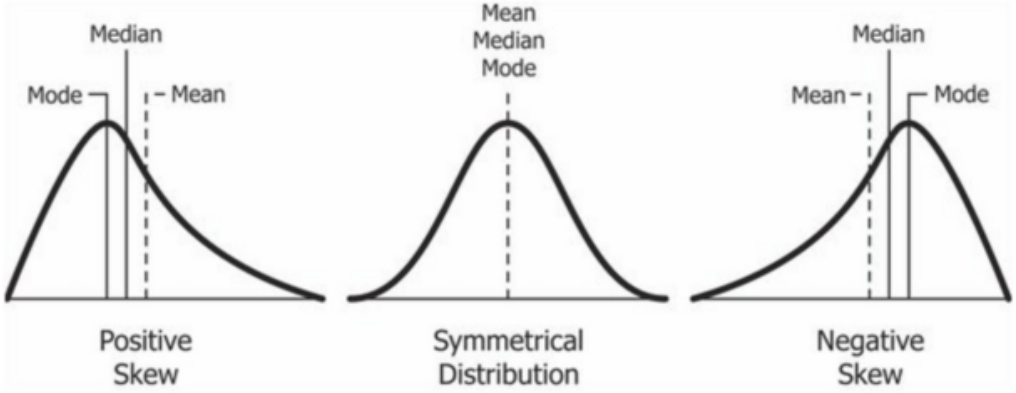
Data Analyst Perspective

STATISTICS

DESCRIPTIVE STATISTICS

Measure Of Shape : Part 1

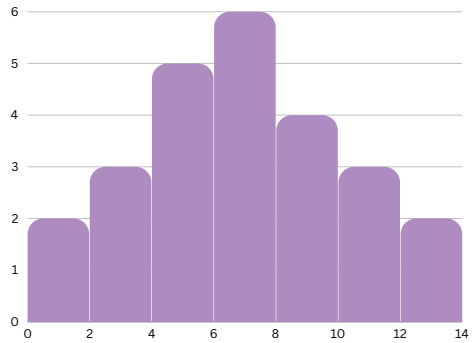


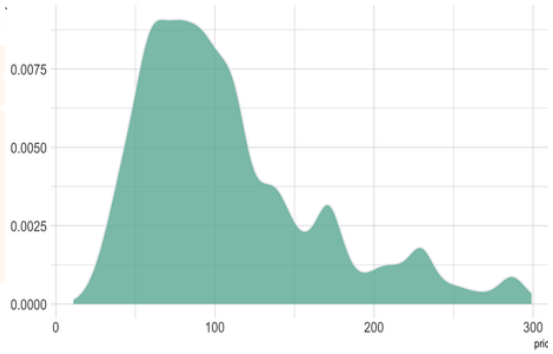
Agenda

- WHAT IS SHAPE
- KEY COMPONENTS OF SHAPE
- GET TO KNOW EACH COMPONENTS
- CHARACTERISTICS & APPLICATIONS OF EACH COMPONENTS

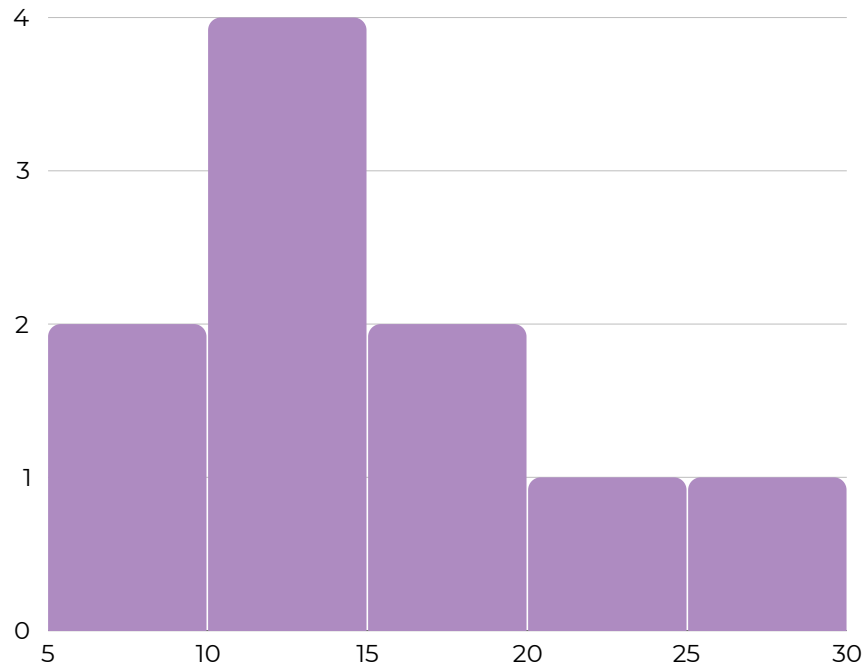
WHAT
IS SHAPE

- **Data distribution** is created by the way individual data points are spread out across different values, and it represents **Shape of the data.**
- This shape can be visualized using graphical representations like **histograms** or **density plots.**





Data Points :5,11,20,12,15,10,9,19,14,25



Shape & Distribution

KEY COMPONENTS
OF SHAPE

- Skewness (Asymmetry)
- Kurtosis (Tailedness)

SKEWNESS

- Skewness measures the **asymmetry of the distribution** around its mean.
- In simpler terms, it tells us whether the **data is tilted** more towards one side (left or right) rather than being evenly distributed around the mean.
- It helps to understand the shape of the data's distribution relative to the **normal distribution.**

