Rank starts from 1.

Mean & std.
Median & interquartile.

Date

Q. From a dataset containing 20 students find the 75th percentile.

→ first sort the data and assign rank.

$\frac{20 \times 75}{100} \Rightarrow 15^{th}$ (at the 15th value of the dataset we have 75th percentile.)

* visualize from slides.

Q. From a dataset containing 20 students find the 25th percentile.

$\frac{21 \times 25}{100} \Rightarrow 5.25^{th}$.

↓

Now this rank is not available in our table.

∴ to cater this we will use the interpolation method.

Percentile $\Rightarrow X_5 + (X_6 - X_5) \times \underline{0.25}$ → multiply with whats in points.

$\Rightarrow 68 + (72 - 68) \times 0.25 \rightarrow$ (69)

↓

it mean 69 students measured 25th Percentile or below.

INTERQUARTILE RANGE:-

The IQR measures the spread of the middle 50% of the data.

$iqr = (75^{th} \text{ percentile}) - (25^{th} \text{ percentile})$

✓ $89 - 72 \Rightarrow 17$ it means that the middle 50% of the dataset varies 17 marks

▸ Scaling the data scales interquartile Range

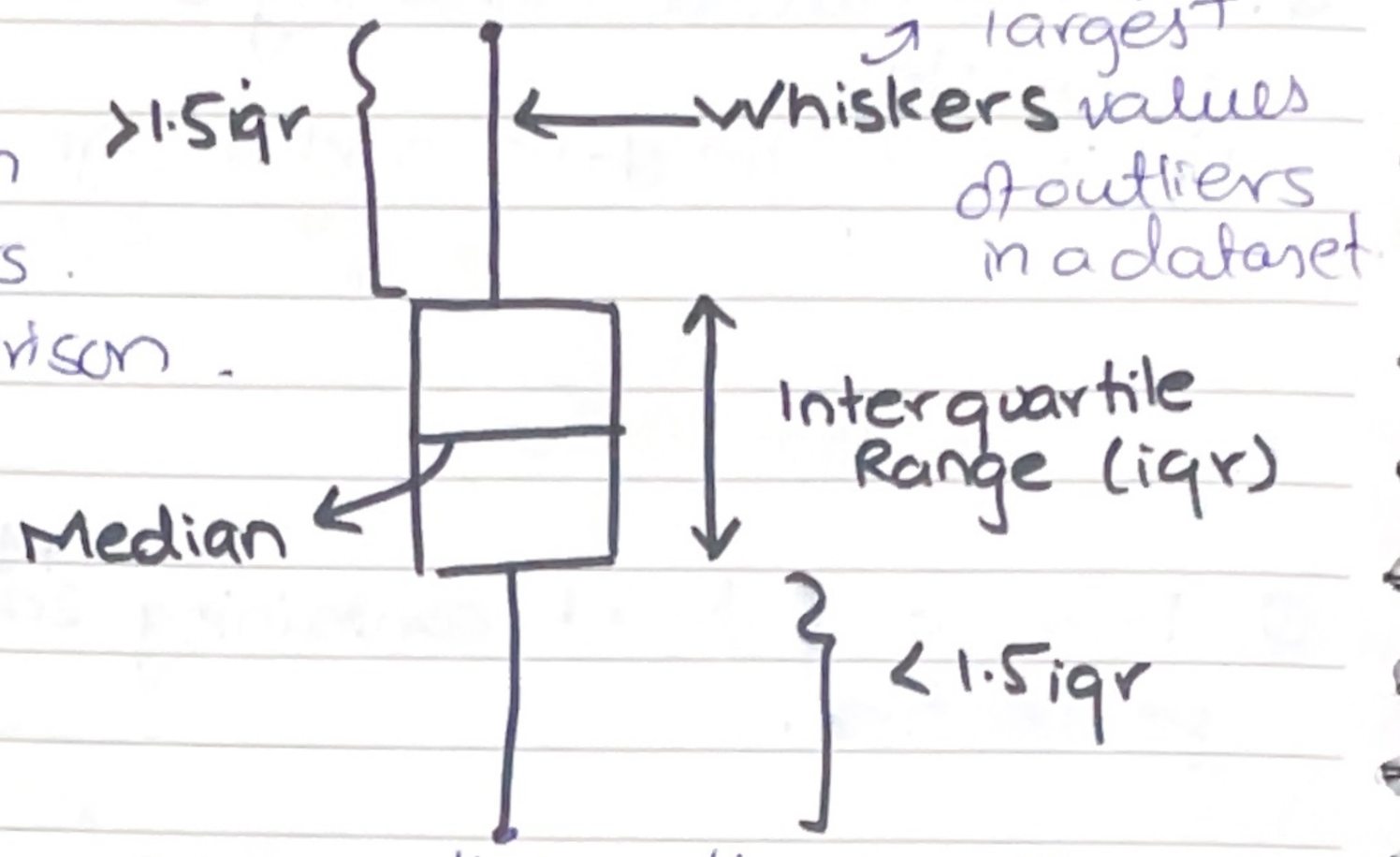
But translation does not change it.

Date

* if there are more peaks or modes then it is apossibility that the data can categorized.

Box Plots

Simpler than histogram
Good for knowing outliers.
Easier to use for comparison.



