What do you mean?

Many a times, we want to describe or summarize data in a meaningful way. This is where descriptive statistics help us. Descriptive statistics summarize and organize data so that they can be easily understood. For example, if we had academic results for 100 students, we might be interested in the overall performance of students. We would also be interested in the overall spread of the marks, i.e., how does the performance of one student differ from the others. We will understand different types of descriptive statistics measures with an example.

Nine of your friends, along with you went to Afro Deli in Stadium Village to grab lunch after school one day and gave different orders. The prices for the ten orders come out to be 7, 10, 5, 12, 10, 26, 14, 9, 9 and 18. As a group you spent $120 in all.

1. Can you represent these prices on a number line?

2. Your friend tells you that on an *average*, each of your friend spent $ 10 on their meal. Do you think they are correct in stating that?

3. Average / arithmetic mean (also written as $\bar{x}$) = $\frac{sum\ of\ all\ observations}{Total\ number\ of\ observations}$. Does this tally with your friend's assertion?

4. Often times, knowing the mean is not enough. E.g., 5, 7 and 26 would give us a mean of 12.6. But there is a big difference between spending $5 and $26. To account for this big difference, we calculate something called *range*. Range = Maximum value - minimum value. What will be the range here?

5. There is a more common measure of spread called *standard deviation*. While range and standard deviation both tell you how spread out your data is, standard deviation tells you about the spread relative to the mean. Standard deviation = $\sqrt{\frac{\sum_{i=1}^{10}(x_i-\bar{x})^2}{10}}$. Here, $\sum_{i=1}^{10}(x_i-\bar{x})^2$ is 336. What is the standard deviation?

6. Can you tell which price occurs most frequently in your data? (This is called the mode.)

7. What value divides the the observations into two equal parts (i.e, number of terms on the left side and the right side are the same)? (This is called the median.)

If you decided not to eat out that day, and you were going to be one of the people spending $10 on your meal, would this value (median) change?

These are some of the measures you used above:

1. Arithmetic Mean / Average $= \frac{\sum_{i=1}^{n} x_i}{N} = \frac{sum \ \ of \ \ all \ \ observations}{Total \ \ number \ \ of \ \ observations}$

2. Median $=$ Mean of two middle values (even number of observations) or the middle value itself (odd case) dividing data into two equal parts each having 50% of observations. (You sort the observations in ascending order before computing median)

3. Mode $=$ Most frequently occuring observation

4. Range $=$ Maximum observation - Minimum observation

5. Standard Deviation $(\sigma) = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N}}$

6. Variance $(\sigma^2) =$ (Standard Deviation)$^2 = \frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N}$