SECTION REPORTS

MATH 12 // STATISTICS // SPRING 2019

Overview

This document is contains all of the problems that your group must complete through the duration of this class. Submit these problems in "Section Reports". The content of each section report is detailed below.

After you hand in a section report, you may move onto the next. In the meantime, I will provide comments on your previous section report. You are then asked to rewrite the section report, addressing my comments. You may then hand in the next draft. This process is repeated until I mark it satisfactory. After all section reports in a chapter are marked satisfactory, you may submit section reports for the next chapter.

Each section report should be professional, neat, clear, and written in complete sentences as if you were writing a lab report. You may handwrite the section reports, but if your writing is not neat, at least ensure that the formatting is clean.

An example section report is posted at

`https://danielmichaelcicala.github.io/teaching/example-section-report.pdf`

Note: Many of the statistics used in this document are fabricated for the purpose of teaching statistical concepts.

Chapter 1

Chapter 1, Section 1.

- Define a *population* and a *sample*. What is the difference between the two? Create your own example to illustrate the difference.
- Describe the difference between descriptive statistics and inferential statistics. Invent three examples of a situation in which one would use descriptive statistics and another three in which one would use inferential. Justify each example.
- A (fake) study conducted at Norco Community College determined that student who attended class between 90% and 100% of the time usually received A's in the class. Many students who attended between 75% and 90% of the time received a B in the class. Many students who attended less than 75% of classes failed.
  (a) What *variables* were considered under this study?
  (b) What are the *data* in this study?
  (c) Which kind of statistics was used: descriptive or inferential?
  (d) What is the *population* under study?
  (e) Was a *sample* collected? If so, from where?
  (f) Based on the information given, comment on the relationship between the two variables.

Chapter 1, Section 2.

- What is the difference between a *quantitative* variable and *qualitative* variable? Create three examples of each type and justify why your example is correct.
- What is the diffence between *discrete* variables and *continuous* variables? Create three examples of each type and justify why your example is correct.
- Imagine you are at the doctors and the nurse is taking your weight. As you step onto the scale, you notice that it only gives weights in hundreth's of a pound. For example, the scale will round a weight of 155.246 up to 155.25 and a weight of 123.423 down to 123.42. Is weight a discrete variable? Justify your answer.
- All measuring devices necessarily round. Otherwise, they'd have to display infinitely many decimal places, which is absurd. Can we ever actually determine a piece of continuous data exactly? What is the *boundary* of a number? If I had a thermometer round temperatures to one decimal point, what is the boundary of 34.3°C? If I had a watch that measured seconds, but not fractions of a second, what is the boundary of 45 seconds?
- There are four levels of measurement described in this section. What are the called? Which level would letter grades given in a class fall under? IQ? Weight? Model of car? Describe what is meant by a *true zero*.

Chapter 1, Section 3.

- Define six sampling methods. Create a situation in which each method would be used. Give a negative aspect for each method and describe why it is negative.
- What is a *sampling error*? Can we always determine a sampling error exactly? Why or why not? Give an example of when we can determine a sampling error exactly.
- Define the population considered in the following situations
  (a) Monmouth University reported that 57% of people with a college degree are homeowners.
  (b) CNN found that 85% of federal inmates are serving less than 10 years
  (c) Eating eggs in the morning may contribute to higher cholesterol
  (d) Indoor cats live, on average, three years longer than outdoor cats.
- Identify the sampling method used
  (a) The Dean of RCC wants feedback from faculty. So the Dean randomly selects 30 teachers to interview.
  (b) Between 12:30pm and 12:45pm, every cashier at Target asks their customers whether they live in Riverside County.
  (c) Every 100th Hersey Kiss made at the factory is weighed to ensure accuracy of the machine.
  (d) To ensure student satisfaction, the principal at a local high school surveys ten freshmen, ten sophmores, 15 juniors, and 15 seniors about their experiece.

Chapter 2

Chapter 2, Section 1.

