Summary Statistics

Mean, Median, Mode, Standard Deviation… (and all other horrible things which belong to this topic)

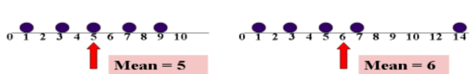
Why do we need these measurements?

- No fear, these are only different ways of calculating an average of the average
- Taken step by step everybody can do it … promised
- Examples from your future work:
  - Stock exchange market
  - Manufacturing of goods (quality control, stock management)
  - Customer survey analysis
  - Negotiations of salaries
  - and many more examples which I couldn't think of …

- Before you can begin to understand statistics , there are three terms you will need to fully understand
  - Mean
  - Median
  - Mode
- These terms will help us to understand the concept of standard deviation

The Mean

- The first term '**average**' is something we have been familiar with from a very early age when we start analyzing our marks on report cards. We add together all of our test results and then divide it by the sum of the total number of marks there are. We often call it the average. However, statistically it's the **Mean!**

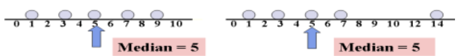


The Mean (Example)

- Four test results: 15, 18, 22, 20
- The sum is: 75
- Divide 75 by 4: 18.75
- The 'Mean' *(Average)* is 18.75 *(Often rounded to 19)*

The Median

- The Median is the **'middle value'** in your list. When the totals of the list are odd, the median is the **middle entry** in the list after sorting the list into **increasing order**. When the totals of the list are even, the median is equal to the **sum of the two middle** (after sorting the list into increasing order) **numbers divided by two**. Thus, remember to line up your values, the middle number is the median! Be sure to remember the odd and even rule.

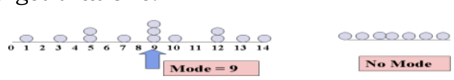


The Median (Example)

- Find the Median of: 9, 3, 44, 17, 15
  (Odd amount of numbers)
- Line up your numbers: 3, 9, 15, 17, 44 (*smallest to largest*)
- The Median is: 15 (The number in the middle)
- Find the Median of: 8, 3, 44, 17, 12, 6
- *(Even amount of numbers)*
- Line up your numbers: 3, 6, 8, 12, 17, 44
  Add the 2 middles numbers and divide by 2
- (8+12) ÷ 2 = 20 ÷ 2 = 10
  The Median is 10.

The Mode

- The mode in a list of numbers refers to the list of numbers that occur most frequently. A trick to remember this one is to remember that mode starts with the same first two letters that most does. **Mo**st frequently - **Mo**de. You'll never forget that one!



The Mode (Example)

- Find the mode of:
  9, 3, 3, 44, 17 , 17, 44, 15, 15, 15, 27, 40, 8,
- Put the numbers is order for ease:
  3, 3, 8, 9, 15, 15, 15, 17, 17, 27, 40, 44, 44,
- The Mode is 15 *(15 occurs the most at 3 times)*
- *It is important to note that there can be more than one mode and if no number occurs more than once in the set, then there is no mode for that set of numbers.*

Average/Mean has a Problem though

... which is that if there are some exceptionally small or large values in our data set, they can have an undue influence on the values in the middle of our data set.

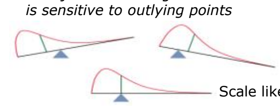
Just imagine my husband will buy his shares only because he calculated an average!!!

Stock market Example

Value	Telecom	SAP
1	92	56
2	50	42
3	49	51
4	44	50
5	48	54
Average	**56.6**	**50.6**

outlier

A major disadvantage of the mean is that it is sensitive to outlying points



Helping hand comes from ***Standard deviation***

standard deviation tells us instantly whether we are dealing with data like set 1 or set 2.

Here is the prove - the standard deviation for Telecom stock is **19.92**, for SAP stock it is **5.37**;

Example (Booklet p.39)

The weekly mileage for 40 company cars has the distribution shown below. We wish to calculate the arithmetic mean mileage, $\bar{X}$.

Mileage	No. of cars (f)	X	fX
100 & < 120	4	110	440
120 & < 140	10	130	1300
140 & < 160	19	150	
160 & < 180	5	170	850
180 & < 200	2	190	380
Total	40	-	16470

$\frac{120-110}{2}$

$4 \cdot 110$

$\bar{X} = \frac{\sum fX}{n}$, here $n = \sum f$

$\bar{X} = \frac{16470}{40} = 411{,}75$ miles

Example (Booklet p.40)

9	15	18	19	19	19	20	20	22	22
25	26	26	26	27	27	28	29	29	29
30	30	30	30	31	31	33	33	34	35
37	38	39	40	41	41	43	43	44	45

19 = III
20 = II
22 = II
26 = III
27 = II
29 = III
30 = IIII
31 = II
33 = II

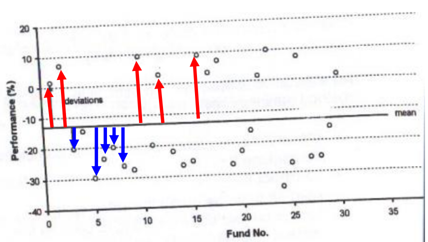
We wish to calculate the median number of clients seen per day, M.

We find

As the totals of the list are even, the median is equal to

$\frac{29+30}{2} = 29{,}5$

Example (Booklet p.43)



Remember ↑ + ↓ = 0

Example (Booklet p.45)

9 5 1 6 10 5

We wish to calculate the standard deviation, s, of this data set.

squared in order to get rid of negative values

Actual sales X	Mean $\bar{X}$	Deviation $X - \bar{X}$	Squared deviation $(X - \bar{X})^2$
9	6	$9-6=3$	$3^2 = 9$
5	6	$5-6=-1$	$(-1)^2 = 1$
1	6	-5	$(-5)^2 = 25$
6	6	0	0
10	6	4	16
5	6	-1	1
		Σ 0	Σ 52

$\bar{x} = \frac{\sum x}{n} = \frac{36}{6} = 6$

the sum of deviations is always 0

$s = \sqrt{\frac{\sum (x-\bar{x})^2}{n}} = \sqrt{\frac{52}{6}} = \sqrt{8{,}6} = 2{,}9$

Example (Booklet p.47)

Bonus (£)	No. of employees (f)	X	X^2	$f \cdot X$ ⇓ fX	fX^2
20 & < 40	4	30	900	120	3600
40 & < 60	7	50	2500	350	17500
60 & < 80	5	70	4900	350	24500
80 & < 100	4	90	8.100	360	32400
Total	20	-	16400	1180	78000

$\bar{X} = \frac{\sum fX}{n} = \frac{1180}{20} = 59$; $n = \sum f = 20$

$S = \sqrt{\frac{\sum fx^2}{n} - \bar{X}^2} = \sqrt{\frac{78000}{20} - 59^2} = \sqrt{419} =$
$= 20{,}469 \approx 20{,}47$