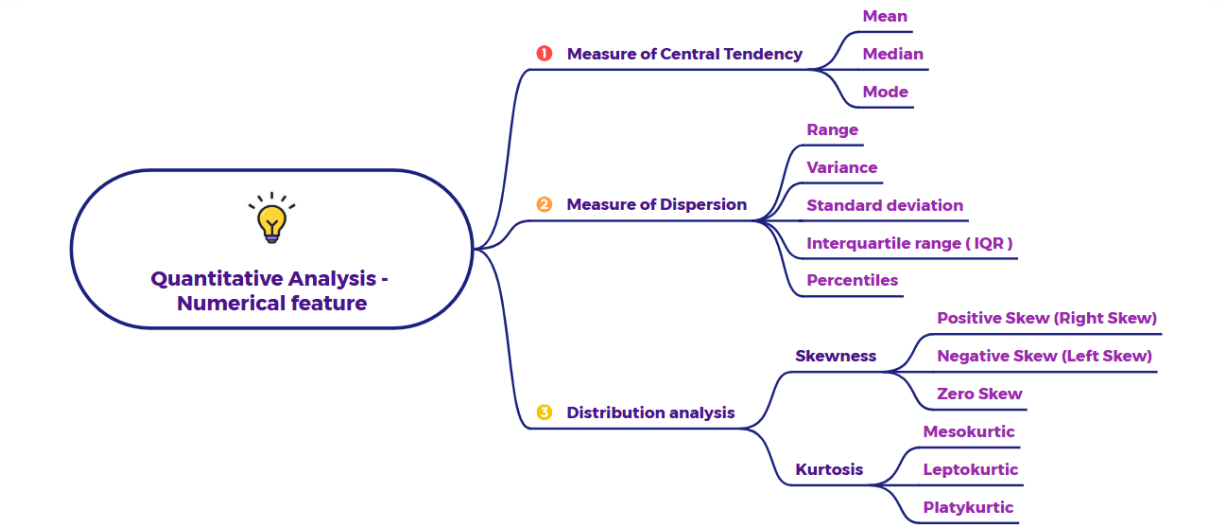
Different types of Quantitative analysis for Numerical variable



1. Measure of Central Tendency: Finding the "Average"

Imagine you have a bunch of scores from a test. Measures of central tendency help you find a typical or central value within that set of scores. It's like trying to find the "middle ground" or the most representative value.

- **Mean:** This is what most people think of as the "average." You add up all the numbers and then divide by how many numbers there are. It gives you a balanced central point.
- **Median:** This is the middle value when you arrange all the numbers in order from smallest to largest. If you have an even number of values, it's the average of the two middle ones. The median is useful because it's not as affected by extreme high or low values (outliers) as the mean.
- **Mode:** This is the value that appears most frequently in your dataset. There can be one mode, more than one mode, or no mode at all if all values are unique. It tells you what's the most common occurrence.

So, central tendency helps us get a quick sense of the "typical" value in our numerical data.

2. Measure of Dispersion: Understanding the Spread

Now, just knowing the "average" isn't the whole story. We also want to know how spread out or varied the numbers are. Measures of dispersion tell us about this variability.

- **Range:** This is the simplest measure – it's just the difference between the highest and the lowest value in your dataset. It gives you a basic idea of the total spread.
- **Variance:** This gives you an idea of how much the individual data points deviate from the mean, on average. A higher variance means the data points are more spread out.
- **Standard Deviation:** This is closely related to variance. It's essentially the square root of the variance. The standard deviation is often preferred because it's in the same units as your original data, making it easier to interpret the spread. A small standard deviation indicates that the data points are clustered closely around the mean, while a large standard deviation [1] suggests they are more spread out.
- **Interquartile Range (IQR):** Imagine dividing your ordered data into four equal parts (quartiles). The IQR is the range of the middle 50% of the data – it's the difference between the value at the 75th percentile and the value at the 25th percentile. Like the median, the IQR is less sensitive to outliers.
- **Percentiles:** These divide your data into 100 equal parts. For example, the 25th percentile is the value below which 25% of the data falls. Percentiles help you understand the relative standing of a particular value within the dataset.

So, dispersion helps us understand how much the data points differ from each other and from the central tendency.

3. Distribution Analysis: Looking at the Shape

Finally, distribution analysis helps us understand the overall shape of our data when we visualize it (like in a histogram). It tells us about the pattern of how the values are spread.

- **Skewness:** This tells us about the symmetry of the distribution.

    - **Positive Skew (Right Skew):** The tail on the right side of the distribution is longer. This often means there are some unusually high values pulling the mean to the right of the median.
    - **Negative Skew (Left Skew):** The tail on the left side of the distribution is longer. This often indicates some unusually low values pulling the mean to the left of the median.
    - **Zero Skew:** The distribution is roughly symmetrical, like a bell curve. The mean, median, and mode are usually close to each other.
- **Kurtosis:** This describes the "tailedness" or the peakedness of the distribution compared to a normal distribution.
    - **Mesokurtic:** The distribution has a shape similar to a normal distribution – not too peaked and not too flat tails.
    - **Leptokurtic:** The distribution is more peaked around the mean and has heavier tails (more extreme values) than a normal distribution.
    - **Platykurtic:** The distribution is flatter around the mean and has thinner tails (fewer extreme values) than a normal distribution.

Distribution analysis gives us a visual and descriptive understanding of the pattern and shape of our numerical data.

In a nutshell, quantitative analysis of numerical variables involves looking at where the center is, how spread out the data is, and what the overall shape of the data looks like. These different types of analysis provide a more complete picture of the information contained within the numbers.