

@letthedataconf

Everything to
know about

Skewness

-letthedataconf



What is *Skewness* ?

- Skewness means “**lack of symmetry**”.
- It refers to a distortion or asymmetry that deviates from the symmetrical bell curve, or normal distribution, in a set of data.
- We study skewness to have an idea about the shape of the curve which is drawn with the help of the data.

@letthedataconfess



When is a distribution said to be skewed?

A distribution is said to be skewed if:

- Mean, Median, and Mode fall at different points, i.e. $\text{Mean} \neq \text{Median} \neq \text{Mode}$
- Quartiles are not equidistant from Median.
- The curve drawn from the data is not symmetrical but stretched more towards one side than to the other.



@letthedataconfess



What are the Measures of skewness?

Karl Pearson's coefficient of skewness

1. Formula = $(\text{Mean} - \text{Mode}) / \text{Standard Deviation}$
2. Limits for Karl Pearson's coefficient of skewness are -3 to +3.
3. If Mean = Median = Mode; Skewness = 0. Hence, for a symmetrical distribution, i.e. for which Skewness is zero, mean median and mode coincide.

@letthedataconfess



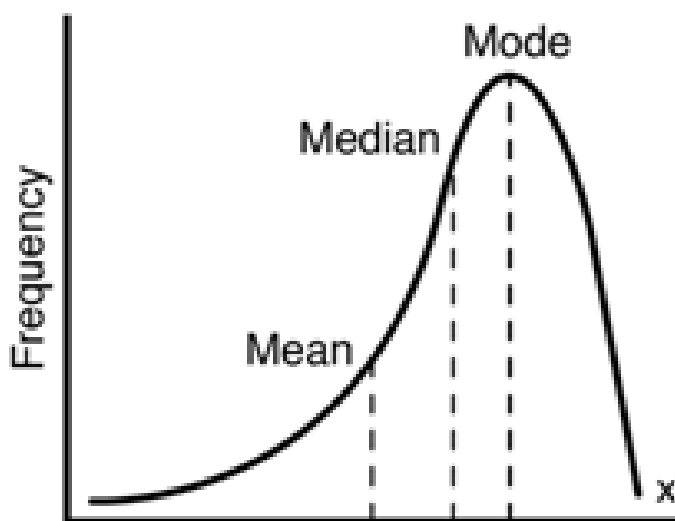
Bowley's coefficient of skewness

1. Formula =
$$\frac{(Q3 - \text{Median}) - (\text{Median} - Q1)}{(Q3 - \text{Median}) + (\text{Median} - Q1)}$$
2. Limits for Bowley's coefficient of skewness are -1 to +1.
3. If $Q3 - \text{Median} = \text{Median} - Q1$; Skewness = 0.
Hence, for a symmetrical distribution, the median is equidistant from the upper and lower quartiles.
4. Used when the mode is ill-defined or extreme observations are present in the dataset.

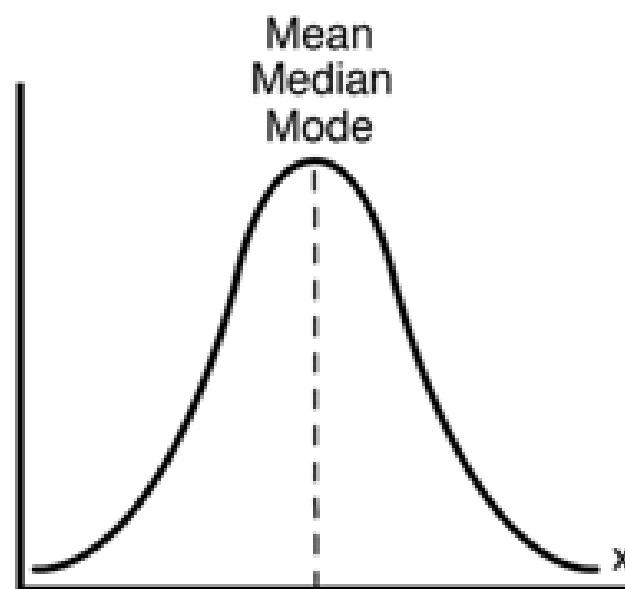
@letthedataconfess



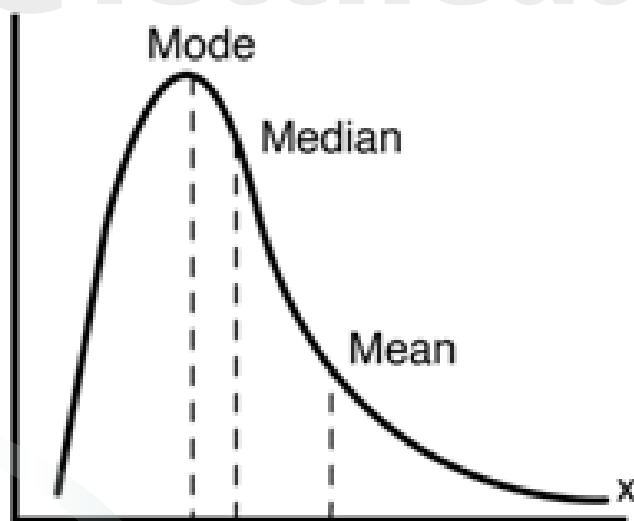
Right and Left Skewness



(a) Negatively Skewed



(b) Normal (no skew)



(c) Positively skewed



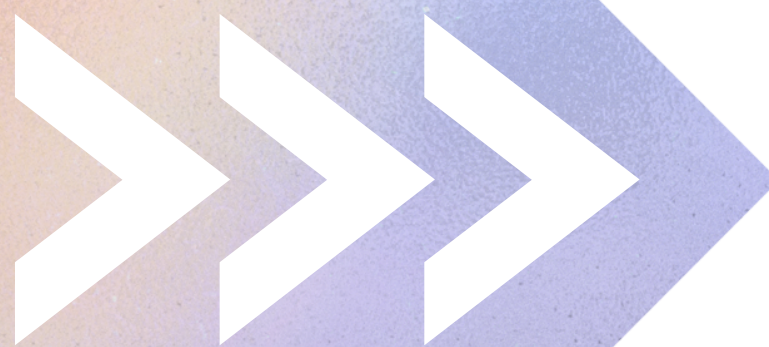
Right Skewness

Also known as **Positive Skewness**

Happens when the larger tail of the distribution lies towards the higher values of the variate, i.e. if the curve drawn with the help of the data is stretched more towards the right than to the left.

In this case, $\text{Mean} > \text{Median} > \text{Mode}$

@letthedataconfess



Left Skewness

Also known as **Negative Skewness**

Happens when the larger tail of the distribution lies towards the lower values of the variate, i.e. if the curve drawn with the help of the data is stretched more towards the left than to the right.

In this case, $\text{Mean} < \text{Median} < \text{Mode}$

@letthedataconfess



www.letthedataconfess.com



@letthedataconfess

LET THE DATA CONFESS

Understand | Learn | Code | Implement



Ace your next data science interview

Practice Now: <https://www.letthedataconfess.com/mock-interview/>

