STatistics → البيانات علم جمع
تحليل البيانات وتفسيرها وعرضها

Descriptive الإحصائيات الوصفية

1- organizing and summarizing data using numbers and graphs:
تنظيم البيانات وتلخيصها باستخدام الأرقام والرسوم البيانية

2- Data summary: Bar Graphs, Histograms, Pie charts, etc.
slope of graph & skewness

3- Measures of central Tendency: Mean, Median & mode
مقاييس النزعة المركزية: المتوسط، الوسيط، والمنوال

4- measures of variability: Range, Variance & standard deviation
تباين النظام والانحراف المعيارى

Mean → $\bar{X} = \dfrac{\sum X}{n}$ جمع القيم
عدد القيم
عدد البيانات

Median الرقم الأوسط

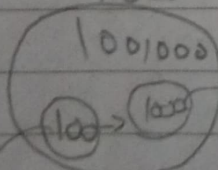## Inferential: الإحصائيات استنتاجية

1. using sample data to make inference or draw a conclusion of the Popula

2. uses probability to determine how confident we can be that the conclusions
We make are correct (Confidence Intervals & Margins of error) هامش الخطأ

XYZ
أخذت عينات مختارة عشوائياً
لأنها أكثر دقة وموثوقية

100,000

6 Descriptive → (100) → (1000)   أنا بقيس عينة من السيارات
blue cars

20/100 —   بختار عينة بشكل عشوائي أو أخترها

17 ± 1 %

20% ± 2 Inferential

12, 7, 14, 5, 7, 11, 9

mean $= \dfrac{sum}{n} = \dfrac{12+7+14+5+7+11+9}{7} = \dfrac{65}{7} = 9.286$

5, 7, 7, 9, 11, 12, 14   منتصف

median = 9

5, 7, 7, 9, 11, 12, 14

mode = 7   الرقم الى ظهر كثير

5, 7, 7, 9, 11, 12, 14

range = 14 − 5 = 9

3, 5, 6, 8, 9, 11, 4

mode = None

5, 5, 6, 8, 8, 8, 9, 10

median = 8

4, 11, 12, 13, 15, 21, 28, 36, 38, 48

median $= \dfrac{15+21}{2} = \dfrac{36}{2} = 18$

How to calculate variance البيانات

مقياس لانتشار البيانات قد تكون بعض بتنشر بشكل متباعد جداعن بعض

البيانات هذاكون اغرب اك المتوسط، وبالتالي التباعد اكبر

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

يشير اك انتشار اكبر في $\overset{mean}{\underline{\quad}}$ البيانات

6, 9, 14, 10, 5, 8, 11

Data	$X_i - \overset{9}{\bar{X}}$	$(Y_i - \bar{X})^2$
5	-4	16
6	-3	9
8	-1	1
9	0	0
10	1	1
11	2	4
14	5	25
63		56

$$\bar{X} = \frac{\sum X_i}{n} = \frac{63}{7} = 9$$

$$S^2 = \frac{56}{6} = 9.3$$

| 6 7 8 9 10 | 4 6 8 10 12 |

$$\bar{X} = \frac{6+7+8+9+10}{5} = 8 \qquad\qquad \bar{X} = \frac{4+6+8+10+12}{5} = 8$$

$$S^2 = \frac{4+1+0+1+4}{4} = \boxed{2.5} \qquad S^2 = \frac{16+4+0+4+16}{4} = \boxed{10}$$

عشان الداتا منتشرة من mean ‏ ← التباين هنا بكبير

التباين هنا أقل ← يعني الداتا اقل تركيزاً أو أكثر انتشاراً من mean

الفرق المعياري ← How to calculate The standard Deviation

82, 93, 98, 89, 88

$$\bar{X} = \frac{450}{5} = 90$$

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

$$S = \sqrt{\frac{(82-90)^2+(93-90)^2+(98-90)^2+(89-90)^2+(88-90)^2}{5-1}}$$

$$S = 5.958$$
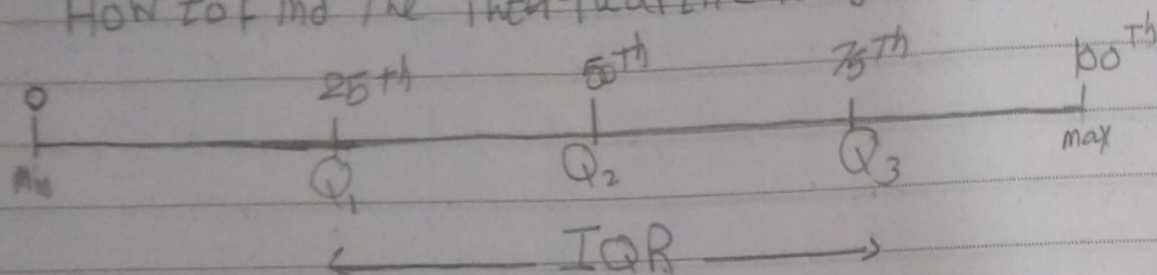
76, 84, 69, 92, 58, 89, 73, 97, 85, 77

$$\bar{X} = \frac{800}{10} = 80$$

$$S = \sqrt{\frac{(80-76)^2 + (84-80)^2 + (69-80)^2 + (92-80)^2 + (58-80)^2 \ldots}{10-1}}$$

$$S = 11.709$$

How to find The interquartile range & any outliers - Descriptive

$$IQR = Q_3 - Q_1$$

outliers $\rightarrow [Q_1 - 1.5\,IQR \;,\; Q_3 + 1.5\,IQR]$

outliers لو قيم خارج هذا المجال تكون خارج عن النطاق و نقول عنها

| 5 | 8 | 15 | 26 | 10 | 18 | 3 | 12 | 6 | 14 | 11 |

3 5 6 8 10 $\boxed{11}$ 12 14 15 18 26
                    Q_2

median = 11

3 5 $\boxed{6}$ 8 10 | 12 14 $\boxed{15}$ 18 26
      Q_1 Q_2 Q_3

$$IQR = Q_3 - Q_1 = 15 - 6 = 9$$

outliers $= [-7.5 \;,\; 28.5]$