

SKEWNESS AND KURTOSIS

Skewness is the measure that indicates asymmetry distribution of data in the given dataset. When skewness is “Zero”, it indicates symmetrical distribution. Meaning, the dataset is balanced. If skewness is less than zero, it is negatively skewed where $\text{Mode} > \text{Median} > \text{Mean}$. If skewness is greater than zero, it is positively skewed where $\text{Mean} > \text{Median} > \text{Mode}$.

Kurtosis is statistical measure that describes shape of distribution, specifically the tailedness and peakness of a probability distribution. It helps to measure mean and standard deviation by providing insights into the distribution’s peak and tail behaviour.

If Kurtosis=3, it is called Mesokurtic with moderate tail and peak. If Kurtosis<3, it is called Platykurtic with lighter tail and flatter peak indicating fewer extreme values. If Kurtosis>3, it is called Leptokurtic with heavier tails and sharper peak indicating more extreme values

For “Placement.csv” dataset,

	sl_no	ssc_p	hsc_p	degree_p	etest_p	mba_p	salary
Kurtosis	-1.2	-0.60751	0.0869008	-0.0974897	-1.08858	-0.470723	-0.239837
Skew	0	-0.132649	0.162611	0.204164	0.282308	0.313576	0.8067

ssc_p, hsc_p, degree_p, etest_p, mba_p and salary, Kurtosis < 3. This indicates that these columns are platykurtic containing fewer extreme values.

For Skewness, ssc_p < 0, which indicates it is negatively skewed $\text{Mode} > \text{Median} > \text{Mean}$. hsc_p, degree_p, etest_p, mba_p and salary are approximately equal to 0, which indicates balanced distribution $\text{Mean} = \text{Median} = \text{Mode}$.