Assignment-2

Ques 1. How can one figure out what the interquartile range is?

Ans It is a measure of where the "middle fifty" is in a dataset

Formulae for $[IQR = Q_3 - Q_1]$

$$\begin{cases} Q_3 = \frac{75}{100} \times (n+1) \\ Q_1 = \frac{25}{100} \times (n+1) \end{cases}$$

Ques 2 What exactly is the value of the 5 number theory?

Ans 5 number summary is used to remove outliers in the data
A summary consist of 5 values

1. Maximum value
2. Minimum value
3. Lower Quartile
4. Upper Quartile
5. Median

All value given summary of a data because each value describes a specific part of a dataset.

1. Median gives center of dataset
2. Quartile span the middle half of the dataset
3. Highest & lowest value provides additional data about actual dispersion of the data

All values can also be used to form Box and whisker plots

Ques 3 What is the relationship b/w standard deviation & variance?

Ans Variance represents the average squared deviation from the mean value of data, while standard Deviation represents square root of that number.

Both measure variability in a distribution.

$$\text{Variance} = (\text{Standard deviation})^2$$

$$V = \sigma^2$$

Ques 4 What does the difference b/w variance & standard deviation mean?

Ans

	Standard Deviation	Variance
①	Square root of variance	The average squared differences from the mean
②	Spread b/w numbers in a data set.	The average degree to which each point differs from the mean
③	↓ Standard duration ↓ volatility ↑ " " ↑ "	Degree to which returns vary or change over time

Ques 5 When is it appropriate to refer skewed data distribution.

Ans When Data is neither symmetric nor normal because the values trail off more sharply on one side than on the other. Right skew Eg wealth Distribution, Left Skew → Eg → Life Span