Date: \ \

2 PROCESS :

01 - Descriptive Statistics

Measures of Central tendency , Measures of Dispersion or Variability percentiles and quantiles

Objective :

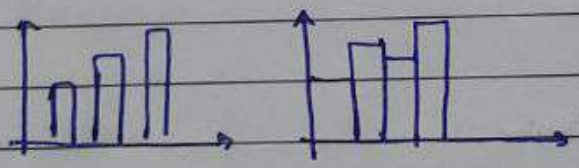
- Descriptive Statistics vs Inferential Statistics
- Mean, Median, Mode, and Range How to find it
- How to calculate Variance
- How to calculate the Standard Deviation
- How to find the Interquartile Range & any Outliers

Video (1)

Statistics

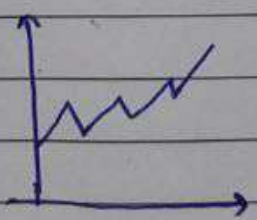
Descriptive

① Organizing and summarizing data
[1] using numbers & graphs
Bar Graphs , Histograms



Pie charts

line Graph



~~Inferential~~

② shape of graph & skewness
symmetrical or skewed to right or left

② Central tendency
Mean, Median, Mode
the averge, Middle, most frequently

③ Variabilty
Range, vaciance , Standard deviation

the range is a measure of Dispersion

Inferential

Sample data to make an inference
or draw a conclusion of the population

Confidence Intervals & Margins of error

Video ② Mean, Median, Mode and Range

[12, 7, 14, 5, 7, 11, 9]

$$\text{Mean} = \frac{\text{Sum}}{\text{num}} = \frac{12+7+14+5+7+11+9}{7} = 9,286$$

Median sort first [5, 7, 7, 9, 11, 12, 14]
Median = middle = 9

Mode highest frequency 7 appears twice
Mode = 7

Range difference between highest and lower
Range = 14 - 5 = 9

Date: \ \

another example

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

$\text{Mean} = \frac{6+14+8+5+3+11+9}{7} = 8$

Median = 8 [3, 5, 6, (8), 9, 11, 14]

Mode = None

Range = 14 - 3 = 11

even num example

[6, 8, 5, 5, 9, 8, 10, 8]

$\text{Mean} = \frac{59}{8} = 7,375$

$\text{Median} = \frac{8+8}{2} = 8$ [5, 5, 6, 8, 8, 8, 9, 16]

Mode = 8

Range = 10 - 5 = 5

[12, 15, 21, 4, 36, 15, 11, 48, 29, 38]

$\text{Mean} = \frac{229}{10} = 22,9$ [4, 11, 12, 15, 15, 21, 29, 36, 38, 48]

$\text{Median} = \frac{15+21}{2} = 18$

Mode = 15

Range = 48 - 4 = 44

Date:

Video ③ How to Calculate Variance

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

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

first sort

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

Mean $= \bar{x} = \frac{63}{7} = 9$

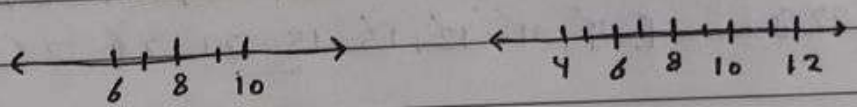
$S^2 = \frac{56}{7-1} = 9,3$

the variance is a measure of the spread of the data

spread far apart or closer to the mean

high variance ⇒ larger spread in the data

example [6, 7, 8, 9, 10] [4, 6, 8, 10, 12]



higher variance

$\bar{x}_1 = \frac{40}{5} = 8$ $\bar{x}_2 = \frac{40}{5} = 8$

$s_1^2 = \frac{10}{5-1} = 2,5$

$s_2^2 = \frac{40}{5-1} = 10$ higher variance #

①	②
$(6-8)^2$	$(4-8)^2$
$(7-8)^2$	$(6-8)^2$
$(8-8)^2$	$(8-8)^2$
$(9-8)^2$	$(10-8)^2$
$(10-8)^2$	$(12-8)^2$
10	40

Date: \ \

Video ④ How to calculate the Standard Deviation

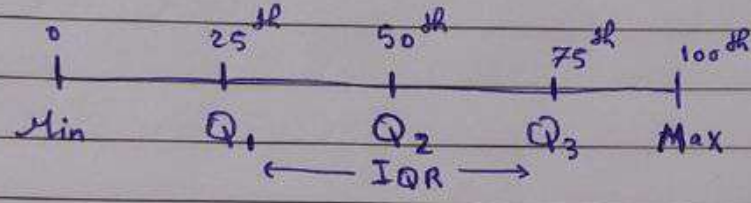
[82, 93, 98, 89, 88]

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

$S = \sqrt{\frac{\Sigma (X - \bar{X})^2}{n-1}}$ the square root of variance

$S = \sqrt{\frac{(-8)^2 + (3)^2 + (8)^2 + (-1)^2 + (-2)^2}{4}} = 5,958$

Video ⑤ How to Find Interquartile Range & any outliers



quartiles divide the data to Four equal parts

★ Interquartile Range the middle 50% of the data

$IQR = Q_3 - Q_1$

★ Outliers out side $[Q_1 - 1,5 * IQR , Q_3 + 1,5 * IQR]$

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

first sort [3, 5, 6, 8, 10, 11, 12, 14, 15, 18, 26]

Q_1 Q_2 Q_3

Median = 11

Median of lower part = 6 IQR = 15 - 6 = 9

Median of upper part = 15

Outliers [6 - 1,5 * 9 , 15 + 1,5 * 9]

[-7,5 , 28,5]