MEASURES OF CENTRAL TENDENCY – MEDIAN AND MODE

MEDIAN

The middle value in a data set when its components are ordered sequentially—that is, in either ascending or decreasing order of magnitude—is known as the median.

COMPUTATION OF MEDIAN

Exclusive Series

The following procedures must be followed in order to determine the median in a single series:

Step 1: Sort the data either ascending or descending.
Step 2: Use the following equation:

$$\text{Median (Me)} = \text{Size of } \left[\frac{N+1}{2}\right]$$

Discrete Series

The values of the variable are provided together with their frequencies in a discrete series or frequency array.

Continuous Series

Finding the median in a continuous series is not an easy task. The median in this instance is located between the class interval's upper and lower bounds.

ASPECTS OF THE MEDIAN

1. The minimum is the total of all item departures from the median, ignoring signs.
2. Because the median is a positional average, high values have no bearing on it.

MEDIAN VS MEAN

1. Ease of Calculations: The median is simpler to compute than the mean.
2 Sample Fluctuations: Generally speaking, sampling fluctuations have a bigger impact on the median than the mean; occasionally, though, the mean may be more impacted than the median.
3. Algebraic Treatment: When it comes to additional algebraic treatment, mean is unquestionably better than median. The combined mean can be determined, but not the combined median.

4. Open-end classes: The median can be computed with ease in an open-end distribution, but the mean cannot be ascertained.

5. Impact of Extreme values: Because the median is unaffected by the values of extreme items, it might be more representative than the arithmetic average.

6. Graphic presentation: The mean value cannot be graphically ascertained, however the median value can be determined visually.

THE GOODS AND THE BAD IN THE MEDIAN

Merits of Median

1. Simplicity: The median is straightforward to compute and comprehend. The median can often be found with only a quick visual assessment.

2. Ideal average: The median has a fixed and defined value since it is determined firmly.

3. Graphic presentation: Ogive curves can also be used to determine the median value graphically.

4. Unaffected by extreme values: The computation of the median value is not impacted by the extreme values present in the data set.

5. Possible even with incomplete data: Incomplete data does not prevent the calculation of the median.

6. Fit for qualitative data: The median can be used to address qualitative traits that are difficult to quantify.

Demerits of Median

1. The median is not based on every observation: because it is a positional average and is not derived from every item in the distribution.

2. Affected by sampling variations: It is impacted by sampling fluctuations. Therefore, if the class-intervals are not uniform, the median value is no longer applicable.

3. Absence of additional algebraic processing: The median cannot be processed algebraically.

4. Requires arrangement: Since the median is an average of the data position, it takes time to arrange the magnitudes in an ascending or descending order when there are a lot of observations.

5. Uniform distribution of observations: in the median class is the basis for the computation of the median formula in the case of a clustered frequency distribution, which is an unrealistic assumption. In real life, this expectation is rarely fulfilled.

APPLICATIONS OF MEDIAN

QUARTILES

A value known as the median divides a series into two equal sections. Other positional values also exist that split a series into many segments. Quartiles are the most often used positional values in addition to median.

A series is divided into four equal pieces by quartiles.

1. First Quartile, or Lower Quartile (Q_1): Q_1 splits the distribution so that 25% of the total items fall below it and 75% fall above it.

2. Median, or Second Quartile (Q_2): It's been talked about before.

3. Third or Upper Quartile (Q_3): Q_3 splits the distribution so that one-fourth (25%) and three-fourths (75%) of the total items lie above and below it.

MODE

Another significant and philosophically helpful indicator of central tendency is the mode. The value that appears in a set of observations the most frequently and around which the other items in the set cluster the densest is called the mode.

Key Information regarding Mode

- Mode is widely used to gauge people's preferences and taste for a specific brand of product.
- When it comes to frequency distribution, the value that corresponds to the maximum frequency determines the mode.
- Circumstances in which the "Maximum Frequency" Rule is broken: The concentration of observations around a value with the highest frequency may in some circumstances be less than the concentration of observations around a different value. The maximum frequency criterion cannot be applied in this case to define the mode.
- The letter 'Z' stands for the value of mode.
- When determining the most typical value, the mode is preferred over the mean and median.

CALCULATION OF MODE

1. INDIVIDUAL SERIES

There are two ways to determine the mode of a single series:
1. By observation
2. By transforming discrete series—that is, by frequency distribution—from individual series.

(i) Mode via Observation Method

It is possible to observe the occurrence of things in a distribution.

Step 1: Put the information in either descending or ascending order.
Step 2: Mode ode is the item that appears most frequently in the series.

(ii) Mode via Discrete Series Conversion of Individual Series

If a series has more items than that, it can be divided into discrete items, with the value corresponding to the highest frequency being calculated.

2. DISCRETE SERIES (or Frequency Array)

There are two methods to determine mode in a discrete series or Frequency Array:

(i) Mode by Observation, known as Inspection Method

(ii) Mode by Grouping Method.

Let us discuss these two methods in detail:

(i) Mode by Observation

The mode can be determined by inspection if:

- Frequencies are regular and homogeneous; and
- The item with the highest frequency is only one.
- In this scenario, the modal value would be the value that corresponds to the highest frequency.

(ii) Mode by Grouping Method

It is not necessary that the value that occurs most frequently or whose frequency is maximum be always the mode if the frequency distribution is irregular and heterogeneous. The grouping method is typically applied in these situations to determine the mode.

