Math 1111 Written Assignment 1

Sampling and Descriptive Statistics

Instructions

Please answer each of the following questions in order.

You may write your answers neatly by hand or type them up, or a mix of the two. You will need to submit your assignment as a PDF file. You can use your phone to scan written work and convert to PDF.

You may use any software or technology to answer the questions below. You may also work by hand. I recommend (but do not require) that you use Microsoft Excel.[i]

Your assignment must be submitted on Moodle. Emailed submissions will not be accepted.

Questions

1. [3 marks] At Onesizefitsall College, every first-year student takes MATH 1111. There are thirty sections of the course to accommodate all the students. <u>Could it possibly be fair to choose a section at random and treat it as an SRS</u> (simple random sample) of the entire first-year population at that college? If so, under what circumstances or assumptions? When would a randomly chosen section <u>not</u> be a good way to sample students? Answer in a brief paragraph.

2. Use the **birthweight** data file for this question. This file contains the birth weight (in kg) of 42 infants.

    a. [3 marks] Find the following values:
        i.   Min
        ii. Q1
        iii. Median
        iv. Q3
        v.   Max
        vi. IQR

    b. [2 marks] Construct a box plot of the data.

    c. [1 mark] What does the shape of the box plot tell you about the data in this case? Comment on the concentration (equivalently, spread) and on the symmetry or skew of the data set.

    d. [3 marks] Construct a frequency table for the data with 5 to 6 intervals (bins).

    e. [3 marks] Construct a histogram based on the frequency table. Give the graph a meaningful title. Label the axes appropriately.

    f. [1 mark] Describe the shape of the histogram.

    g. [1 mark] Find the sample mean and sample standard deviation.

h. [1 mark] How do the standard deviation and interquartile range help you to determine concentration of the data and whether or not there are potential outliers?

i. [2 marks] Are there any potential outliers? List the value(s) that could be outliers. Use a formula to check the end values to determine if they are potential outliers. Include your work and formula to support your conclusion.

---

[i] I recommend Excel because you have access to Office as a VCC student, and because there are so many tutorials online for using Excel for descriptive statistics. We don't have any specific Excel-related learning outcomes in this course, so you can feel free to use any app that works for you.