Final Project

STAT 011 with Prof Suzy

Spring '23

1 Overview

Your final project will consist of a poster presentation, shared and presented during the last week of classes. The topic of your final project is open-ended so feel free to choose something that interests you. Your project and poster presentation will be completed in groups of 3-4 people and should broadly fit into one of the three categories below:

1.1 Option A: Research report on a topic that interests you

This could be a topic we have touched on in class, or something you learned about in another class or through your own reading or experience. Such a project might involve reading articles in journals, newspapers, magazines, and books; summarizing and discussing the statistical issues involved; and drawing conclusions. Examples of possible topics include the use of DNA fingerprinting in court cases, the use of statistics in standardized testing, the use of election polls in the media, the role of statistics and data science in epidemiology.

1.2 Option B: Analysis of an existing data set

You might apply some exploratory or inferential technique we have learned in class to a data set that has already been collected and is publicaly available. You can find data sets on the web using sites such as www.census.gov and www.city-data.com for demographic data, www.leesmovieinfo.net for movie earnings, etc., and use it to answer some question you're interested in. The CDC also has extensive data sets on the Coronavirus at https://www.cdc.gov/DataStatistics/.

1.3 Option C: Original survey or experiment

You might conduct a virtual survey on some topic that interests you (e.g. twitter poll, email poll, Swarthmore poll/experiment). Such a project might involve developing a survey or set of questions, selecting a strategically random sample, and gathering the data to analyze. You'll have to identify and define your variables of interest before collecting data. If you are going to ask your peers to participate, be considerate and make the activity as quick and painless as possible.

2 Requirements/Rubric

You and your group mates will create a poster summarizing your project and will present this to Prof Suzy on one of the last two days of class. You should use large print on your poster and aim for an attractive and well-organized display. Your poster and presentation must reflect you and your group's own original work. Plagiarism will be checked with software and a strong indication of plagiarized content will result in a grade of zero for all group members.

Please submit your final project idea/topic to Prof Suzy's for approval by the **end of March**. You are encouraged to consider connecting your knowledge of statistics to the topics of another course in your project topic. For projects (B) or (C), the type of data you use should involve at least one quantitative variable and one categorical variable. Typically people doing surveys will collect 5–8 variables on each subject.

Your final project will be graded on a scale of 0-20 points possible. All group members will receive the same grade *unless* the majority of the group members bring a specific complaint to Prof Suzy about a member not participating. In this case, a student who is not participating may receive a lower grade than their group mates.

Rubric	Component	Points Available	Description
Poster	Title	0-1	Is the title short, catchy, and relevant to the topic?
	Technical specifications	0-1	Is the poster neatly organized and clearly legible? Was the topic approved on time? Was the poster printed on time? Does the poster contain sections for Background, Methodology, Results, and Discussion? Does the poster include at least one relevant graphic?

Rubric	Component	Points Available	Description
	Main content	0-5	How clear is the statement of question or purpose? Can a reader understand what data (or information) was analyzed? Are the methods adequate to address the main statement of question or purpose? Are the conclusions over- or under- stated?
	References	0-5	Are references cited properly at the bottom of the poster? Are they referenced throughout the poster, where appropriate?
Presentation	Duration	0-2	Is the presentation finished within a 3-5 minute window?
	Participation	0-1	Did all group members contribute to the presentation and/or poster?
	Use of poster	0-5	Is the poster referenced during the presentation? Are presenter(s)' comments consistent with the information on the poster?

3 Example Topics

3.1 Option A: Relate Statistics to Another Class

Have you heard any mention of important statistical studies in your other courses? Would you like to dig into a particular application of statistics to, say, a humanities course? Can you find a reliable source and can you explain some of the statistical components?

3.2 Option B: Stat Problem Solving Data Challenge

Professor Suzy Thornton will accompany interested students through an exploration of this year's Computing, Government, and Graphics sections of the American Statistical Association's annual Data Challenge Expo.

Where: Math/Stat Lounge

When: 5-8pm on March 20 and 21 RSVP for one or both days here

The challenge data set we'll be working with is the National Crime Victimization Survey (NCVS) of 2020. The NCVS has gathered data on personal and household crimes since 1973. The primary goals of the survey are to collect information about victims, to explore the consequences of crime, and to estimate the number and types of crimes that go unreported.

You are welcome to additionally (or alternatively) consider related data sets:

- NCVS Series and Supplements data
- Annual Survey of Jails 2020
- National Prisoner Statistics
- American Community Survey

In addition to the data for download, the NCVS website has links to studies that utilize the NCVS data. These studies may serve as inspiration or background for your final project.

3.3 Option C: Design and analyze an experiment

You can be as creative as our practical constraints allow in designing your own experiment. The minimal requirements will be to ensure there is a treatment with at least two levels and at least one numerical variable and to strategically incorporate randomness in subject selection and/or treatment assignment. A few examples could be things like:

- How much of a beverage is lost to the packaging? (E.g. Compare the left over amounts of liquid in cans, bottles, squeeze pouches, etc amounts after the product is consumed.)
-
-

3.4 Option C: Conduct and analyze a sample survey

You can be as creative as our practical constraints allow in designing your own survey. The minimal requirements will be to ensure responses to your survey provide information on at least one categorical variable and one numeric variables and to strategically incorporate randomness in subject selection. A few examples could be things like:

-
-
-