Data handling

CLERICAL MEDICAL

Why collect data?

Data is usually collected to answer a specific question. It is important that you know what question you wish to answer before you start collecting, otherwise you may not collect the right data.

You will need:
- a ruler
- a protractor
- a calculator
- a compass

Population and sample

- Population is a term used to describe the set or collection of objects that you are interested in investigating
- Sample refers to the part of the population that you collect data on.

Sample size is an important consideration for any investigation, it must be big enough to represent the whole population accurately, but small enough for you to measure effectively.

The investigation

You will investigate (for your year group at school):

1. What mode of transport is the most popular for transport into school
2. Is there any correlation between distance from school and duration of journey.

Definition
A correlation means that there is a relationship between the two variables.

You will need to record (for each individual):

- What mode of transport is used. You may wish to use these categories:
  - walk
  - cycle
  - bus
  - train journey
  - car journey
  - other

  If there are individuals that use a combination of transport, record the type that represents the majority of the duration of their journey.
- How far away from school they live, as the crow flies.
- How long a typical journey into school takes.

Data handling

Bristol CM 1824 CLERICAL MEDICAL

Data collection

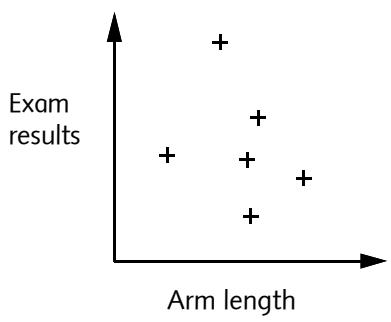
Name	Mode of transport	Distance (miles)	Duration (minutes)

Data handling

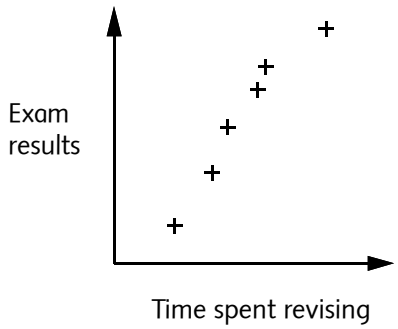
Scatter diagrams

A scatter diagram is good for showing if there is any correlation between different sets of data.

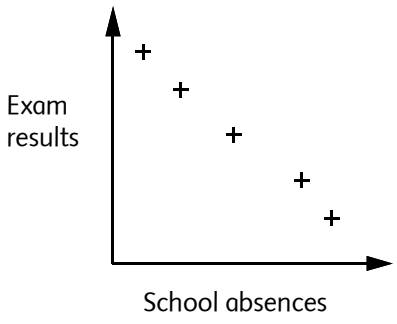
Three types of correlation can occur



No correlation



Positive correlation: This is where as when one variable increases so does the other.



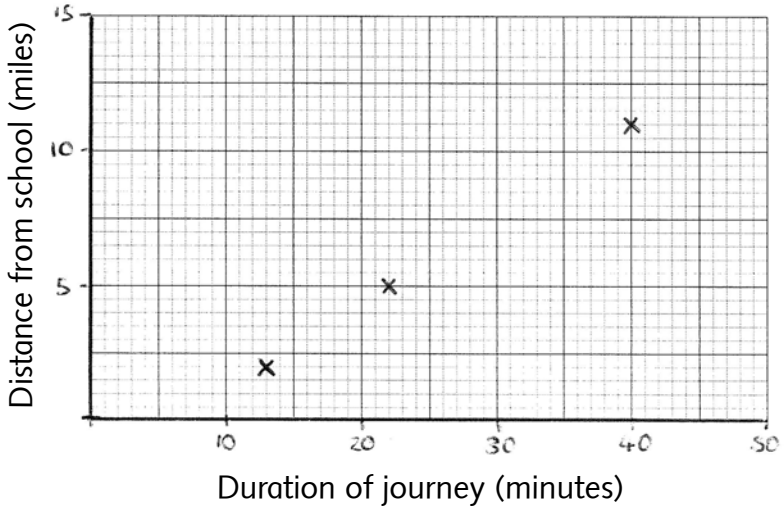
Negative correlation: This is where one variable decreases whilst the other increases.

A strong correlation between the two variables results in a straight line on the graph.

