Describing Data Key Terms:

- **Measures of Central Tendency**: Values that describe the center of a distribution. The mean, median, and mode are 3 measures of central.
- **Mean**: A measure of central tendency that is determined by dividing the sum of all values in a data set by the number of values.
- **Median**: The value of the middle term in a set of organized data. For a set of data with an odd number of values, it is the value that has an equal number of data values before and after it, or the middle value. For a set of data with an even number of values, the median is the average of the 2 values in the middle positions.
- **Mode**: The value or values that occur with the greatest frequency in a data set.
- **Outliers**: Extreme values in a data set.
- **Broken Line Graph**: A graph that is used when it is necessary to show change over time. A line is used to join the values, but the line has no defined slope.
- **Continuous Data**: Data for which the plotted points can be joined.
- **Correlation**: A statistical method used to determine whether or not there is a linear relationship between 2 variables.
- **Data Set**: A collection of observations of a variable.
- **Dependent Variable**: The variable represented by the values that are plotted on the y-axis.
- **Discrete Data**: Data for which the plotted points cannot be joined.
- **Five-Number Summary**: 5 values for a data set that include the smallest value, the lower quartile, the median, the upper quartile, and the largest value.
- **Frequency Distribution**: A table that lists all of the classes and the number of data values that belong to each of the classes.
- **Histogram**: A graph in which the classes are on the horizontal axis and the frequencies are plotted on the vertical axis. The frequencies are represented by vertical bars that are drawn adjacent to each other.
- **Independent Variable**: The variable represented by the values that are plotted on the x-axis.
- **Interquartile Range (IQR)**: The difference between the third quartile and the first quartile.
- **Left-Skewed Distribution**: A distribution in which most of the data values are located to the right of the mean.
- **Line of Best Fit**: A straight line drawn on a scatter plot such that the sums of the distances to points on either side of the line are approximately equal and such that there are an equal number of points above and below the line.
- **Midpoint**: The value obtained by adding the lower and upper limits of a class and dividing the sum by 2.
- **Qualitative Variable**: A variable that can be placed into specific categories according to some defined characteristic.
- **Quantitative Variable**: A variable that is numerical in nature and that can
be ordered.

- **Right-Skewed Distribution** - A distribution in which most of the data values are located to the left of the mean.

- **Scatter Plot** - A graph used to investigate whether or not there is a relationship between 2 sets of data. The data is plotted on a graph such that one quantity is plotted on the x-axis and one quantity is plotted on the y-axis.

- **Symmetric Histogram** - A histogram for which the values of the mean, median, and mode are all the same and are all located at the center of the distribution.

- **Variable** - A characteristic that is being studied.

- **Categorical Data** - Data that are in categories and describe characteristics, or qualities, of a category.

- **Double Box-and-Whisker Plots** - 2 box-and-whisker plots that are plotted on the same number line.

- **Numerical Data** - Data that involves measuring or counting a numerical value.