9 Representing data sets

Data collected during numeracy activities can often be represented in a variety of ways: as tables of numerical values, as graphs, maps or scale drawings. It is sometimes possible to identify patterns in collections of data.

Patterns are important, as they can be used in new situations to make predictions about similar sets of data. For example, we are familiar with the formulae for the properties of circles:

$$C = 2\pi r \qquad A = \pi r^2$$

If the radius **r** of any circle is known, then the circumference **C** or the area **A** can be easily calculated.

Another familiar formula is Ohm's Law, relating the current **I**, voltage **V** and resistance **R** in an electrical circuit. This can be written in three equivalent forms:

$$V = IR \qquad I = \frac{V}{R} \qquad R = \frac{V}{I}$$

If any two of the quantities **current**, **voltage** or **resistance** are known, then the third can be calculated.

Sometimes it is convenient to produce algebraic expressions to represent patterns in data as a way of saving time with future calculations. For example, engineering students might use trigonometry to devise a formula which links the side **L** of a hexagon bolt or nut with the width across the flats **W.** If one of these quantities is known, the other can then be quickly calculated.
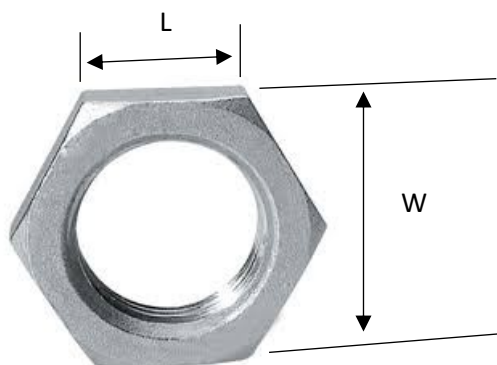

Figure 236: Dimensions of a hexagon nut

A useful skill in numeracy is the ability to move easily between numeric, algebraic and geometric representations of particular collections of data. Patterns in the data may be identified by plotting graphs, and it may then be possible to represent the patterns as algebraic expressions for use in solving similar problems.

As a first example of analysing sets of data, we will look at the factors which affect the stopping distance of a car. This can be an important way of increasing safety awareness for young drivers.

Stopping distances

Braking distance for a vehicle depends on two sets of factors: those affecting the reaction time of the driver, and those affecting the deceleration rate for the vehicle when the brakes are applied. A typical set of stopping distances are included in the Highway Code:
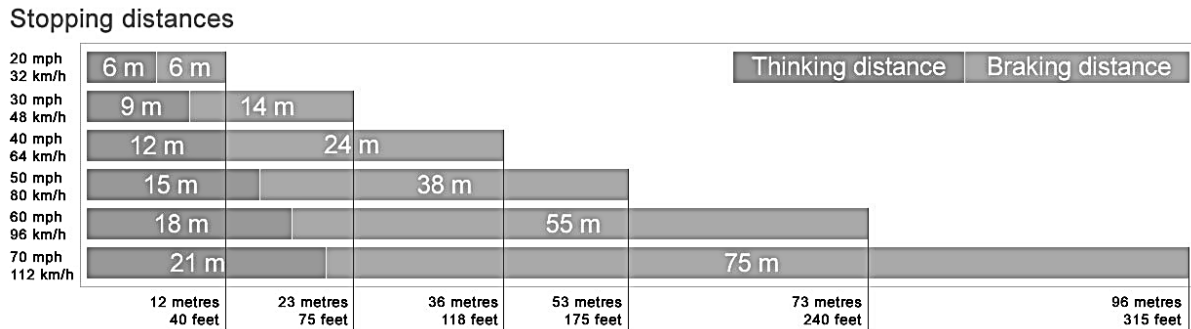
Stopping distances

	Thinking distance	Braking distance
20 mph / 32 km/h	6 m	6 m
30 mph / 48 km/h	9 m	14 m
40 mph / 64 km/h	12 m	24 m
50 mph / 80 km/h	15 m	38 m
60 mph / 96 km/h	18 m	55 m
70 mph / 112 km/h	21 m	75 m

| 12 metres / 40 feet | 23 metres / 75 feet | 36 metres / 118 feet | 53 metres / 175 feet | 73 metres / 240 feet | 96 metres / 315 feet |

Figure 237: Stopping distances for cars travelling at different speeds

We can enter these figures in a spreadsheet and plot graphs of the data:

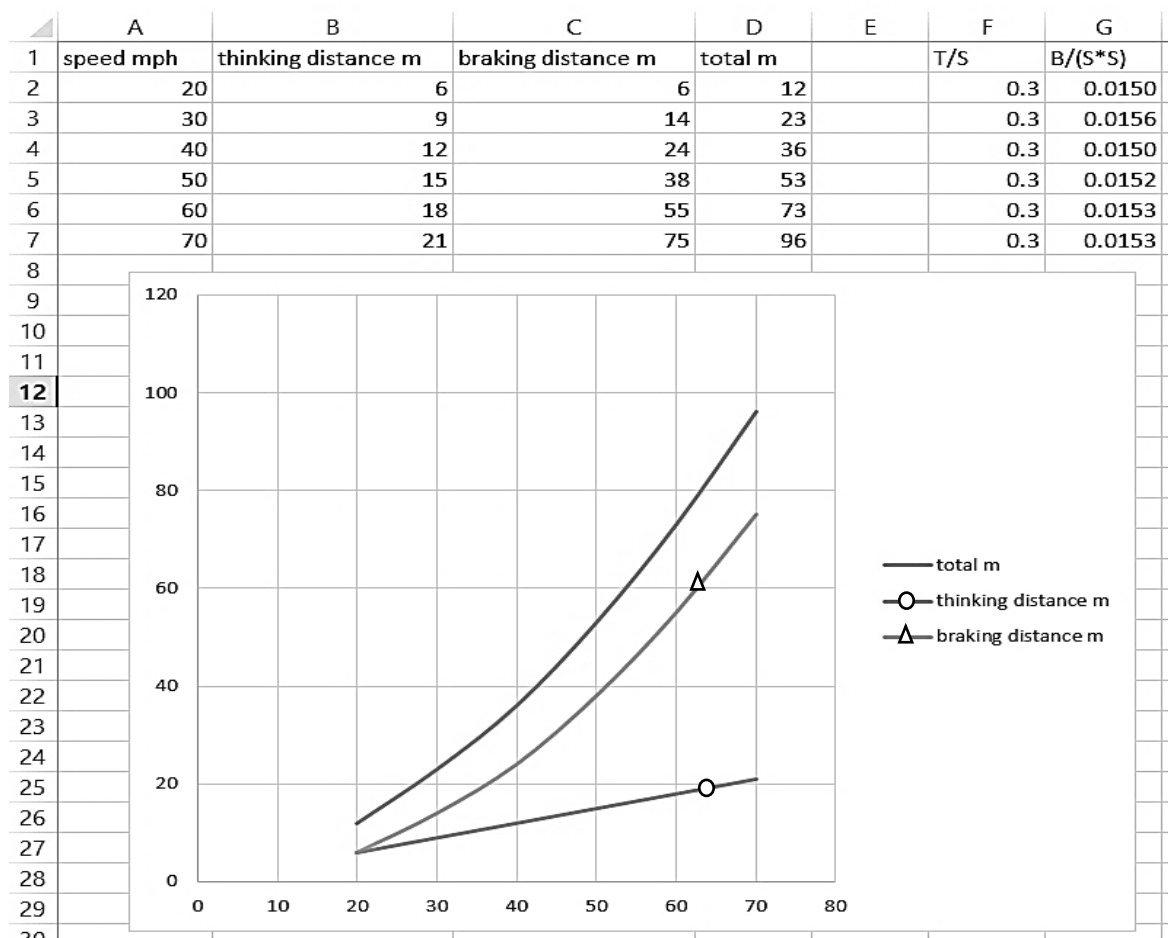|	A	B	C	D	E	F	G
1	speed mph	thinking distance m	braking distance m	total m		T/S	B/(S*S)
2	20	6	6	12		0.3	0.0150
3	30	9	14	23		0.3	0.0156
4	40	12	24	36		0.3	0.0150
5	50	15	38	53		0.3	0.0152
6	60	18	55	73		0.3	0.0153
7	70	21	75	96		0.3	0.0153

Figure 238: Thinking and braking distances at different speeds

Columns have been added to the spreadsheet in figure 238 to explore the relationship between speed and thinking distance, and the relationship between speed and braking distance. It is found that, within reasonable accuracy, two constant ratios exist:

$$\frac{thinking\ distance: m}{speed: mph} = 0.3$$

$$\frac{braking\ distance: m}{(speed: mph)^2} = 0.015$$

We can combine these results to produce an overall formula for total stopping distance:

$$total\ distance: m = (0.3 \times speed: mph) + (0.015 \times speed: mph^2)$$

We might point out that thinking distance is based on a sudden unexpected event occurring, and represents the delay as the driver becomes aware of the situation and begins to apply the brakes.   If the driver is already anticipating the actions of other road users and can spot potentially dangerous situations developing, the thinking time will be reduced or eliminated completely. This considerably improves the safety margin.

Figure 239: A potential road hazard developing

In this example situation, an alert driver would anticipate that the cyclist will move out into the car lane to avoid parked vehicles ahead. The cyclist may not be aware of cars following him, and may give no signal before turning into the traffic.

The stopping distances given in figure 237 are a general average under normal driving conditions, and can change as a result a variety of factors such as weather conditions, the condition of a car's tyres, and the level of concentration of the driver. We will investigate these factors further:

Figure 240 provides more detailed data on braking distances at different speeds for both dry and wet road conditions. As we might expect, stopping distance will be greater on a wet road.
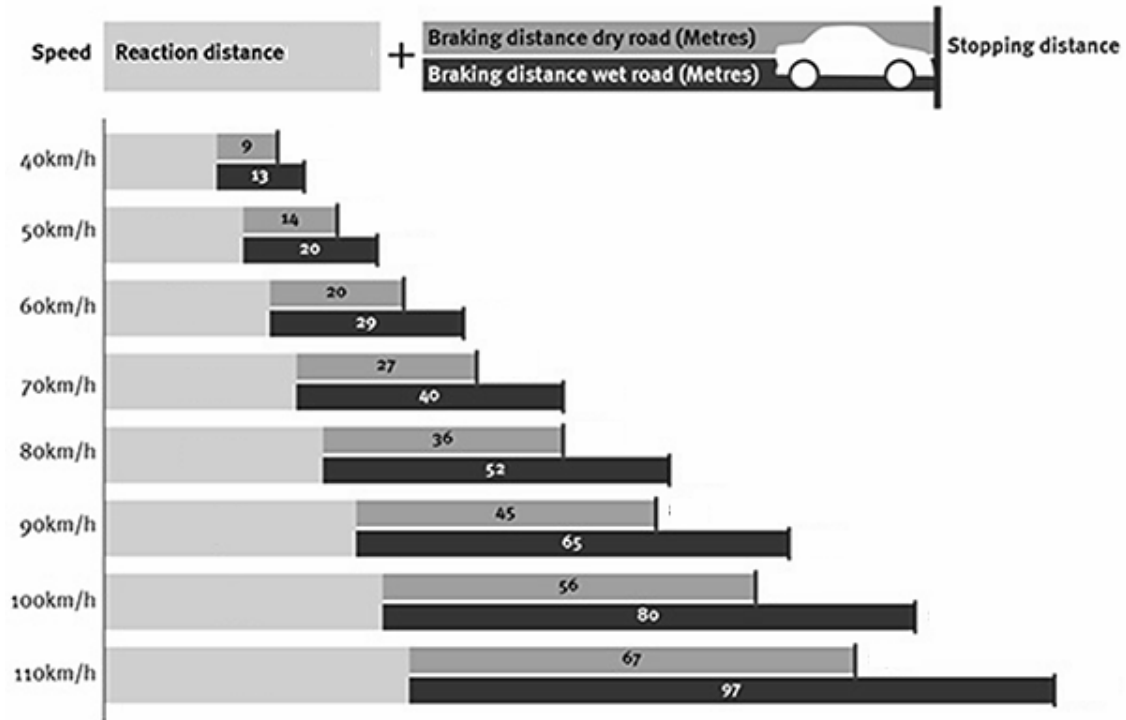
Figure 240: Effect of weather conditions on braking distances

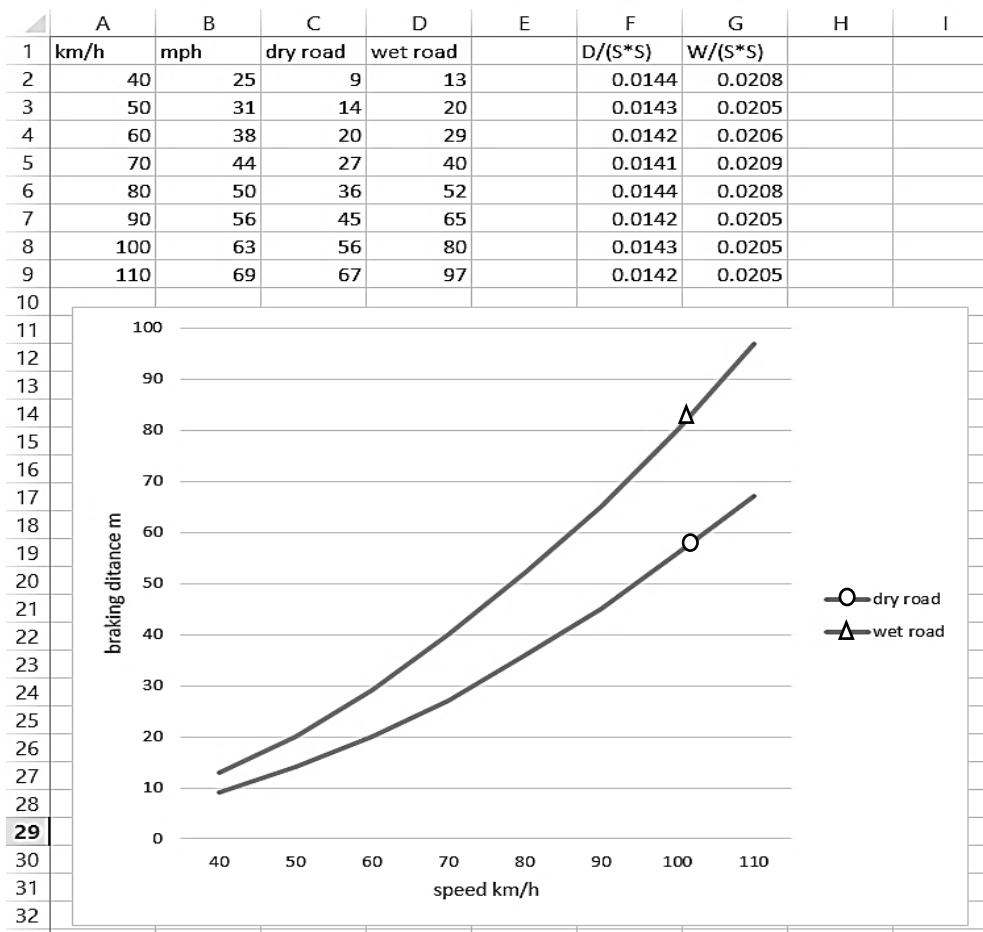The figures in figure 240 have been entered into a spreadsheet and plotted as a graph.

