STAT 383: Probability and Statistics

Statistical Research Project Due: 13 December, 2019 11:59 PM

1 Problem Statement

As an adult human being, I suspect that there are questions you might have about life, the universe, and everything. Or, at the very least, I expect that you have some questions about some things that you would like to answer. Your first task for your project for STAT 383 will be to brainstorm a few ***specific*** and ***related*** questions that you could answer by using statistical analyses on attainable data. The questions you are trying to answer should be of interest to a broader audience than just yourself.

I have found a nice repository provided by Eastern Michigan's library that has links to freely available data sources (open source data sets!) that you could use to help you answer your questions. You may find these at `https://guides.emich.edu/data/free-data`. These include data pertaining to the economy, politics, health, climate, energy. Some are specifically sorted by the types of analyses you might perform! It is by no means an exhaustive collection. For instance, there is not much in the way of sports statistics provided. Additionally, you may have collected data through your own research or in some other class that may be worthy of consideration.

Come up with (at least) three testable questions that you could answer using statistical analysis. Again, they need to be related to the same general topic. You should try to incorporate several different types of analyses that we have looked at in the class (and possibly some that we have not) in answering these questions. For instance, you might want to use a t-test to answer one question, an ANOVA procedure for another, and a regression to answer the third.

The last component to the project is to write up a report that summarizes your statistical findings. The report must contain writing to explain: what you did, why you did it, and what you discovered; a mathematical summary of your statistical processes; and some figures and tables that help illustrate your findings.

2 General Report Guidelines

- Your report should have a cover page that includes each person's name and student ID. You may work with a partner or on your own.
- Your report should consist of an Introduction, Methods, Data, Analysis, and Conclusion. The Introduction should provide an overview of the topic. What is the topic? What is already known and established? Why might someone be interested in answering the questions you are considering? The Introduction provides background information and motivates the research questions. Citations should be given where appropriate. The Methods should detail what statistical analyses will be performed and explain how each method could be used to answer your research questions. Your hypotheses should be explicitly stated in the Methods section. The Data section should be a relatively brief description of what data you have obtained. The Analysis section is where you present your findings, including tests, statistics, p-values, confidence intervals, figures, etc. It should include some brief interpretations and descriptions as well. The Conclusion section is where you state your overall conclusions and draw attention to potential issues with your analysis as well as indicating what future work might be conducted (Was there other/better data that you didn't have access to? Are there other methods you could have used? Etc.).
- This is a report in which you are expected to describe the data and your results. Do not narrate your activities. I don't need you to describe how you make a particular calculation. That should be clear from the Methods section in your formulae there.

- All images, graphs, and tables should be labeled and captioned. They should be discussed and described in the text. I should be able to figure out what useful information is contained within a figure or table. I should not have to stare at it until I figure out what is important. It needs to be clear.
- Spelling, grammar, and punctuation are all important. Your grade will be penalized for basic writing mistakes. This is an exercise in presenting scientific work professionally. It will factor into the grade calculation.
- Formulas and equations should be formatted properly. Use an equation editor. Do not write something like "xˆ2+yˆ2=rˆ2" or put an image of an equation in your report. This looks tacky and unprofessional.

3 Grading Rubric

1. Cover Page
   - Date, Name(s), ID Number(s), Other relevant info? (3 points)
2. Introduction
   - Does it provide relevant background information and sufficient motivation? (5 points)
3. Methods
   - Are hypotheses to research questions clearly stated? (5 points)
   - Are at least three DIFFERENT methods outlined? (5 points)
   - Are methods suitable for addressing each research question? (5 points)
   - Are relevant formulae given for each method? (3 points)
4. Analysis
   - Are all methods given in the previous section used appropriately/accurately for data? (10 points)
   - Does the paper include at least three graphical depictions of the data? (10 points)
   - Are the results correct? (10 points)
   - Are descriptions and explanations regarding the results provided? (10 points)
5. Conclusion
   - Are each of the research questions clearly addressed? (5 points)
   - Is attention drawn to any questions that might be raised based on results? (2 points)
   - Is possible future work specified? (2 points)
6. Transitions, Grammar, and Linguistic Style
   - Does the narrative flow smoothly? Are there abrupt changes from one topic to another? Are sections clearly labeled? (5 points)
   - Is the grammar and punctuation in the paper correct? (5 points)
   - Are words spelled correctly? (5 points)
7. Notations and Mathematical Style
   - Are figures appropriately labeled and captioned? (3 points)
   - Are equations numbered and properly punctuated? They should be part of a sentence! (3 points)
   - Are equations properly formatted? (4 points)