Semester Project Paper Guidelines

Due Date: December 10, 2019 (Tuesday)

Your paper will be a statistical analysis of a topic that you choose. Your topic must be a **quantitative** phenomenon/characteristic, one which you think might be "normally distributed." After deciding on the topic, it is **necessary** to obtain a **sample of size n $\geq$ 30**. It is strongly suggested that the data you collect contains at least twelve (12) distinct values[*]. In your paper (typically, 3-5 pages in length) you need to include each of the following items:

10pts 1. Introduction, detailing the phenomenon/characteristic being analyzed. Describe how the sample was obtained and what is the population which you imagine that it represents. Was the data obtained with a simple random, stratified, systematic, clustered, multistage, or convenience sampling technique? Is your data discrete[*] or continuous? Discuss any other relevant matters that may have either affected how your sample was obtained or that could impact the outcome(s) of the statistics to be analyzed. Be mindful of the units of measurement that apply throughout your paper.

6pts 2. List the raw data in ascending order.

15pts 3. Determine the values for the arithmetic mean, the five number summary, the mode, and the midrange.

8pts 4. Determine the values for the range and standard deviation.

20pts 5. Group the data according to the class criteria discussed on pages 41-45, and use it to create a...
   a. frequency distribution (with midpoints & relative frequencies)
   b. histogram corresponding to the frequency distribution (label the mean)
   c. cumulative frequency polygon (*a.k.a.* an "ogive" or line graph)
   d. determine the grouped data's mean and standard deviation

9pts 6. Determine the Percent of the raw data in each of the following intervals:
   a. P(x-s < x < x+s)
   b. P(x-2s < x < x+2s)
   c. P(x-3s < x < x+3s)

10pts 7. Compare "6a, b, and c" (above) with the theoretical percentages for a normally distributed population (as cited in the Empirical Rule).

10pts 8. Describe your insights and/or any observations you deem significant from the analysis (items 1-7). *E.g.*, is there any detectable bell-shape with the data, or is it saliently skewed to the left/right, bimodal, etc.? Do you think the data represents a population that is normally distributed or not? Explain/justify your assertions.

2pts 9. List the source(s) of your data, and any other information you may have used from an outside source.

10pts Legibility, coherency, and overall quality of presentation will be considered in the grading of your paper. Handwritten work while accepted is not recommended.

Any tables, graphs, etc. should occupy no more than 1/3 page each. If you are unsure of what topic (phenomenon) to select for your project, feel free to consult with the instructor who may suggest examples. Heights and/or weights (of humans) will not be allowed as a suitable topic of analysis. If after obtaining your data, it does not appear to conform to various preconceptions you may have about the nature of this assignment, do NOT attempt to massage or "fudge" your info as presenting and analyzing data contrary to expectations is an extremely important aspect of study whenever it does occur.

Late papers received on December 12th (Thursday) will be subjected to a penalty of ten points (-10). No papers will be accepted thereafter (a grade of zero will be issued).