PSTAT 126 Final Project Details

Final Project Report Due Date: Sunday, June 9, 10pm

About the Final Project

The final project is an opportunity to put all that you have learned in PSTAT 126 to the test. Your task is to conduct a start-to-finish regression analysis of a real dataset. This will require all of the skills we have learned over the quarter: from plotting the data, exploring regression models to answer scientific questions, and presenting the results in a clear, concise and engaging manner. This is a **group project**, where groups can be **no larger than 3 people**.

You are given two options for your final project. Whenever you work on the project you should write down what you have done, and save any `R` code you may generate along the way. You are required to write your final report using `R` markdown, including relevant code, summaries, and plots. However, unlike a homework assignment, the report should be written in narrative format, clearly explaining to the reader what analysis you are doing, why you are doing it, and what the results tell you.

Structure of the Final Report

Each group is submit one report, which should be prepared using `R` markdown and must be structured as follows:

1. **Title Page:** Must include the title of your project, and the names of all group members.
2. **Abstract (100 words):** Brief summary of your project, including what questions are investigated and your findings.
3. **Problem and Motivation (200 words):** In this section you should describe (i) the relevant background of the data set and (ii) the motivation for your project i.e., why your readers should be interested.
4. **Data:** Briefly describe the data set and relevant variables that you will use in your project.
5. **Questions of Interest:** Describe in plain English the questions that your analyses will answer. Scientific, not statistical, terminology should be used here. For example, words like 'association', 'effect', or 'relationship' are okay while 'p-value', 'coefficient,' or 'regression' are not.
6. **Regression Methods:** Briefly outline (in a logical order) the different regression methods that you will use in the project to answer the questions of interest.
7. **Regression Analysis, Results and Interpretation:** In separate subsections, you should carry out the analysis for each part of your chosen project. Your narrative should include:
    - Code: You are writing this report in `R` markdown, so relevant code and its output should be included as part of the report.
    - Important Details of the Analysis: If you did a hypothesis test, then state your null and alternative hypothesis, the value of your test-statistic, p-value and your decision. Provide similar detail for confidence intervals, submodel tests, ANOVA tables, etc. **You can pull this information from `R` output, but it should be stated in the text so the reader doesn't have to go looking for it.**

- Plot: All exploratory and diagnostic plots should be shown. Plots should be readable, but should not take up an entire page.

8. **Conclusions (200 words):** Summary of your findings and any comments you may have about the reliability or generalizability of your analysis.