**

**Assignment 2: Data Fundamentals with Excel – Modules 1-5**

Please complete Data Fundamentals using the LinkedIn Learning video link posted in Week 4. Download the exercise files to follow along and learn while watching, listening and practicing.

1. **Foundational concepts of data analysis – measures of central tendency.**
2. How do you measure the centrality of your data?

The mean is from averaging the sample by taking the sum of all values divided by the number of values in the dataset. Median would be the middle value after lining all the values in order. Mode would be the value with the most occurrences in the dataset.

1. What is the difference between variance and standard deviation?

The variance takes all the distances each data point is from the mean and then squares and averages the result while the standard deviation is the square root of the variance which can give us a standard to know what is our normal range

1. How is the sample standard deviation calculated? What is its value in the example given?

The standard deviation is calculated by squaring the variance. The Standard deviation gives us a reference into what we can expect to be normally occurring or rare. If the data is normal distributed we should see 68% of our values in the middle within one standard deviation from the mean showing likely events while those past 3 standard deviations are rare events containing .3% of values.

1. **Visualize data**
2. What is the difference between a histogram and a scatter chart?

A histogram uses intervals or bins to separate the data and visualizes that data using bar columns. Histograms are best for seeing the frequency within variables. A scatterplot uses dots to indicate each occurrence and are best to see the trends between variables.

1. When would you decide to use a scatter chart instead of a histogram?

I would use a scatterplot if I am interested in viewing the trend or relationship between two variables such as shoppers distance and money spent in a store.

1. **Test a hypothesis**
2. What is the difference between a null and alternative hypothesis?

The null hypothesis is the point you are trying to prove while the alternative is the counter or doubt to the claim the hypothesis is making.

1. What is a null hypothesis?

The null hypothesis the main idea or point you are trying to prove. It is the quantifiable or observable research problem.

1. What defines a “significant” result?

I significant result is one that has an alpha or a confidence ratio of 95% or higher.

1. **Utilize data distributions**
2. What is a “normal” distribution?

A continuous distribution of data points and has a bell-shaped curve. Can have an infinite amount of data points.

1. How is a binominal distribution different than a normal distribution?

A discrete type of distribution sample with a finite amount of data points. Usually a fixed sample.

1. **Measure covariance and correlation**
2. What does a covariance of zero imply?

A zero covariance indicates that there is no relationship whether it be positive or negative between the two variables.

1. What does a correlation of -.50 imply?

A correlation of -.5 implies a moderate negative correlation between two variables.