

### Essential knowledge

- Construct & interpret frequency tables & polygons, 2-way tables, line, bar, & pie charts
- Find and interpret averages from a list and a table
- Construct and interpret time series graphs, stem and leaf diagrams and scatter graphs

### Key Vocabulary

- Population:** the whole group that is being studied.
- Sample:** a selection taken from the population that will let you find out information about the larger group.
- Representative:** a sample group that accurately represents the population.
- Random sample:** a group completely chosen by chance. No predictability to who it will include.

### Prior learning links

- Representing data (Y8)
- The data handling cycle (Y8)
- Measures of location (Y8)

### Stratified sampling

A stratified sample divides a population into smaller groups known as strata. In stratified sampling, the strata are formed based on members' shared features (e.g. year group, gender, etc.)

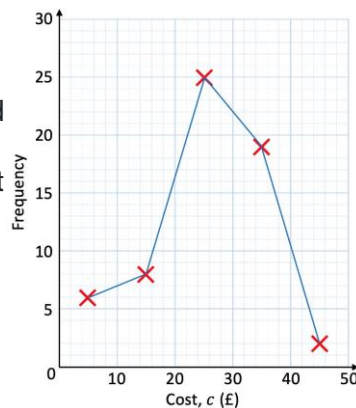
### Frequency tables

A frequency table has two main columns. The first column has all the outcomes as individual values or in the form of class intervals. The second column tells us about the frequency of each outcome. For example, the frequency table to the right indicates that there were 6 items (second column) that cost more than £0 but less than £10 (first column).

Cost of item, $c$ (£)	No. of items
$0 < c \leq 10$	6
$10 < c \leq 20$	8
$20 < c \leq 30$	25
$30 < c \leq 40$	19
$40 < c \leq 50$	2

### Frequency Polygon

A frequency polygon is a graph that shows the frequencies of grouped data. It is a type of frequency diagram that plots the midpoints of the class intervals against the frequencies and then joins up the points with straight lines.



### Averages from lists

24, 8, 4, 11, 8,

#### The mean

To find the mean divide the overall total by how many pieces of data you have.

**Mean  $(55 \div 5) = 11$**

#### The mode

This is the number OR the item that occurs the most (it does not have to be numerical).

**Mode = 8**

#### The median

The value in the centre (in the middle) of the data. This can be easier to calculate if the data is ordered first. **NOTE:** If there is no single middle value find the mean of the two numbers left.

**Median = 8**

### Stem and leaf diagrams

A stem and leaf diagram is a method of organising numerical data based on place value.

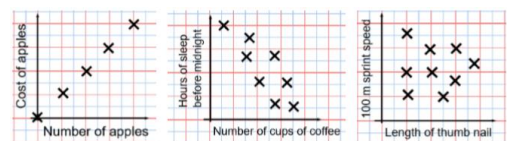
Each number is split into two parts.

- The first digit(s) form the stem,
- The last digit forms the leaf (this should only contain one number).

A stem and leaf diagram must have a key.

Key : 2		0 means 20	
Stem	Leaf		
0	1 4		
1	3 6 6 7		
2	0 2 5		
3	6 7 7 7 8		
4	0 1 3		

### Linear Correlation



**Positive**
**Negative**
**No**  
**Correlation    Correlation    Correlation**

A scatter graph illustrates the relationship between two variables.

### Prior learning links

A group of 20 students were asked how many siblings they have. The results are shown in the table. How many siblings in total?

Number of siblings	Frequency
0	3
1	7
2	5
3	3
4	1
5	1

Here are the heights (in cm) of 8 towers. Calculate the mean, median, mode, and range of the data.

6   14   28   18   10   15   17   28

### Key Vocabulary

Use cover, look, write, check to write the definitions ...

**Population:**

**Sample:**

**Representative:**

**Random sample:**

### Stratified sampling

List three strata you could use to divide Unity's population of pupils into different groups.

Would the same strata be applicable for the staff?

### Frequency tables

Which class interval contains the median?

Lifetime (months)	Frequency
$0 < t \leq 12$	1
$12 < t \leq 24$	9
$24 < t \leq 36$	13
$36 < t \leq 48$	56
$48 < t \leq 60$	21

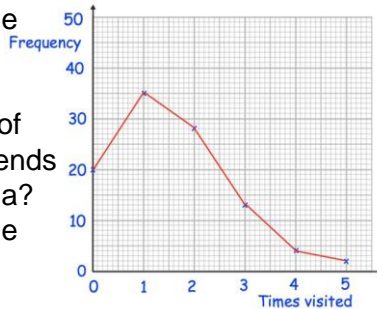
There are 30 students in a class. Miss Ash knows that the median shoe size is 5. Fill in the frequency table with two possible values.