Plotting a scatter diagram

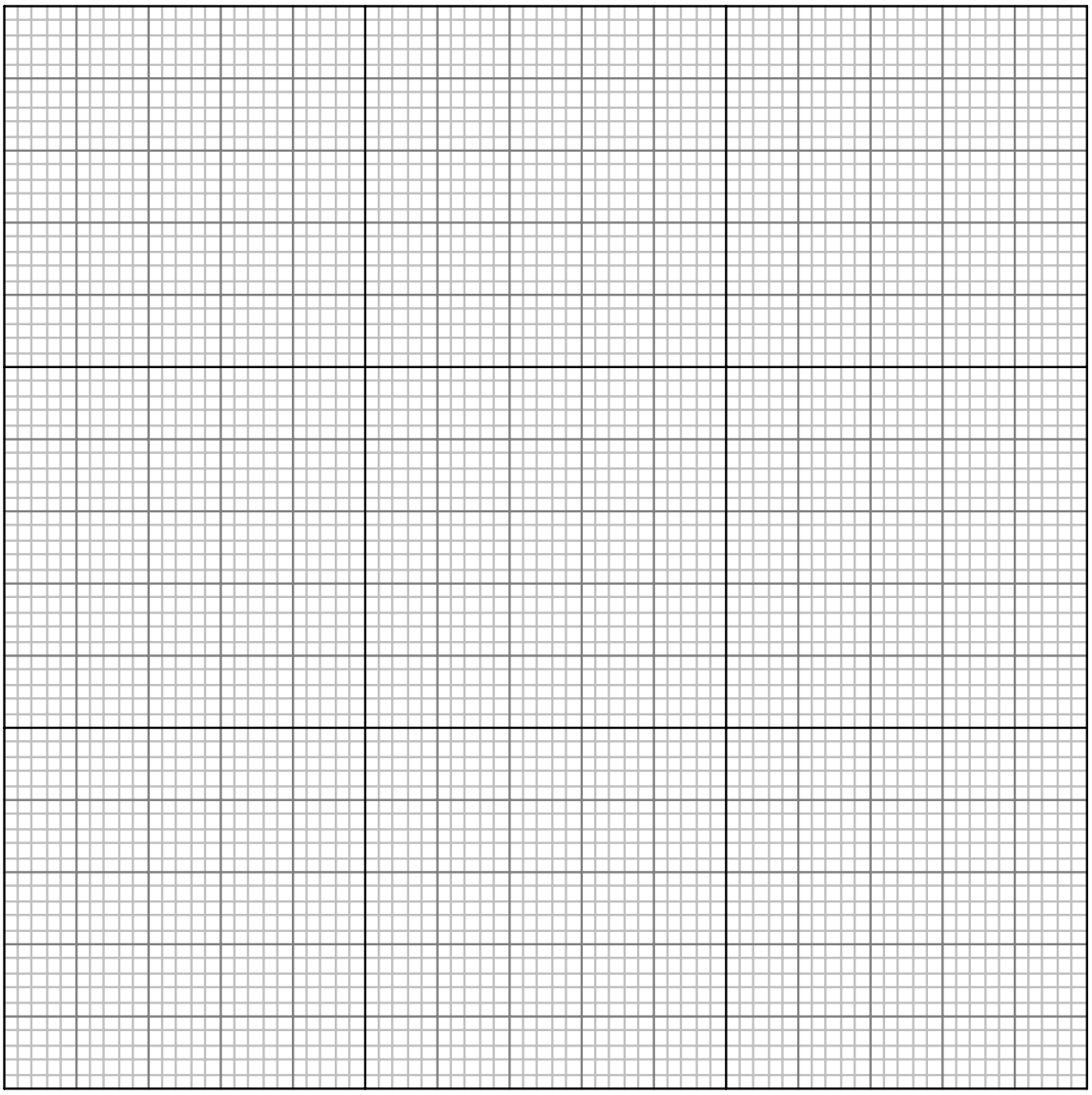
- You need to create a graph with one data set (e.g. distance from school) along the X axis and one data set (e.g. duration of journey) along the Y axis.
- In this example the data on distance from school and duration of journey is used to plot each individual as a point on the graph.
- The more points are plotted on the graph the clearer the relationship between the two variables becomes.

Name	Distance (miles)	Duration (minutes)
John	5	22
Anna	11	40
Geoff	2	13



Data handling

In the space below draw a scatter diagram showing the relationship of 'distance from school' to 'duration of journey'



Is there any correlation between the two data sets? If so, is it a strong correlation?

Data handling

Pie chart

A pie chart is a circle divided into sections. The size of the sections is proportional to the value they represent.

To work out what area of a pie chart will be taken up by a particular result you need to convert your results as a fraction of total responses into a fraction of 360 (the number of degrees in a circle). To do this multiply the results, as a fraction of total responses, by 360.

For example: 24 people out of 30 walk to school.

$$\frac{24}{30} \times 360 = 288$$

288 is the same proportion of 360 as 24 was of 30. 288 is the size the angle should be for the section representing the proportion of pupils that walked.

Represent the data you have on the mode of transport used when travelling to school in the form of a pie chart. Write the percentage each section represents of the whole of your sample on the pie chart