part I**

*Please exchange with your group by 11:59 PM on March 13th.*

*Please turn in a revised draft to me on Sakai by 11:59 pm on March 19th.*

***Assignment Details***

For your final project, you will analyze your dataset in R and discuss your results. Specifically, you will run a linear regression model (or several) to test expectations you make about the patterns you expect to observe in the data. Then, after conducting the analyses in R, you will discuss your findings. You should include an RMD file with the code you use, an Excel/CSV file with your final data that you use to run the models and a knitted html file with your project write-up. It is important to write in a manner that is clear, concise, and statistically correct. We will spend time in class discussing best practices for statistical writing.

This portion of the project will focus on project description, variable introduction, and data visualization.

***For the first part of the project, you should include the following parts (55 points):***

* **Before Chunk One:** An introduction where you introduce your topic to the reader and briefly discuss your expectations based upon your research. Find 2-3 existing studies to cite here. (estimate: approx. 300-400 words- these may vary; 10 points)
* **Chunk One:** A discussion your variables of interest. (approx. 100-150 words, 15 points)
  + After this chunk, please discuss measures of **central tendency** and **spread** for your main dependent and independent variables.
* **Chunk Two**: A visualization of your variables of interest. (30 points, 20 for data visualization, 10 for data visualization description.)
  + Construct **at least three** graphs in R that visualize your data effectively for the reader. These can be any of the plots we have learned in class or others (not pie charts) that you believe effectively visualize the data for the reader.
  + After this chunk, please **also** provide additional context to help the reader better understand the visualizations. (approx. 100-150 words)

***You should also include (15 points):***

* An RMD file with all the code you used to conduct the analysis and plenty of comments so that the replicator can see what you did.
* Please also include your data as an Excel file or CSV when you upload on Sakai.

***Logistics (15 points, 10 for writing quality and 5 for other logistics)***

Your final draft section should:

* incorporate the feedback from the proposal
* include an RMD file, Excel/CSV file with your data, and a knitted html file
* be between 1800 and 2100 words long (hard cap at 2500 words); please cite any works you include and include a word count with your submission
* be well-written; [here](https://www1.cmc.edu/pages/faculty/JPitney/writing.htm) is a link to a guide to writing that Professor Jack Pitney of Claremont McKenna College designed.
* use a consistent citation style of your choice throughout

***Peer Review Component***

You will be participating in a peer review workshop with members of your lab group (more details at the start of the week of 3/8). 15 points of your grade will be based upon making a good faith effort to participate in this workshop. You will be receiving an email from your TA about how to show engagement and should follow the directions in that email.

***Rationale***

This assignment will bring together the statistical concepts that we have learned over the course of the semester thus far and give you practice honing your statistical writing skills. You will also have the opportunity to revise your project based on my feedback on your exam question and your group’s suggestions on the project.

***Rubric***

* 55 points: Project write-up sections (point breakdown above).
* 15 points: Inclusion of RMD file with comments and Excel/CSV file.
* 15 points: Engagement in the peer review process
* 10 points: Writing quality and the overall readability and flow of the project.
* 5 points: Following of other above logistics.