Year 9 Unit 1 Statistics

TYPES OF DATA

data	a collection of **information**
qualitative	data that can only be **written in words**, not numbers, e.g. eye colour, favourite animal
quantitative	**numerical** data, e.g. shoe size, height of a plant
continuous	**numerical data** that can be **measured**, e.g. height of a plant, it has an infinite number of possible values within a selected range, it is **on a scale**
discrete	data which can only take **certain values**, e.g. eye colour, shoe size (categorical in science)
grouped	numerical data that has been ordered and **sorted into groups** called classes
data representation	a **table or chart or graph** which gives more meaning to a set of data these include **bar charts, line graphs, pictograms, pie charts, stem and leaf diagrams, two-way tables, scatter graphs, frequency polygons** and **histograms**

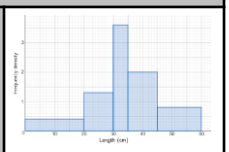
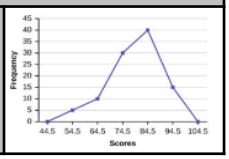
COMPARING DATA

comparing data	**compare averages** to say who is **better**/faster/taller **compare ranges** to say who is **more consistent** / less varied

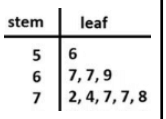
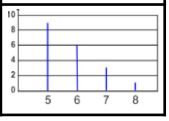
AVERAGES AND RANGE FROM A FREQUENCY TABLE

mean	method: **multiply the variables by their frequencies** (fx column), **total the fx column**, **divide by total frequency**
mode / modal class	the most frequent value or class; the one with **the highest frequency**
median	**use half the total frequency** to find the **middle position**, then **locate the row** this occurs in using the 'subtotal' column
range	**difference between** the **largest** and **smallest values of the variable (first column)**

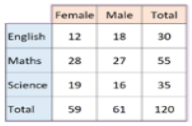
DISPLAYING GROUPED DATA

class width	the **range** of a **group** (class) i.e. aged 15-20 has a class width of **5**	
histogram	the **area** of the bars represents the frequency, there are **no gaps** between bars	
frequency density	the **heights** of the bars on a histogram $frequency\ density = \frac{\boldsymbol{frequency}}{\boldsymbol{class\ width}}$	
frequency polygon	a line graph made by plotting the **frequency against** the **midpoints** of each group	

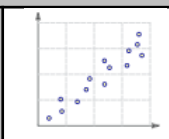
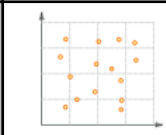
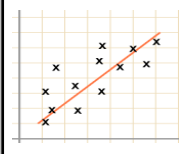
DISPLAYING UNGROUPED DISCRETE NUMERICAL DATA

stem and leaf diagram	a way of displaying a **list of numbers** the **stem goes down** and the **leaves go out** to the right, It has a **key**	
vertical line graph	**like** a **bar chart**, but the **bars have no width**, they are just **straight lines** up the page	

DISPLAYING BIVARIATE DATA

bivariate data	data containing **two variables**	
variable	something that can **change or vary**	
two-way table	shows information about **two variables** which **do not overlap**, the **numbers represent frequencies**	

	Female	Male	Total
English	12	18	30
Maths	28	27	55
Science	19	16	35
Total	59	61	120

scatter graph	a **graph** to show **bivariate** data	
correlation	when there is a **relationship** between two sets of data, but we don't know if one caused the other	
causation	when the independent variable **causes** the dependent variable	
positive correlation	as one variable increases, the other **increases**	
negative correlation	as one variable increases, the other **decreases**	
no correlation	there is **no relationship** between the two variables	
line of best fit	a line that **best represents** the data on a scatter graph In maths GCSE it is **always straight**, but in science it can be curved	
outlier	a value that '**lies outside**' most of the other values in a set of data, it is **much smaller** or **much larger** than the other values in a set of data	

MISLEADING REPRESENTATIONS

misleading representations	Look for: • **frequency scales: too large**, or **too small**; has **missing numbers**; **doesn't start at zero**; • the **axes** are **incorrectly labelled**; • **data** is **missing**; • **bar charts** with **varying width bars** or **varying space between** them; • **proportions** for **pie charts not adding** up to **100%**