

# Skewness and Kurtosis

## Example: Placement data analysis

We have a list of 215 students details like marks(percentage), course details, work experience, placement details like placed/ not placed and salary.

With the above details we took the Quantitative(numeric) columns and found the **skewness** and **kurtosis** from the Placement list.

	sl_no	ssc_p	hsc_p	degree_p	etest_p	mba_p	salary
<b>Mean</b>	108.0	67.303395	66.334744	66.358558	72.100558	62.278186	277648.648649
<b>Median</b>	108.0	67.0	65.0	66.0	71.0	62.0	265000.0
<b>Mode</b>	1	62.0	63.0	65.0	60.0	56.7	300000.0
<b>Q1:25%</b>	54.5	60.6	60.9	61.0	60.0	57.945	240000.0
<b>Q2:50%</b>	108.0	67.0	65.0	66.0	71.0	62.0	265000.0
<b>Q3:75%</b>	161.5	75.7	73.0	72.0	83.5	66.255	300000.0
<b>99%</b>	212.86	87.0	91.129	83.86	97.0	76.1142	NaN
<b>Q4:100%</b>	215.0	89.4	91.15	88.5	98.0	77.89	390000.0
<b>Min</b>	1	40.89	42.75	50.0	50.0	51.21	200000.0
<b>Max</b>	215	89.4	91.15	88.5	98.0	77.89	390000.0
<b>IQR</b>	107.0	15.1	12.1	11.0	23.5	8.31	60000.0
<b>1.5*IQR</b>	160.5	22.65	18.15	16.5	35.25	12.465	90000.0
<b>Lesser</b>	-106.0	37.95	42.75	44.5	24.75	45.48	150000.0
<b>Greater</b>	322.0	98.35	91.15	88.5	118.75	78.72	390000.0
<b>Skewness</b>	0.0	-0.132649	0.162611	0.204164	0.282308	0.313576	0.8067
<b>Kurtosis</b>	-1.2	-0.60751	0.086901	-0.09749	-1.08858	-0.470723	-0.239837

## Skewness:

- “SSC\_P” percentage is less than zero so it is **negative** skewness (Mean > Median > Mode)
- HSC\_P, DEGREE\_P, ETEST\_P, MBA\_P, SALARY is greater than zero so it is **positive** skewness (Mode > Median > Mean)

There are three types of skewness:

- **Right skew** (also called positive skew). A right-skewed distribution is longer on the right side of its peak than on its left.
- **Left skew** (also called negative skew). A left-skewed distribution is longer on the left side of its peak than on its right.
- **Zero skew**. It is symmetrical and its left and right sides are mirror images.

## Kurtosis:

- “HSC\_P” percentage is greater than zero so it has **positive** excess kurtosis. It is a **Leptokurtic** distribution which has fat-tailed, meaning that they have **many outliers**.
- SSC\_P, DEGREE\_P, ETEST\_P, MBA\_P, SALARY is less than zero so it has **negative** excess kurtosis. It is a **Platykurtic** distribution which has thin-tailed, meaning that they have **few outliers**.

The three categories of kurtosis are:

- **Mesokurtosis**: An excess kurtosis of 0. Normal distributions are mesokurtic.
- **Platykurtosis**: A negative excess kurtosis. Platykurtic distributions are thin-tailed, meaning that they have few outliers.
- **Leptokurtosis**: A positive excess kurtosis. Leptokurtic distributions are fat-tailed, meaning that they have many outliers.

	Mesokurtic	Platykurtic	Leptokurtic
<b>Tailedness</b>	Medium-tailed	Thin-tailed	Fat-tailed
<b>Outlier frequency</b>	Medium	Low	High
<b>Kurtosis</b>	Moderate (3)	Low (< 3)	High (> 3)
<b>Excess kurtosis</b>	0	Negative	Positive
<b>Example distribution</b>	Normal	Uniform	Laplace