DA 6823
Kilger
Module 1: Part #2 (35 points)
**The Power of Statistics + the Levels of Measurement + the Different Classes of Variables and Determining Appropriate Statistical Technique + Basic Descriptive Measures**

**General Instructions:** In your own words, answer each of the following questions - don't copy (e.g. cut and paste) some definition out of a book word for word. This is not a group project – you are expected to complete this module on your own. You may refer to text books, online or other sources but not your fellow classmates. If you don't understand the question, feel free to ask the instructor in class, in office hours or in an email.

1. The first couple of questions deal with the concepts of population and sample.

    a. What is a population? (3 points)
    Is a big group that a statistic generalizes by surveying a smaller group of the group.

    b. What is a sample? (3 points)
    A small group statistics use to generalize a population with.

    c. What is the objective of inferential statistics in terms of sample and population? (4 points)
    To understand a population by only studying a sample instead of the whole population.

2. Name two common measures of central tendency (4 points)
    • Mean
    • Median

3. Describe one situation where one measure of central tendency might be better than another measure (2 points)
    • Mean: Finding the classes average test score rating.
    • Mode: Collecting data on people's ratings of the top 5 pizza places in town.

4. Variance and standard deviation are two of the most commonly used measures of variability. Take a look at the formula for variance below:

$$\text{Sample Variance} = s^2 = \frac{\Sigma(X - \overline{X})^2}{n - 1}$$

a. Looking at the variance formula, if the data points (X) are closely packed around the sample mean, what happens to the sample variance? (3 points)

==The sample variance is very close to zero.==

b. The sample standard deviation is closely related to the sample variance.  How? (2 points)

==Standard deviation is the square root of the sample variance.==

5. Many sample statistics you encounter have standard errors associated with them. Imagine that you are studying the heights of the undergraduate student body of UTSA. The total number of students is 30,000 (e.g. the population) and you randomly pull 10 samples of 100 students each from that population.

a. Will the sample means from each of the 10 samples be the same? ( 2 points)

==No==

b. The standard error of the mean represents the variation in sample means that you find in different samples.  The formula below is the standard error of the mean.

$$SE = \frac{s}{\sqrt{n}} = \sqrt{\frac{s^2}{n}}$$

i. What happens to the standard error of the mean if there is a lot of variation in the data? (3 points)

==Standard error increase.==

ii. What happens to the standard error of the mean as sample size increases? (3 points)

==The Standard error decreases.==

6. The sum of squares is a statistical concept that measures variation in data that you will find in many different statistical techniques.  Here is the formula for sum of squares below:

$$SS = \sum (X - \overline{X})^2$$

    a. This formula should look familiar.  What other statistical measures of variation are related to this? (3 points)
    Sample variance.

    b. As the data points (X) get spread out farther and farther from the sample mean, what happens to the sum of squares? (3 points)
    The sum of squares becomes larger.