**Project Instructions**

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

*Please turn in a revised draft to me on Sakai by 11:59 pm on April 20th.*

*Please note: the exchange is happening before we get to multiple regression, if you only have up through the start of chunk four done at the time you exchange, that is fine.*

*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. You should revise the first part of your project that you already turned in and received feedback on.

Your project write-up should include:

* A revision of the part of the project that you already turned in. (15 points)
  + **Before Chunk One:** An introduction where you introduce your topic to the reader and briefly discuss your expectations based upon your research (estimate: approx. 300-400 words- these may vary)
  + **Chunk One:** A discussion your variables of interest. (approx. 100-150 words)
  + 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.
    - 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 provide additional context to help the reader better understand the visualizations. (approx. 100-150 words)
* New sections (40 points):
  + **Chunk Three:** Conduct a simple hypothesis test (e.g., difference of means test, Chi-Square Test). After this chunk, please discuss and interpret your results. (approx. 200-250 words, 10 points.)
  + **Between Chunks Three and Four:** Discuss the control variables you will include in your regression model. You should include at least three control variables beyond the focal independent variable(s) in the model. Then, please tell the reader what evidence you need to support your expectations. (approx. 300 words, 5 points)
  + **Chunk Four:** Please run your model. After running this model, please discuss your results. Did you find evidence in support of your expectations? In discussing your results, please interpret the coefficients and p-values and discuss your R-squared value. Also, please discuss whether there are any influential points or collinear predictors. (approx. 500-600 words, 20 points)
  + **After Chunk Four:** Write a brief conclusion where you discuss the implications of your results and any next steps you would take in the analysis. (approx. 150-200 words, 5 points)

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

*Rationale*

This assignment will bring together the statistical concepts that we have learned over the course of the semester 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 proposal and your group’s suggestions on the project.

*Rubric*

* 15: points: revision of previous sections
* 40 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.