Shoe Size	Frequency
4	4
5	
6	
7	10

### Frequency Polygon

Tia asked her friends how many times they visited the cinema last month.

- How many people visited twice?
- What is most popular number of times that her friends visited the cinema?
- How many people did Tia survey in total?



### Averages from lists

Calculate the mean median and mode for the following sets of data:

5, 7, 3, 5, 8, 9, 10, 2

-2, -1, 5, 8, -2, 2, -1, 9, -1, 1, 2, -1

2.3, 2.6, 2.8, 2.7, 2.8, 2.7, 2.4, 2.3, 2.1, 2.3

The tally chart shows the favourite sport of the students in a class.

What is the modal sport?  
How many students are in the class?  
How many more students liked football than rugby?

Sport	Tally
Rugby	
Football	+++ +
Hockey	+++ +
Cricket	

### Stem and leaf diagrams

The stem and leaf diagram below shows heights of Mr Bold's flowers.

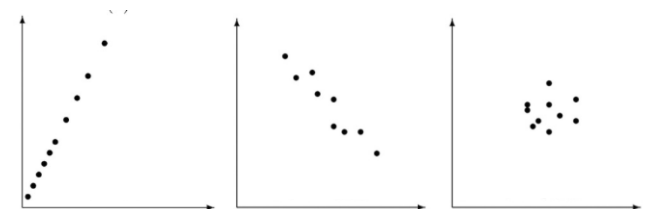
- How many flowers does he have?
- What is the height of the tallest flower?
- How many flowers are taller than 40cm?
- How many flowers are 14cm tall?
- What fraction of the flowers are under 20cm?

**Key: 0|9 means 9cm**

0	9
1	2 4 4 4 8 9
2	0 4 5 8
3	2 4 9
4	1 6 8 8
5	3

### Linear Correlation

What type of correlation does each scatter graph show?



### Essential knowledge

- Types of data (e.g. discrete, continuous)
- Population and sampling
- Frequency tables and polygons
- Two-way tables
- Time series graphs
- Stem and leaf diagrams

### Key Vocabulary

Use cover, look, write, check to write the definitions ...

**Bias:**

**Primary data:**

**Secondary data:**

**Outlier:**

### Prior learning links

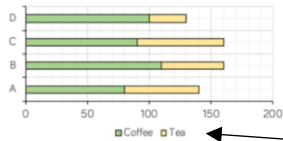
- Representing data (Y8)
- The data handling cycle (Y8)
- Measures of location (Y8)

### The Range

The range of a data set is the difference between the greatest value and lowest value within a collection of numbers. Subtract the smallest from the greatest value in the set to find the range of given data points.

### Bar and line charts

#### Composite bar charts

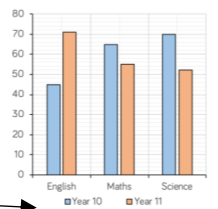


Compare the bars green compared to yellow.  
The size of each bar is the frequency.  
Overall total easily comparable

Categories clearly indicated

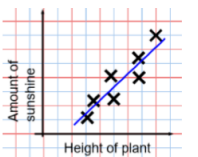
Bars are compared side by side  
Easier to compare subgroups

#### Dual bar charts



### The line of best fit

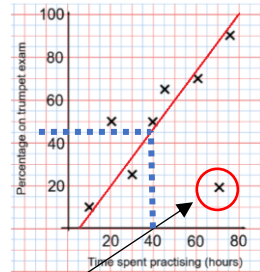
The Line of best fit is used to make estimates about the information in your scatter graph



It is only an estimate because the line is designed to be an average representation of the data.

It is always a **straight line**.

- The line of best fit **DOES NOT** need to go through the origin (The point the axes cross)
- There should be approximately the same number of points above and below the line. The line extends across the whole graph



This point is an **"outlier"**. It doesn't fit this model and stands apart from the data.