- Write down each name in your group, first and last. Place all of the letters used in your names (count repetitiions) into a categorical frequency distribution.
- Choose any sports team. Look up the roster and record the number of every player. Make two grouped frequency distribtutions for these numbers. One with class width of 5 and another with class width of 10.
- What is the point of class boundaries? How are they different from class limits?
- Look up the average high temperature (in Farenheit) of each capital city in North and South America in January of any year. Place the temperatures into a cumulative frequency distribution. Choose your own class width, and describe why you choose that for the class width. Are there any outliers? If yes, justify why you think it's an outlier. If not, say why not.
- List five reasons for constructing frequency distributions

Chapter 2, Section 2.

- Look up the total number of points scored for every Super Bowl (that is, add the points that each team scored). Construct a histogram with class boundaries of your choice. Describe why you chose the class boundaries as you did. Why do you think the shape of the histogram is like it is?
- Look up the age that every Canadian Prime Minister took office. Construct a frequency polygone of the ages.
- Construct two ogives. One for the total number of points scored every Super Bowl and another for the age that every Canadian Prime Minister took office. Compare and contrast how the ogives look relative to the histogram and to the frequency polygone.

Chapter 2, Section 3.

- Define a time series graph and a pie graph. Invent your own example of a situation in which to use a time series graph and another for a pie graph. You can be general.
- Suppose there are 250 people in a movie theater. Of them, 25 are wearing red shirts, 57 are waering blue shirts, 71 are wearing black shirts, 61 are wearing orange shirts, and 36 are wearing beige shirts. What graph is captures this information better: time series or pie? Construct that graph. *(If you answer a pie graph, be sure to compute the necessary percentages)*
- Carla is using a FitBit to keep track of her steps. She wants to graph her steps each day of the previous week. Which graph is better suited for this purpose: time series or pie? Use your answer to graph the data

Day	Steps
Monday	7,514
Tuesday	8,385
Wednesday	10,014
Thursday	4,598
Friday	12,345

(If you answer a pie graph, be sure to compute the necessary percentages)

Chapter 3, Section 1.

- There are two different formulas to compute a mean. We use the notation $\overline{X}$ to refer to the mean of a *sample*, which is given by the formula

$$\overline{X} = \frac{X_1 + X_2 + \cdots + X_n}{n} = \frac{\sum X}{n}.$$

  We use the notation μ to refer to the mean of a *population*, which is given by the formual

$$\mu = \frac{X_1 + X_2 + \cdots + X_N}{N} = \frac{\sum X}{N}.$$

  You can find these in the text. Explain the meaning of the following notations
  (a) X
  (b) $\sum X$
  (c) n
  (d) N
  (e) X_1, X_2, ..., X_n, X_N

  What are the difference between these two equations?
- Suppose we are given a bunch of data. If we are told nothing about its source, do we assume it comes from a sample or population? *Hint: the answer to this is contained in a single sentence in the textbook.*
- In your section report for Chapter 2 Section 1, you wrote down the average high temperatures for national capital cities in North and South America. Write all of those same numbers in a table (don't write the city names, just the numbers in any order). What is the average of these temperatures? You don't need to write down the entire equation. Instead describe in a sentence *how* you computed the average. What value is n? What is the meaning of n in this example?
- Go to the website

  https://weathlygorilla.com/richest-singers-world/

  and make a table containing the net worth of the singers (there's no need to record the names of the singers). From what population are the singers a sample (there are multiple correct answers).
  (a) Is this data discrete or continuous?
  (b) What is the average net worth of the singers? Do you think this average gives a fair representation of the population? Why or why not?
- Place the same dollar amounts into a (grouped) frequency distribution with 10 classes.
  (a) Based on your answer above, should your frequency distribution contain a column for class limits? Why or why not?
  (b) What is the mean of this grouped data?
  (c) Is this the same as the mean of the net worths above? Why or why not?
- Next, write down any 30 numbers between 10,000 and 100,000. Think of these as the net worth of non-famous singers. Combine into a single data set these 30 numbers with the 20 net worths above. What is the mean of

this data set? What is the median? Which is more representative of the general Californian? Why?

- STILL NEEDED: mode, weighted mean.