

Name:

Date:

Class:

Vocabulary List Answer Key

Provide definitions for the vocabulary terms listed below; use equations where appropriate.

Mean: The average value in a data set. $\frac{\text{sum of the data}}{\text{number of data points}} = \text{mean}$

Median: The middle number (or the average of the two middle numbers) after a data has been arranged from least to greatest. The 50th percentile.

Mode: The most frequently occurring number in a data set.

Range: The difference between the largest and smallest values in a data set.

$$\max(\text{data}) - \min(\text{data}) = \text{range}$$

Standard deviation: The variance in the data, or how “spread out” the data is.

Population vs. sample: A population is every possible data point, whereas a sample is a reasonable subset of the population. Samples are used when it is unreasonable to measure an entire population.

Histogram: A type of plot used to understand the distribution of a data set. All possible data points are plotted along the x-axis, and the y-axis is the frequency with which the points occur.

Time series plot: A graph with time on the x-axis and the values of interest on the y-axis. Useful for examining *temporal* trends (or trends in time).

Scatter plot: A graph that enables the examination of the relationships between variables. Each variable is plotted on an axis, and patterns or fitted models suggest relationships in the data.

Linear regression (or linear fit): A method by which a standard “ $y = m*x+b$ ” model is fit to a data set to solve for the coefficients (m and b) that result in the best fit, or the line with the least distance to each data point. Used to explain relationships in data or predict future data.

Coefficient of determination (R^2): A number that explains how well the data fits the model (such as a linear model). A value of 1 indicates a perfect fit, while a value of 0 indicates a poor fit. If data fits a model well, all of the points are close to the line or pattern; if the data is a poor fit, the points are far away or spread out from the line or pattern. (Refer to the example plots in the *Pre-Activity Reading* for more explanation.)