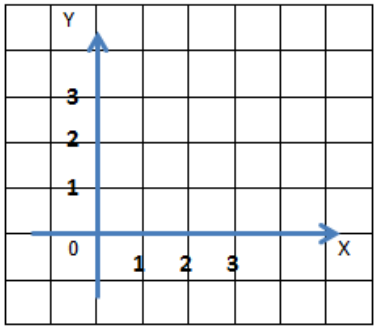
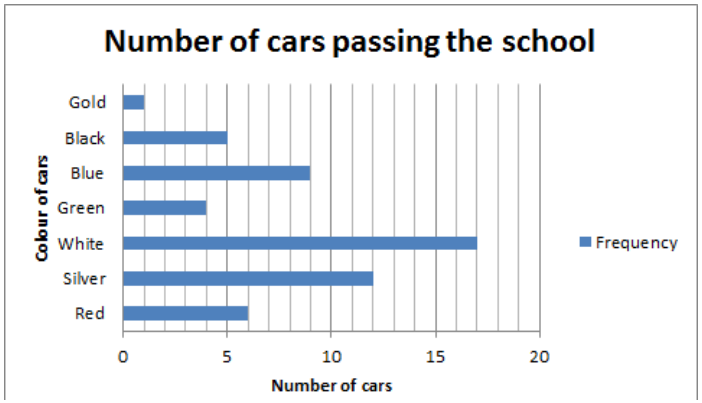
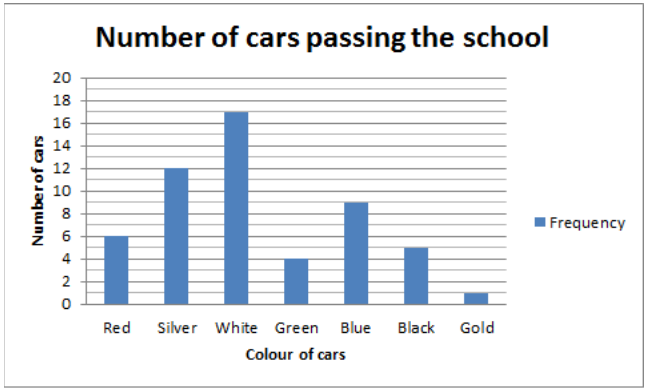
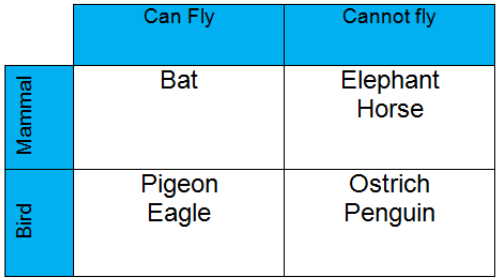
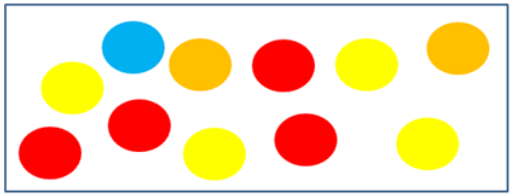
# Data and analysis

| Terms | Illustrations | Definitions |
|---|---|---|
| **Analysis of data** | | To make statements about a set of data based on interpretation of the results. |
| **Average** | | The average is known as the number typical of a set of numbers. It is also used as another term for the mean. |
| **Axis** |  | A fixed, reference line from which locations, distances or angles are taken. Usually grids have an x axis and y axis |
| **Bar chart / Bar graph** |  | A bar graph (also bar chart) is a graphical display of data using bars of different heights. They can also be displayed horizontally. |