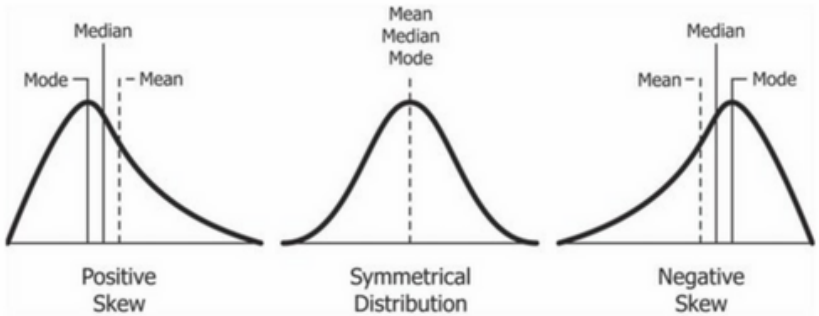
SKEWNESS

- **Methods to Find out Skewness of a Dataset :**

1. Visually Inspecting the plot/visualization
2. Using Formula

- **Types of Skewness :** On the basis of shape or value of Skewness we have 3 types of Skewness:

1. Positive Skewness (Right-Skewed)
2. Negative Skewness (Left-Skewed)
3. Zero Skewness(Symmetrical



SKEWNESS
FORMULA

$$\text{skewness} = \frac{\sum_{i=1}^{N}(x_i - \bar{x})^3}{(N-1)s^3}$$

where:

- s is the standard deviation
- $\bar{x}$ is the mean of the distribution
- N is the number of observations of the sample

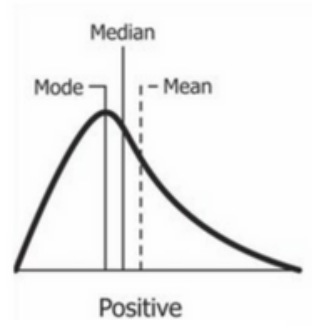
- Formula looking dangerous, **trust me it's not** ,just you have to combine the basic measures which we studied previously .
- Also ,don't worry we have predefined functions for all the statistical measures

Interpretation.....

SKEWNESS
INTERPRETATION

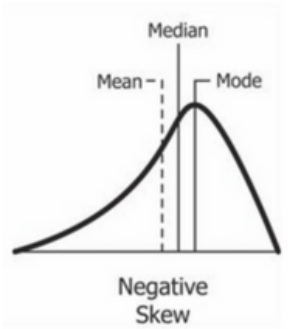
Skewness > 0:

- Positive skewness(right-skewed).
- A longer or fatter right tail.
- There are a few very large values in the data.
- The mean is greater than the median.



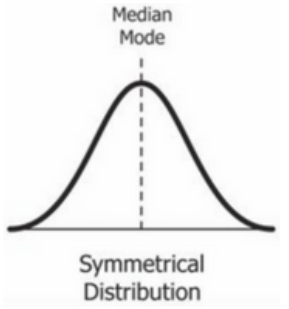
Skewness < 0:

- Negative Skewness (Left-Skewed).
- The tail on the left side is longer or fatter.
- Indicates that there are a few very small values in the data.
- The mean is less than the median.



Skewness = 0:

- Indicates a perfectly symmetrical distribution.
- **0 skewness alone does not necessarily mean that the distribution is a normal distribution.(Imp)**



SKEWNESS
IN DATA ANALYSIS

Outlier Detection:

- Extreme skewness might indicate the presence of outliers that can be investigated separately.

Model Selection :

- Skewness affects the assumptions of many statistical models.
- For instance, linear regression assumes normally distributed residuals; if skewness is present, this assumption might be violated.

Data Transformation:

- If skewness is extreme, data transformations (e.g., log transformation) may be applied to normalize the data for further statistical analysis.

SKEWNESS
APPLICATIONS

Financial Markets:

- Skewness is crucial in finance, where it helps assess the **risk of investment returns**.
- A right-skewed return distribution implies more chances of unusually high returns, but a left-skewed distribution suggests a higher likelihood of extreme losses.

Healthcare:

- In medical studies, skewness can reveal the distribution of a variable like **cholesterol levels**, indicating if most patients have low levels with a few having dangerously high levels.

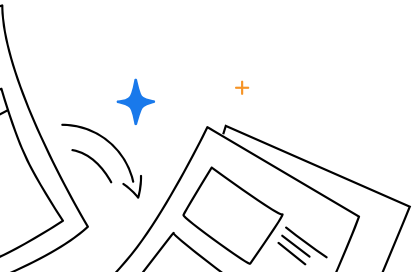
Marketing/Product Launch:

- Skewness in customer ratings after a smartphone launch reveals **overall satisfaction trends**, highlighting niche enthusiasm (positive skew) or broad appeal with a few dissatisfied customers (negative skew).

SUMMARY

- Skewness is a vital concept in statistics, revealing the asymmetry of data distributions. Its practical applications span **finance, quality control, environmental science, healthcare, and more.**

- By understanding skewness, analysts can make more informed decisions, **apply appropriate transformations, and better communicate data insights.**



- Well done friends, we learned a very crucial statistics measure, be **proud of yourself and pat your back.**

- A fun challenge waiting for all the **ice-cream lovers**.

ICECREAM CHALLENGE

Imagine you're hosting an ice cream party and have two flavors of ice cream: Vanilla and Chocolate. You've asked your guests to scoop out their favorite flavors, but they've been scooping unevenly.

Scenario 1:
Vanilla scoops: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10

Scenario 2:
Chocolate scoops: 1, 1, 2, 2, 3, 3, 4, 4, 5, 6

Questions:

1. If you had to guess, which flavor's scoops are more "skewed" and why? (Think about which flavor might have had more scoops clustered around one end or the other.)
2. How would you explain to your guests what skewness means using this ice cream scenario?

THANK YOU

Share your thoughts and feedback !!