Course Project - Phase 1

Shaun Pritchard

Rasmussen College

STA3215

Fie Wang

April,11 2020

Course Project – Phase 1

**Introduce your scenario and data set.**

Provide a brief overview of the scenario you are given above and the data set that you will be analyzing.

The data set is based on a new infectious disease in which will determine if the age of patients correlates or facilitates for the rise, spread, and possible cause of the new disease. Determining if the increase in patients is related to this new infectious disease. Calculating if it is targeting specific age groups.

The data set consists of 65 patients that have the infectious disease with ages ranging from 35 years of age to 81 years of age for NCLEX Memorial Hospital. Remember this assignment will be completed over the duration of the course.

**Classify the variables in your data set.**

* Which variables are quantitative/qualitative?
  + Quantitative variable would be the age variable
  + Qualitative would be the patient numbers of the sample
* Which variables are discrete/continuous?
  + Patient numbers are discrete
  + Patient age is continuous
* Describe the level of measurement for each variable included in your data set.
  + There are 65 patients being used in this sample training from ages 35 to 81.
* Discuss the importance of the Measures of Center and the Measures of Variation.
  + the measure of center summarizes in variable a representation of all the data from its center. Variation uses this center to find normal distributions as a way to expresses the standard deviation as a percentage of the mean.
* What are the measures of center and why are they important?
  + Mean and median both try to measure the "central tendency" in a data set. The goal of each is to get an idea of a "typical" value in the data set. The mean is commonly used, but sometimes the median is typically preferred. It is resistant to the presence of extreme values
* What are the measures of variation and why are they important?
  + The measures are the range, standard deviation, and variance the range of a set of data values is the difference between the maximum data value and the minimum data value. The standard deviation of a set of sample values, denoted by s, is a measure of how much data values deviate away from the mean. basically, a transformation or velocity from the mean. The variance is a measure of how spread out the set of data is from the average (mean, median).

Calculate the measures of center and measures of variation. Interpret your results in context of the selected topic. Based off the age being they all have the infection and the numbers to identify them are not quantitative.

Mean 62.54

Median 64

Mode 69

Midrange 75.5

Range 46

Variance 85.69

Standard Deviation 9.26

**Conclusion**

Recap your ideas by summarizing the information presented.

It seems that the age average age range of patients at a median of 64 years old and the mean of 63 years of age have the average infection rate meaning this age range could more susceptible to the infection.it seems the standard deviation is spread apart at 9.26 which means based on the sample of patients that age range of infected patients varies more than those of the median or mean age range of patients at 63-64.

Resources:

Mario F. Triola (2018). Elementary Statistics, America, Pearson

References

1