	A	B	C	D	E	F	G	H	I
1	km/h	mph	dry road	wet road		D/(S*S)	W/(S*S)		
2	40	25	9	13		0.0144	0.0208		
3	50	31	14	20		0.0143	0.0205		
4	60	38	20	29		0.0142	0.0206		
5	70	44	27	40		0.0141	0.0209		
6	80	50	36	52		0.0144	0.0208		
7	90	56	45	65		0.0142	0.0205		
8	100	63	56	80		0.0143	0.0205		
9	110	69	67	97		0.0142	0.0205		

Figure 241: Braking distances at different speeds for dry and wet road conditions

Columns have again been added to the spreadsheet explore the relationships between speed and braking distance for both dry and wet road conditions. Speeds were converted from km/hour to miles/hour, to allow a direct comparison with our earlier formulae. It is found within reasonable accuracy that two ratios exist:

$$dry\ road: \qquad \frac{braking\ distance:m}{(speed:mph)^2} = 0.0143$$

$$wet\ road: \qquad \frac{braking\ distance:m}{(speed:mph)^2} = 0.0206$$

This leads to two alternative equations for total stopping distance:

Dry conditions:
$$total\ distance = (0.3 \times speed) + (0.0143 \times speed^2)$$
Wet conditions:
$$total\ distance = (0.3 \times speed) + (0.0206 \times speed^2)$$

The next factor we might consider is the state of the car's tyres. We have one set of data from tests carried out at 50mph:
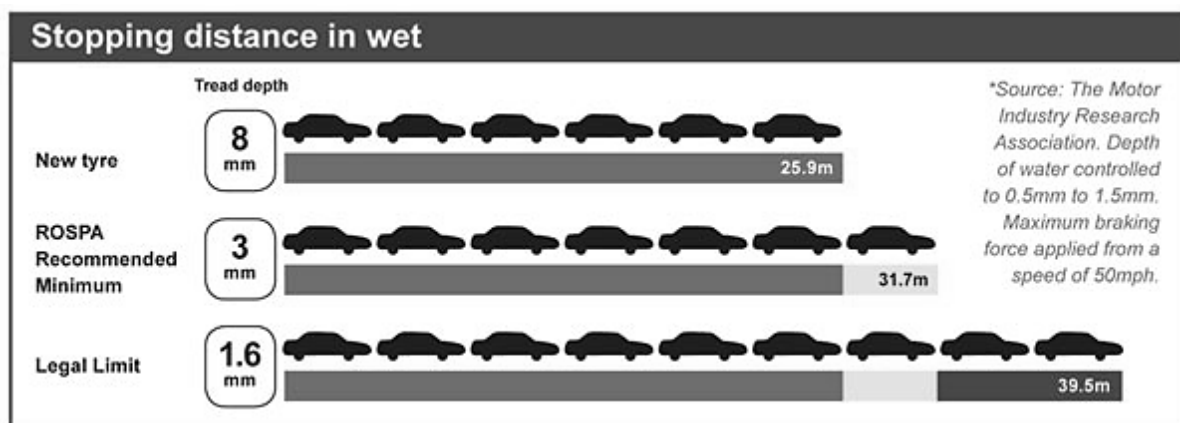
Figure 242: Braking distance at 50mph, in relation to tyre tread depth

Let us assume that many cars have tyres with around the minimum recommended tread depth of 3mm. According to the tests in figure 242, the braking distance would increase from 31.7m to 39.5m if the car tyres were worn to the legal limit. This represents an increase in braking distance of:

$$\frac{(39.5 - 31.7)}{31.7} \times 100\% = 25\%$$

We might reasonably suppose that this 25% increase in braking distance would also apply at other speeds. We are not told the road conditions for the tests, but supposing that a car with heavily worn tyres has to brake on a wet road:

$$total\ distance:m = (0.3 \times speed:mph) + (0.0206 \times speed:mph^2 \times 1.25)$$

We will now turn our attention to the reaction time of the driver. A small research project was carried out to investigate the relative effects of drinking alcohol up to the legal limit, reading and texting while driving. Figure 243 shows the increases in stopping distances for two different drivers.
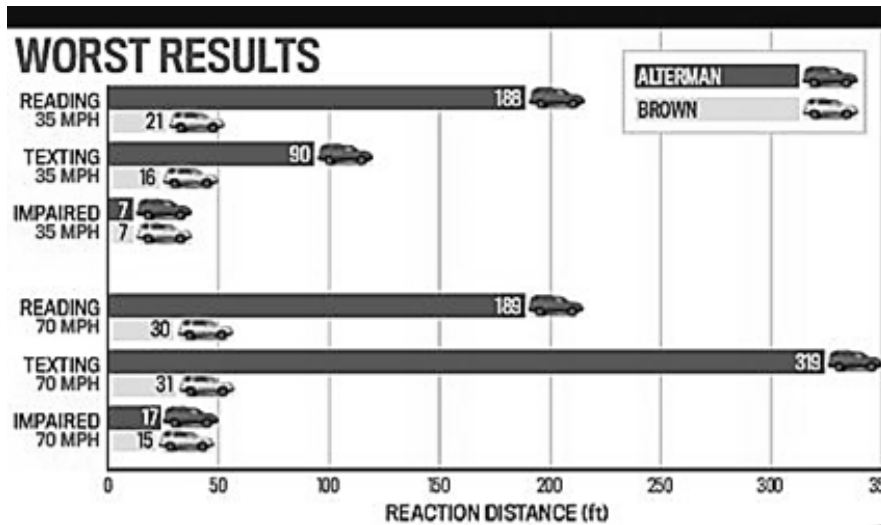

WORST RESULTS

		ALTERMAN	BROWN

READING 35 MPH 188 / 21

TEXTING 35 MPH 90 / 16

IMPAIRED 35 MPH 7 / 7

READING 70 MPH 189 / 30

TEXTING 70 MPH 319 / 31

IMPAIRED 70 MPH 17 / 15

REACTION DISTANCE (ft) 0 50 100 150 200 250 300 350

Figure 243: Increases in stopping distances for two drivers when distracted by reading or texting, or when impaired by drinking alcohol up to the legal limit

It is clear that reading or texting while driving has a much more serious effect on reaction time than a moderate amount of alcohol consumption.   The results show large variations between the two drivers, perhaps depending on their different abilities to concentrate on multiple tasks at the same time.

To obtain useable data, the thinking distances have been converted from feet to metres, then averaged between the two drivers to obtain values for the increase in response to the three distracting factors.

	A	B	C	D	E	F	G	H	I	J
1			ft			m			mean m	
2		Drink	Reading	Texting	Drink	Reading	Texting	Drink	Reading	Texting
3	35 mph									
4	A	7	188	90	2.1336	57.3024	27.432	2.1336	31.8516	16.1544
5	B	7	21	16	2.1336	6.4008	4.8768			
6										
7	70 mph									
8	A	17	189	319	5.1816	57.6072	97.2312	4.8768	33.3756	53.34
9	B	15	30	31	4.572	9.144	9.4488			
10										
11										
12			increased distance m per 10mph					0.653143	6.9342	6.117771
13										
14										

Figure 244: Estimates of average increase in thinking distance for drivers distracted by reading or texting, or influenced by alcohol up to the legal limit

Whilst these results are very approximate and should be treated with caution, we have some indication of the relative effects on drivers.

These results provide further formulae for total stopping distance:

For a driver at the legal alcohol limit:

$$total\ distance = ((0.3 + 0.065) \times speed) + (0.0143 \times speed^2)$$

For a driver who is distracted by reading or texting:

$$total\ distance = ((0.3 + 0.65) \times speed) + (0.0143 \times speed^2)$$

The results we have collected can now be plotted in a spreadsheet to compare stopping distances under different situations which might affect the driver and vehicle:
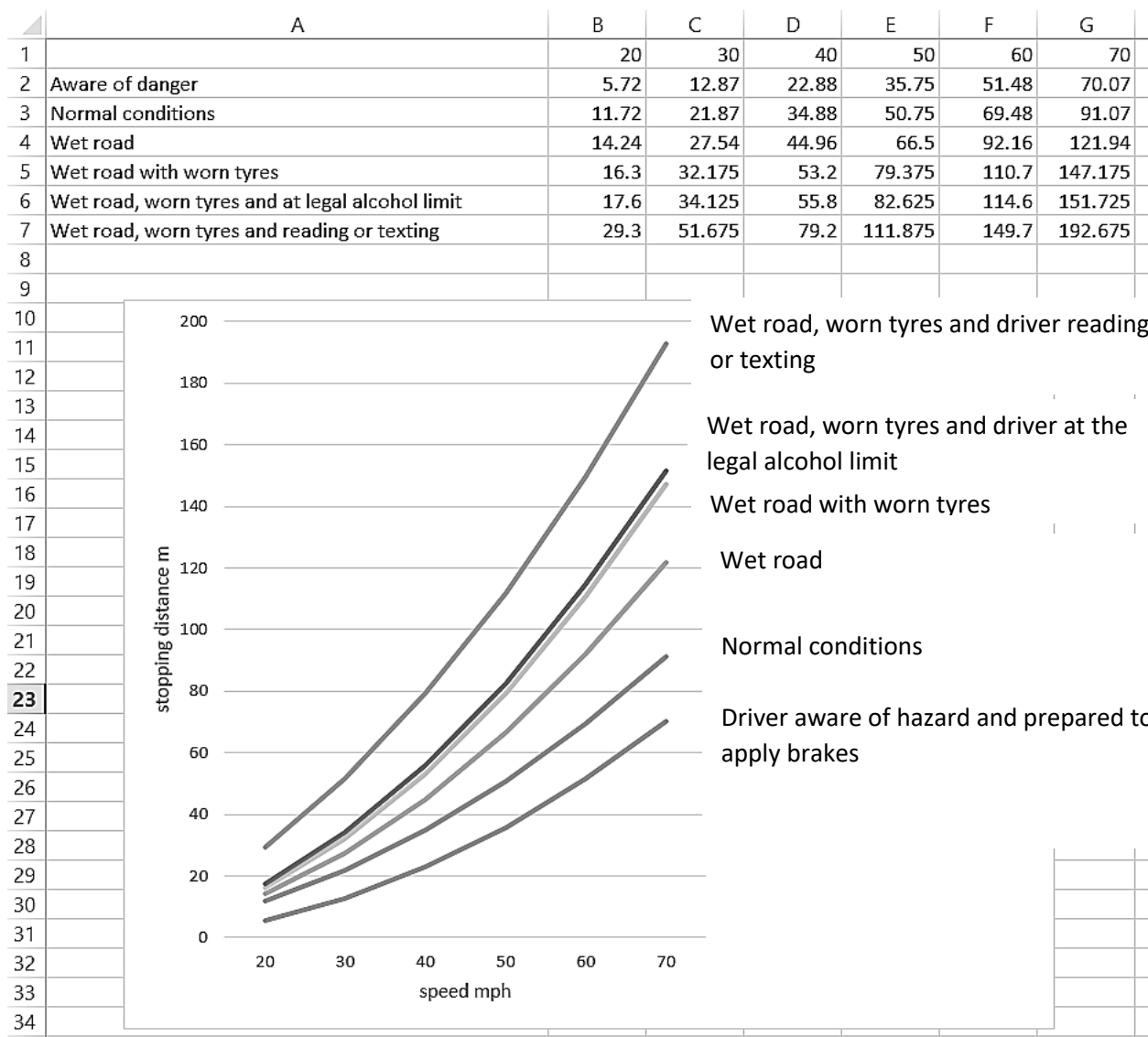
	A	B	C	D	E	F	G
1		20	30	40	50	60	70
2	Aware of danger	5.72	12.87	22.88	35.75	51.48	70.07
3	Normal conditions	11.72	21.87	34.88	50.75	69.48	91.07
4	Wet road	14.24	27.54	44.96	66.5	92.16	121.94
5	Wet road with worn tyres	16.3	32.175	53.2	79.375	110.7	147.175
6	Wet road, worn tyres and at legal alcohol limit	17.6	34.125	55.8	82.625	114.6	151.725
7	Wet road, worn tyres and reading or texting	29.3	51.675	79.2	111.875	149.7	192.675
8							
9							

Figure 245: Stopping distances under different conditions which might affect a driver and vehicle

Stopping distances clearly increase when the driver is distracted, or when road and tyre conditions are poor. To see how this might matter, we will look at an example:

Figure 246 shows a section of road in Bangor, North Wales. A series of distance measurements have been added.

Figure 246: Ffordd Deiniol, Bangor

If the car ahead brakes very suddenly in an emergency, then about 20 metres of stopping distance will be needed to avoid a collision. Referring to the graph in figure 245:

> A driver would be able to stop safely from a speed of 20 mph, except in a situation where they were distracted by reading or texting!

> At 30 mph, the driver could possibly stop in time if the emergency was totally unexpected. However, they would probably stop safely if aware of traffic conditions and anticipating a possible hazard ahead.

> At speeds above 30 mph, particularly in wet road conditions or with worn tyres, it is unlikely that the driver would avoid a collision.

Lessons from this work are clear: Drivers should avoid distractions, and maintain a correct speed according to the traffic and weather conditions. Keep a safe separation distance from the vehicle ahead, and try to anticipate road hazards which are developing.

River flow

The next example of identifying a pattern in a data set examines river flood levels for the town of Dolgellau in North Wales. Dolgellau has historically experienced flooding from the River Wnion, with the most recent serious flood in the centre of the town occurring in the 1960's (figure 247).

Figure 247:

Flooding in Bridge Street, Dolgellau

Students have carried out a project to investigate whether the flood level in the town could be predicted by monitoring the water level upstream in the headwaters of the River Wnion.

A first step was to construct electronic water level monitoring equipment. Following the designs of Keeland, Dowd and Hardegree (1997), barometric water depth recorders were assembled. These make use of pressure transducers which are encased in plastic tubes and attached to the river bed.
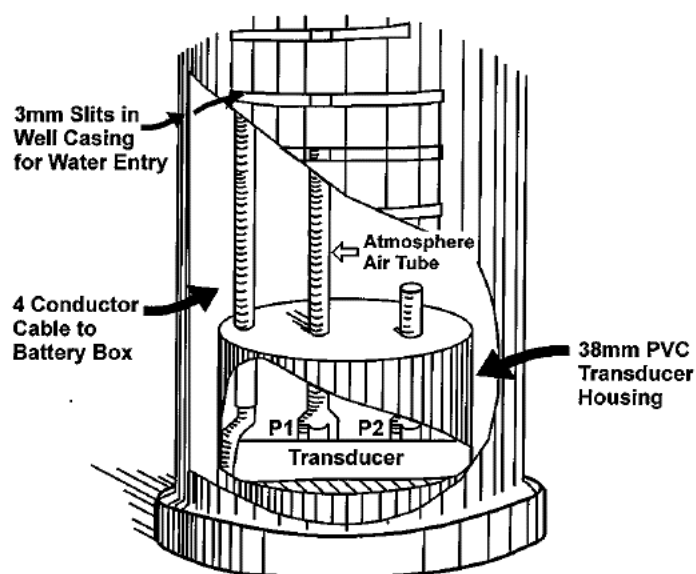
Figure 248: Pressure transducer assembly for monitoring water depth (Keeland et al., 1997)

The pressure transducer consists of a piezo-electric plate which changes its resistance when it is flexed. One side of the plate is exposed to water at the river bed, whilst the other is connected to a plastic tube running up to the river bank where it is open to atmospheric air pressure. As the depth of water in the river changes, the differential pressure across the plate changes. The piezo-electric resistance is incorporated into an electronic circuit so that changes in pressure produce a varying output voltage.

The electronic circuitry and power supply are housed in a metal container which can be secured on the river bank. An electronic data logger records the output voltage at fixed intervals of 5 minutes. The data could then be downloaded to a laptop computer and processed by spreadsheet.

Figure 249: Electronic circuitry, power supply and data logger for the water depth recorder.

After downloading data, the recorded voltages are converted to water depths in metres. Voltage is linearly related to water depth, so a simple formula can be used:

$$water\ depth: m = A + (B \times voltage)$$

where A and B are two constants which are determined experimentally during the calibration of the equipment.