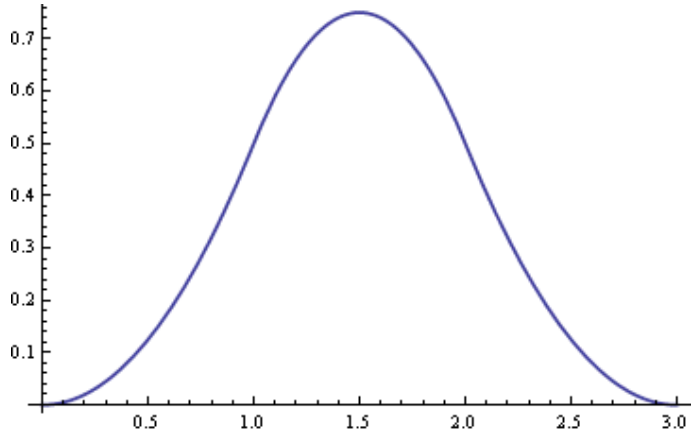
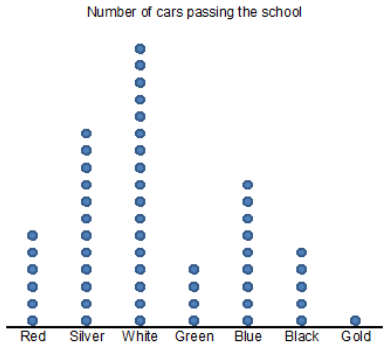
# Data and analysis

| | | |
|---|---|---|
| |  | |
| **Bias** | | A 'false' or 'invalid' result e.g. when collecting data on Scottish people's favourite supermarket, it would be biased if you conducted the research outside a certain supermarket as it could encourage them to say the one they have just been in to.<br>A systematic built-in error which makes all values wrong by a certain amount, e.g. Always measuring own height wearing shoes with thick soles. |
| **Carroll diagram** |  | A two way table used for grouping items according to characteristics. |

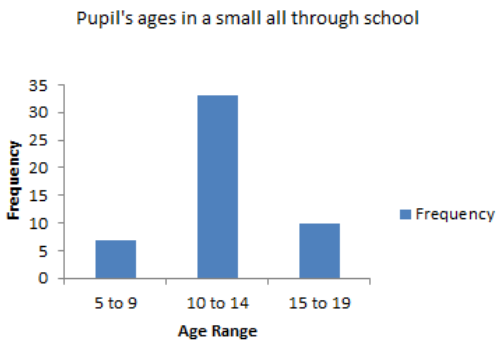
# Data and analysis

| | | |
|---|---|---|
| **Census** | | When data is collected for every member in a group. |
| **Certainty** | | The probability that an event will definitely happen. |
| **Chance** |  <br>There are 11 balls in this box.<br>The chances of pulling out a red ball is 4/11<br>The chances of pulling out a yellow ball is 4/11<br>The chances of pulling out a blue ball is 1/11<br>The chances of pulling out an orange ball is 2/11<br>It could be estimated from the calculated chances that;<br>• There is an equal chance of pulling out a red or yellow ball<br>• You are most likely to pick out a red or yellow ball<br>• You are least likely to pick out a blue ball | The number of times an event is likely to happen compared to the number of times it could happen.<br>For example;<br>There is a 1 in 6 chance of throwing a 3 on a dice labelled 1-6. It is likely to happen once as there is only 1 number 3 on the dice but it could happen 6 times. |
| **Consequences** | | The impact a decision can make on yourself and on others. For example; Reading food labels when shopping for the family – The majority of the food items state it is high in sugar, fat and calories. If this food is eaten each night, consequences for the family may be tooth decay long term, gradual weight gain etc. |
| **Continuous data** | | Continuous data is measured and can be any value within a range *e.g. the length of a leaf.*<br>The time taken to run a race is continuous as all measurements have meaning. |

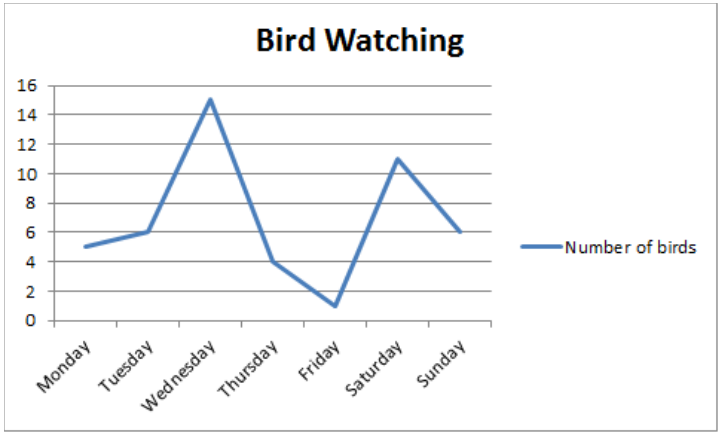
# Data and analysis

| | | |
|---|---|---|
| **Data** | | A collection of facts, such as numbers, words, measurements, observations |
| **Discrete data** | | Discrete data is counted and can only take certain values - like whole numbers *e.g. the number of cars passing by a school.*<br><br>Shoe size is an example of discrete data as size 6 and 7 have a meaning however size 6.2 does not. |
| **Distribution** |  | The distribution is a description of the overall shape of the data when displayed graphically. |
| **Dot plots** |  | A graphical display of data using dots. |

