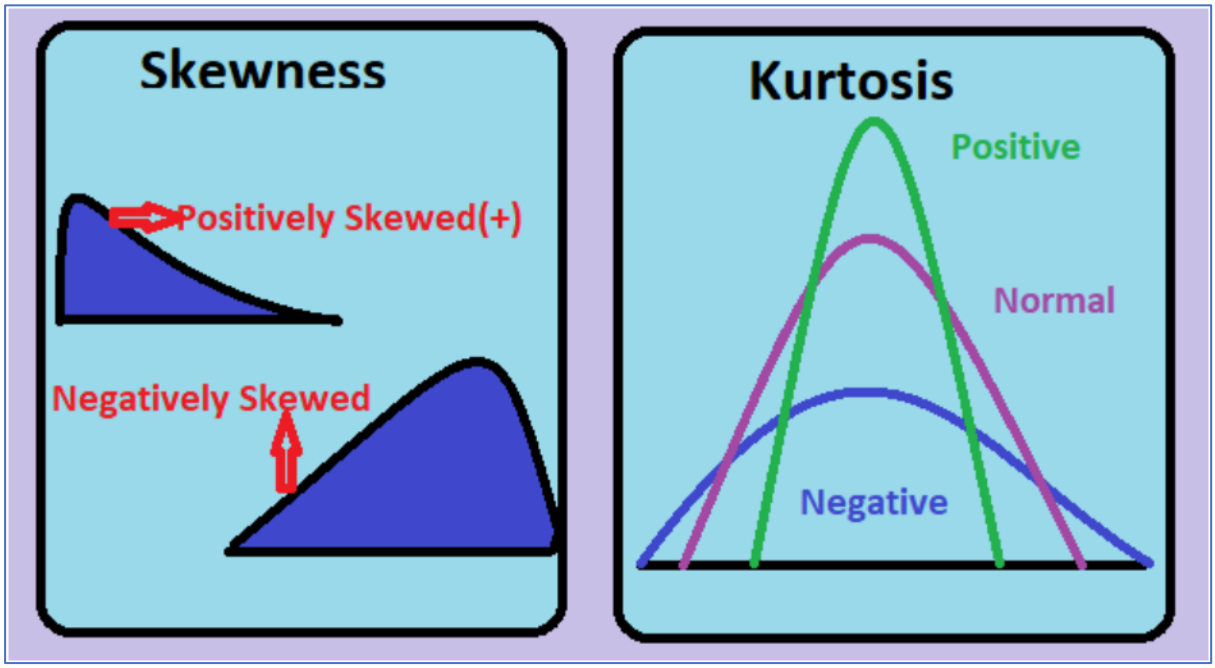
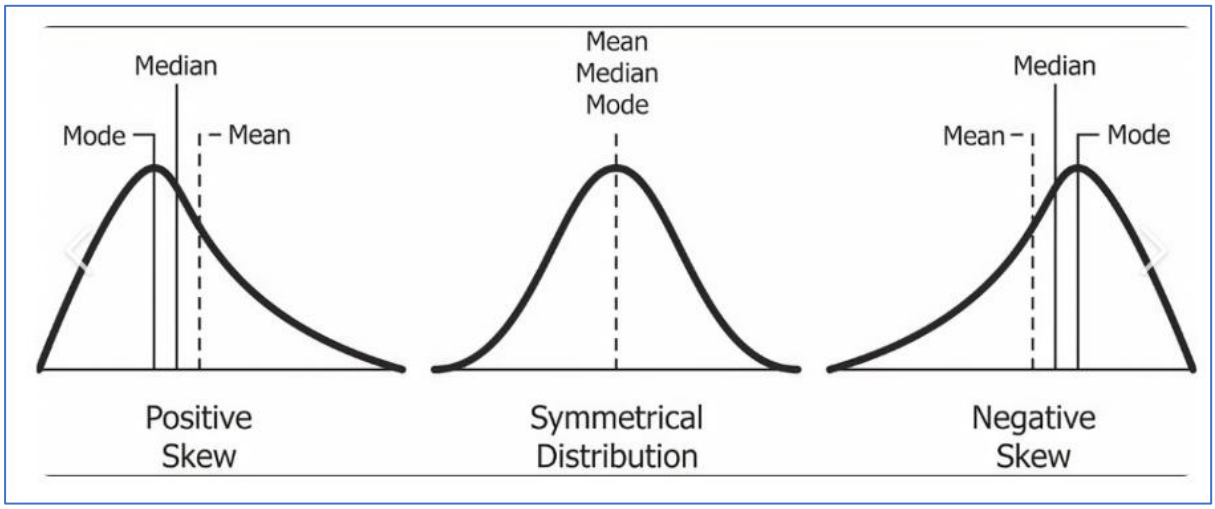
Skewness and kurtosis

Skewness and kurtosis are statistical measures describing a data distribution's shape, with **skewness** measuring asymmetry (left/right leaning) and **kurtosis** measuring peaked Ness and "tailed Ness" (outlier presence) compared to a normal distribution. Positive skew means a long tail to the right (more high values), negative skew means a long tail to the left (more low values), while zero skew is symmetrical. High kurtosis (leptokurtic) indicates heavier tails and a sharper peak (more outliers), while low kurtosis (platykurtic) means lighter tails and a flatter peak.

Diagram:





Skewness: Asymmetry

- **Definition**: How lopsided the data is.
- **Positive Skew (Right-Skewed)**: Tail extends to the right; mean > median > mode (e.g., income data).

- **Negative Skew (Left-Skewed)**: Tail extends to the left; mean < median < mode (e.g., test scores where most do well).
- **Zero Skew**: Symmetrical distribution (e.g., normal distribution).

Kurtosis: Peaked Ness & Tails

- **Definition**: How heavy the tails are, indicating outlier presence.
- **Leptokurtic (Positive Kurtosis)**: Peaked, heavy tails (more extreme outliers).
- **Platykurtic (Negative Kurtosis)**: Flat, light tails (fewer extreme outliers).
- **Mesokurtic**: Same peak/tails as a normal distribution (kurtosis = 3, or 0 in excess kurtosis).

Why They Matter

- **Data Understanding**: Reveals patterns beyond mean/dispersion.
- **Model Reliability**: Checks for normality, vital for many statistical tests.
- **Risk Assessment**: In finance, high kurtosis signals higher risk of extreme events.