River depth was has been monitored for a number of months in the headwater stream of the River Wnion at Pared yr Ychain (figure 250) and below the town of Dolgellau (figure 251). Typical results are shown in figure 252, where separate graph lines represent the upstream and downstream hydrographs. Water depths are very much greater in Dolgellau than in the headwaters stream, but the graph lines seem to show a similar sequence of peaks and troughs.

Figure 250:

Headwater stream at Pared yr Ychain

Figure 251:

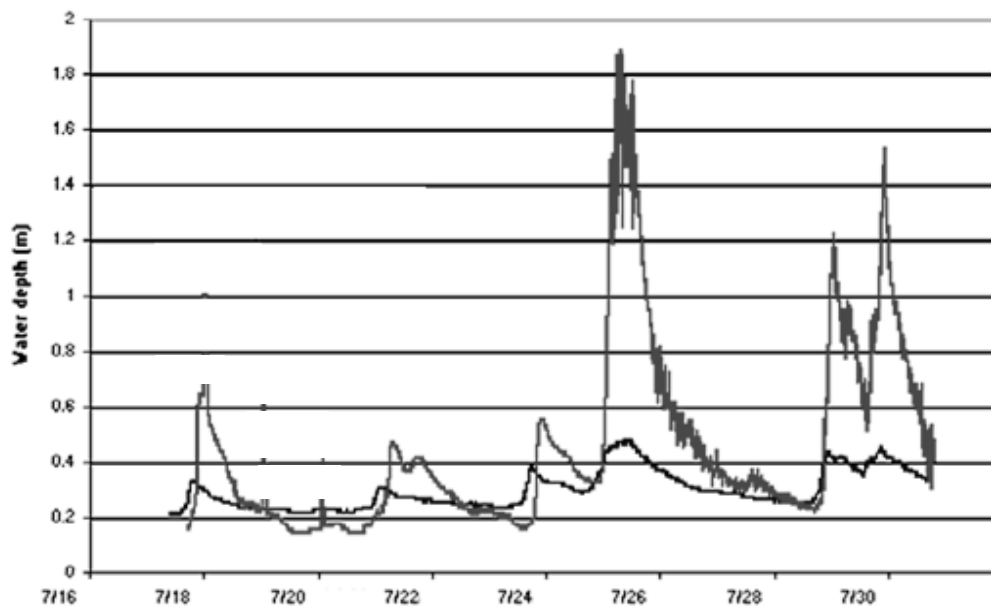River Wnion in Dolgellau under flood conditions

Figure 252: Hydrographs for the River Wnion at Pared yr Ychain and Dolgellau

Experiments with the spreadsheet led to a discovery that the upstream hydrograph line could be transformed to closely fit the downstream hydrograph using a function:

$$D = Ae^{dB} + C$$

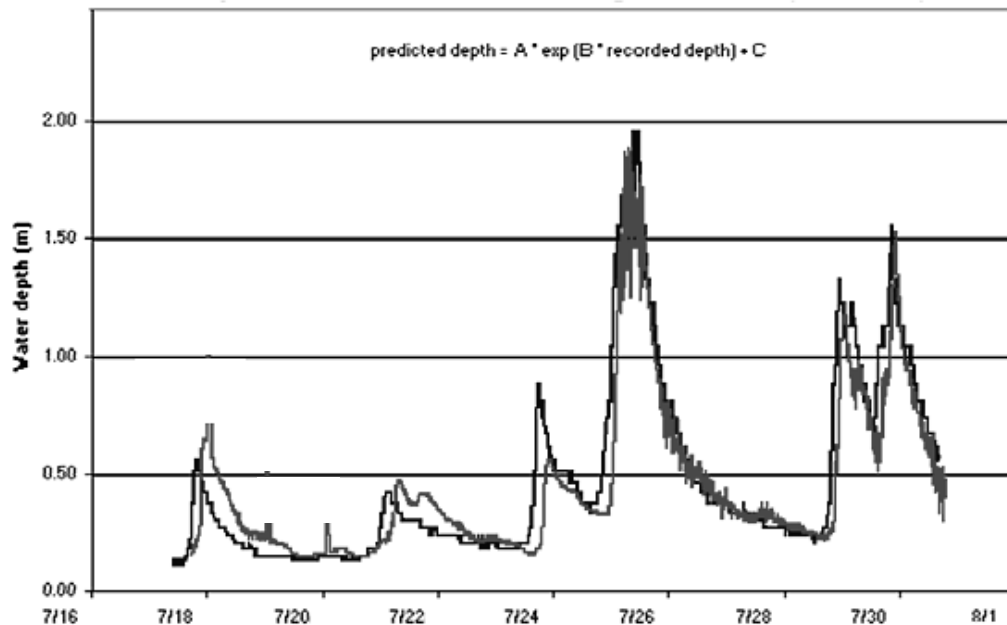where **D** is the downstream depth and **d** is the river depth monitored in the headwaters.

Figure 253: Upstream hydrograph transformed to fit the downstream hydrograph

The equation parameters were determined empirically as:

A = 0.05, B = 8.2, C = -0.15

The significance of this work is that the transformed upstream hydrograph is 3 hours ahead of the downstream hydrograph, providing a very accurate prediction of flood levels to be expected 3 hours later in the town.

This flood prediction system could be automated, with upstream river depth values transmitted in real time by mobile phone link to a processing computer.

In the next section we will look at another area of study, Health and Social Care, where patterns in data are becoming increasingly important for decision making with patients. We will examine calculations of human body parameters:

- Body mass index, and Body fat percentage. These values may be used to identify obesity and malnutrition
- Body surface area, and Blood volume. These values may be used to determine drug dosage

Human body parameters

Body Mass Index

Body Mass Index is a measure of body weight relative to height, and indicates whether a person is underweight, overweight or in the healthy weight range.

Body Mass Index is calculated using the formula:

$$BMI = \frac{bodyweight: kg}{(height: m)^2}$$

Body Mass Index is an estimate of body fat and provides a guide to the risk factor for diseases associated with excess body fat. A higher BMI can indicate a higher risk from heart disease, high blood pressure, type 2 diabetes, gallstones, breathing problems, and certain cancers. The relative levels of risk have been assessed by calculating average mortality ratios for groups of randomly selected individuals within particular ranges of Body Mass Index. Mortality ratio represents the ratio of observed deaths in each study group to expected deaths in the general population.
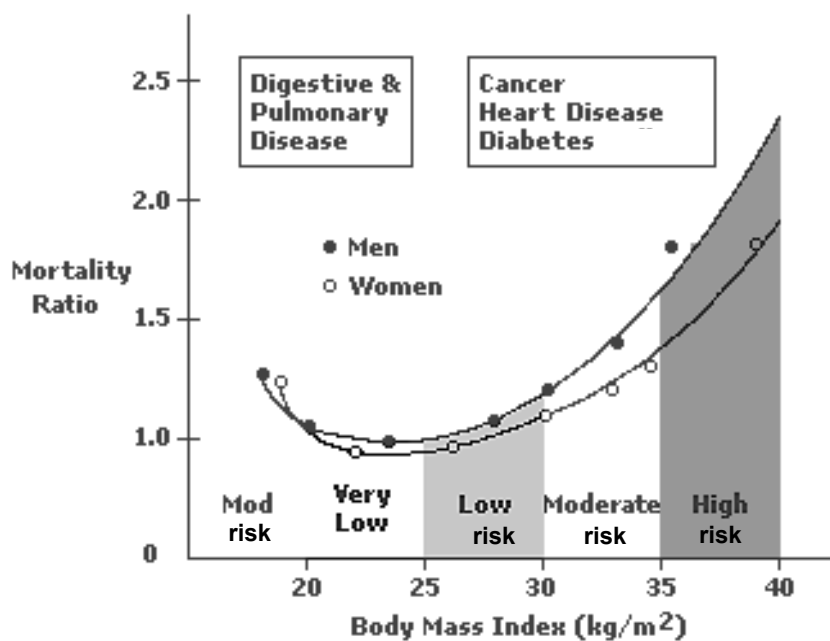
www.biodyncorp.com/product/450/bmi_450.html

Figure 254: Mortality ratio associated with different ranges of Body Mass Index

Based on this work, a classification of Body Mass Index values which is commonly accepted is:

Underweight	Below 18.5
Normal	18.5–24.9
Overweight	25.0–29.9
Obesity	30.0 and Above

Although Body Mass Index measurements can be applied to most men and women, the method does have some limitations for determining healthy weights. The standard BMI ranges may be inappropriate in the cases of:

- Competitive athletes and bodybuilders, because heavier muscle weight may affect the results.
- Pregnant or nursing women, because they need more fat reserves than usual.
- People over 65, because BMI values as high as 29 may not be unhealthy at this age.

Measurement of waist circumference can help with screening for possible health risks associated with excess weight and obesity. If most fat is around the waist rather than at the hips, there is a higher risk for heart disease and type 2 diabetes. This risk increases with a waist size greater than 35 inches for women or greater than 40 inches for men. To correctly measure the waist, a tape measure should be placed around your middle, just above the hipbones. The measurement is made after breathing out.

Classification of Overweight and Obesity by BMI, Waist Circumference, and Associated Disease Risks

	BMI (kg/m^2)	Obesity Class	Disease Risk Relative to Normal Weight and Waist Circumference	
			Men 102 cm (40 in) or less Women 88 cm (35 in) or less	Men > 102 cm (40 in) Women > 88 cm (35 in)
Underweight	< 18.5		-	-
Normal	18.5–24.9		-	-
Overweight	25.0–29.9		Increased	High
Obesity	30.0–34.9	I	High	Very High
	35.0–39.9	II	Very High	Very High
Extreme Obesity	40.0 +	III	Extremely High	Extremely High

www.nhlbi.nih.gov/health/educational/lose_wt/BMI/bmi_dis.htm

Figure 255: Combining BMI and waist measurements to evaluate risk from obesity

Weight loss would be recommended for people with a BMI greater than or equal to 30, or those with a BMI of 25 to 29.9 and an increased risk factor due to fat around the waist. Even a small weight loss of between 5 and 10 percent of current weight can help lower the risk of developing diseases associated with obesity. People who are overweight but do not have a high waist measurement may just need to avoid further weight gain, rather than lose weight.

Sometimes it is not possible to make direct measurements of a person's height and weight, for example: if a patient is confined to bed and it is undesirable for them to stand upright to use a weighing scale. In these cases, estimates can be made from other body parameters.

- Tables are available for estimating height from a measurement of the ulna bone in the forearm (figure 256).
- Body Mass Index can be roughly estimated by measuring the upper arm circumference, at a point half way between the shoulder and elbow. The tape measure should be snug but not tight.
  - If the mid upper arm circumference is less than 23.5cm, the BMI is likely to be less than 20 kg/m^2.
  - If the mid upper arm circumference is more than 32.0cm, the BMI is likely to be more than 30 kg/m^2.

Estimating height from ulna length

Measure between the point of the elbow (olecranon process) and the midpoint of the prominent bone of the wrist (styloid process) (left side if possible).

Height (m)	men (<65 years)	1.94	1.93	1.91	1.89	1.87	1.85	1.84	1.82	1.80	1.78	1.76	1.75	1.73	1.71
	men (≥65 years)	1.87	1.86	1.84	1.82	1.81	1.79	1.78	1.76	1.75	1.73	1.71	1.70	1.68	1.67
	Ulna length (cm)	32.0	31.5	31.0	30.5	30.0	29.5	29.0	28.5	28.0	27.5	27.0	26.5	26.0	25.5
Height (m)	Women (<65 years)	1.84	1.83	1.81	1.80	1.79	1.77	1.76	1.75	1.73	1.72	1.70	1.69	1.68	1.66
	Women (≥65 years)	1.84	1.83	1.81	1.79	1.78	1.76	1.75	1.73	1.71	1.70	1.68	1.66	1.65	1.63

www.bapen.org.uk/screening-and-must/must/introducing-must

Figure 256: Estimating height from a measurement of the ulna bone in the arm

Body Mass Index for children and teenagers is calculated the same way as for adults. However, the BMI value and age then need to be located on a specific male or female age chart to determine whether the child is within a healthy weight range (figure 257).

Figure 257: Body Mass Index and age charts for boys and girls

Girls and boys develop differently and have different amounts of body fat at different ages, so a child's age is taken into account when considering their Body Mass Index. There is an assumption that at any age, children will have a range of BMI values which follow a normal distribution. Most will be close to the mean, with relatively small numbers of children with BMI values much smaller or larger than the mean. The 5% of children with the lowest BMI form the left tail of the normal distribution, and the 5% of children with the highest BMI form the right tail. The percentages are proportional to the shaded areas beneath the normal distribution histogram.
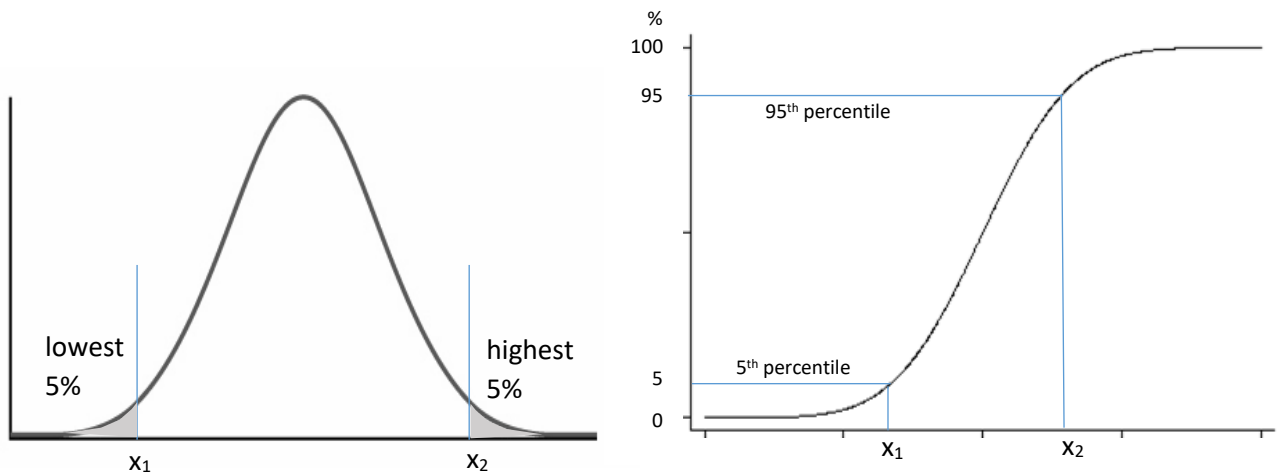
Figure 258: Normal distribution of BMI values, represented (left) as a histogram and (right) as a cumulative curve

It is often convenient to plot the normal distribution as a cumulative curve. This has a characteristic S-shape. The curve rises only slowly at first, as few children are added with low BMI values. The curve then rises steeply for BMI values around the mean as many children with these common readings are added, then the curve falls off more gently to add the last few children with high BMI values.

The point on the cumulative curve at which 5% of the children have been added is called the **5th percentile**. The point, much further on, where 95% of the children have now been added is called the **95th percentile**.

The charts shown in figure 257 summarise the percentile values obtained from the BMI data of a large number of children, giving an overview of the BMI distributions within the populations of boys and girls of different ages. To use the chart, a point is plotted for the child's age and BMI, then interpreted as:

Underweight	Less than the 5th percentile
Healthy weight	5th percentile to less than the 85th percentile
Overweight	85th to less than the 95th percentile
Obese	Equal to or greater than the 95th percentile

Maintaining a healthy weight during childhood is especially important for heart health.

Obese children have a high probability of staying obese throughout their entire lives.