**Interpolation** is using the line of best fit to estimate values inside our data point.

**Extrapolation** is where we use our line of best fit to predict information outside of our data.

### Draw and interpret Pie Charts

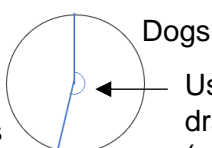
Type of pet	Dog	Cat	Hamster
Frequency	32	25	3

There were 60 people asked in this survey (Total frequency)

"32 out of 60 people had a dog"

$$\frac{32}{60} \times 360 = 192^\circ$$

This fraction of the 360 degrees represents dogs

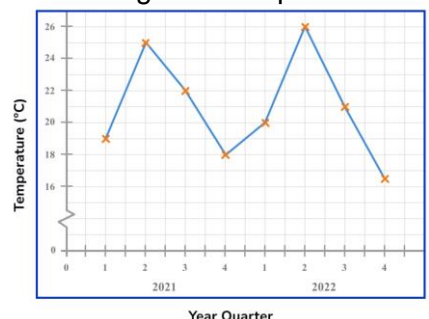


Use a protractor to draw the angle (e.g. 192°)

### Time series graphs

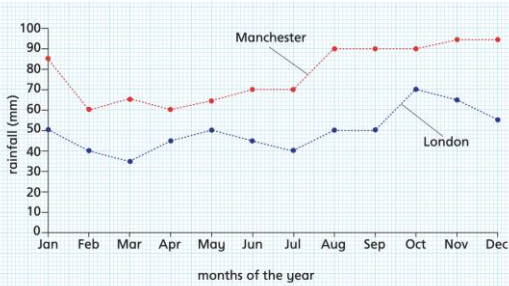
A time series graph is a line graph that shows data such as measurements (e.g. temperature), sales or frequencies over a given time period

They can be used to show a pattern or trend in the data and are useful for making predictions about the future.



### Prior learning links

Is the data shown in the chart qualitative, discrete or continuous?



- Describe the difference between discrete and continuous data.
- Describe the difference between qualitative and quantitative data.

### Key Vocabulary

Use cover, look, write, check to write the definitions ...

- Population:**  
**Sample:**  
**Representative:**  
**Random sample:**

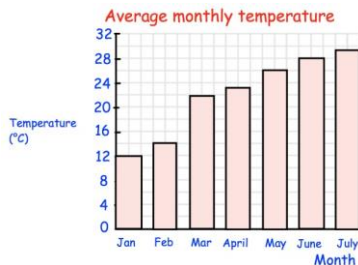
### The Range

Find the range for each of the following:

- 5, 9, 1, 5, 7, 4, 3
- 6.2, 7.3, 8.8, 1.5, 4.1
- 7, 9, -2, 13, 9, 8, 20, -8, 1

### Bar and line charts

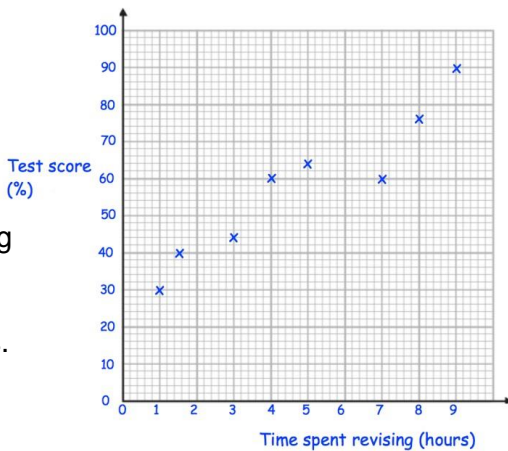
The bar chart shows information about the average temperature on an island.



- What was the average temperature in March?
- Which month had an average temperature of 26°C?
- What is happening to the average temperatures between Jan and July?
- Between which two months was there the greatest rise in temperature?

### The line of best fit

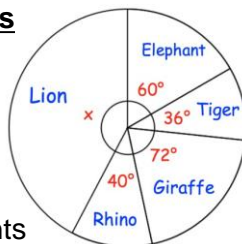
The scatter graph (right) shows information about the number of hours spent revising for a test and the test result for a group of 8 students.



- Daisy spent 7 hours revising for the test. What is Daisy's test score?
- Harry's test score was 30%. How many hours did Harry spend revising?
- Draw a line of best fit.
- Another student spent 6 hours revising for the test.
- Find an estimate of their test score.
- Explain why it might not be sensible to use the scatter graph to estimate the score for a student that spent 15 hours revising.

### Draw and interpret Pie Charts

90 students went on a school trip to Longleaf Safari Park. They were asked their favourite animals.



- What fraction of the students chose elephant?
- What fraction of the students chose tiger?
- What fraction of the students chose giraffe?
- What fraction of the students chose rhino?
- Find x
- How many students chose elephant?

### Time series graphs

