

Y10 Maths Knowledge Organiser Higher Tier: Statistical Diagrams and Averages

What must I be able to do?	Key vocabulary
New content: <ul style="list-style-type: none"> Be able to draw and interpret pie charts ➤ Mathswatch 128a Draw and interpret line graphs and other simple charts Solve problems which involve averages ➤ Mathswatch 62 Calculate averages from frequency tables ➤ Mathswatch 130a, 130b Draw scatter graphs and use lines of best fit ➤ Mathswatch 129 	Discrete data Data which takes <u>fixed values</u> . E.g. Shoe sizes, number of people. Continuous data Data which can take <u>any value</u> . Usually any data which is measured, e.g. height, weight. Bivariate data Data for <u>two variables</u> . Correlation A measure of <u>relationship</u> between <u>two variables</u> .

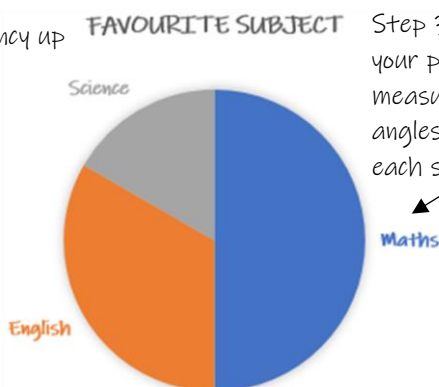
Pie Charts

Favourite Subject	Number of students	Angle calculation	Angle to draw
Maths	30	$30 \div 60 \times 360 =$	180°
English	20	$20 \div 60 \times 360 =$	120°
Science	10	$10 \div 60 \times 360 =$	60°
Total = 60			

Step 1: Work out the total number of students by adding the frequency up

Step 2: For each frequency divide it by the total and multiply by 360 (as the total angles in a circle = 360°)

Step 3: Now draw your pie chart, measuring these angles and labelling each sector



Types of averages

e.g. Given this list of numbers 3, 7, 5, 4, 7

Mean: $3 + 7 + 5 + 4 + 7 = 26$

$26 \div 5 = 5.2$

The mean value is 5.2

Median: First, write in ascending order

3, 4, 5, 7, 7

Only 5 in the middle

The median value is 5.

Mode: The number which appears the most is 7. (7 appears twice)

The modal value is 7.

e.g. Given this list of numbers 7, 9, 3, 5

Mean: $7 + 9 + 3 + 5 = 24$

$24 \div 4 = 6$

The mean value is 6

Median: First write in ascending order

3, 5, 7, 9

5 and 7 in the middle.

The median value is 6

$5 + 7 = 12$

$12 \div 2 = 6$

Mode: Each number appears an equal number of times (only once)

There is no mode.

Range

e.g. Given the list of numbers 3, 6, 10, 3, 5, 8 the range is the largest (10) subtract the smallest (3). $10 - 3 = 7$. The range is 7.

The range is not an average but instead is a measure of spread. In general, a lower range is better as it implies the data is more consistent.

Reverse mean questions

With these you need to find the total amount by working backwards. It usually involves multiplying the mean by the number of values.

e.g. The mean height of 4 basketball players is 1.88m. A 5th player joins who is 1.96m tall. What is the new mean height of all 5 players?

First, find the total height of the original 4.

$$4 \times 1.88 = 7.52\text{m}$$

New total height is therefore $7.52 + 1.96 = 9.48\text{m}$

$$\text{New mean height is } 9.48 \div 5 = 1.896\text{m}$$

Comparing data

When comparing data you should write two statements, one comparing an average (mean, median or mode) and another comparing the spread (range).

E.g. Joe and Emma are testing frisbees. They each throw their frisbee 3 times and measure how far it travels in metres.

Joe's results are: 13.2, 17.6 and 11.5

Emma's results are: 14.5, 13.9 and 14.8. Compare the results.

Joe's mean average: $13.2 + 17.6 + 11.5 = 42.3$ $42.3 \div 3 = 14.1\text{m}$

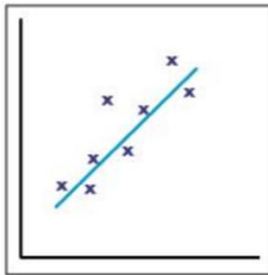
Joe's range is $17.6 - 11.5 = 6.1\text{m}$

Emma's mean average: $14.5 + 13.9 + 14.8 = 43.2$ $43.2 \div 3 = 14.4\text{m}$

Emma's range is $14.8 - 13.9 = 0.9$.

So, on average Emma's frisbee went further as $14.4 > 14.1$. Emma's frisbee was also more consistent as her range was only 0.9 while Joe's range was 6.1m.

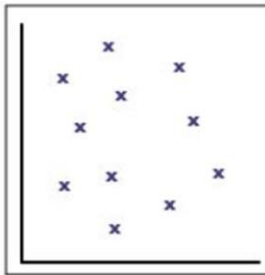
Types of correlation



Positive correlation:
As one value increases,
so does the second.



Negative correlation:
As one value increases,
the second decreases



No correlation.

