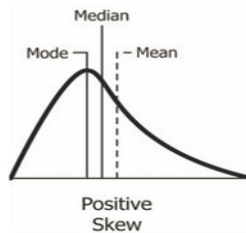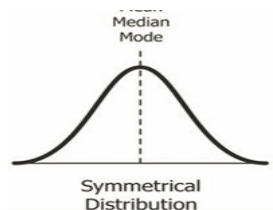# SKEWNESS & KURTOSIS

Skewness:

- It is a measure of symmetry, or more precisely asymmetry of data distribution, indicates how the data is spread out in a dataset.
- If skew > 0 , it is right-skewed or that right tail is longer or the peak value lies in the left-hand side.
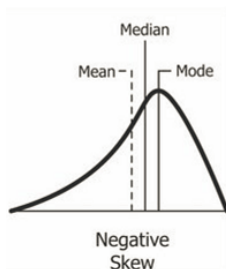

Mean > Median > Mode

- If the skew $= 0$, then the peak value is lie in the center.

Mode = Median = Mean

- If the skew < 0, it is left skewed and the peak value lies in the right-hand side.

Mean < Median <Mode

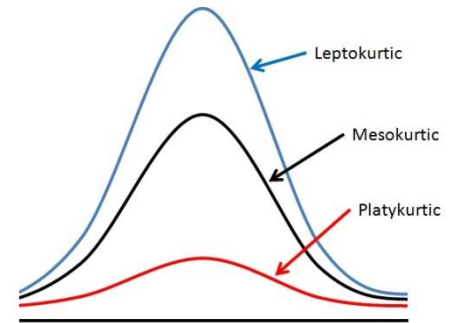|	sl_no	ssc_p	hsc_p	degree_p	etest_p	mba_p	salary
Mean	108	67	66	66	72	62	277648
Median	108.0	67.0	65.0	66.0	71.0	62.0	265000.0
Mode	1	62.0	63.0	65.0	60.0	56.7	300000.0
Q1:25%	54.5	60.6	60.9	61.0	60.0	57.945	240000.0
Q2:50%	108.0	67.0	65.0	66.0	71.0	62.0	265000.0
Q3:75%	161.5	75.7	73.0	72.0	83.5	66.255	300000.0
Q4:100%	215.0	89.4	91.15	88.5	98.0	77.89	390000.0
IQR	107.0	15.1	12.1	11.0	23.5	8.31	60000.0
1.5 Rule	160.5	22.65	18.15	16.5	35.25	12.465	90000.0
Lesser	-106.0	37.95	42.75	44.5	24.75	45.48	150000.0
Greater	322.0	98.35	91.15	88.5	118.75	78.72	390000.0
Min	1	40.89	42.75	50.0	50.0	51.21	200000.0
Max	215	89.4	91.15	88.5	98.0	77.89	390000.0
Skewness	0.0	-0.132649	0.162611	0.204164	0.282308	0.313576	0.8067
Kurtosis	-1.2	-0.60751	0.086901	-0.09749	-1.08858	-0.470723	-0.239837

From the above table of Placement data,

- The skewness lies in positive value. ie. Skew >0 for columns hsc_p, degree_p, etest_p, mba_p, salary. So for these columns Mean > Median > Mode. Peak value lies in the left side of the curve so
  * The highest value score in hsc_p columns lies before the mid value of 65 marks.
  * The highest value score in degree_p columns lies before the mid value of 66 marks.
  * The highest value score in etest_p columns lies before the mid value of 71 marks.
  * The highest value score in mba_p columns lies before the mid value of 62 marks.
  * The highest value score in salary_p columns lies before the mid value of 265000 marks.

- The skewness lies in negative value. ie. Skew <0 for column ssc_p. So for these column Mean< Median < Mode. Peak value lies in the right side of the curve so the highest marks score in ssc_p columns is after mid point of 67 marks.

Kurtosis:

- It is measure of Peakness or measure of tailedness of a distribution.
- Kurtosis (K) < 3 it is said to be Platykurtic.
- If K = 3, then it is Mesokurtic.
- If K >3, then it is Leptokurtic.

	sl_no	ssc_p	hsc_p	degree_p	etest_p	mba_p	salary
Mean	108	67	66	66	72	62	277648
Median	108.0	67.0	65.0	66.0	71.0	62.0	265000.0
Mode	1	62.0	63.0	65.0	60.0	56.7	300000.0
Q1:25%	54.5	60.6	60.9	61.0	60.0	57.945	240000.0
Q2:50%	108.0	67.0	65.0	66.0	71.0	62.0	265000.0
Q3:75%	161.5	75.7	73.0	72.0	83.5	66.255	300000.0
Q4:100%	215.0	89.4	91.15	88.5	98.0	77.89	390000.0
IQR	107.0	15.1	12.1	11.0	23.5	8.31	60000.0
1.5 Rule	160.5	22.65	18.15	16.5	35.25	12.465	90000.0
Lesser	-106.0	37.95	42.75	44.5	24.75	45.48	150000.0
Greater	322.0	98.35	91.15	88.5	118.75	78.72	390000.0
Min	1	40.89	42.75	50.0	50.0	51.21	200000.0
Max	215	89.4	91.15	88.5	98.0	77.89	390000.0
Skewness	0.0	-0.132649	0.162611	0.204164	0.282308	0.313576	0.8067
Kurtosis	-1.2	-0.60751	0.086901	-0.09749	-1.08858	-0.470723	-0.239837

- From the above table , all the columns lies K < 3, So it is Platykurtic.