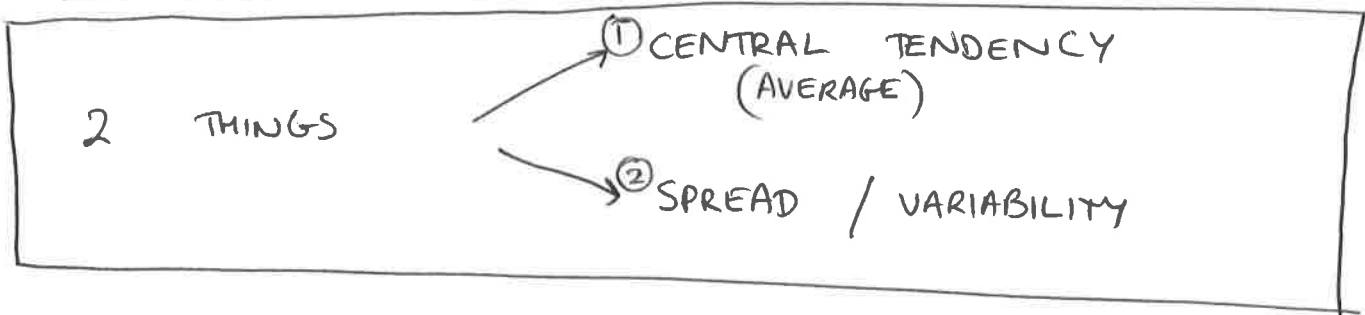


ANALYSING DATA



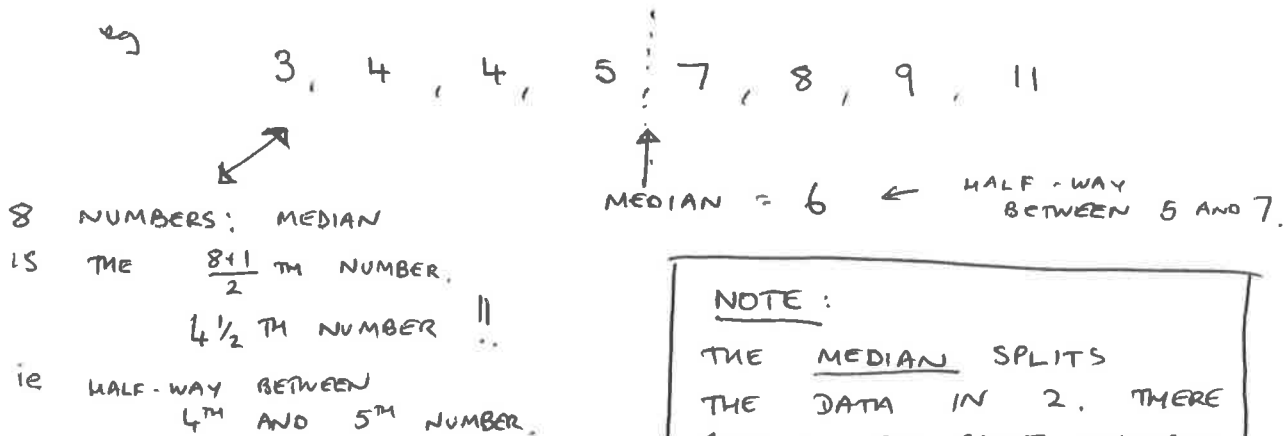
① AVERAGES MEAN / MEDIAN / MODE.

• MEAN = $\frac{\text{SUM OF THE NUMBERS}}{\text{NUMBER OF NUMBERS}}$

MOST COMMONLY USED AVERAGE.

• MEDIAN = THE "MIDDLE" NUMBER
(NUMBERS MUST BE LISTED IN ORDER)

- CAN BE HALF-WAY BETWEEN 2 MIDDLE NUMBERS



NOTE:

THE MEDIAN SPLITS THE DATA IN 2. THERE SHOULD BE SAME NUMBER OF NUMBERS LESS THAN THE MEDIAN AS THERE ARE NUMBERS BIGGER THAN THE MEDIAN

- YOU CAN BE ASKED TO FIND MEDIAN OF A STEM AND LEAF DIAGRAM.

2 WAYS. → CROSS OFF BIGGEST + SMALLEST, THEN 2ND BIGGEST + 2ND SMALLEST, UNTIL YOU ARE LEFT WITH THE MEDIAN.

→ COUNT THE NUMBER OF LEAVES.

• ADD 1, THEN DIVIDE BY 2.

• YOU ARE LOOKING FOR THIS "TERM" OF THE STEM + LEAF, STARTING WITH SMALLEST

• MODE → MOST OFTEN

- MOST COMMON VALUE IN THE SET OF DATA.
- USED FOR CATEGORICAL DATA.
- eg CAN'T GET MEAN / MEDIAN OF COLOUR OF CAR / EYES ETC.

HOW DO WE KNOW WHICH AVERAGE TO USE?