In contrast to problems of obesity, **malnutrition** leading to excessively low body weight is a common health problem of particular concern amongst elderly people. Risk of malnutrition should be monitored by staff of care homes.

Malnutrition is characterised by a combination of low Body Mass Index and weight loss. The British Association for Parenteral and Enteral Nutrition (BAPEN) has produced a methodology for identifying patients at risk from malnutrition. This involves a series of stages:

Step 1: Body Mass Index is calculated, then a score is allocated:

        Greater than 20 kg/m^2:   score 0
        18.5 – 20 kg/m^2:   score 1
        Less than 18.5 kg/m^2:   score 2

Step 2: Weight loss during the past 3-6 months is assessed. Any planned weight loss on the recommendation of medical staff is excluded.

        Less than 5% weight loss: score 0
        5% - 10% weight loss: score 1
        More than 10% weight loss: score 2

Step 3: If the patient is acutely ill and is likely to have had no nutritional intake for more than 5 days, then a score of 2 is allocated.

The scores from steps 1 – 3 are then added, to determine the risk of malnutrition:

        Score 0:   low risk
        Score 1:   medium risk
        Score 2 or more: high risk

The outcome can then be used to plan appropriate patient care:

For **low risk** patients, it is simply necessary to continue with malnutrition screening at regular intervals. This might be carried out monthly in a care home.

For **medium risk** patients, dietary intake should be monitored for a period of 3 days. If this is adequate, then no further action is needed other than regular screening. If, however, the patient's diet is found to be inadequate, then a care plan to improve nutrition needs to be put in place.

For **high risk** patients, professional help should be sought from a dietician, and a plan put in place to improve nutrition as a matter of urgency.

Although Body Mass Index is used to screen for overweight and obesity, the technique has limitations. The parameter that BMI attempts to estimate is body fat. However, this estimate may not be accurate. For example, a child may have a high Body Mass Index for their age and sex, but to determine if excess fat is really a problem requires an assessment of **body fat percentage**. A person's total body fat percentage is the weight of the person's fat divided by the person's weight.

Body fat has various functions and includes **essential body fat** and **storage body fat**.

A certain percentage of body fat is essential to human functions such as: insulating internal organs and tissues, regulating body temperature, and for reproductive functions. The percentage of essential fat is 3 to 5% in men, and 8 to 12% in women.

Storage fat consists of fat accumulated in adipose tissue, whose main role is to store energy.

Typical percentages of body fat are:

	Females (% fat)	Males (% fat)
Athletes	14-20%	6-13%
Good fitness	21-24%	14-17%
Acceptable	25-31%	18-25%
Obese	32%+	25%+

A simple method of estimating body fat percentage is to use a Body Fat Calliper. Readings are then interpreted according to age and sex, as in the example chart in figure 259.

Body Fat % Measurement Chart for Men

Accu-Measure™ Reading in Millimeters

AGE	2-3	4-5	6-7	8-9	10-11	12-13	14-15	16-17	18-19	20-21	22-23	24-25	26-27	28-29	30-31	32-33	34-36
18-20	2.0	3.9	6.2	8.5	10.5	12.5	14.3	16.0	17.5	18.9	20.2	21.3	22.3	23.1	23.8	24.3	24.9
21-25	2.5	4.9	7.3	9.5	11.6	13.6	15.4	17.0	18.6	20.0	21.2	22.3	23.3	24.2	24.9	25.4	25.8
26-30	3.5	6.0	8.4	10.6	12.7	14.6	16.4	18.1	19.6	21.0	22.3	23.4	24.4	25.2	25.9	26.5	26.9
31-35	4.5	7.1	9.4	11.7	13.7	15.7	17.5	19.2	20.7	22.1	23.4	24.5	25.5	26.3	27.0	27.5	28.0
36-40	5.6	8.1	10.5	12.7	14.8	16.8	18.6	20.2	21.8	23.2	24.4	25.6	26.5	27.4	28.1	28.6	29.0
41-45	6.7	9.2	11.5	13.8	15.9	17.8	19.6	21.3	22.8	24.7	25.5	26.6	27.6	28.4	29.1	29.7	30.1
46-50	7.7	10.2	12.6	14.8	16.9	18.9	20.7	22.4	23.9	25.3	26.6	27.7	28.7	29.5	30.2	30.7	31.2
51-55	8.8	11.3	13.7	15.9	18.0	20.0	21.8	23.4	25.0	26.4	27.6	28.7	29.7	30.6	31.2	31.8	32.2
56 & UP	9.9	12.4	14.7	17.0	19.1	21.0	22.8	24.5	26.0	27.4	28.7	29.8	30.8	31.6	32.3	32.9	33.3
	LEAN					IDEAL				AVERAGE				ABOVE AVERAGE			

www.accumeasurefitness.com/body-fat-measurement-charts-for-men-and-women.html

Figure 259: Body fat measurement by calliper

Body fat percentage naturally increases with age, so that the healthy range for older people is higher than for teenagers.

Alternative measurement methods are possible to determine body fat percentage. An approximate method for males is to make measurements of total body weight (kg) and waist measurement (cm).

Factor 1 is calculated as: (Total body weight x 1.082) + 94.42
Factor 2 is calculated as: Waist measurement x 4.15
Lean Body Mass is then equal to Factor 1 - Factor 2
Body Fat Weight is given by: Total bodyweight - Lean Body Mass

We than calculate Body Fat Percentage as: (Body Fat Weight x 100) / total bodyweight

Body surface area

Body surface area can be an important parameter for calculating drug dosage. Body surface area can be measured directly, but this is understandably a very laborious process. Livingston and Lee (2001) have shown that there is a good correlation between body surface area and body weight using the formula:

$$body\ surface\ area = 0.1173\ \times weight: kg^{0.6466}$$
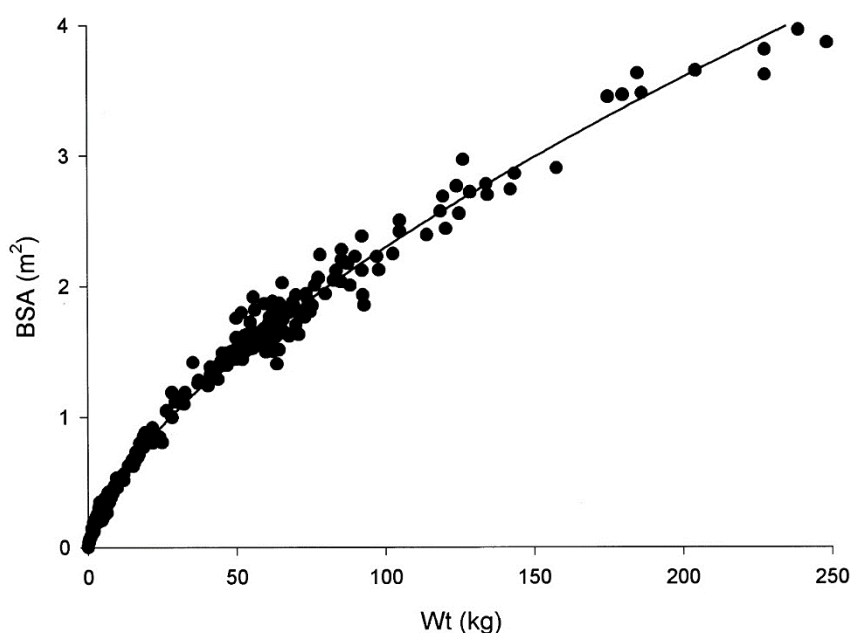
Figure 260: Relation between body surface area and weight

It is therefore possible to estimate body surface area directly from body weight, to a reasonable degree of accuracy.

An equation using the additional parameter of height in the calculation of body surface area was developed by Gehan and George (1970):

$$body\ surface\ area{:}\ m^2 = 0.0235\ height{:}\ cm^{0.42246} \times weight{:}\ kg^{0.51456}$$

The Gehan-George equation is based on a correlation with direct measurements of over 400 subjects, including both children and adults.

Blood volume

Another body parameter which is also of use in calculating drug dosage is blood volume. Blood volume can be estimated using Nadler's equation (Nadler, Hidalgo and Bloch, 1962). Two different versions are provided:

Males:

$$blood\ volume{:}\ ml = \ 0.0003668\ \times (height{:}\ cm)^3 + 32.2 \times weight{:}\ kg + 604$$

Females:

$$blood\ volume{:}\ ml = \ 0.000356\ \times (height{:}\ cm)^3 + 33 \times weight{:}\ kg + 183$$

An alternative method simply uses the average number of millilitres of blood per kilogram of body weight to calculate total blood volume. For males, this is 75 millilitres of blood per kilogram. It may be necessary to make adjustments using **Gilcher's Rule**. Not all types of tissue contain the same amount of blood. If a person is obese or very thin, this will influence the amount of blood per kilogram of body weight.

Muscular men have 75 ml of blood per kilogram of body weight

Normal men have 70 ml/kg

Thin men have 65 ml/kg

Obese men have 60 ml/kg

A similar method can be used to calculate blood volume in women.

In this section we have seen that an analysis of body measurements from many individuals has allowed health professionals to develop mathematical formulae for a range of body parameters which are of value in diagnosis and treatment procedures.

Three dimensional landscape modelling

A very large data set which is useful for many purposes is the Digital Elevation Model for Britain. This data set, produced by the Ordnance Survey and freely available for download from the Internet, gives the height of each point on the land surface to an accuracy of 0.1m. Points are arranged on a 50 metre grid covering the whole of Britain.

Data files consist of an array of height values for each 10km grid square. An extract is shown below.

```
       ----+----5----+----6----+----7----+----8----+----9----+----(
127  -0.9 -0.9 -0.9 -0.9 -0.9 -0.9 -0.9 -0.9 -0.9 -0.9 21 20.1 2
128  -0.9 -0.9 -0.9 -0.9 -0.9 -0.9 -0.9 -0.9 -0.9 26 23.4 21 20.
129  -0.9 -0.9 -0.9 -0.9 -0.9 -0.9 -0.9 -0.9 34.7 30.8 27.5 24 2
130  -0.9 -0.9 -0.9 -0.9 -0.9 -0.9 40.5 43 40.6 32.5 29.2 26.2 2
131  -0.9 -0.9 -0.9 -0.9 -0.9 -0.9 42.1 41.1 38.8 33.8 29.7 28.1
132  -0.9 -0.9 -0.9 -0.9 -0.9 -0.9 41.5 41.3 36.9 31.9 29.6 29 2
133  -0.9 -0.9 -0.9 -0.9 -0.9 32.3 40 39.8 38 31.2 29 28.9 28.6
134  -0.9 -0.9 -0.9 -0.9 -0.9 36.6 39.1 39.2 36.8 34.9 30.6 30.6
135  -0.9 -0.9 -0.9 -0.9 -0.9 -0.9 35 39.8 40.4 39.6 35.3 34.5 3
136  -0.9 -0.9 -0.9 -0.9 -0.9 -0.9 31.4 38.3 41.2 40.6 40.2 37.1
137  -0.9 -0.9 -0.9 -0.9 -0.9 -0.9 -0.9 1.2 6.9 41.5 41.8 41.1 3
138  -0.9 -0.9 -0.9 -0.9 -0.9 -0.9 -0.9 -0.9 -0.9 41.6 42 41.1 3
139  -0.9 -0.9 -0.9 -0.9 -0.9 -0.9 -0.9 -0.9 -0.9 42.8 42.8 41.1
140  -0.9 -0.9 -0.9 -0.9 -0.9 -0.9 -0.9 -0.9 41.2 41.5 41.9 41.1
141  -0.9 -0.9 -0.9 -0.9 -0.9 -0.9 -0.9 -0.9 38.2 40.5 40.7 39.1
```

Figure 261: Example of a Digital Elevation Model file

To demonstrate how the content of the file can be used, the row of values highlighted above has been plotted as a cross section in figure 262. This represents a section of cliff coastline.

Figure 262: Land surface cross section plotted from Digital Elevation Model data

One of the most useful ways of showing height data on a map is as a series of contour lines linking points of equal altitude. An isolated hill summit would be represented by a series of closed contours inside one another.
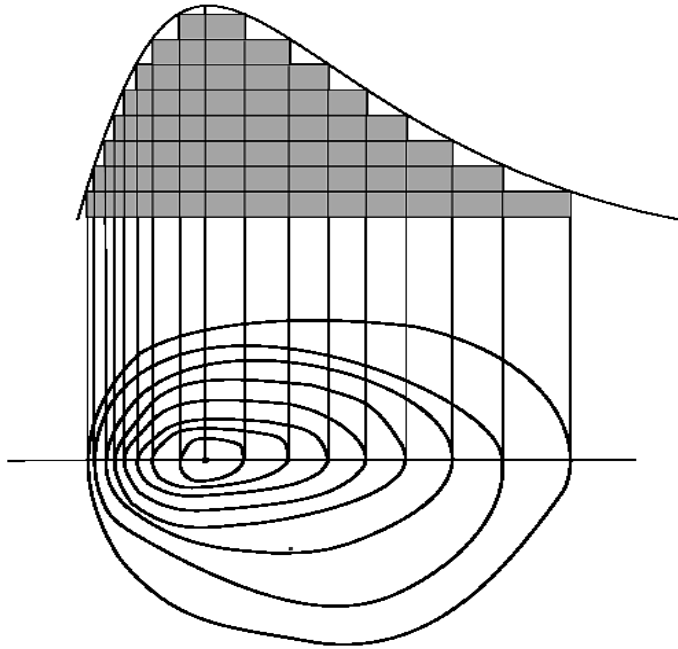
Figure 263:

Relation of land surface cross section to contour pattern for an isolated hill

Contour maps can be drawn using the height data in the Digital Elevation Model. We will look at a procedure to do this which can be implemented as a computer program.

The first step is to obtain the four heights for the corners of a 50m grid square. We then calculate the height of the middle point of the grid square as an average of the corner heights. The grid square has now been divided into four triangles A - D, and the heights of the triangle corners are known.
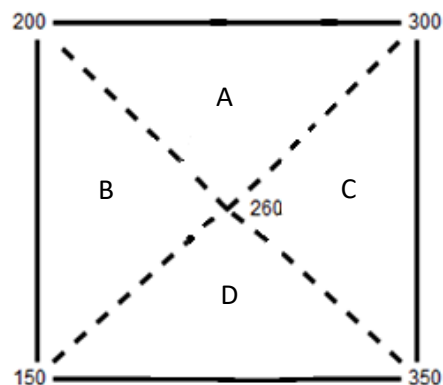
Figure 264:

Example grid cell of the Digital Elevation Model

We can select suitable contour values to be plotted. Let us suppose that contours are required at heights of 250m and 280m. We will take each contour height in turn, beginning with 250m.

For each of the four triangles of the grid square, we check whether the contour line should cut through the triangle. This will be true if the contour height is between the maximum and minimum corner heights of the triangle. The 250m contour will cross triangles A, B and D, but is too low to cross triangle C.

