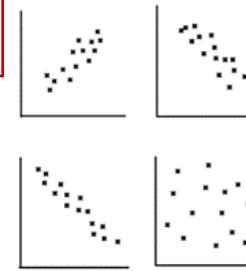


VOCABULARY

Data, primary, secondary, qualitative, quantitative, continuous, discrete, mode, median, mean, grouped, estimated mean, range, correlation, relationship, outlier, line of best fit, composite,

**GCSE HIGHER YEAR 9 AUTUMN TERM UNIT 3
AVERAGES, RANGE, DATA COLLECTION AND REPRESENTATION**



Ensure that you always draw your line of best fit! Use a pencil.

Types of Data	Qualitative Data – non-numerical data	Qualitative Data – eye colour, gender
	Quantitative Data – numerical data	
	Continuous Data – data that can take any numerical value within a given range.	
	Discrete Data – data that can take only specific values within a given range.	Discrete Data – number of children, shoe size

Grouped Data	Data that has been “bundled” in to categories.	<table border="1"> <tr> <th>Foot length, <i>l</i>, (cm)</th> <th>Number of children</th> </tr> <tr> <td>$10 \leq l < 12$</td> <td>5</td> </tr> <tr> <td>$12 \leq l < 17$</td> <td>53</td> </tr> </table>	Foot length, <i>l</i> , (cm)	Number of children	$10 \leq l < 12$	5	$12 \leq l < 17$	53
	Foot length, <i>l</i> , (cm)	Number of children						
$10 \leq l < 12$	5							
$12 \leq l < 17$	53							
	Seen in grouped frequency tables, histograms, cumulative frequency							

Primary / Secondary Data	Primary Data – collected yourself for a specific purpose.	Primary Data – data collected by a student for their own research project.
	Secondary Data – collected by someone else for another purpose.	

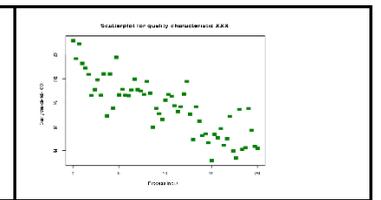
Mean	Add up the values and divide by how many values present.	The mean of 3, 4, 7, 6, 0, 4, 6 is $\frac{3 + 4 + 7 + 6 + 0 + 4 + 6}{7} = 5$
-------------	--	--

Median Value	The middle value. Put the data in order and find the middle one. If there are two middle values , find the number half way between them by adding them together and dividing by 2 .	Find the median of: 4, 5, 2, 3, 6, 7, 6 Ordered: 2, 3, 4, 5 , 6, 6, 7 Median = 5
---------------------	--	---

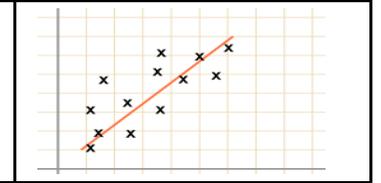
Mode /Modal Value	Most frequent/common. Can have more than one mode (called bi-modal or multi-modal) or no mode (if all values appear once)	Find the mode: 4, 5, 2, 3, 6, 4, 7, 8, 4 Mode = 4
--------------------------	---	--

Range	Highest value subtract the Smallest value Range is a ‘ measure of spread ’. The smaller the range the more consistent the data.	Find the range: 3, 31, 26, 102, 37, 97. Range = $102 - 3 = 99$
--------------	--	---

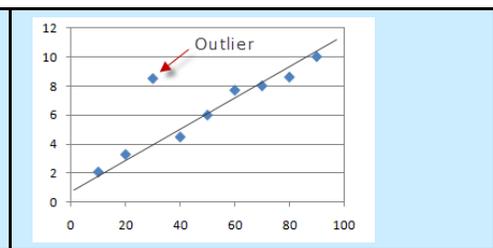
Scatter Graph
A graph in which values of **two variables** are plotted along two axes to **compare** them and see if there is any **connection** between them.



Line of Best Fit
A **straight line** that **best represents the data** on a scatter graph.



