Graphical Techniques, Skewness & Kurtosis, and Box Plot

# Graphical Techniques

## Overview

Graphical techniques are visual methods used to present data. They help in understanding the distribution, trends, and patterns within the data, making it easier to draw insights and communicate findings.

## Types

1. Bar Charts:  
- Used for comparing categorical data.  
- Each bar represents a category, and the height of the bar represents the value.  
  
2. Histograms:  
- Used for displaying the distribution of a continuous variable.  
- Data is divided into intervals (bins), and the height of each bar shows the frequency of data points within each interval.  
  
3. Pie Charts:  
- Used for showing the proportions of categorical data.  
- Each slice represents a category's proportion of the total.  
  
4. Line Graphs:  
- Used for showing trends over time.  
- Data points are connected by lines, illustrating changes over intervals.  
  
5. Scatter Plots:  
- Used for showing relationships between two continuous variables.  
- Each point represents a pair of values, revealing correlation patterns.

# Skewness & Kurtosis

## Overview

Skewness and kurtosis are statistical measures that describe the shape of a data distribution.

## Skewness

Definition: Skewness measures the asymmetry of a distribution.  
Types:  
- Positive Skew (Right Skew): The right tail is longer; most data points are on the left.  
- Negative Skew (Left Skew): The left tail is longer; most data points are on the right.  
Interpretation: Skewness helps in identifying the direction and extent of skew.

## Kurtosis

Definition: Kurtosis measures the "tailedness" of a distribution.  
Types:  
- Leptokurtic: High kurtosis; data has heavy tails and a sharp peak.  
- Mesokurtic: Normal kurtosis; similar to a normal distribution.  
- Platykurtic: Low kurtosis; data has light tails and a flat peak.  
Interpretation: Kurtosis helps in understanding the presence of outliers and the extremity of values.

# Box Plot

## Overview

A box plot (or box-and-whisker plot) is a graphical representation that summarizes a dataset's distribution.

## Components

Box: Represents the interquartile range (IQR) from the first quartile (Q1) to the third quartile (Q3).  
Median Line: A line inside the box showing the median (Q2).  
Whiskers: Lines extending from the box to the smallest and largest values within 1.5 \* IQR from Q1 and Q3, respectively.  
Outliers: Data points outside the whiskers, plotted as individual points.

## Interpretation

Center: The median line indicates the dataset's central tendency.  
Spread: The length of the box and whiskers shows the variability.  
Skewness: The position of the median line within the box indicates skewness.  
Outliers: Points outside the whiskers highlight potential outliers.