Explore your Data: Mode, median and mean

Mode, Median and Mean

Mean, median, and mode are the three measures of central tendency used in statistics to describe the central or typical value of a set of data.

In "Graphs and shapes of Distributions" section it is explained how to **summarizing a distribution** of your data in terms of graphs. Now it's time to measure the **center of your distribution**. Once we talk about measuring central tendency of a variable then 3 M's come into picture.

- **Mode**
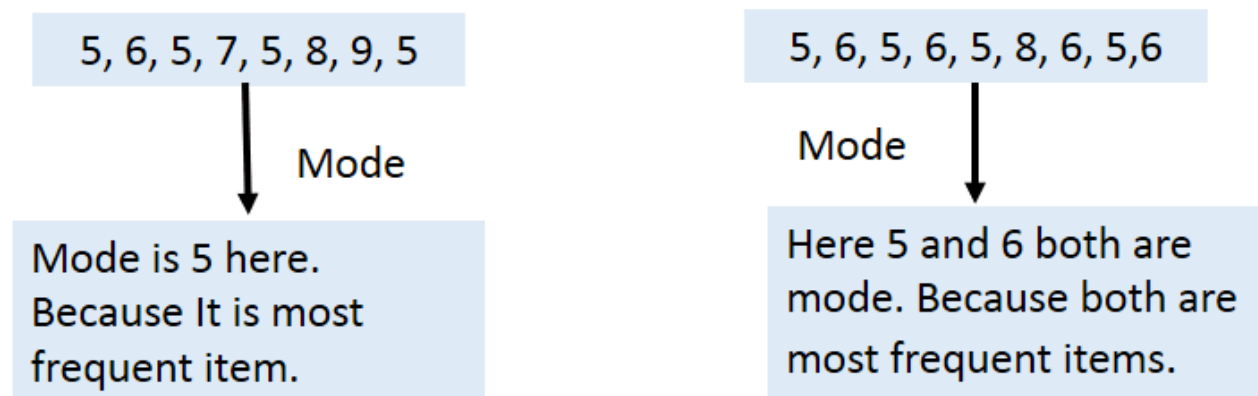- **Median**
- **Mean**
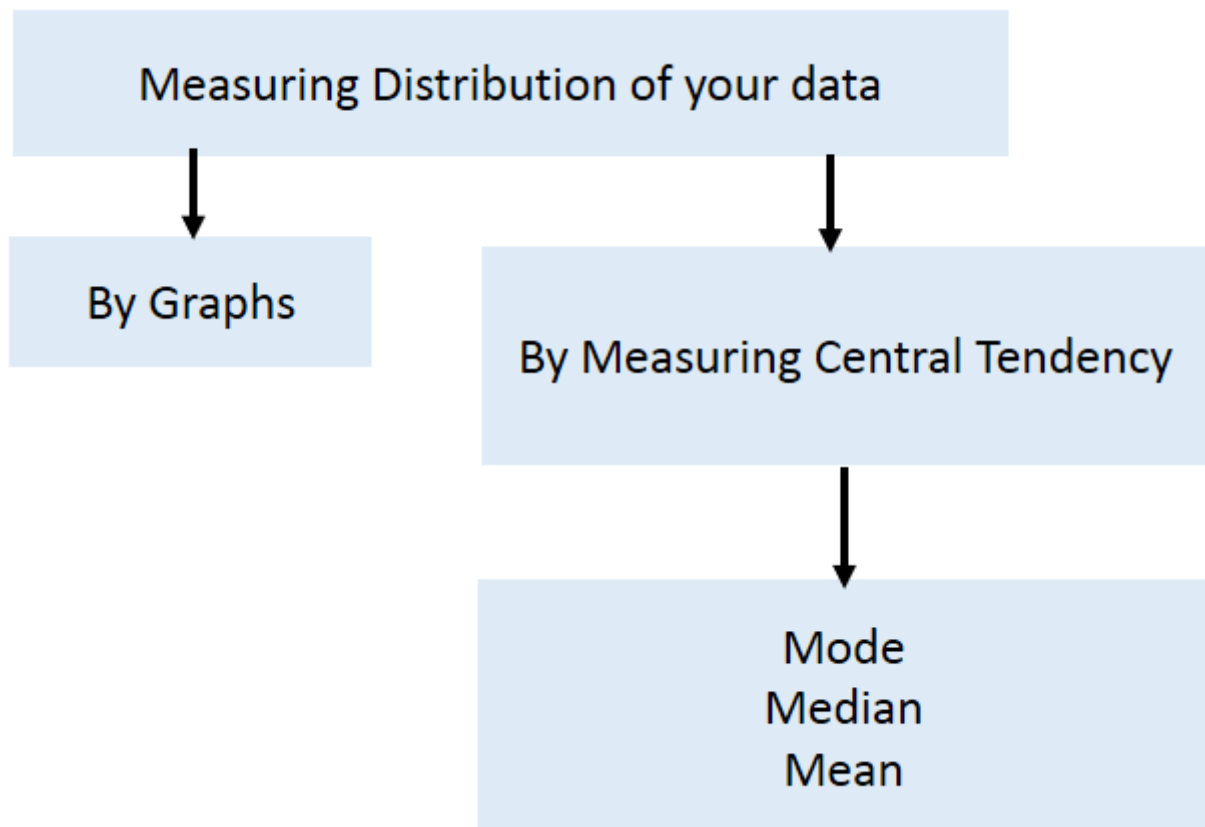
Mode