Outlier
A value that ‘**lies outside**’ most of the other values in a set of data.
An outlier is **much smaller** or **much larger** than the other values in a set of data.



Mean from a Table

1. Find the **midpoints** (if necessary)
2. Multiply Frequency by values or midpoints
3. Add up these values
4. Divide this total by the Total Frequency

If **grouped data** is used, the answer will be an **estimate**.

Height in cm	Frequency	Midpoint	F × M
$0 < h \leq 10$	8	5	$8 \times 5 = 40$
$10 < h \leq 30$	10	20	$10 \times 20 = 200$
$30 < h \leq 40$	6	35	$6 \times 35 = 210$
Total	24	Ignore!	450

Estimated Mean height: $450 \div 24 = 18.75\text{cm}$

Median from a Table
Use the formula $\frac{(n+1)}{2}$ to find the position of the median.
n is the total frequency.

If the total frequency is 15, the median will be the $\left(\frac{15+1}{2}\right) = 8\text{th}$ position

Pictogram Uses **pictures** or symbols to **show the value** of the data.
A pictogram must have a **key**.

Black
 Red
 Green = 4 cars
 Others

Two Way Tables A table that **organises data** around **two categories**.
Fill out the information step by step using the information given.
Make sure all the totals add up for all columns and rows.

Question: Complete the 2 way table below.

	Left Handed	Right Handed	Total
Boys	10		58
Girls			
Total		84	100

Answer: Step 1, fill out the easy parts (the totals)

	Left Handed	Right Handed	Total
Boys	10	48	58
Girls			42
Total	16	84	100

Answer: Step 2, fill out the remaining parts

	Left Handed	Right Handed	Total
Boys	10	48	58
Girls	6	36	42
Total	16	84	100

Frequency Table (including tally column) A record of **how often each value** in a set of data **occurs**.

Number of marks	Tally marks	Frequency
1		7
2		5
3		6
4		5
5		3
Total		26

Stem and Leaf Statistics represented in a table. Only units displayed in the leaf column.
Remember you **must have a key**

Here are the marks gained by 30 students in an examination:
63 58 61 52 59 65 69 75 70 54 57 63 76 81 64
68 59 40 65 74 80 44 47 53 70 81 68 49 57 61

STEM	LEAF
4	0 4 7 9
5	2 3 4 7 7 8 9 9
6	1 1 3 3 4 5 5 8 8 9
7	0 0 4 5 6
8	0 1 1

5 | 2 = 52

The times taken (in minutes) to complete a task by a group of boys and girls are shown in the back-to-back stem and leaf diagram.

Boys	Stem	Girls
4	1	5 9 9
9 8 0 0	2	0 2 7 7
8 7 3 0	3	0 2 4
3	4	5 6

Key: Boys Girls
4 | 1 represents 14 mins 1 | 5 represents 15 mins

- How many boys are there in the group?
- What is the shortest time for the boys?
- How many girls took longer than 40 minutes to complete the task?
- What is the longest time overall? Is this achieved by a boy or a girl?

Bar Chart Represents data as vertical blocks.
x-axis shows the **type** of data
y-axis shows the **frequency** for each type of data
Each bar should be the **same width**
There should be **gaps** between each bar
Remember to **label** each axis.

Types of Bar Chart **Compound/Composite Bar Charts** show data stacked on top of each other.
Comparative/Dual Bar Charts show data side by side.

Pie Chart Used for showing **how data breaks down into its constituent parts**.
When drawing a pie chart, **divide 360° by the total frequency**. This will tell you how many degrees to use for the frequency of each category.
Remember to **label** the category that each sector in the pie chart represents.

If there are 40 people in a survey, then each person will be worth $360 \div 40 = 9^\circ$ of the pie chart.

Line Graph A graph that uses **points connected by straight lines** to show how data changes in values.
This can be used for **time series data**, which is a series of data points spaced over uniform time intervals in **time order**. Often used to display financial/sales data.