If the contour crosses a triangle edge, its cutting position is found on the linear gradient. For example, the 250m contour cuts the top edge of the grid square at a point mid-way between the corners at heights of 200m and 300m. Once the cutting points for a contour have been found, the contour line itself can be drawn.
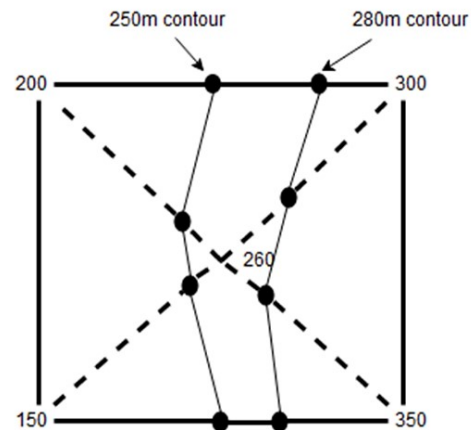
Figure 265:

Example of plotting contour lines across a grid cell of the Digital Elevation Model

If the procedure is repeated for each grid square of a selected area, a full contour map can be created. It is useful to add a map image below the contours, so that the accuracy of the computer program can be checked, and to help with locating points on the map area.
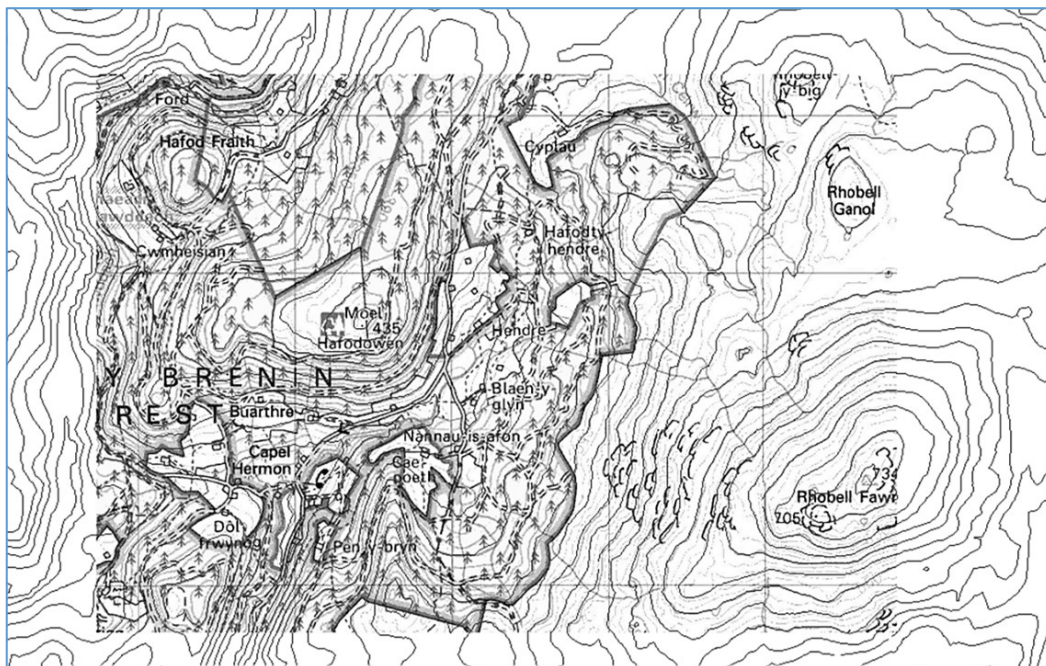
Figure 266: Contour map created from Digital Elevation Model data

For many geographical applications, it is useful to produce a three dimensional representation of the land surface, rather than a two dimensional map. The Digital Elevation Model again provides the necessary data.

The simplest three dimensional representation method uses isometric projection, as outlined previously in chapter 7. The basic approach for this technique is illustrated in figure 267 below.
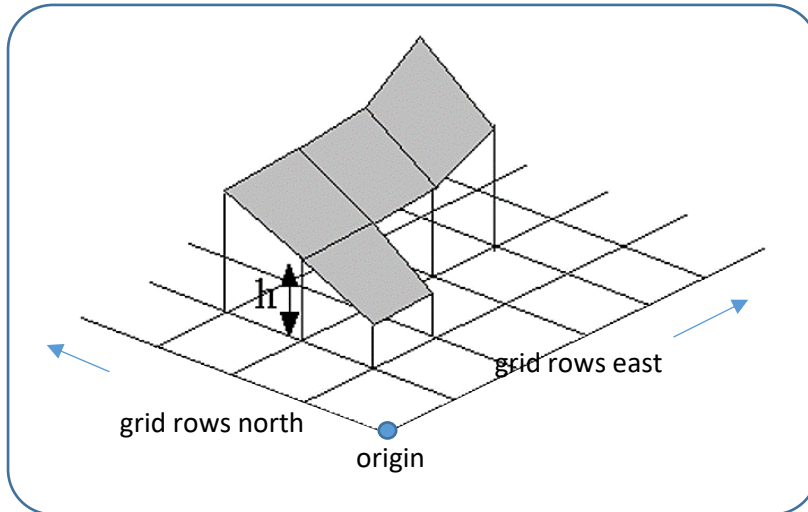
Figure 267:

Construction of a land surface model in isometric projection

The position on the screen of the origin point is chosen.

Each grid square of the digital elevation model is considered in turn. The position on screen of the grid square base plane is calculated by applying the appropriate offset for the numbers of grid rows north and east of the origin.

Vertical lines are constructed from the corners of the grid square to represent the heights of the four corners of the grid square. An opaque surface patch is then plotted to create an element of the ground surface.

The procedure is repeated for all grid squares of the area to be represented.

A correct land surface model will be created, but this will appear very plain and featureless. A better representation is created if the surface patches can be shaded according to the direction of the incident light, as in figure 268.
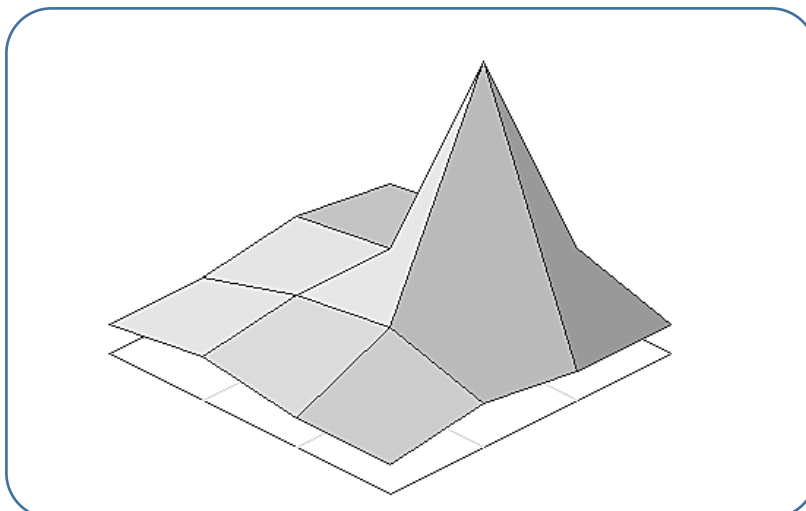
Figure 268:

Land surface model with light shading

To apply light shading, we find the angle ϴ between the **incident light** and a **surface normal** perpendicular to the surface patch. Where this angle is small, the light will be shining directly onto the surface and the surface patch should be shaded in a lighter colour.

surface normal

ϴ
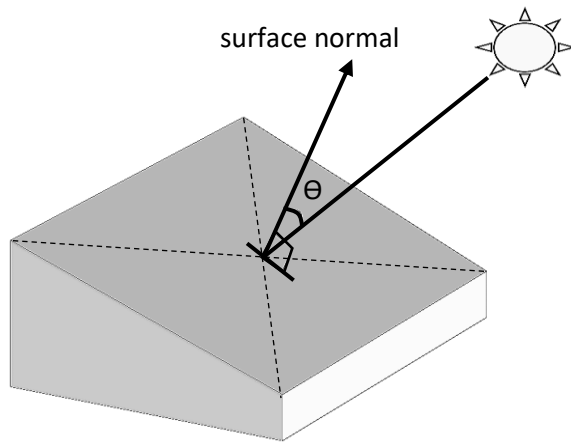
Figure 269:

Technique to determine the level of light shading for a surface patch

We begin by calculating the vector components of the diagonals of the surface patch. These will be the cell dimensions in the x and y horizontal plane, and the height in the z vertical direction.

a_z **a** a_y a_x b_z **b** b_y b_x
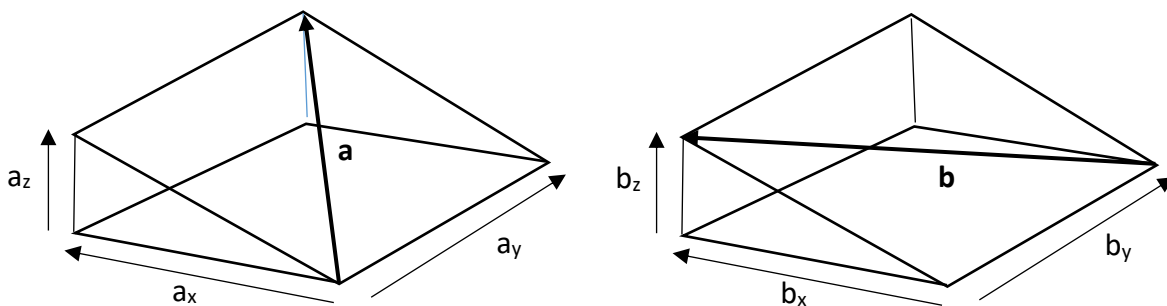
Figure 270: Vector components of the surface patch diagonals

The next step is to determine the vector components of the surface normal **c**, which will project from the surface at right angles to both of the diagonals.

a $(a_x \cdot a_y \cdot a_z)$

a x **b** = **c** $(c_x \cdot c_y \cdot c_z)$

b $(b_x \cdot b_y \cdot b_z)$
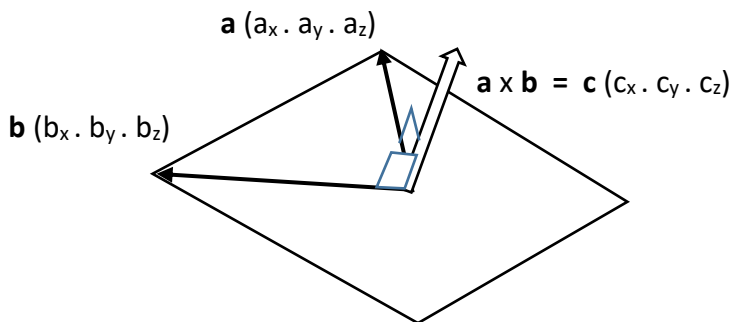
Figure 271:

Calculation of the surface normal vector

The surface normal vector is given by the **vector cross product** of the two diagonal vectors.

Components of the vector cross product are calculated from the components of the diagonal vectors by means of the formulae:

$$c_x = a_y b_z - a_z b_y$$
$$c_y = a_z b_x - a_x b_z$$
$$c_z = a_x b_y - a_y b_x$$

Once we know the direction of the surface normal, we can move on to determine the angle which this makes with the incident light. It is first necessary to decide on the direction of the light source. We will define this as the vector **d**.
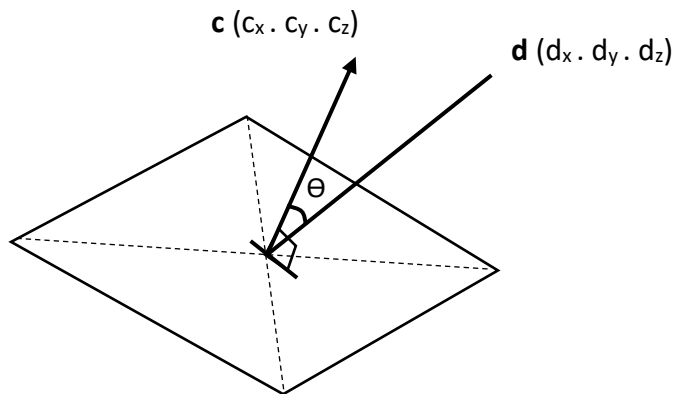
Figure 272:

Angle Ɵ between the surface normal vector **c** and the incident light vector **d**

The angle Ɵ can be found using the **vector dot product** formula, which gives the cosine of the angle between two vectors.

$$\boldsymbol{c} . \boldsymbol{d} = |\boldsymbol{c}|\,|\boldsymbol{d}| \cos \theta$$

We calculate the dot product **c.d** from the components of the two vectors:

$$\boldsymbol{c} . \boldsymbol{d} = c_x d_x + c_y d_y + c_z d_z$$

Rearranging the dot product formula shows that cos Ɵ can be obtained by dividing this quantity by the product of the vector magnitudes.

$$\cos \theta = \frac{c_x d_x + c_y d_y + c_z d_z}{|\boldsymbol{c}|\,|\boldsymbol{d}|}$$

The magnitude of each vector can be found using an expression of the form:

$$|\boldsymbol{c}| = \sqrt{c_x{}^2 + c_y{}^2 + c_z{}^2}$$

Now that we know the angle between the surface patch and the incident light, this can be used to set the light level for the cell. If the angle is small, the light will be falling directly onto the surface and a high light level can be used. If the incident angle is large, the surface patch should be shown in shadow.

After creating the surface geometry, we can improve the three dimensional model by adding either a section of map or air photograph to each surface patch, as shown below.
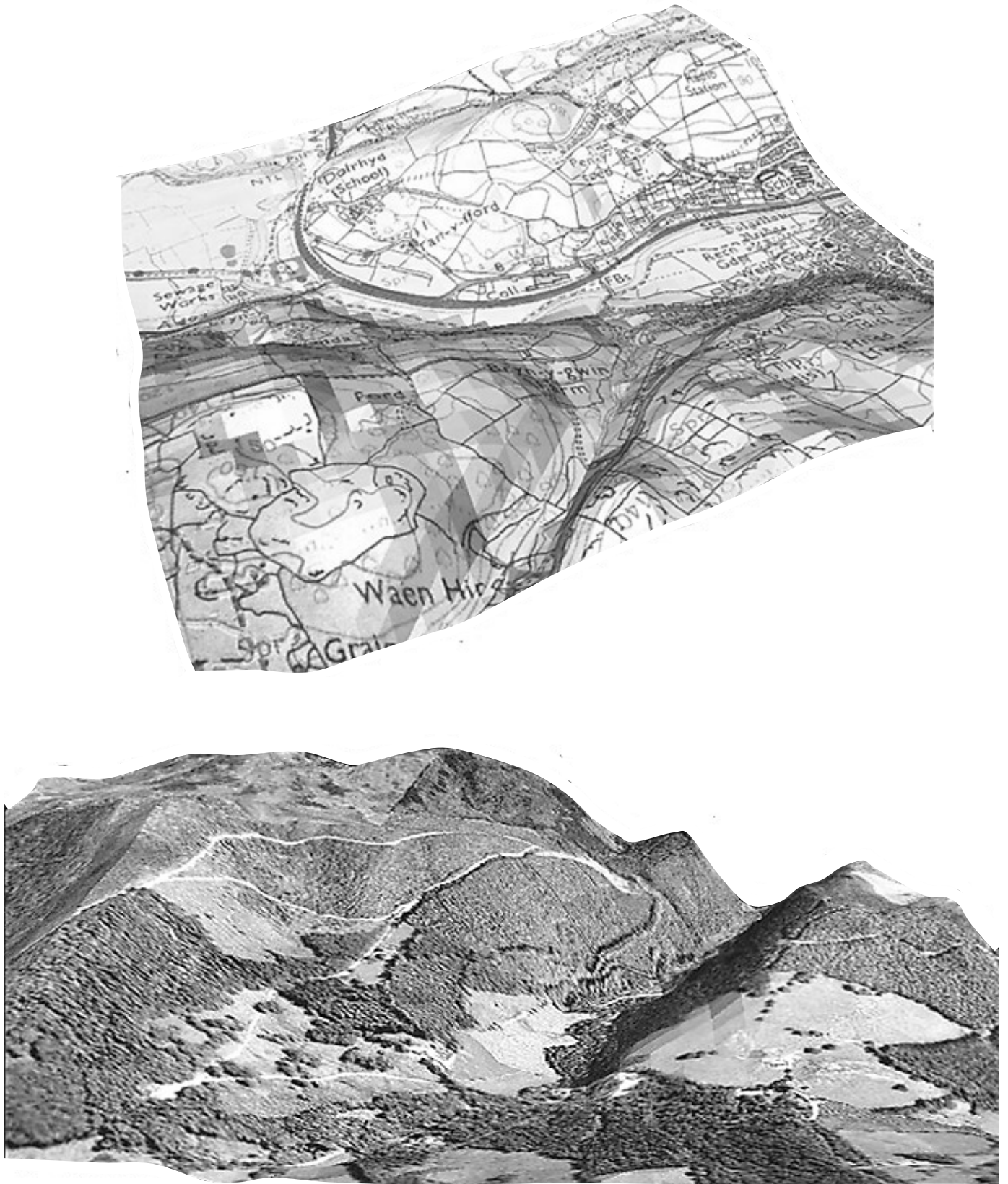
Figure 273: Isometric landscape models produced by computing students, using (above) map and (below) air photograph images