The mode is the value that occurs most frequently in a dataset. It is the most common value in the dataset. A dataset can have multiple modes, or no mode at all if there is no value that occurs more than once.

If your variable of interest is measured in nominal or ordinal (Categorical) level then Mode is the most often used technique to measure the central tendency of your data.

Finding the mode is easy. Basically, it is the value that occurs most frequently. In other words, mode is the most common outcome. Mode is the name of the category that occurs more often.

There is a chance of having more than one mode in your variable.

5, 6, 5, 7, 5, 8, 9, 5

↓ Mode

Mode is 5 here. Because It is most frequent item.

5, 6, 5, 6, 5, 8, 6, 5,6

Mode ↓

Here 5 and 6 both are mode. Because both are most frequent items.

Here you have two modes.

Mode Example

If you still don't understand and want to know you to calculate mode step by step then follow the below link http://www.purplemath.com/modules/meanmode.htm

Median

The second measure of central tendency is the median. The median is nothing more than the middle value of your observations when they are order from the smallest to the largest.
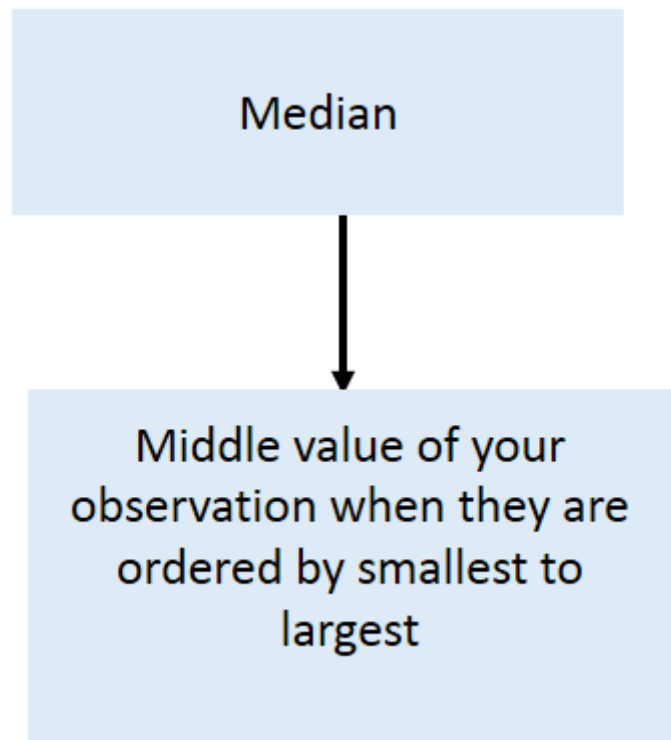
So, the median is the middle value in a sorted dataset. It is the value that separates the highest 50% of the data from the lowest 50%. To find the median, you need to sort the data in order and find the value that is exactly in the middle. If there are an even number of values, the median is the average of the two middle values.
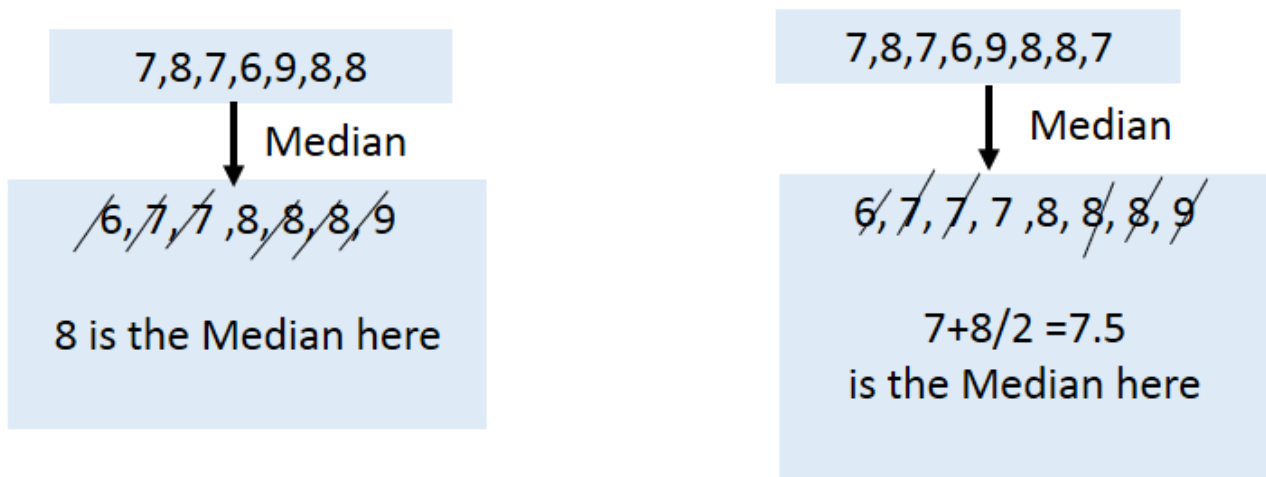
It involves two steps:

Step1: Oder your cases from smallest to largest

