BSTA 630 Project 2022 Fall

Each student will choose one to work on from the four datasets provided by the instructor. You will determine which of the techniques introduced in the course are appropriate for describing the data and answering the scientific questions of interest for the project you decide to work on. Each student will be required to submit a written report at the end of the semester.

Introduction (Please limit this step to no more than **1 page**)

- Describe the background, motivation, and justification to conduct the current study.
- Determine one or more specific scientific questions of interest.

Data Description (Limit to no more than **1 page**, excluding tables and figures)

- Describe the data collection and study design.
- Choose and describe the primary and secondary outcomes (at least one continuous and one binary/categorical) and other relevant variables.
- Consider the relationship among the selected variables.
- Evaluate the potential distribution of your outcome variables, using **graphical methods** and **summary statistics**. If there is any missing data, they should be described in this section.
- Comment on normal and binomial/multinomial distribution assumptions.

Point Estimates and Confidence Intervals (Limit to no more than **1 page**)

- Obtain point estimates for population parameters of interest.
- Obtain confidence intervals for those parameters, using both parametric (e.g., asymptotic) and nonparametric (e.g., bootstrap) methods. Indicate which interval you used. Interpret your results.

Hypothesis Testing (Limit to no more than **1 page**)

- Translate the scientific questions of your interest into null and alternative hypotheses.
- Choose appropriate approaches for hypothesis testing. Justify your choice of test and discuss potential issues that you might have in making your choice.
- Carry out the test and interpret your results. Check assumptions if applicable.

Linear Model (Limit to no more than **2 pages,** excluding figures and tables)

- Construct linear model(s) to answer ONE OR MORE of the research questions of interest.
- Justify the procedures or reasons for model and variable selection.
- Fit the model and interpret parameter estimates and confidence intervals if applicable.
- Do model diagnosis using graphical and statistical tools and discuss appropriate remedy procedures if needed.

Categorical Data Analysis (Limit to no more than **2 pages,** excluding figures and tables)

- Perform simple association test and interpret the results.
- Construct a simple generalized linear regression model.

Future Study Planning (Limit to no more than **1 page**)

- Imagine a further study that extends the questions of interest from your current study.
- Use the information from the current study to plan how many samples you will need for the new study.

Conclusion and Discussion (Limit to no more than **1 page**)

- Summarize the findings of your study.
- Discuss the strength and weakness of the study design.
- Discuss the potential finding impact and the possible study extensions.

Final report is due on Dec 16, 2022

- Submit your report on Canvas (**1.5 or double spacing, font size 12**).