Three dimensional landscape models can be made even more realistic by using projective geometry to create perspective images. The principle of this technique is to imagine that the solid landscape model is being viewed through a window. Imaginary lines are constructed from points on the model to the viewer. Where these lines pass through the window, the corresponding feature of the image is drawn.
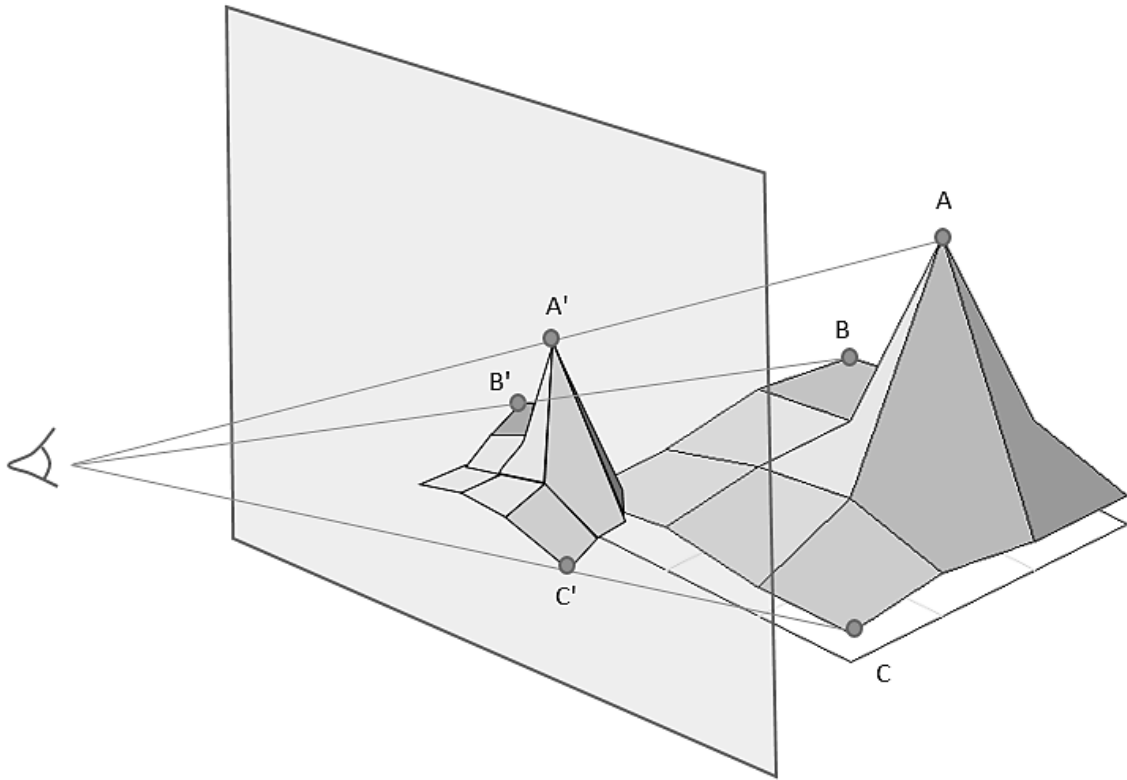
Figure 274: Construction of a perspective image using projective geometry

Projective geometry is used in landscape modelling software such as Google Earth, as in the example below.

Figure 275: Perspective image of the Mawddach Estuary and Rhinog mountains, created with Google Earth landscape modelling software

Powers of numbers

When data sets are collected, it is often important to identify a mathematical function which describes the data. This then allows estimates to be made for additional data points. As an example, consider the population data collected in censuses in England and Wales at ten year intervals over the period 1801 to 2011. This data is plotted as a scatter graph in figure 276 below.

Year	Residents
1801	8,872,980
1811	10,150,037
1821	11,978,875
1831	13,897,187
1841	15,906,741
1851	17,927,609
1861	20,066,224
1871	22,712,266
1881	25,974,439
1891	29,002,525
1901	32,527,843
1911	36,070,492
1921	37,886,699
1931	39,952,377
1941	
1951	43,757,888
1961	46,104,548
1971	48,749,575
1981	48,521,596
1991	50,748,000
2001	52,360,000
2011	56,075,900
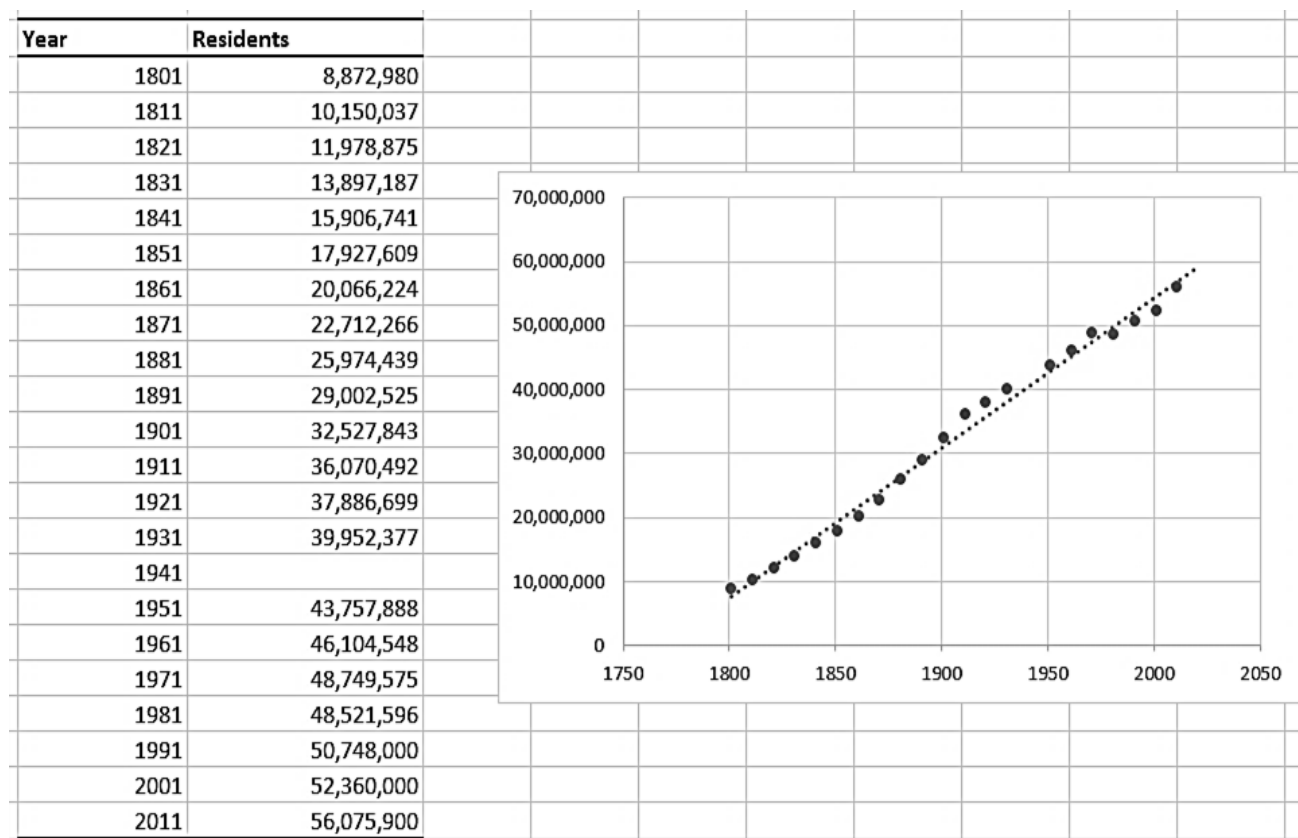
Figure 276: Population of England and Wales

The plotted points appear reasonably close to a straight line, so can be approximated by an equation of the form:

$$y = ax + b$$

where **a** is gradient of the line, and **b** is the intercept with the vertical axis when x = 0. The Excel spreadsheet has a function to calculate the equation of the best fit straight line through the points. This is displayed as:

$$y = 234,110.024255 \, x - 413,954,770$$

By substituting x =2021 into this equation, we can obtain a prediction for the population of England and Wales at the 2021 census as 59,181,590.

Data can easily be analysed if it has a linear trend, but often the data points follow some other pattern. We will look at examples:

In 1619, the astronomer Johannes Kepler published a formula relating the size of the orbit of a planet to the time taken to orbit around the sun. This formula is known as Kepler's third law, and we will look at how this can be derived.

Planetary orbits are generally not circular, but instead have the shape of an ellipse. The Sun is at one of the two foci of the ellipse. The sum of the distances from the two foci to any point on the ellipse is constant, such as the total of the distances **a** and **b** in figure 277.
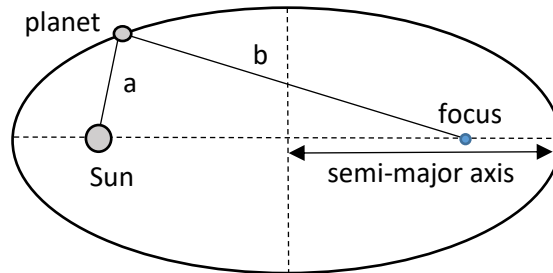
Figure 277: The elliptical orbit of a planet

The measurement used by Kepler for the size of the planetary orbits is the length of the semi-major axis of the ellipse.

planet	orbit semi-major axis, relative to Earth's orbit	orbital period, Earth years
Mercury	0.39	0.24
Venus	0.72	0.62
Earth	1.00	1.00
Mars	1.52	1.88
Jupiter	5.20	11.86
Saturn	9.54	29.46

This data has been plotted in figure 278. The data points appear to follow a curve, so may be related by an equation of the form:

$$\text{period} = \text{distance}^x$$

where x is a power that we need to determine.
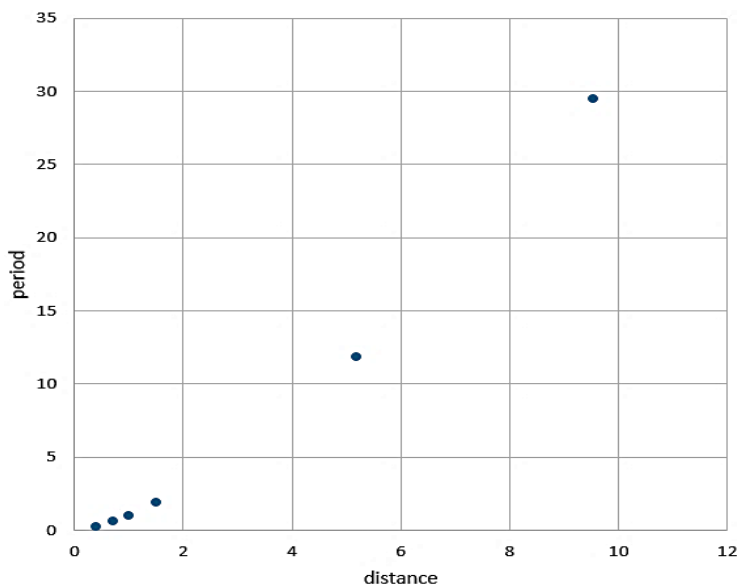
Figure 278:

Scatter plot of planet orbital distances and periods

To analyse functions involving powers of numbers, we will need to use **logarithms**. Logarithms are created for particular bases. The logarithm of a number is the **power** of the **base** which produces that number. For example, using base 2:

$$\log_2(32) = 5, \quad \text{because } 32 = 2^5$$

We can *multiply* numbers by *adding their logarithms*. Again using base 2:

$$\log_2(4) = 2 \quad \log_2(32) = 5$$

$$4*32 = 128 \quad \text{so adding logarithms: } 2 + 5 = 7 \quad 2^7 = 128$$

In general, if $C = A * B$,   then $\log(C) = \log(A) + \log(B)$

To raise a number to a *power*, we *multiply the logarithm* by the power. For example:

$$4^3 = 64$$

$$\log_2(4) = 2 \quad \log_2(4^3) = 3 \times 2 = 6 \quad 2^6 = 64$$

In general, if $C = A^B$,   then $\log(C) = B \log(A)$

These relations form the basis for the method to investigate functions involving powers of numbers. We begin by plotting a graph for the logarithms of the two variables **distance** and **period**. Logarithms to any base can be used, but it is often convenient to use natural logarithms to the base 2.718... (this number, known as *e*, will be considered in more detail when we discuss calculus in chapter 12). If the two variables are related by a power function, the points will lie on a straight line. This does happen, as shown in figure 279.
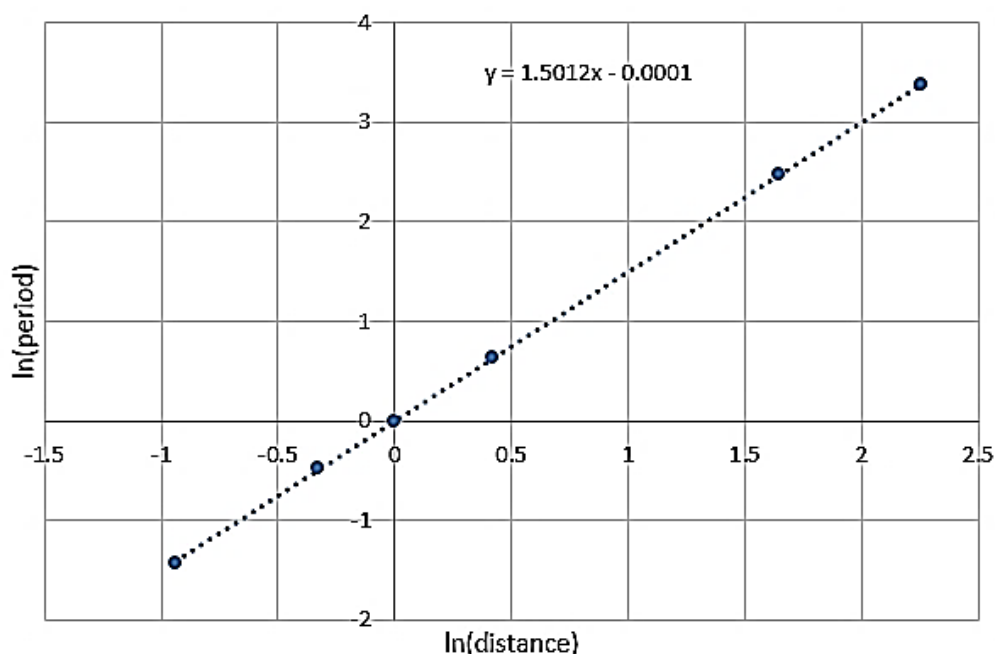
Figure 279: Logarithmic plot of planet orbital distances and periods

We find that the equation of the best fitted straight line, within experimental error, is:

$$y = 1.5012 \, x - 0.0001$$

We now need to interpret this result. If the equation that we are looking for has the form:

$$T = D^n$$

where T is the orbital time and D is the orbital distance, then

$$\ln(T) = n \ln(D)$$

Since our log plot shows *ln(D)* as the horizontal **x** coordinate, and *ln(T)* as the vertical **y** coordinate, then **n** will be the gradient of the line. The value for **n** was found to be 1.5 within reasonable experimental error. We can therefore write:

$$T = D^{1.5}$$

This can be expressed more conveniently by squaring both sides

$$T^2 = D^3$$

We have found Kepler's third law:

'The square of the orbital period for a planet is proportional to the cube of the semi-major axis of its orbit'.