IF THERE ARE EXTREME VALUES USE THE MEDIAN

OTHERWISE USE THE MEAN

THIS IS OUR FAVOURITE. WE WANT TO USE THE MEAN

UNLESS IT'S CATEGORICAL DATA WHEN WE HAVE TO USE MODE

WHY ??

EXTREME VALUES MAKE THE MEAN WAY TOO BIG OR TOO SMALL. THE MEDIAN IS NOT AFFECTED BY EXTREME VALUES.

eg A SMALL COMPANY EMPLOYS 5 PEOPLE, WITH THE FOLLOWING SALARIES :
€ 20,000 € 30,000 € 31,000 € 32,000 € 200,000

"THE BOSS" !!

MIDDLE NUMBER

MEDIAN = € 31,000 → WE USE THIS AS A GOOD AVERAGE

MEAN = $\frac{20\,000 + 30\,000 + 31\,000 + 32\,000 + 200\,000}{5}$ = € 62,600

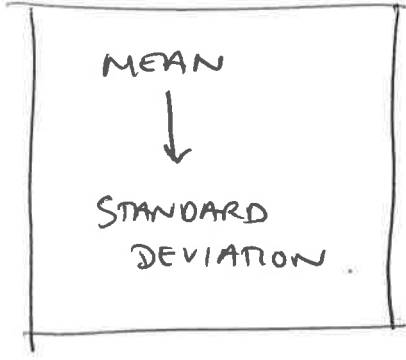
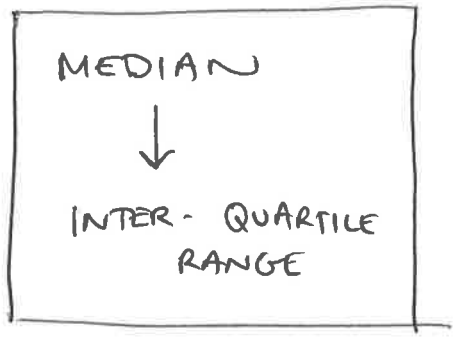
↑
WAAAY TOO HIGH DOESN'T REPRESENT "AVERAGE" SALARY VERY WELL.

② SPREAD [OR VARIABILITY]

- HOW "SPREAD OUT" IS OUR DATA ?
- THIS CAN BE IMPORTANT TO US IF WE'RE DOING A SURVEY / STATISTICAL ANALYSIS. WE CAN SEE HOW CONSISTENT / AGREED THE DATA IS.
- HOW DO WE KNOW WHICH TO USE ?
IT DEPENDS WHICH AVERAGE WE'RE USING

AVERAGE :

SPREAD :



INTER-QUARTILE RANGE

• RANGE = BIGGEST — SMALLEST

i.e. DIFFERENCE BETWEEN TOP + BOTTOM

• INTER-QUARTILE RANGE →

THE "MIDDLE HALF"

- SPLIT DATA UP INTO QUARTERS / QUARTILES
- THE INTER-QUARTILE RANGE IS THE DIFFERENCE BETWEEN THE 3RD QUARTER AND THE 1ST QUARTER.

OR $IQR = Q_3 - Q_1$

INTER-QUARTILE RANGE CONTINUED

SPLIT INTO
TOP HALF
+
BOTTOM HALF.

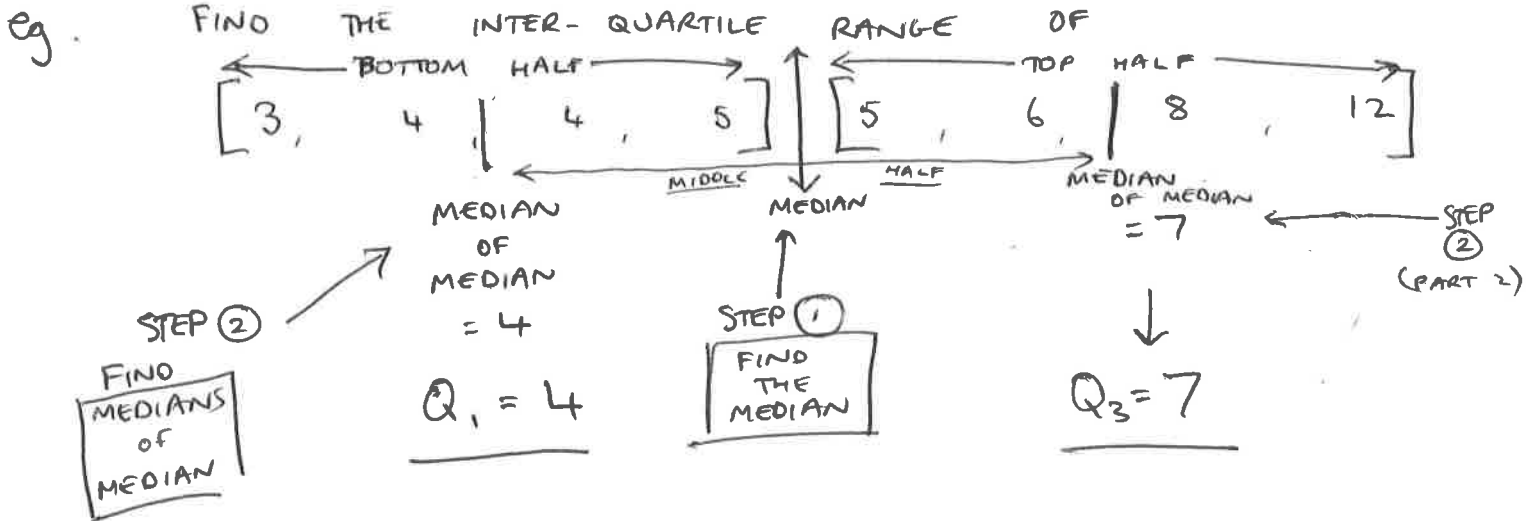
1st → GET THE MEDIAN.
DRAW A BOX AROUND EVERYTHING TO
THE LEFT (AND RIGHT) OF THE MEDIAN.

