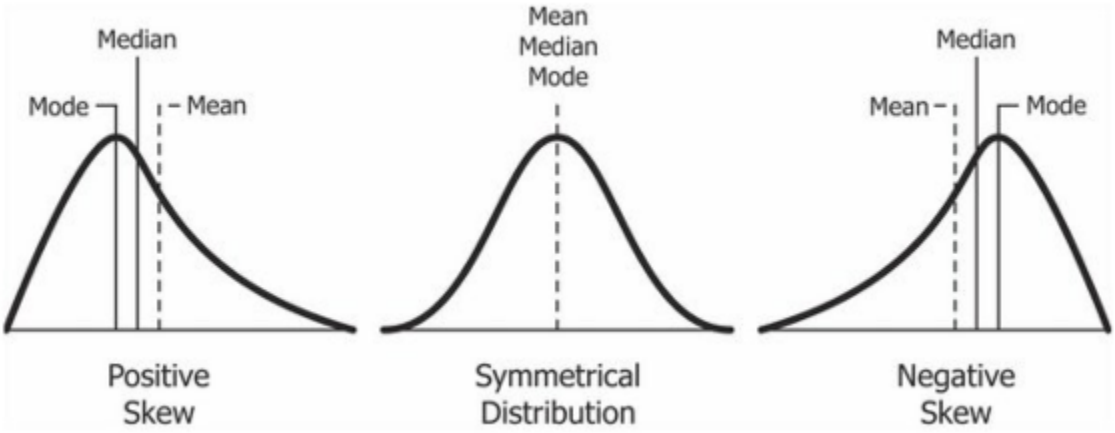
HOPE AI

Skew | Kurtosis Analysis Document

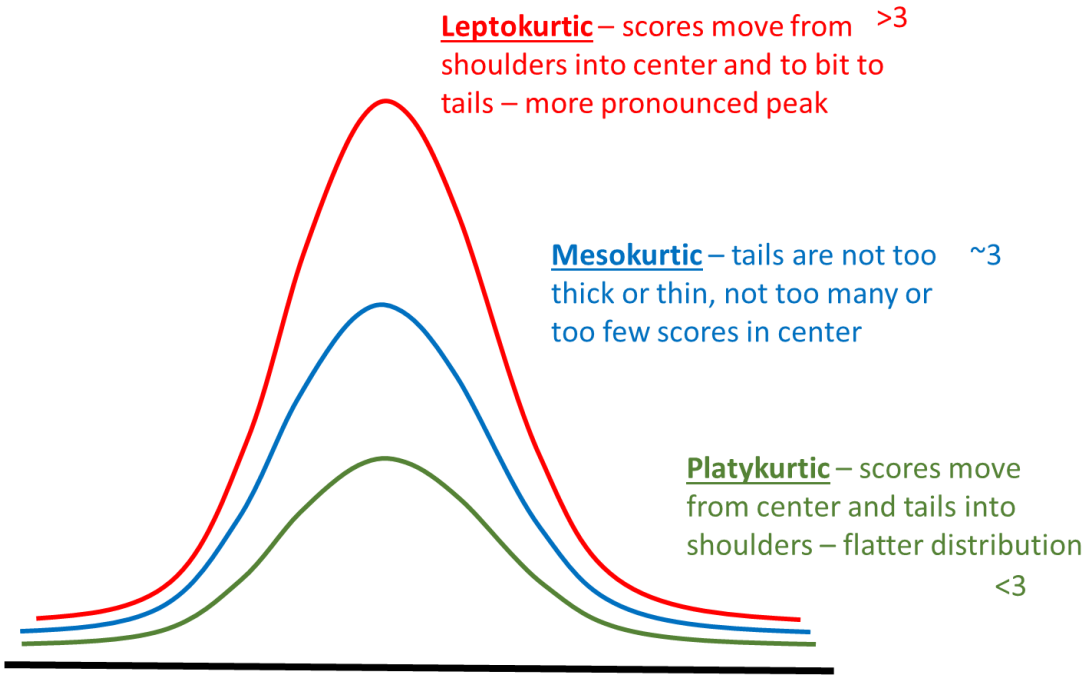
Github link of Data:

https://github.com/JayachandraPrabha/3.Data-Science/blob/main/Placement.csv

1) **Skew:**
   - Skew is nothing but the peak position on a bell curve.
   - When data points are not distributed symmetrically to the left/right sides of the median.
   - If the bell curve is shifted to the left or the right, it is said to be skewed.



2) **Kurtosis:**
   - Happenings/broadness in the bell curve
   - Frequency distribution is peaked in comparison with a normal curve.
   - It is the degree of peakedness of a distribution



Let us consider the given data, some of the basic informations about the given data are,

1) The overall shape of the dataset: **215 rows × 15 columns.**

2) The data basically deals with the survey of the candidates who have **placed and not placed** in the placement and the columns in the data are,
   - 'sl_no', 'gender', 'ssc_p', 'ssc_b', 'hsc_p', 'hsc_b', 'hsc_s', 'degree_p', 'degree_t', 'workex', 'etest_p', 'specialization', 'mba_p', 'status', 'salary'
   - In the above mentioned columns there were qualitative columns / variables (Categorical data) and quantitative columns / variables (Numerical data), hence both were separated.

3) Below is the obtained data from the dataset:

	sl_no	ssc_p	hsc_p	degree_p	etest_p	mba_p	salary
Mean	108.0	67.303395	66.334744	66.358558	72.100558	62.278186	277648.648649
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
min	1.0	40.89	42.75	50.0	50.0	51.21	200000.0
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
99%	212.86	87.0	91.129	83.86	97.0	76.1142	NaN
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
1.5Rule	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
kurtosis	-1.2	-0.60751	0.086901	-0.09749	-1.08858	-0.470723	-0.239837
skew	0.0	-0.132649	0.162611	0.204164	0.282308	0.313576	0.8067

4) From the obtained insights, let us segregate the need informations:

colName	Skew Value	Skew Info	Kurtosis Value	Kurtosis Info
ssc_p	-0.1326	Negative	-0.6075	Platykurtic
hsc_p	0.1626	Positive	0.0869	Platykurtic
degree_p	0.2041	Positive	-0.0974	Platykurtic
etest_p	0.2823	Positive	-1.0885	Platykurtic
mba_p	0.3135	Positive	-0.4707	Platykurtic
salary	0.8067	Positive	-0.2398	Platykurtic

Result:

Hence the skew & kurtosis in the given data were as follows:

- In the given data,all are positive skew except in ssc_p **(-0.1326)** is considered as negative skew.
- In kurtosis, all the values are <3 so, all lies in platykurtic kurtosis.