As another example of the use of logarithms in analysing a function involving a power, we will look at an experiment in physics to determine the value of **g**, the acceleration due to gravity. This involves measuring the time taken for a metal ball to fall freely under gravity through various known distances.

An accurate measurement of times can be made electronically using apparatus of the type shown in figure 280.
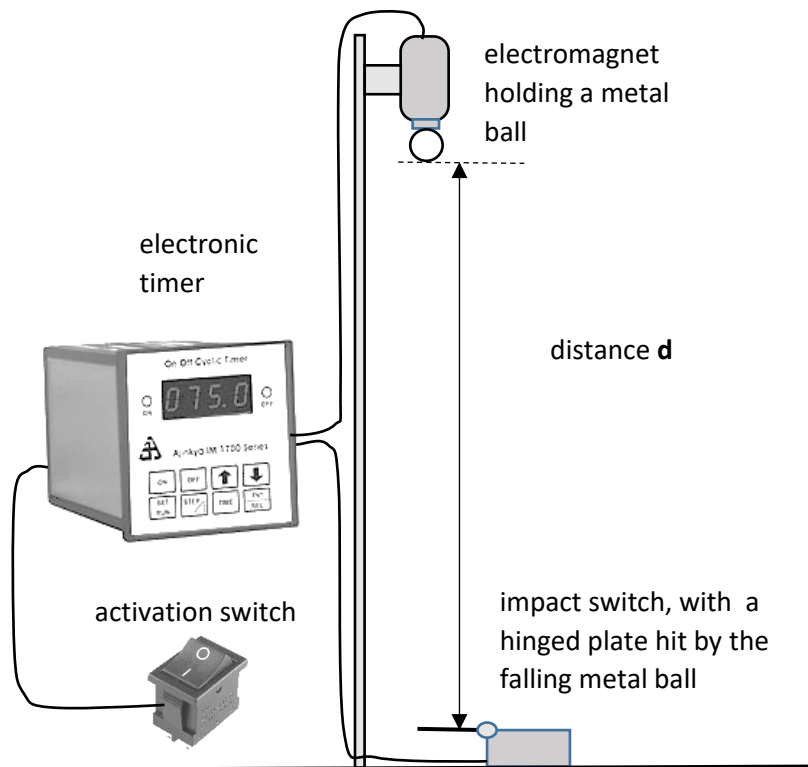
Figure 280: Experiment to determine acceleration due to gravity

An electromagnet holding a metal ball is positioned above a switch plate. When the system is activated, the electromagnet releases the ball and instantly starts an electronic timer. After falling a measured distance **d**, the ball impacts the switch plate and the timer is immediately stopped. The experiment can be repeated a number of times with different distances **d**.

We begin by plotting a scatter graph of the experimental results. Typical values are shown in figure 281.

d	t
0.5	0.319
0.6	0.350
0.8	0.404
1.0	0.452
1.2	0.495
1.4	0.535
1.6	0.571
1.8	0.606
2.0	0.639
2.2	0.670
2.4	0.700
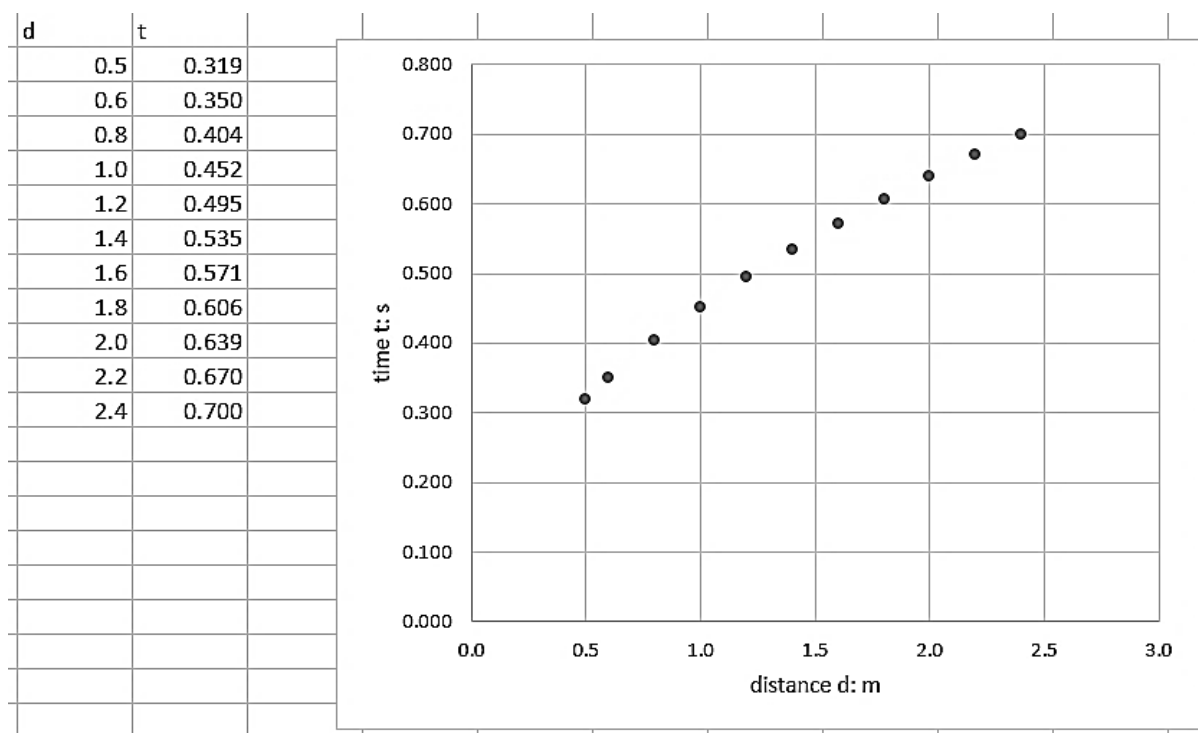
Figure 281: Scatter plot of distances and times for free fall under gravity

It is known that the equation for free fall under gravity is:

$$d = \frac{1}{2}gt^2$$

where **d** is distance in metres, **t** is time in seconds, and **g** is the acceleration due to gravity in ms^{-2}. Equations of motion will be discussed further in calculus chapter 12.
Rearranging to make **t** the subject of the equation:

$$t^2 = \frac{2d}{g}$$

$$t = \sqrt{\frac{2d}{g}} \quad = \quad \sqrt{\frac{2}{g}} \ . \ \sqrt{d}$$

Taking logarithms of both sides of the equation, remembering that we can multiply numbers by adding their logarithms, and that a square root is equivalent to a power of ½ :

$$\ln(t) = \frac{1}{2}\ln\left(\frac{2}{g}\right) + \frac{1}{2}\ln(d)$$

We can plot a graph for the logarithms of the two variables **distance** and **time**. The experimental points fall close to a straight line.

ln(d)	ln(t)
-0.69315	-1.14256
-0.51083	-1.04982
-0.22314	-0.90634
0	-0.79407
0.182322	-0.7032
0.336472	-0.62549
0.470004	-0.56037
0.587787	-0.50088
0.693147	-0.44785
0.788457	-0.40048
0.875469	-0.35667

y = 0.5004x - 0.7947
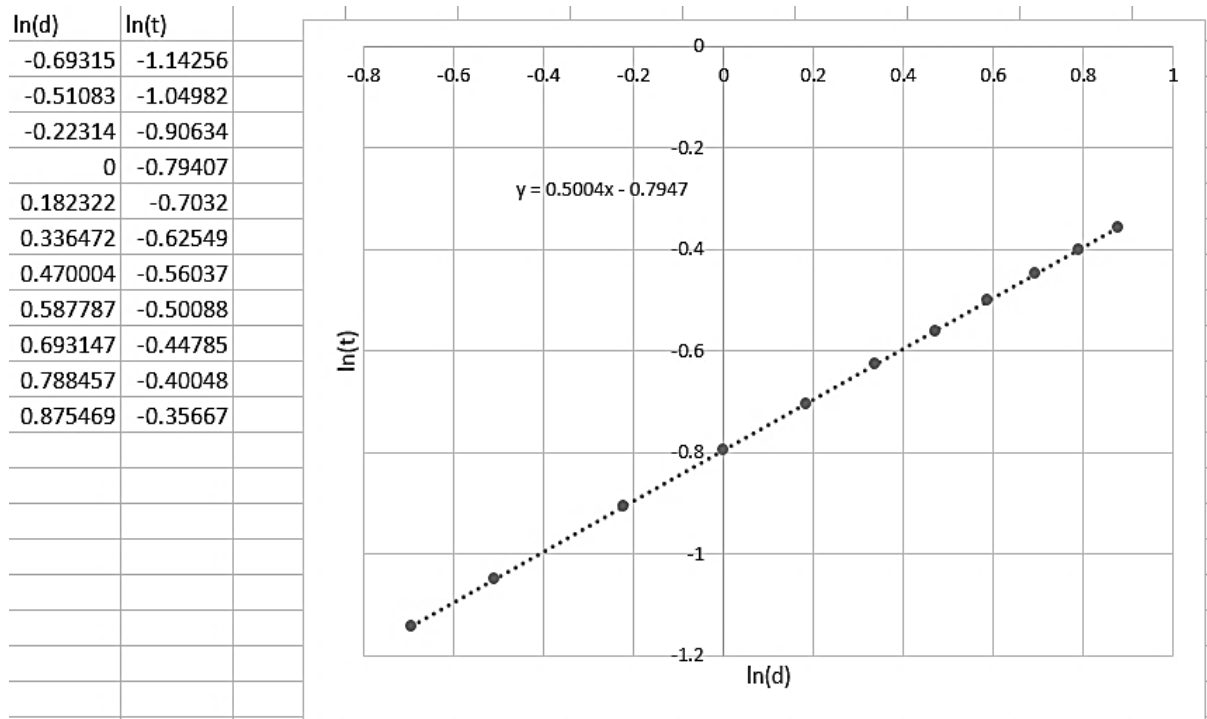
Figure 282: Logarithmic plot of distances and times for free fall

Within experimental error, the gradient of the line is 0.5. This is consistent with the term:

$$\frac{1}{2}\ln(d)$$

in the equation above. The intercept of the logarithmic graph represents the added term:

$$\frac{1}{2}ln\left(\frac{2}{g}\right) = -0.7947$$

so:

$$ln\left(\frac{2}{g}\right) = -1.5894$$

We can reverse the natural logarithm by applying an exponential function:

$$\frac{2}{g} = \exp(-1.5894)$$

giving:

$$g = 9.80\ ms^{-2}$$

Logarithms provide a useful way of analysing a wide range of data sets where we suspect that one property varies as a power of another property.   The first step is to plot a graph of natural logarithms of the two variables, to test whether a linear relationship exists. If so, then the techniques demonstrated above can generate a solution to the problem.

Kirchhoff's laws

The current or voltage in a simple Direct Current circuit can be found using Ohm's Law. For example:
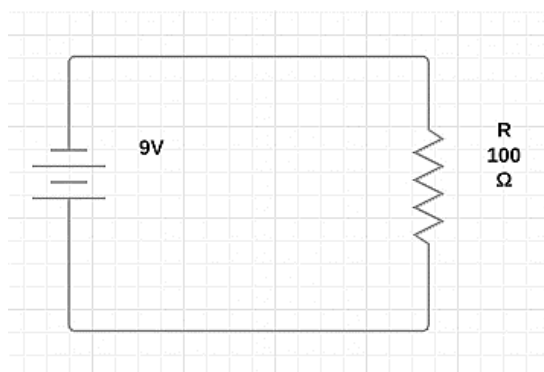
Figure 283:

Simple Direct Current and resistance circuit

The current flowing in the circuit is given by the formula:

$$I \ = \ \frac{V}{R} \ = \ \frac{9V}{100\Omega} \ = \ 0.09 \ A$$

However, more complex circuits with multiple power sources and resistances can be more complex to analyse. Here we have two power sources and three resistances arranged in two separate loops in the circuit. The objective is to calculate the currents flowing in each branch of the circuit, and the voltages at the points B and E:
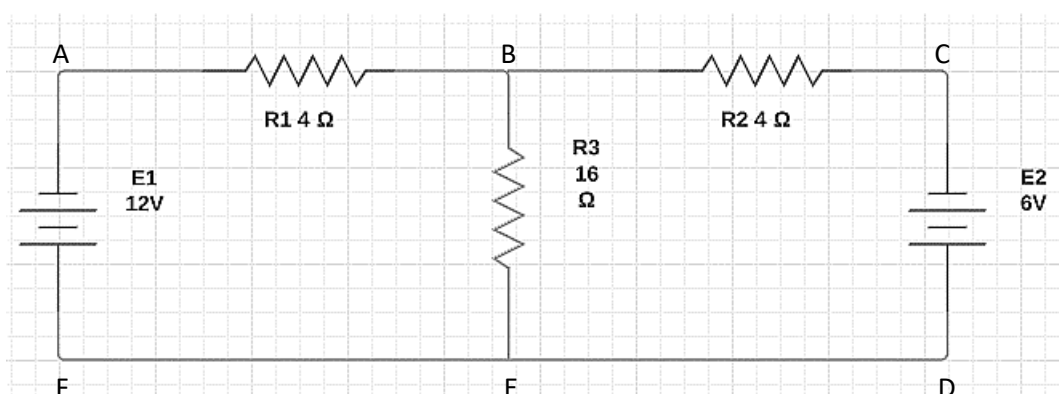
Figure 284: Direct Current circuit with multiple loops

Kirchhoff's Laws provide a means of analysing more complex electrical networks, and are an important application of simultaneous equations.

Kirchhoff's First law: for a given junction or node in a circuit, the sum of the currents entering equals the sum of the currents leaving.

Kirchhoff's Second law: around any closed loop in a circuit, the sum of the potential differences across all elements is zero.

Both of these laws are common sense conservation principles, which essentially say that electrical energy cannot be created from nowhere, and cannot disappear for no reason.