3. CONTINOUS SERIES

The modal class in continuous series is the specific class or group in which the mode is found. Two techniques are employed to ascertain mode:

(i) Inspection or Observation Method
(ii) Grouping Method

(i) Method of Observation

We can use the observation approach to identify Mode if the frequencies are homogeneous, regular, and have a single maximum frequency.

(iii) Grouping Method

As was previously said, the Inspection Method is only useful when the series is homogeneous and regular. It is best to use the grouping method when there is any irregularity.

RELATIONSHIP BETWEAN MEAN, MEDIAN AND MODE

1. Symmetrical Distribution:

For symmetrical curves, Mean (X) = Median (Me) = Mode (Z), since the values of the mean, median, and mode are equivalent in a symmetrical distribution.

2. Asymmetrical Distribution:

The majority of distributions in real life are not symmetrical. The mean, median, and mode of an asymmetrical series have distinct values. The height of the curve in the life distribution is not at the middle, making it non-bell shaped. A distribution that is asymmetrical (skewed) might be either favorably or negatively skewed.

- The majority of the values of observations in a distribution that is positively skewed fall to the right of the mode. These measurements will be in the following order of magnitude: Mean > Median > Mode.
- Values of smaller magnitude are concentrated more to the left of the mode in a negatively skewed distribution. These measurements will have the following order of magnitude: Mean < Median < Mode.

MERITS AND DEMERITS OF MODE

Merits of Mode

1. Simplicity: Calculating it is quite easy and straightforward. It doesn't require a lot of mathematical computation to calculate. Inspection can be used to determine the mode in a single, discrete series.
2. Not Affected by Extreme Item Values: The values of extreme items have no effect on it. Even in the absence of knowledge about these extreme data, it can be computed.
3. Possibility in Open-end Series: Without determining the class boundaries, it can be ascertained in open-end distributions.
4. Beneficial for both qualitative and quantitative information: Both quantitative and qualitative data can be described by a mode.
5. Graphic determination: A histogram can be used to make this judgment.
6. Representative value: The mode is the value that best captures the data since it is the one around which observations are concentrated the most.

Demerits of Mode

1. The mode is not a strictly defined : measure because there are multiple ways to determine its value.
2. Not Based on Every Observation in a Series: The mode value is not derived from every observation in the series.

3. Not Algebraically Treatable: The mode is unsuitable for additional mathematical manipulation. For instance, we are unable to ascertain the overall mode of the combined data from the modal values and the sizes of two or more series.

4. Indeterminate: It's possible that the mode's value will never be known. Finding the modal class in the case of multi- and bi-modal distributions is challenging.

5. Affected by sample fluctuations: The mode is far more impacted by sampling fluctuations than the mean is.

COMPARISON BETWEEN MEAN, MEDIAN AND MODE

The notions of mean, median, and mode have all been thoroughly covered. Nonetheless, the decision of which approach to apply for a particular set of data is based on a number of factors that fall under the following general categories:

1. rigorously defined: The mode is not rigorously defined in all circumstances, but the mean and median are.

2. Based on all observations: Every observation should serve as the foundation for a suitable average. The mean alone satisfies this requirement; the median and mode fall short.

3. Have sampling stability: When the least amount of sampling changes is required, the mean should be preferred.

4. Additional algebraic processing: It ought to be amenable to additional mathematical processing. Since only mean can satisfy this property, mean is used as a measure of central tendency in the majority of statistical theories.

5. Easy to compute and comprehend: An average need to be simple to compute and comprehend. All three of the averages meet this requirement.

6. Not impacted by extreme values: The extreme observations shouldn't have an excessive impact on it.

FORMULAE AT A GLANCE

1. MEDIAN	
Individual Series	Me = Size of $[\frac{N+1}{2}]$ Average of two items lying on either side of $[\frac{N+1}{2}]$

Discrete series	Me = size of $[\frac{N+1}{2}]$ item
Continuous series	Determine Median Class as $[\frac{N}{4}]$ item and apply the formula: Me $= l_1 + \frac{\frac{N}{2} - c.f.}{f} \times i$
2. LOWER QUARTILE	
Individual series	$Q_1 = Size\ of\ (\frac{N+1}{4})$ item
Discrete series	Q_1 = $Size\ of\ (\frac{N+1}{4})$ item
Continuous series	Determine Quartile Class as $[\frac{N}{4}]$ item and apply the formula: Me $= l_1 + \frac{\frac{N}{4} - c.f.}{f} \times i$
3. UPPER QUARTILE	
Individual series	$Q_3 = Size\ of\ 3\ (\frac{N+1}{4})$ item
Discrete series	$Q_3 = Size\ of\ 3\ (\frac{N+1}{4})$ item
Continuous series	Determine Quartile Class as 3 $[\frac{N}{4}]$ item and apply the formula: $Q_3 = l_1 + \frac{\frac{3N}{4} - c.f.}{f} \times i$
4. MODE	
Individual series	Mode is the value, which occurs largest number of times.
Discrete series	If the frequencies are regul;ar and homogeneous and there is a single maximum frequency, then mode is the value corresponding to the highest frequency

	(otherwise use grouping method)
Continuous series	Step 1 : Determine the modal class : (i) by inspection, if frequencies are regular homogenous and there is a sngle maximum frequebncy. Otherwise (ii) Grouing method Step 2: Apply the following formula : Mo = $l_1 + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times i$