UNIT 4 VOCABULARY

1. **Box and Whisker Plot**: A diagram that summarizes data using the median, the upper and lowers quartiles, and the extreme values (minimum and maximum). Box and whisker plots are also known as box plots. It is constructed from the five-number summary of the data: Minimum, Q1 (lower quartile), Q2 (median), Q3 (upper quartile), Maximum.
2. **Frequency**: the number of times an item, number, or event occurs in a set of data
3. **Grouped Frequency Table**: The organization of raw data in table form with classes and frequencies
4. **Histogram**: a way of displaying numeric data using horizontal or vertical bars so that the height or length of the bars indicates frequency
5. **Inter-Quartile Range (IQR**): It is a measure of variation in a set of numerical data, the interquartile range is the distance between the first and third quartiles of the data set (sometimes called upper and lower quartiles). Example: For the data set {1, 3, 6, 7, 10, 12, 14, 15, 22, 120}, the interquartile range is 15 – 6 = 9.
6. **Maximum value**: The largest value in a set of data
7. **Mean Absolute Deviation**: the average distance of each data value from the mean (𝑥̅). The MAD is a gauge of “on average” how different the data values are form the mean value.

𝑀𝐴𝐷=

𝑡𝑜𝑡𝑎𝑙 𝑑𝑖𝑠𝑡𝑎𝑛𝑐𝑒 𝑓𝑟𝑜𝑚 𝑡ℎ𝑒 𝑚𝑒𝑎𝑛 𝑓𝑜𝑟 𝑎𝑙𝑙 𝑣𝑎𝑙𝑢𝑒𝑠 ÷ 𝑛𝑢𝑚𝑏𝑒𝑟 𝑜𝑓 𝑑𝑎𝑡𝑎 𝑣𝑎𝑙𝑢𝑒𝑠

1. **Mean*:*** A measure of center in a set of numerical data, computed by adding the values in a list and then dividing by the number of values in the list. Example: For the data set {1, 3, 6, 7, 10, 12, 14, 15, 22, 120}, the mean is 21.
2. **Measures of Center**: The mean and the median are both ways to measure the center for a set of data.
3. **Measures of Spread**: The range and the mean absolute deviation are both common ways to measure the spread for a set of data.
4. **Median**: A measure of center in a set of numerical data. The median of a list of values is the value appearing at the center of a sorted version of the list—or the mean of the two central values, if the list contains an even number of values. Example: For the data set {2, 3, 6, 7, 10, 12, 14, 15, 22, 90}, the median is 11. Note: The median is a good choice to represent the center of a distribution when the distribution is skewed or outliers are present.
5. **Minimum value**: The smallest value in a set of data.
6. **Mode**: The number that occurs the most often in a list. There can more than one mode, or no mode.
7. **Mutually Exclusive**: two events are mutually exclusive if they cannot occur at the same time (i.e., they have not outcomes in common).
8. **Outlier**: A value that is very far away from most of the values in a data set.
9. **Range**: A measure of spread for a set of data. To find the range, subtract the smallest value from the largest value in a set of data.
10. **Sample**: A part of the population that we actually examine in order to gather information.
11. **Simple Random Sampling**: Consists of individuals from the population chosen in such a way that every set of individuals has an equal chance to be a part of the sample actually selected. Poor sampling methods, that are not random and do not represent the population well, can lead to misleading conclusions.
12. **Stem and Leaf Plot**: A graphical method used to represent ordered numerical data. Once the data are ordered, the stem and leaves are determined. Typically the stem is all but the last digit of each data point and the leaf is that last digit.