Solutions to Homework 1

Prof. Jordan C. Hanson

July 15, 2020

Exercises: 42, 44, 46, 51, 52, 54, 56, 58, 60, 66, 76, 90

1. Exercise 42:
   - a) The population is the set of all clients
   - b) The sample is a subset of the population actually measured
   - c) The parameter is the average of the population: the mean time spent exercising
   - d) The statistic is the sample mean of times spent at the gym
   - e) The variable is the time per trip to the gym
   - f) The data is a list of times spent at the gym, forming the sample
2. Exercise 44:
   - a) The population is all of the patients who have had heart attacks
   - b) The sample is the subset of the population actually measured who have had heart attacks
   - c) The parameter is the mean time of recovery for the whole population
   - d) The statistic is the mean recovery time found in the sample
   - e) The variable is the amount of time needed for recovery
   - f) The data is the list of recovery times for each patient in the sample
3. Exercise 46:
   - a) All voter's in the politicians district
   - b) A random selection of voters in the district
   - c) The proportion of voters in this district who think the politician is doing a good job
   - d) The same propotion as (c) but for the sample
   - e) The number of voters in the sample who check yes divided by the total number
   - f) Qualitative or categorical data
4. Exercise 51: The correct choice is **a**, a variable.
5. Exercise 52: The correct choice is **c**, a statistic.
6. Exercise 54: Percent body fat is an example of quantitative continuous data (e.g. 15.1 percent).
7. Exercise 56: Time in line is an example of quantitative continuous data (e.g. 7 minutes).
8. Exercise 58: Most watched television show is an example of qualitative or categorical data
9. Exercise 60: Distance is quantitative continuous (e.g. 4.5 km)
10. Exercise 66: **One example given here:** *Use simple random sample to choose 25 colleges in the state. Use all statistics classes from each college. List all the colleges together with a two-digit number (`http://en.wikipedia.org/wiki/List_of_colleges_and_universities_in_California`) Use a random number generator to pick 25.*
11.
    - a) During the Great Depression, many people could not afford the items that would have placed them on those lists, so the high N value incurred sampling error.
    - b) Samples that are too small can have sampling error or bias

- c) Sampling error
- d) Stratified sampling

12.
- a) 4 percent
- b) 13 percent
- c) Not necessarily. The N values could be different, but moreoever, the population demographics could be different.