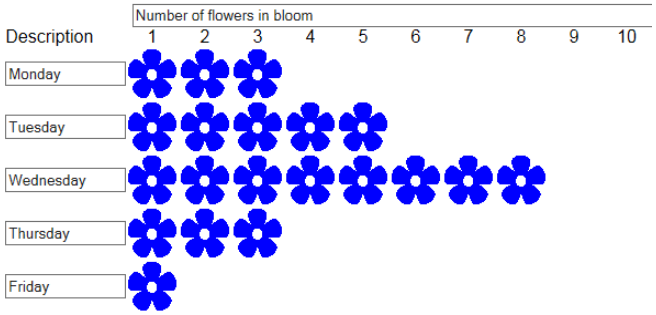
# Data and analysis

<table>
<tr><td>Draw conclusions</td><td></td><td>To make statements about a set of data based on results.</td></tr>
<tr><td>Event</td><td></td><td>A single result of an experiment.</td></tr>
<tr><td>Frequency table</td><td>
<table>
<tr><th colspan="3">Number of cars passing the school</th></tr>
<tr><th>Colour</th><th>Tally Marks</th><th>Frequency</th></tr>
<tr><td>Red</td><td>𝍸 I</td><td>6</td></tr>
<tr><td>Silver</td><td>𝍸 𝍸 II</td><td>12</td></tr>
<tr><td>White</td><td>𝍸 𝍸 𝍸 II</td><td>17</td></tr>
<tr><td>Green</td><td>IIII</td><td>4</td></tr>
<tr><td>Blue</td><td>𝍸 IIII</td><td>9</td></tr>
<tr><td>Black</td><td>𝍸</td><td>5</td></tr>
<tr><td>Gold</td><td>I</td><td>1</td></tr>
</table>
</td><td>A table used to note tally marks and show frequencies of each item.</td></tr>
<tr><td>Grouped data</td><td>For example: Data sorted into classes e.g. 11-15, 16-20.</td><td>Data is grouped together into intervals.<br>For example: Data sorted into classes e.g. 11-15, 16-20.</td></tr>
<tr><td>Histogram</td><td> 
</td><td>A graphical display of data using bars of different heights, it is similar to a bar chart but a histogram groups numbers into ranges.</td></tr>
</table>

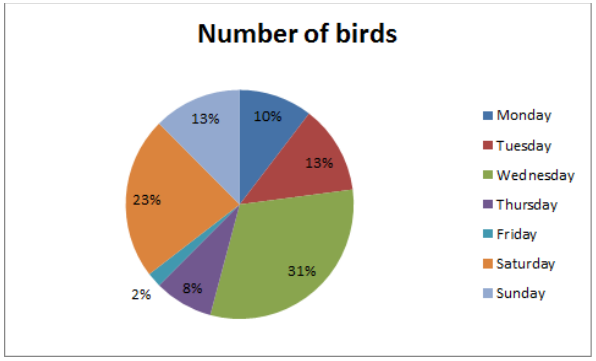
# Data and analysis

| | | |
|---|---|---|
| **Labelling** | | The labels on a graphical representation which give further information about the data.<br>E.g. if gathering data on minibeasts in the school garden, the y axis would be labelled 'number of minibeasts' and the x axis would be labelled 'types of minibeasts'. |
| **Least common / least popular** | | The number or item which appears least often in a set of data. |
| **Likelihood** | | The chance that an event will happen. |
| **Line Graph** |  | A graph that shows information that is connected in some way – such as change over a period of time |
| **Make predictions** | | Use data available to suggest what the future may be. |
| **Mean** | | The mean is the average of the set of data – it is the sum of the numbers divided by how many numbers there are.<br>For example: In the set of numbers 5, 5, 6, 7, 8, 12, 13, 15, 16 the mean would be (5+5+6+7+8+12+13+15+16) / 9 = 87/9 = 9.67 to 2 d.p. |