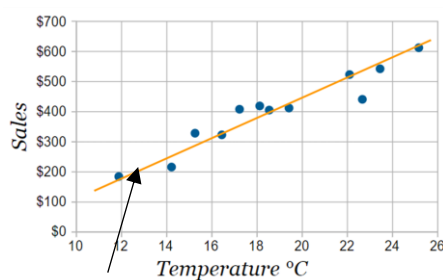
Drawing and using scatter graphs

e.g.

Ice Cream Sales vs Temperature	
Temperature °C	Ice Cream Sales
14.2°	\$215
16.4°	\$325
11.9°	\$185
15.2°	\$332
18.5°	\$406
22.1°	\$522
19.4°	\$412
25.1°	\$614
23.4°	\$544
18.1°	\$421
22.6°	\$445
17.2°	\$408

Each pair of values is plotted as a point on the scatter graph

e.g. (17.2, \$408)



Line of best fit.

Drawn by hand **using a ruler** to fit the data as best as possible.

Shows the general trend and can be used to make predictions if you only knew one value

e.g. if the temperature was to be 21°C you would predict sales of about \$480 by reading up to the line of best fit from 21°C.

The line of best fit does not usually go through (0, 0).

Averages from tables

e.g.

# of people	Frequency
1	8
2	6
3	3
4	4
Total = 21 cars	

The **mode** will be the group with the largest frequency. The highest frequency is 8 so the mode is 1 person in a car.

The **median** is the middle value. There are 21 values in total (the sum of the frequency) so the middle value will be the 11th. The first 8 values are all 1s, the next 6 values are all 2s which is 14 values in total. So the 11th value was a 2. The median is 2 people in a car.

The **mean** is the average number of people per car:

# of people	Frequency	Total
1	8	$8 \times 1 = 8$
2	6	$6 \times 2 = 12$
3	3	$3 \times 3 = 9$
4	4	$4 \times 4 = 16$
21 cars		45 people

8 cars have 1 person. $8 \times 1 = 8$.

6 cars have 2 people. $6 \times 2 = 12$.

3 cars have 3 people. $3 \times 3 = 9$.

4 cars have 4 people. $4 \times 4 = 16$.

So the total is $8 + 12 + 9 + 16 = 45$ people.

$45 \div 21 = 2.14$ people per car (2d.p.)

Estimating the mean from a grouped frequency table

e.g.

Test Score	Frequency	Midpoint (of test score)	Estimated Total
0-10	5	$(10 + 0) \div 2 = 5$	$5 \times 5 = 25$
11-20	4	$(20 + 11) \div 2 = 10.5$	$4 \times 10.5 = 42$
21-30	8	$(21 + 30) \div 2 = 25.5$	$8 \times 25.5 = 204$
31-40	12	$(40 + 31) \div 2 = 35.5$	$12 \times 35.5 = 426$
Total = 29 people			697

Estimated mean is:

$$\begin{aligned} & \text{Estimated total} \div \text{total frequency} \\ & = 697 \div 29 = 24.03 \text{ (2dp)} \end{aligned}$$

In a grouped frequency table you do not know the actual values, e.g. we know 5 people scored between 0 and 10 but not their actual scores. So we cannot add up their scores to find an accurate total. The way around this is to estimate their scores and we use the **midpoint** of the values for this estimation. The rest of the question follows the same pattern as a normal frequency table.

Note: For this estimation to be accurate we assume that the groups are **evenly distributed** (this means that there are approximately the same amount of values above the midpoint as there are below the midpoint in each group).

Estimating the median from a grouped frequency table

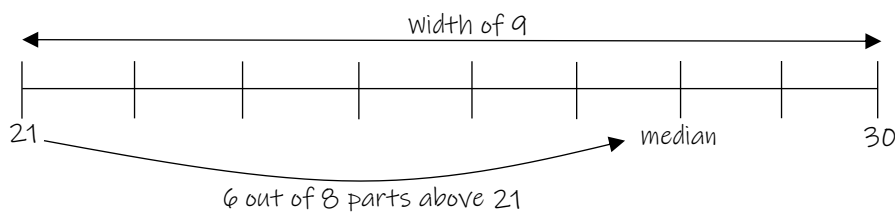
e.g.

Test Score	Frequency	Running total
0-10	5	5
11-20	4	$5 + 4 = 9$
21-30	8	$9 + 8 = 17$
31-40	12	$17 + 12 = 29$
Total = 29 people		

The first step is to find which group contains the median. The median position is $\frac{n+1}{2} = \frac{29+1}{2} = 15^{\text{th}}$ value.

The running total column shows where the 15th value is. After the first group there were only 5, the first + second group has 9 people in total, the first + second + third is 17 and we have gone past the 15th. The 15th person is in the group 21 - 30.

We split the group 21-30 into 8 equal parts (because it has a frequency of 8). To get from the 9th person (running total up to the end of the previous group) to the 15th person (median) we need to go 6 more people ($15 - 9 = 6$).



So we need 6 out of 8 equal parts into the group 21 - 30

$$\frac{6}{8} \times 9 + 21 = 27.75$$

← The estimated median

9 is the width of the group
 $31 - 21 = 9$

The group starts at 21 so the median must be between 21 and 30

GLUE

HERE