Step 2: Find the middle Value

- If you have odd number of cases then finding middle value is easy. Let's think you have 5 cases. So, after ordering always 3rd position is the middle value.
- If you have even number no cases (let's think 6 cases). In this case there is no single middle value. Then how do we calculate median? Well, we just take the average of the two middle values.

Median

Middle value of your observation when they are ordered by smallest to largest

Median Example

7,8,7,6,9,8,8

Median

6,7,7 ,8,8,8,9

8 is the Median here

7,8,7,6,9,8,8,7

Median

6,7,7, 7 ,8, 8, 8, 9

7+8/2 =7.5
is the Median here

Mean

The mean, also known as the arithmetic mean or average, is the sum of all values in a dataset divided by the number of values. It is a measure of the central tendency that reflects the center of mass of the data. The formula for the mean is:

Mean = (sum of values) / (number of values)

It's the third measure of central tendency is the most often used one, and also the one you most probably already know quite well: the mean. The mean is the sum of all the values divided by the number of observations. It is nothing but the average value.

$$\overline{X} = \frac{\Sigma X}{n}$$

Now the question is When to use what measurement of central tendency?

- If data is Categorical (Nominal or Ordinal) it is impossible to calculate mean or median. So, go for mode.
- If your data is quantitative then go for mean or median. Basically, if your data is having some influential outliers or data is highly skewed then median is the best measurement for finding central tendency. Otherwise go for Mean.

7,8,7,6,9,8,8

↓ Mean

Sum(53) / #cases (7)
=7.6