

Range = Range is the difference between the max bin vs min bin val . value.

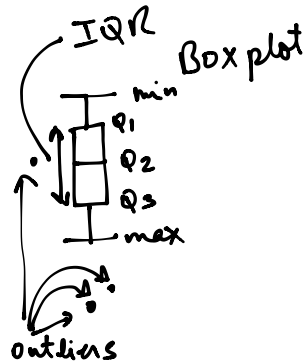
IQR \rightarrow Interquartile range. $Q_3 - Q_1$

Outliers: data far away from the population

Outlier \Rightarrow Outlier $< Q_1 - 1.5(IQR)$

\therefore Outlier $> Q_3 + 1.5(IQR)$

Box plot: we visualize outliers by using Box plot



*mean usually lies inbetween $Q_1 \rightarrow Q_3$, unless there are outliers



standard deviation :

$$\bar{X} \Rightarrow \text{mean} = \bar{X} = \frac{\sum X}{\#}$$

Std. Deviation \Rightarrow

Deviation from the mean $(X_i - \bar{X})$

Average deviation \Rightarrow just an average of the sum of deviations

Avg deviation \Rightarrow is 0, use it for checking work

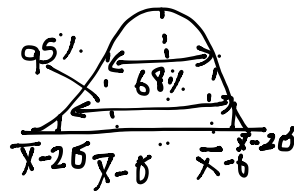
* we can take absolute values or square of $(X_i - \bar{X})$ to compute avg deviation

$$\text{Sum of squared dev} = \sum (X_i - \bar{X})^2 = SS$$

$$\text{variance: } \frac{\sum (X_i - \bar{X})^2}{\# \text{ of items}} \Rightarrow$$

$$\text{standard deviation} \Rightarrow \sqrt{\text{variance}} = \sqrt{\frac{\sum (X_i - \bar{X})^2}{\#}}$$

68% of data falls in between 1 std. deviation.



for sample it is used population

$$S = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n-1}}$$

sample standard deviation
corrected standard deviation

population standard deviation

$$\sigma = \sqrt{\frac{\sum (X_i - \bar{X})^2}{\#}}$$

used for entire data set.