

# Math 140 Project 1 Instructions

## Spring 2020 / Teachout

### Project 1

**Collect Data**  
**Create Excel Spread Sheet**  
**Bias Essay**

#### Collecting your Data

Collect data yourself. If it is a group project, then each member of your group needs to collect data and should ask the same two questions. This cannot be data found on the web. Your data needs to have one categorical question with only two possible categorical answers and one quantitative question where the answer is a number. Do not pick a quantitative question where the answer is always zero. For each person or object ask both questions. For example, we might ask people the following two questions. "Do you smoke any type of cigarettes, yes or no?" (Categorical question with only two possible answers.) "How old are you, in years?" (Quantitative question where the answer is a number.)

Data Collecting Method: You can use any data collecting method. It is ok to use put up a voluntary response survey on social media or ask people in your class or at your work conveniently. It is ok if the data is biased. Part of your project is to analyze how bad your data is.

Sample Size: Keep collecting data until you have at least 30 for each of your categorical variables. (For example, if your categorical question was "Do you smoke any type of cigarettes, yes or no?", then you should have 30 or more people that said they smoke and 30 or more people that do not smoke.)

Note: This does not mean that you have to have 30 exactly. The actual data will probably be 31 smokers and 93 non-smokers. More data is better!!

#### Create Excel Spread Sheet

Put the categorical data into one column of an excel spreadsheet. Put the corresponding quantitative data the next column next to the categorical data. Do not put units in quantitative columns. This will mess up computer programs trying to analyze the data. Put the units in the title at the top of the quantitative columns. Be careful to not mess up the order of the values. Remember the columns came from the same people or objects. Now separate the quantitative data into groups. The category makes up the two groups. For example smokers and non-smokers. If the quantitative variable is age in years, put the ages of the smokers in the third column and the ages of the non-smokers in the fourth column. When you have only two groups (like smokers and nonsmokers), you should have a total of four columns of data in your excel spreadsheet.

Column 1: Raw categorical data (smoke or not)  
Column 2: Raw quantitative data (age in years)  
Column 3: Separated Quantitative Data (age in years for smokers)  
Column 4: Separated Quantitative Data (age in years for non-smokers)

#### Essay

You will write an essay on how you collected the data. What was your population of interest? Your data cannot be a census. You must have a larger population of interest in mind. Describe the process of collecting your data. Was the data random? What method did you use? Explain all 5 sources of bias (Question, Response, Non-Response, Deliberate and Sampling Bias) and explain why your data does or does not have that source of bias. How well you think this data will represent the population of interest? Most essays are two or three pages single spaced.

#### Turning in your project

- For project 1, you will need to turn in your bias essay as well as a printout of your entire excel spreadsheet.
- If it is a group project, make sure that the first and last names of everyone in your group are on both the bias essay and the excel spreadsheet printout.
- See homework schedule for due date.

## Project 1 Grading Rubric

- Collect Categorical & Quantitative Data (35% of grade)
- Essay (40%)
  - Population of Interest (5%)
  - Describe how data was collected and method (5%)
  - Sampling Bias (explain why it is or is not in your data) (5%)
  - Question Bias (explain why it is or is not in your data) (5%)
  - Response Bias (explain why it is or is not in your data) (5%)
  - Deliberate Bias (explain why it is or is not in your data) (5%)
  - Non-response Bias (explain why it is or is not in your data) (5%)
  - Explain why your data does or does not represent the population of interest. (5%)
- Excel Spread Sheet (25%)
  - Separating Quantitative Data by Group (5%)
  - Column 1 (Original categorical data) (5%)
  - Column 2 (Original quantitative data) (5%)
  - Column 3 (Separated quantitative data for group 1) (5%)
  - Column 4 (Separated quantitative data for group 2) (5%)