# Data and analysis

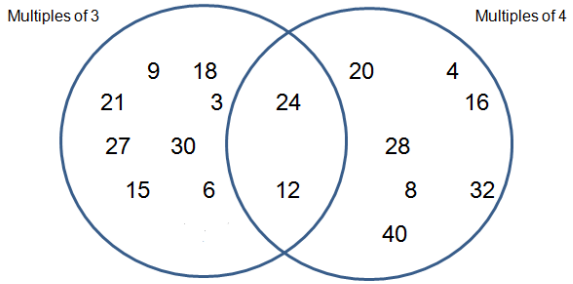
| | | |
|---|---|---|
| **Median** | | The median is the middle value in a sorted list of numbers.<br>For example: In the set of numbers 5, 5, 6, 7, 8, 12, 13, 15, 16 the median would be 8. |
| **Misleading information** | | Information which has been adapted by either presentation or selection to give the wrong impression of the true data. |
| **Mode** | | The mode is the number which appears most often in a set of data.<br>For example: In the set of numbers 5, 5, 6, 7, 8, 12, 13, 15, 16 the mode would be 5. |
| **Most common / most popular** | | The number or item which appears most often in a set of data. |
| **Pictogram** |  | A Pictogram or Pictograph is a way of showing data using images. |

# Data and analysis

| | | |
|---|---|---|
| **Pie chart** |  | A chart which uses 'pie slices' to show relative sizes of data. The sections of the chart can be recorded in percentages, e.g. half of the pie represented 50% of the data collected. |
| **Predictions** | | An educated guess at future events based on past experiences. E.g. predicting the weather in December. |
| **Probability** | | How likely something is to happen – calculated as the number of times an event actually happened divided by the number of possible events. It can be expressed as a fraction, decimal fraction or percentage. |
| **Qualitative** | | Descriptive information. |
| **Quantitative** | | Numerical information. |
| **Questionnaire** | | A set of questions used to gather information during a survey. |
| **Range** | | The range is the highest number in the set take away the lowest.<br>For example: In the set of numbers 5, 5, 6, 7, 8, 12, 13, 15, 16 the range would be (highest – lowest) 16 - 5 = 11. |
| **Raw data** | | Raw data is the data collected for example in a survey. |

# Data and analysis

| | | |
|---|---|---|
| **Robust information** | | Robust information has been gathered and presented in an appropriate way. |
| **Sample** | | A selection taken from a larger group (the "population") so that you can examine it to find out something about the larger group. |
| **Sample size** | | The number of pieces of information gathered from the sample in order to represent the whole “population.”<br>E.g. 100 men were surveyed to find out how many hours they spent exercising each week. (100 is the sample size). |
| **Scale** | | The intervals that are used on a graphical representation of data e.g. a scale which rises in ones or in tens, etc. |
| **Stem and leaf plots** | **Data Set:**<br>11, 12, 13, 13, 14, 18, 23, 24, 27, 27, 31, 34, 36, 42<br>**Stem and Leaf Plot:**<br>Stem \| Leaf<br>1 \| 1 2 3 3 4 8<br>2 \| 3 4 7 7<br>3 \| 1 4 6<br>4 \| 2<br>Key: 1 \| 2 means 12 | A table where each data value is split into a "leaf" (usually the last digit) and a "stem" (the other digits).<br>For example "32" is split into "3" (stem) and "2" (leaf).<br>The "stem" values are listed down, and the "leaf" values are listed next to them. |
| **Survey** | | Gathering information about a certain topic or issue for a particular reason. The information can help people make decisions about topics of interest e.g. most popular holiday destinations for young families. |
| **Tally Marks** | | A visual representation of the number of times an item |

| | | |
|---|---|---|
| | | appears in a set, these are bundled in groups of five.<br>For example: \|\| represents 2 and 𝍸 represents 5 |
| **Trends** | | The overall picture of a set of data over time – e.g. the temperature is rising over time.<br>For example: House prices, over time, in the UK have shown an upward trend. |
| **Uncertainty** | | The probability that an event may not happen. |
| **Vague information** | | Vague information is information which is presented without using all available information. |
| **Venn Diagram** |  | A diagram that shows all possible logical relations between a collection of sets of data.<br>For example: Appropriate question would be:<br>What are the common multiples of 3 and 4? |