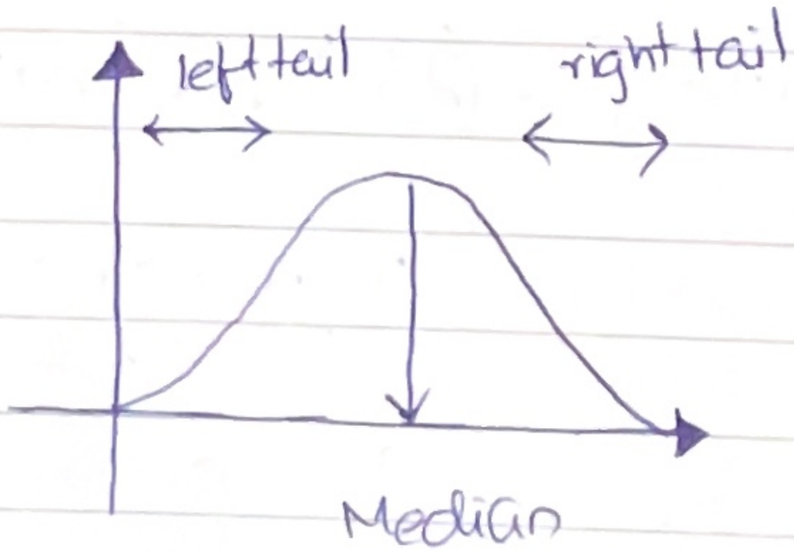
let that interquartile range b/w 25th to 75th percentile is to be founded.

(i) Sort data. ✓
(ii) Ranking data ✓
(iii) Percentile value. ✓
(iv) IQR.
(v) Min & max values to set up ar whiskers.

$$[Q_1 - 1.5(iqr), \; Q_2 + 1.5(iqr)]$$

TYPES OF HISTOGRAM:-

(i) Symmetric

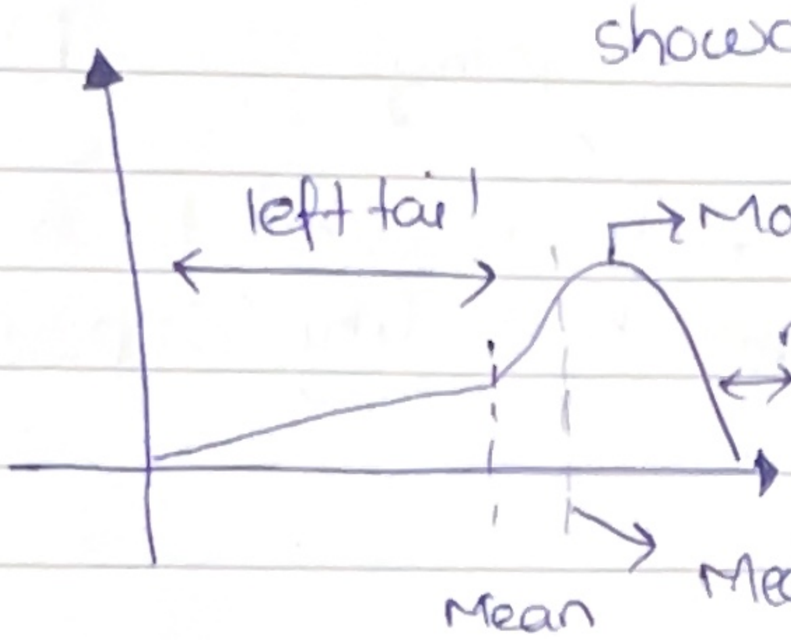


heatmap - 2D data.
bar => count.

3D bar chart: 2D data (hidden structures).

Date

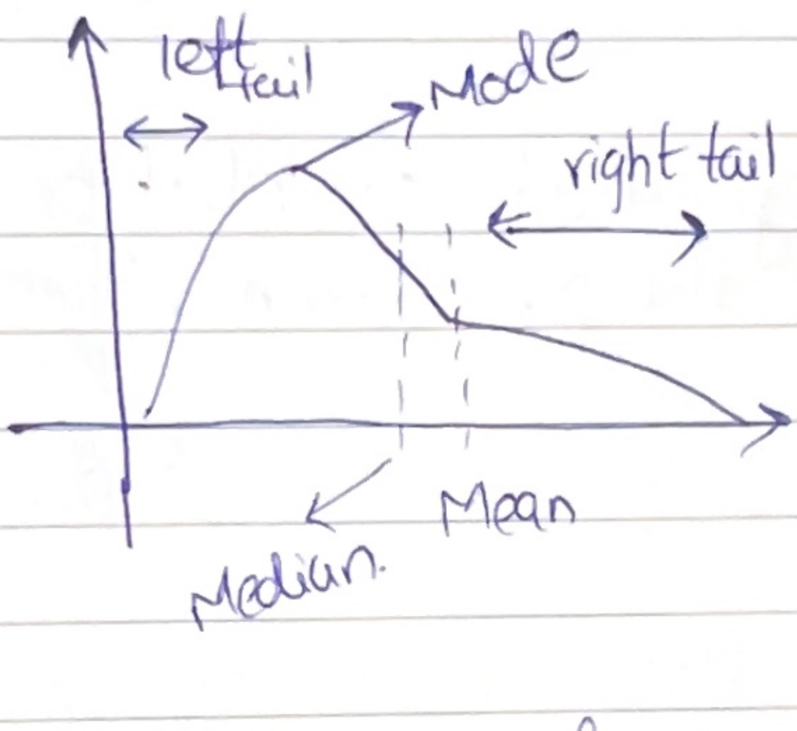
(ii) Left Skew.



heatmaps & 3D bar chart showcase direct functional relationship.

we suse correlation & scatter plot instead for viceversa

(iii) Right Skew.



→ Heatmaps / 3D plots are good for categorical data.
→ Scatterplots are special for describing relation b/w two continuous values. (specially geographical 2D data)