

Final Project (Stage 1)

Stat140-04

Overview

Pick a dataset, and analyze/visualize the data in a creative and insightful way.

You will

- write a paper about ONE research question you find interesting and use data visualization and analysis to support your exploration.
- create a poster and do a 4-minute presentation to communicate the potential of your interested research area to the rest of the class.

Important dates

- The paper is due on Gradescope on Friday Nov 20 6PM EST.
- The poster is due on Gradescope on Sunday Nov 22 6PM EST.
- Presentation in class on Zoom on Monday and Tuesday.

Only 30% of the projects will be selected to go on to the next stage.

1. Proposal Paper

Write a paper addressing ONE research question of interest. The research question can involve either one or two variables. Variables can be categorical or quantitative.

Your paper will be comprised of the following parts:

PART 1: Introduction What is your research question? Why do you care?

PART 2: Data Introduce the dataset you are working with. What are the observational units? What is/are the variable(s)? What is the sample? What is the population? How was the data collected? Are there any bias in the sampling method? What are the implications of the data collection method for inference?

PART 3: Exploratory Data Analysis Perform relevant descriptive statistics, including summary statistics and visualization of the data. Make sure to interpret these results in your own words – don't just give output.

PART 4: Conclusion What can a layperson conclude from your findings?

Instructions

Write up parts 1-4 in the up to 3 pages in length, including all figures. Points will be taken off for longer papers. Include each of parts 1-4 as separate designated sections. Your report should not include any R scripts, include your R script as an appendix. The appendix is not counted in the 3 pages.

Write as if you are explaining your results to whoever would be interested in your research question, whether this is other scholars in your field or peers sharing your interest in the topic. Keep in mind this audience may or may not have taken statistics. You may assume familiarity with basic descriptive statistics and visualizations. You must be statistically accurate and use correct statistical terminology, but must also explain your conclusions in a way that anyone can understand.

Three pages is not a very long paper. You will need to be concise. Every sentence should add something to your paper.

2. Poster Presentation

It is up to your discretion as to what kind of material you would like to put in the poster, but you are required to list at least three other interesting research questions that can be explored with your data set. Use meaningful visualizations, text, and/or other relevant multimedia content to illustrate why should we care about these research questions.

You will have 4-minute in class to guide us read through your poster. The goal of the presentation is to allow you to communicate to the class your findings and why your research project is worth continuing. Think this in-class presentation as a grant panel – you want to pursue the instructor to select your projects and convince your fellow classmates to join your research team.

Instructions

Make a horizontal poster in Latex, Powerpoint, Keynote, etc. Points will be taken off for incorrectly sized poster. Your poster should not include any R scripts.

Honor Code

The first stage of the final project is to be done individually. Think of this as a week-long take-home exam where you have the chance to design the questions yourself and express your understanding about this world through data. Your instructor and TAs are not allowed to help you debug code, review your draft etc. You cannot share or discuss any aspect of the project with anyone else. Do not copy other people's projects on the internet. Failure to abide by this policy will result in a 0 for everyone involved, and will be treated as a case of academic misconduct.

Grading rubrics

I will use the following rubrics to grade your report.

(https://docs.google.com/spreadsheets/d/19IYI0NKS7iFpJAWN4hAkp2T3m5LZRF4llhv_101J_n0/edit?usp=sharing)