

MATH325 Lab 2

Measures of Central Tendency and Variability

The steps required for completing the deliverables for this assignment, including screen shots that correspond to these instructions, are outlined below. Complete the questions below and paste the answers from Excel below each question (type your answers to the questions **where noted**). Therefore, your response to the lab will be this ONE submitted document.

Context: Remember that statistics are far more than numbers or values – you need to know the context to perform a good analysis!

Central Tendency: Defined as statistics that describe the location of the distribution. This includes the mean, median, mode, and sum of all the values.

Mean. A measure of central tendency. The arithmetic average, the sum divided by the number of cases.

Median. The value above and below which half of the cases fall, the 50th percentile. If there is an even number of cases, the median is the average of the two middle cases when they are sorted in ascending or descending order. The median is a measure of central tendency not sensitive to outlying values (unlike the mean, which can be affected by a few extremely high or low values).

Mode. The most frequently occurring value. If several values share the greatest frequency of occurrence, each of them is a mode. The Frequencies procedure reports only the smallest of such multiple modes.

Maximum. The largest value in the dataset.

Minimum. The smallest value in the dataset.

Percentile Values (Q1, Median, Q3). Values of a quantitative variable that divide the ordered data into groups so that a certain percentage is above and another percentage is below. Quartiles (the 25th, 50th, and 75th percentiles) divide the observations into four groups of equal size default in Minitab). If you want an equal number of groups other than four, select Cut points for equal groups. You can also specify individual percentiles (for example, the 95th percentile, the value below which 95% of the observations fall).

IQR (Interquartile Range). This is the middle 50% of the data calculated by $Q3 - Q1$.

Range The difference between the largest and smallest values of a numeric variable, the maximum minus the minimum.

S.E. Mean. A measure of how much the value of the mean may vary from sample to sample taken from the same distribution. It can be used to roughly compare the observed mean to a hypothesized value

(that is, you can conclude the two values are different if the ratio of the difference to the standard error is less than -2 or greater than $+2$).

StDev. A measure of dispersion around the mean. In a normal distribution, 68% of cases fall within one standard deviation of the mean and 95% of cases fall within two standard deviations. For example, if the mean age is 45, with a standard deviation of 10, 95% of the cases would be between 25 and 65 in a normal distribution.

Variance. A measure of dispersion around the mean, equal to the sum of squared deviations from the mean divided by one less than the number of cases. The variance is measured in units that are the square of those of the variable itself.

1.