

Project #3: Chapter 12 Data and Statistics

Instructions:

Statisticians spend a lot of their time analyzing and reporting data. This project asks you to find two sets of data, and to formulate a formal report on each set using what we have studied in class.

Every business day late will result in a 10% deduction in your grade. **Everything should be typed unless the instructions note otherwise.** If you do hand draw something, make sure that it is your neatest work. You may format your reports in any way that you choose, but be sure that they are organized, readable, and easily followed. Also, make sure that every part is completed and can be found in your work. When you are asked a question, answer it in essay format with good sentence structure. When asked to explain something, it should take at least three good sentences. When I grade this, I want to feel like I am reading a formal report on data.

Report I:

1. Find a set of data with at least 50 entries that is of interest to you. You can either find this data set online, or you may collect it. It does need to be numerical data. Write a brief paragraph citing where you found it or how you collected it and what the data set is. Include the data set in your report.
2. What is the population of this data set? What is the sample population? Is your data set an example of a simple random sample? Why or why not? If it is not, how could it have been collected differently in order to be a simple random sample?
3. Construct a frequency distribution and a histogram that includes a frequency polygon of your data. This can be hand drawn or done on software. Make sure that your axes are labeled, and your graph is titled on your histogram.
4. Find the mean, median, mode, and midrange of your data set (you can use Excel/Desmos/online calculators to do this). Explain each measure of central tendency and why or why not you think it is a good representation of the data.
5. Find the range of your data set. Explain what the range means in the case of your data and what it tells you about it.
6. Find the standard deviation of your data set (you can use Excel/Desmos/online calculators to do this). Interpret this standard deviation in relation to your data.

Report II:

1. Find a set of data that has two variables that correlate with at least 25 data points. You can either find this data set online, or you may collect it (it must be numerical). Write a brief paragraph citing where you found it or how you collected it and what the data set is. Include the data set in your report.
2. Before you do anything with the data, make a prediction on whether you think it is going to demonstrate a positive or negative correlation. Explain why you think this way.
3. Construct a scattergram of your data. This needs to be done on software (Excel or Desmos). Make sure that your axes are labeled, and your graph is titled (if you use desmos, you may need to print it out and then hand write these in).
4. Use Excel or Desmos to find the linear regression line. Your print out of your graph should include this line and the equation of it.
5. Analyze your data based on what you have done with it. Explain its correlation (include r), and why you think it may be correlated in that way.