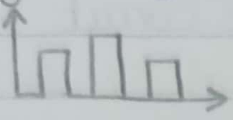
(1) Descriptive

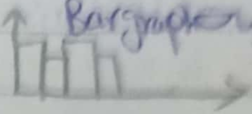
→ Oraganizing and summarizing data using numbers and graphs.

→ data Summary using graphs

(a) Bar graph



(b) histogram



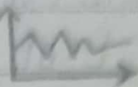
Bar graph

(c) pie charts

مناسب لو انا عايز
اقسم الداتا بناءا على
نسب مئوية

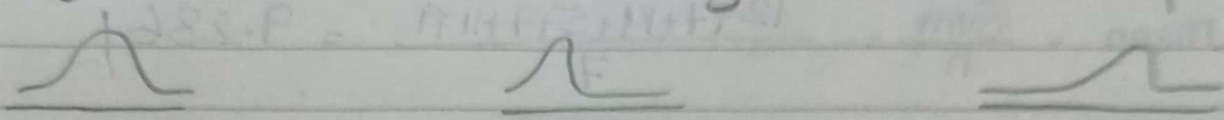
(d) shape of graph and skewness

Line graph



الالتواء
→ skewness ←

data symmatric — skewed to the right — skewed to left



→ describe data using numbers.

→ measures of central of Tendency:

(a) mean sample → $\bar{x} = \frac{\Sigma x}{n}$ ← data sum; n ← # data items

Population mean. $\mu = \frac{\Sigma x}{N}$

"mean is basically the average"

(b) median:- middle number of data set

(c) mode: number occurs most frequently

→ measure of variability: Rang, variance, standard deviation

Inferential

→ using sample of data to make an inference or draw a conclusion of the population.

بستخدمه هنا لو انا عايز اعمل survey معين على عينة كاملة من مجتمع اني

→ uses Probability to determine how confident we can be that the conclusions we make are correct

(Confidence intervals and Margins of errors)

اسأل كل الناس اللي في العينة وهاخد شريحة معينة بس اللي اهمني وابني باقي النتايج عليهم.

"Sample data"

كل ما هنكبر الشريحة (Sample) بتاعتي كل ما هيبقى عندي ثقة اكبر في الداتا بتاعتي

---

②

① mean

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

$$\text{Mean} = \frac{\text{Sum}}{n} = \frac{12+7+14+5+7+11+9}{7} = 9.286$$

② median.

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

ترتيب

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

نحذف من الاقل والاكبر لحد ما نوصل للنص

∴ median = 9

③ Mode = 7 appears twice.

اكتر رقم متكرر

④ Range. = 14 − 5 = 9

difference between highest number and lowest number

~~Ex2:~~

6, 14, 8, 5, 3, 11, 9

① mean = $\frac{6+14+8+5+3+11+9}{7} = 8$

② median : 3, 5, 6, (8), 9, 11, 14

median = 8

③ Mode = none.

④ Range = 14 - 3 = 11

even numbers of data set

Ex1: 6, 8, 5, 5, 9, 8, 10, 8

① mean = $\frac{59}{8} = 7.375$

② median 5, 5, 6, 8, 8, 8, 9, 10

= 8

③ Mode = 8

④ Range = 10 - 5 = 5

Ex2 - 12, 15, 21, 4, 36, 15, 11, 48, 29, 38

① mean = $\frac{229}{10} = 22.9$

② Median 4, 11, 12, 15, 15, 21, 29, 36, 38, 48

median = $\frac{15+21}{2} = 18$

③ Mode = 15

④ Range = 48 - 4 = 44

How to Calculate Variance :- (measure of spread of data)

high variance means high spread of data.

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

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

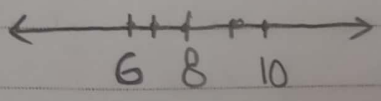
mean = 9

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

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

$$S^2 = \frac{56}{6} = 9.333$$

Ex 2 : 6, 7, 8, 9, 10

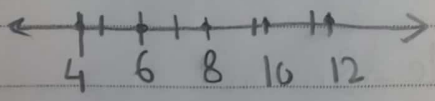


$$\bar{X} = \frac{40}{5} = 8$$

$$S^2 = \frac{(4)+1+0+1+4}{4}$$

$$S^2 = 2.5$$

4, 6, 8, 10, 12



$$\bar{X} = 8$$

$$S^2 = \frac{16+4+0+4+16}{4}$$

$$S^2 = 10$$

How to calculate standard deviation.

82, 93, 98, 89, 88

$$\bar{X} = \frac{\text{Sum}}{n} = \frac{450}{5} = 90$$

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

$$S = \sqrt{\frac{64+9+64+1+4}{4}} = 5.958$$

76, 84, 69, 92, 58, 89, 73, 97, 85, 77

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

$$S = \sqrt{\frac{(-4)^2 + 4^2 + (-11)^2 + (12)^2 + (-22)^2 + (9)^2 + (-7)^2 + (17)^2 + 15)^2 + (-3)^2}{9}}$$

$$S = 11.709$$

Video 5:

How to Find the interquartile Range and any outliers

0	25th	50th	75th	100
min	Q_1	Q_2	Q_3	max

Median [illegible]

$IQR = Q_3 - Q_1$

Outliers → $[(Q_1 - 1.5\, IQR), (Q_3 + 1.5\, IQR)]$

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

① نرتب 3, 5, 6, 8, 10, 11, 12, 14, 15, 18, 26

median = 11

3 5 (6) 8 10 | 12 14 (15) 18 26

Q_1 = 6, Q_2 = 11, Q_3 = 15

$IQR = 15 - 6 = 9$

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

$(6 - 1.5 * 9, 15 + 1.5 * 9]$ $= [-7.5, 28.5]$

26 is not outlier

11, 31, 21, 19, 28, 54, 35, 26, 23, 13, 29, 17

8, 11, 13, 17, 19, 21, 23, 26, 29, 31, 35, 54

15 Q_1 22 Q_2 30 Q_3

$Q_2 = \frac{21+23}{2} = 22$

$IQR = Q_3 - Q_1 = 30 - 15 = 15$

$Outlier = [Q_1 - 1.5 * IQR, Q_3 + 1.5 * IQR]$

$= [15 - 22.5, 30 + 22.5] = [-7.5, 52.5]$

$\therefore 54$ is outlier