**Final Data Exploration Project**

Data Exploration 5101 - Fall 2024

**Due**: Week 14, December 12th, 2024

The most important aspect of a well-written final project is the reader will easily understand the subject matter regardless of their prior knowledge on the subject. Obtain data from the internet, your company, or other unique source about something you are passionate about.

Be specific in all sections of your report, do not assume the reader knows what you are talking about. Proofread! The following is a detailed guide of what is expected for submission:

**Title Page**

* Creative title for your final project (Do not name the project, “Final Project”)
* Name/Due Date/Class/Instructors

**Abstract**

* Collective summary of entire final project – shows a data scientist whether they should read the entire paper.
* By itself on the first page of the final project.
* Contains the most important 1-3 sentences from each section below. 100-250 words. As a collective summary, it may be best to save the abstract for last after you have completed the rest of the report. Major conclusions are discussed in the abstract, including p-values.

**Part 1**. **Introduction**/**Project problem statement** (5% of final grade) - A document that becomes the Introduction to your Final paper. The problem statement should be *at most* two pages in length that at a minimum clearly addresses the following questions:

1. Introduce your data. What data are you exploring?
   * What are your sources of data?
   * What background knowledge is needed to understand your data? Background information (ABSTRACT SECTION) should be obtained from the meta data from your data source and peer-reviewed journal articles.
     + Locating a journal article:
       - Go to Google Scholar (<https://scholar.google.com/>)
       - If you are having trouble accessing articles from home, use a Stockton computer or connect your computer to the Stockton WiFi on campus and you will have access to them.
   * Define your terms. Anything you refer to during the rest of the project should be introduced in the introduction section.
2. What is your research question?
   * Why is it important, interesting (or difficult to answer)?
3. What would you do if the data does not support your expectation?
4. Clearly state your null hypothesis and/or your alternate hypothesis.
   * What did you expect to find out?

Any corrections or suggestions by instructors after submission of the problem statement should be reflected in the final report.

**Part 2**. **Final** **Data Exploration** document (25% of final grade) - .

1. Methodology
   * How was your data obtained?
   * What materials/computer software did you use?
   * How was your data cleaned? Was any data omitted from your study?
   * What calculations or statistical analysis did you do?

Are your methods clearly explained so that someone with no prior knowledge on the subject can understand your thought process?

1. Results
   * Presentation quality graphs, as appropriate
   * ALL graphs should be labeled with a title, axis titles, and units as needed.
   * Label them: **Figure 1:** with a caption**.**
   * Do not include raw data in your results.
   * Do not discuss the results here, but instead we simply just state what they were with graphs, tables, analysis, math, etc. We can provide a statement of what analysis shows, but do not discuss it. Point out the key details.
2. Discussion
   * Start to discuss your results here.
   * Restate your null and/or alternative hypothesis. Do your results support or reject your hypothesis? Why? Be specific.
   * Use appropriate interpretation of statistics and graphs that support the research question
   * Include at least one journal article in your discussion:
     + Go to Google Scholar (<https://scholar.google.com/>)
     + If you are having trouble accessing articles from home, use a Stockton computer or connect your computer to the Stockton WiFi on campus and you will have access to them.
   * Include any limitations or sources of error in your project
     + Was there insufficient data? Did your data not directly address the question?
     + Any other limitations with R?
3. Conclusion or Future Directions
   * What is a future direction to explore this data more?
   * Include a summary of the answer to your research question.
4. Appendix A, which is the R code used to analyze the data and make the graphs
5. References
   * Include at least two peer-reviewed journal articles and ANY other sources you used.

**Presentation**

**Option 1:** A brief class presentation (maximum **two** minutes) on the problem or question you addressed and your findings. You should have **one or two** graphics to accompany your presentation. Your presentation should demonstrate the following:

* Ability to describe succinctly your data question and your findings to a general audience
* Ability to answer audience questions clearly

**Option 2:** Stockton University hosts a [Graduate Research Symposium](https://stockton.edu/graduate/research.html#PastSymposiaGuidelines1-d19e90) during every Fall and Spring semester. The symposium is Thursday, December 5th from 4pm-6pm on Stockton’s Galloway campus. The abstract is due [**November 7th**](https://stockton.qualtrics.com/jfe/form/SV_bkCKiQZMNkzZrF4). If you present a poster or other presentation at the Graduate Research Symposium, then you do not have to present for the class on December 12th.

**\***Confirm with your professors before submitting an abstract**\***

At a minimum, the report should be well-written and include all of the above. The final project is **30%** of the course grade.

The deliverables are:

* Your report in a PDF file uploaded to Blackboard with the filename “<YourLastName>\_Project.pdf”.
* Your data
* Your final poster presentation in .PDF format (If you are presenting at the Graduate Research Symposium)

**Reminder**: You can send us your final project at any time (Not the day it is due!) to look it over for suggestions and pointers before it is graded.