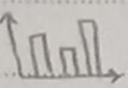
Statistics

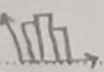
the science of collecting, analyzing and interpreting data.

Descriptive

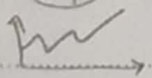
① organizing and summarizing data using numbers & graphs

② Data Summary:

Bar graphs 

Histograms 

pie charts 

line graph 

shape of graph & skewness

Inferential

[1] using sample data to make an inference or draw a conclusion of the population

[2] uses probability to determine how confident we can be that the conclusions we make are correct.

---

[3] - mean ($\bar{x}$) → Average → (Measures of Central Tendency)

$(\bar{x}) = \frac{\sum x}{n}$, $\mu = \frac{\sum x}{N}$ (population)

- median → The middle number
- Mode → the number that occurs most frequently

[4] Range, variance & STD → (Measures of variability)

• mean = $\frac{Sum}{N}$

• median

① برتب الأرقام

② $\frac{n+1}{2}$ → عدد فردي

عدد زوجي ← هنجيب متوسط الرقمين

• Mode → القيمة الأكثر تكراراً

• Range → (الرقم الأعلى - أقل قيمة)

التاريخ: / /

- Variance (s^2)

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$$

Data	$x_i - \bar{x}$	$(x_i - \bar{x})^2$

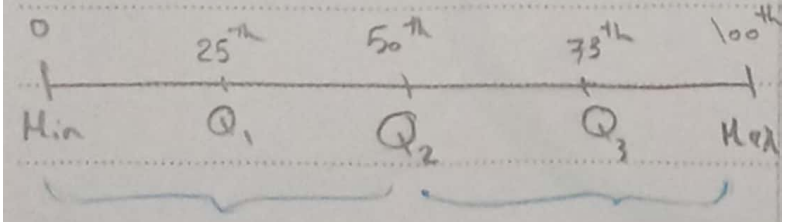
6, 7, 8, 9, 10 $\quad s^2 = 2.5$ $\quad <$ $\quad$ 4, 6, 8, 10, 12 $\quad s^2 = 10$

- Standard deviation (s)

1) $\bar{x}$

2) $s = \sqrt{s^2} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$

- Interquartile Range & Outliers



$$IQR = Q_3 - Q_1$$

Outliers →

$$[Q_1 - 1.5\,IQR\ ,\ Q_3 - 1.5\,IQR]$$

الـ mean بيتأثر بأكبر قيمة في outlier
اللي هي ممكن تكون
على عكس الـ Median اللي مبتأثرش