2nd → FIND THE MEDIANS OF THE MEDIAN

↑
REMEMBER THIS.

THESE ANSWERS ARE Q_1 AND Q_3

3rd → $IQR = Q_3 - Q_1$



$$IQR = Q_3 - Q_1$$
$$= 7 - 4 = \boxed{3}$$

NOTE:

- THERE ARE OTHER WAYS TO CALCULATE THE INTER-QUARTILE RANGE, YOU MIGHT EVEN GET A DIFFERENT ANSWER!

STANDARD DEVIATION

[SYMBOL σ] \bar{x} = MEAN

- WE NEED TO KNOW HOW TO CALCULATE IT USING A CALCULATOR (SEE SEPARATE SHEET)

WHAT IS IT?

- THE AVERAGE DIFFERENCE BETWEEN EACH ITEM OF DATA AND THE AVERAGE OF THE WHOLE SAMPLE

eg IF WE'RE CALCULATING THE STANDARD DEVIATION OF TEACHER'S HEIGHTS: WE FIRST NEED TO KNOW THE AVERAGE HEIGHT. (MEAN)

LET'S PRETEND THE MEAN HEIGHT IS 175 CM

NOW, EVERY TEACHER HAS A DEVIATION FROM THIS MEAN OF 175.

eg MR HILLIARD IS 180 CM.
HIS "DEVIATION" IS 5 CM (180 - 175 = 5)

THE STANDARD DEVIATION IS THE AVERAGE OF ALL THE DEVIATIONS OF ALL THE TEACHERS.

- IT'S COMPLICATED TO CALCULATE MANUALLY, SO WE JUST USE THE CALCULATOR...

How to Calculate the Standard Deviation on a Calculator

(Casio fx-83GT Plus)

Calculating the Standard Deviation of a list of numbers:

Example: Find the Standard Deviation of the following numbers correct to one decimal place: 3, 5, 6, 8, 10, 13.

Step 1: Press **MODE** and then select **2** for statistics. Now select **1** for "1-VAR"

Step 2: Input the numbers, pressing [=] in between

Step 3: Now press [AC] to clear the table

Step 4: We want to get into the "STAT" function of the calculator – We press [Shift] + [1]

Step 5: This brings up a list of options. Press [4] for "VAR" and [3] for " σx ". This is the standard deviation.

In the above example, the answer is 3.3.

Calculating the Standard Deviation from a Frequency Table

Example: The frequency distribution table shows the marks out of five given for a new film by a group of film critics. Use your calculator to find the standard deviation.

Mark	0	1	2	3	4	5
No of Critics	1	3	4	7	4	3

Step 1: Your calculator needs to be set up for handling frequency tables. Press [Shift] and [Setup] (beside the on button). Scroll down to the second screen and press **3** for "STAT", and **1** for "ON"

Step 2: Press **MODE** and then select **2** for statistics. Now select **1** for "1-VAR"

Step 3: Input the marks from the frequency table, pressing [=] in between

Step 4: Using the up and right arrows, go to the top of the next column in the table, and input the frequencies from the "No of Critics", pressing [=] in between

Step 5: Now press [AC] to clear the table

Step 6: We want to get into the "STAT" function of the calculator – We press [Shift] + [1]

Step 7: This brings up a list of options. Press [4] for "VAR" and [3] for " σx ". This is the standard deviation.

In the above example, the answer is 1.3582322982

EMPIRICAL RULE (NORMAL DISTRIBUTION)

DEVIATIONS FROM THE MEAN

REMEMBER
THESE
NUMBERS

• 68% OF POPULATION LIES WITHIN ONE STANDARD DEVIATION OF THE MEAN

• 95% OF POPULATION IS WITHIN 2 STANDARD DEVIATIONS

• 99.7% IS WITHIN 3 STANDARD DEVIATIONS.