If we take currents flowing into a node as positive, and flowing out of the node as negative, the **current** law requires that the algebraic sum is zero. For example:
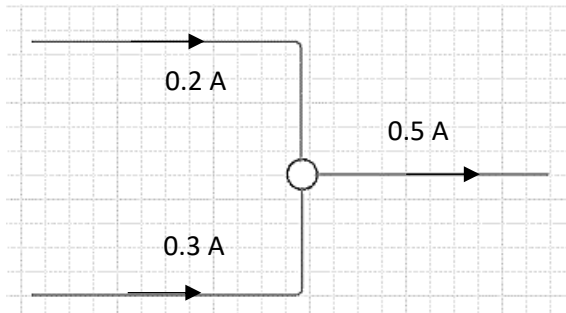
Current law:

0.2 + 0.3 - 0.5 = 0

Figure 285: Representation of Kirchhoff's current law

The **potential difference** law requires that the increase in voltage across power sources are exactly matched by the voltage drops across the resistances around a circuit. For example:
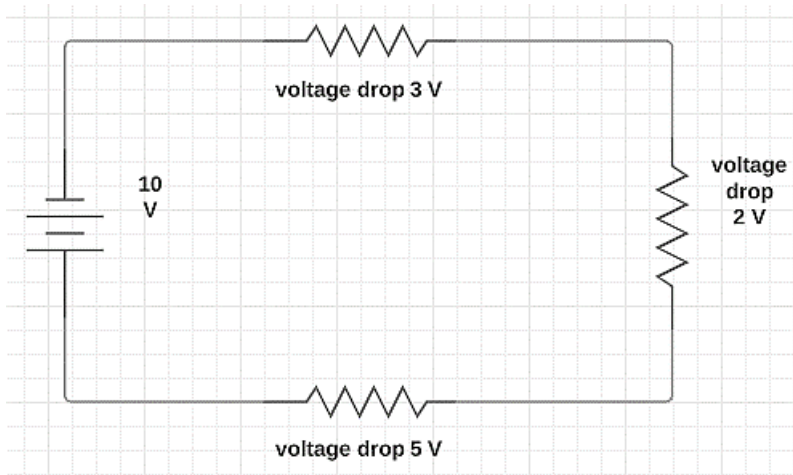
Potential difference law:

$10.0 - 3.0 - 2.0 - 5.0 = 0$

Figure 286: Representation of Kirchhoff's potential difference law

Using these principles, we can now begin an analysis of the multiple loop circuit. A first step is to assume the current flow directions in each of the loops. We will follow the convention of showing current flowing from the positive terminals of the power supplies and returning to the negative terminals:
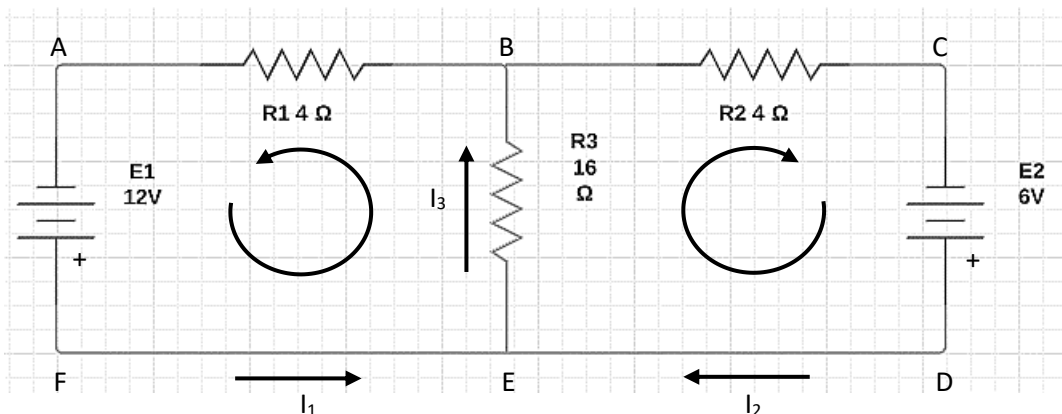
Figure 287: Assumed current flow directions in the loops of the circuit

It is sometimes the case that a strong power supply drives current in a reverse direction through part of the circuit. We will be able to identify this situation during the calculation, as a negative current value will appear in the results.

We will label the currents flowing in the two loops as I_1 and I_2 respectively. Our first task is to find the values of these currents.

The voltage drop across a resistance is given by Ohm's law as:
$$V = I\,R$$
The voltage drop across the 4 Ω resistance R1 will therefore be:
$$I_1 \times 4\ \Omega$$

Both currents flow into node E, so the current leaving the node in the direction of node B will be given by the current law as:
$$I_3 = I_1 + I_2$$
The voltage drop across the 16 Ω resistance R3 will be:
$$(I_1 + I_2) \times 16\ \Omega$$

According to the potential difference law, the algebraic total of changes in potential around the left hand loop must be zero. Therefore:
$$12\,V - (I_1 \times 4\ \Omega) - ((I_1 + I_2) \times 16\ \Omega\,) = 0$$
or:
$$(I_1 \times (4+16)) + (\,I_2 \times 16) = 12$$

Similarly, in the right hand loop:
$$6\,V - (I_2 \times 4\ \Omega) - ((I_1 + I_2) \times 16\ \Omega\,) = 0$$
so:
$$(\,I_1 \times 16) + (I_2 \times (4+16)) = 6$$

We have now produced a set of simultaneous equations:
$$20\,I_1 + 16\,I_2 = 12 \qquad \underline{\hspace{1cm}}(1)$$
$$16\,I_1 + 20\,I_2 = 6 \qquad \underline{\hspace{1cm}}(2)$$

Multiplying equation (1) by 4: $\quad 80\,I_1 + 64\,I_2 = 48$
Multiplying equation (2) by 5: $\quad 80\,I_1 + 100\,I_2 = 30$
Subtracting: $\qquad\qquad\qquad\qquad -36\,I_2 = 18$
$$I_2 = -0.5\ A$$
The minus value indicates that the current direction in the right hand loop is actually in the opposite direction to our original assumption.

Substituting $I_2 = -0.5$ in equation (1) gives:
$$20\,I_1 - 8 = 12$$
$$20\,I_1 = 20$$
$$I_1 = 1.0\ A$$

We can now complete the analysis of the circuit by finding the potential difference at the node B. For the power supply E1, we will assume that the potential falls from 12 V at the positive terminal to 0 V at the negative terminal. This is the result of falls in potential across the two resistances R3 and R1 in the left hand loop of the circuit:
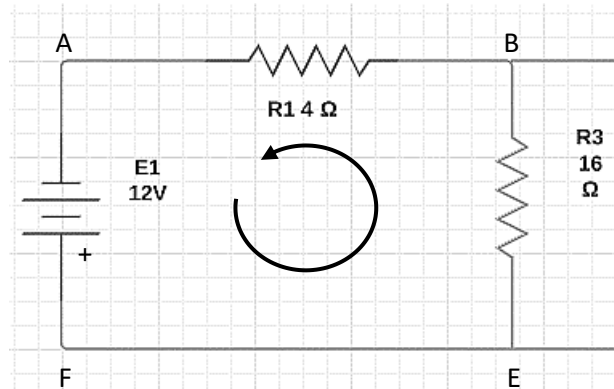
Considering resistance R1, the potential drop is given by Ohm's law:

$$V = I\,R$$

so:
$$V = 1.0\,A \times 4\,\Omega = 4\,V$$

Since node A is at a potential of 0 V, then node B must be at a potential of 4 V.

Calculations using Kirchhoff's laws lead to systems of simultaneous differential equations which must be solved to determine the currents flowing through the separate loops of the circuit. Small systems of simultaneous equations can be solved manually, but it can become more difficult and time consuming to handle the large equations produced for circuits involving many loops. We may then choose to solve the problem with the aid of a computer application, using matrices. This method will be demonstrated with an example:
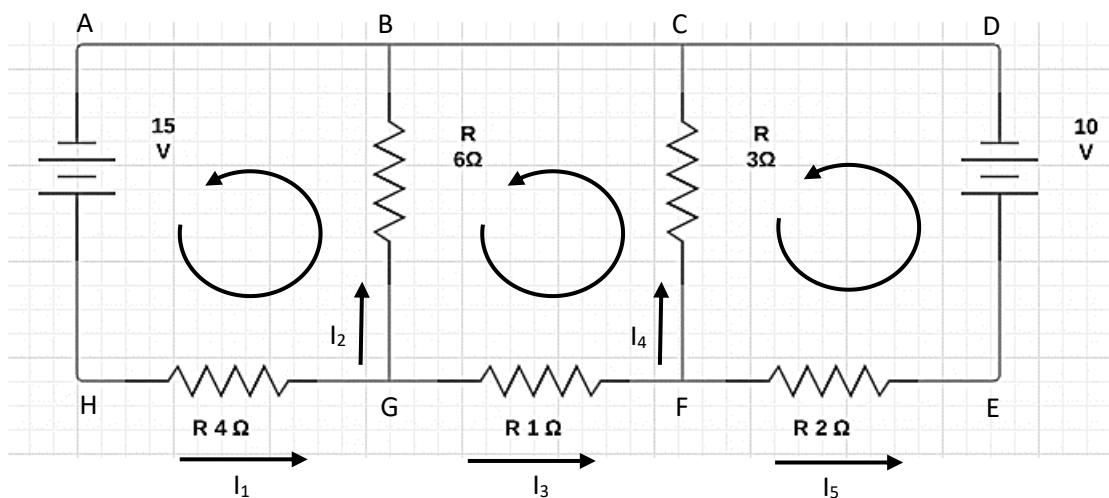
Figure 288: Circuit with assumed current flow directions in the loops

The circuit to be analysed is shown in figure 288. Our objective is to determine the currents $I_1 - I_5$ flowing in the different loops of the circuit. Initial assumptions have been made for the current flow directions, but a negative current value will simply appear in the results if any of these assumptions is incorrect.

We begin by setting up the series of equations. Applying Kirchhoff's current law for the nodes F and G:

$$I_1 - I_2 - I_3 = 0 \qquad \text{so} \qquad I_2 = I_1 - I_3$$

$$I_3 - I_4 - I_5 = 0 \qquad \text{so} \qquad I_4 = I_3 - I_5$$

We can produce an equation for each loop, applying Kirchhoff's potential difference law and calculating voltages by means of Ohm's law:

Left loop

from Ohm's law:	$15 - 4\,I_1 - 6\,I_2 = 0$
substituting for I_2 :	$15 - 4\,I_1 - 6\,(I_1 - I_3) = 0$
so:	$10\,I_1 - 6\,I_3 = 15$

Middle loop:

from Ohm's law:	$6\,I_2 - I_3 - 3\,I_4 = 0$
substituting for I_2 and I_4 :	$6\,(I_1 - I_3) - I_3 - 3\,(I_3 - I_5) = 0$
so:	$6\,I_1 - 10\,I_3 + 3\,I_5 = 0$

Right loop:

from Ohm's law:	$10 + 3\,I_4 - 2\,I_5 = 0$
substituting for I_4 :	$10 + 3\,(I_3 - I_5) - 2\,I_5 = 0$
so:	$-3\,I_3 + 5\,I_5 = 10$

This gives a system of simultaneous equations which can be solved to obtain the current values:

$$10\,I_1 - 6\,I_3 + 0\,I_5 = 15$$
$$6\,I_1 - 10\,I_3 + 3\,I_5 = 0$$
$$0\,I_1 - 3\,I_3 + 5\,I_5 = 10$$

This can be written in an equivalent matrix form as:

$$\begin{bmatrix} 10 & -6 & 0 \\ 6 & -10 & 3 \\ 0 & -3 & 5 \end{bmatrix} \begin{bmatrix} I_1 \\ I_3 \\ I_5 \end{bmatrix} = \begin{bmatrix} 15 \\ 0 \\ 10 \end{bmatrix}$$

When multiplying matrices, each element of the column matrix is applied in turn to each row of the square matrix

$$\begin{bmatrix} 10 & -6 & 0 \\ 6 & -10 & 3 \\ 0 & -3 & 5 \end{bmatrix} \begin{bmatrix} I_1 \\ I_3 \\ I_5 \end{bmatrix} = 10\,I_1 - 6\,I_3 + 0\,I_5$$

$$\begin{bmatrix} 10 & -6 & 0 \\ 6 & -10 & 3 \\ 0 & -3 & 5 \end{bmatrix} \quad \begin{bmatrix} I_1 \\ I_3 \\ I_5 \end{bmatrix} = 6\,I_1 - 10\,I_3 + 3\,I_5$$

If we label the matrices:

$$\begin{bmatrix} 10 & -6 & 0 \\ 6 & -10 & 3 \\ 0 & -3 & 5 \end{bmatrix} = A; \qquad \begin{bmatrix} I_1 \\ I_3 \\ I_5 \end{bmatrix} = X; \qquad \begin{bmatrix} 15 \\ 0 \\ 10 \end{bmatrix} = B$$

then the matrix equation can be written as a multiplication:

$$A \cdot X = B$$

It is known from matrix algebra that:

$$X = A^{-1} \cdot B$$

where A^{-1} is the inverse of matrix A. When a matrix is multiplied by its inverse, the result is the identity matrix.

$$A \cdot A^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The equation $\{\, X = A^{-1} \cdot B \,\}$ provides a method for solving the system of simultaneous equations and finding the values for the currents I_1, I_3 and I_5 :

- obtain the inverse matrix A^{-1}
- multiply this by the column matrix B

Fortunately there are computer applications readily available for finding the inverses of matrices.

Result

You entered matrix

$$A = \begin{bmatrix} 10 & -6 & 0 \\ 6 & -10 & 3 \\ 0 & -3 & 5 \end{bmatrix}$$

The inverse of matrix A is:

$$A^{-1} = \begin{bmatrix} \dfrac{41}{230} & -\dfrac{3}{23} & \dfrac{9}{115} \\[2mm] \dfrac{3}{23} & -\dfrac{5}{23} & \dfrac{3}{23} \\[2mm] \dfrac{9}{115} & -\dfrac{3}{23} & \dfrac{32}{115} \end{bmatrix}$$

---

www.mathportal.org/calculators/matrices-calculators/matrix-calculator.php

Figure 289: Solution for the inverse matrix using an application available on an Internet page

The inverse matrix A^{-1} can be multiplied by matrix B to obtain the result matrix X:

$$\frac{1}{230}\begin{bmatrix} 41 & -30 & 18 \\ 30 & -50 & 30 \\ 18 & -30 & 64 \end{bmatrix}\begin{bmatrix} 15 \\ 0 \\ -10 \end{bmatrix} = \frac{1}{230}\begin{bmatrix} (41 \times 15) + (18 \times -10) \\ (30 \times 15) + (30 \times -10) \\ (18 \times 15) + (64 \times -10) \end{bmatrix}$$

$$= \frac{1}{230}\begin{bmatrix} 435 \\ 150 \\ -370 \end{bmatrix} = \begin{bmatrix} 1.89 \\ 0.65 \\ -1.61 \end{bmatrix}$$

We now have a solution for the currents:

$$I_1 = 1.89 \text{ A} \qquad I_3 = 0.65 \text{ A} \qquad I_5 = -1.61 \text{ A}$$

Substitution of these values in the original Kirchhoff current equations above gives the two remaining values:

$$I_2 = 1.24 \text{ A} \qquad I_4 = 2.26 \text{ A}$$

The method can be readily extended to handle larger circuits with a greater number of loops and power sources.

Boolean algebra

The data patterns which we have examined so far in this chapter have involved numerical values, for example: heights on a land surface, or orbital times for planets. Another set of well-defined patterns relate to the use of logic.

In 1854, the mathematician George Boole published a system of rules for Boolean algebra. This work has important applications in a number of fields where logic expressions are analysed, particularly for computing and electronics.

Boolean algebra is based on defining logic expressions. For example, we might make two statements:

    A: I want a new car
    B: I have enough money to buy a new car

I will only actually buy a new car if both of these expressions are true. This situation can be represented in a Venn diagram.
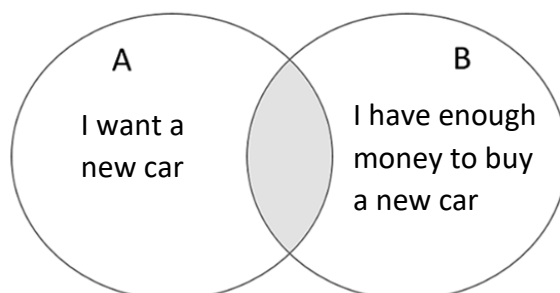

Figure 290: Venn diagram illustrating (A AND B)

The shaded area in figure 290 belongs to both sets A and B, so the two conditions are both met. In Boolean algebra, this is termed (A and B) and is written with a 'dot' symbol:

A . B

Consider now the two statements:

A: My car failed its MOT test and cannot be repaired
B: My car is not big enough to carry the things I need for my job

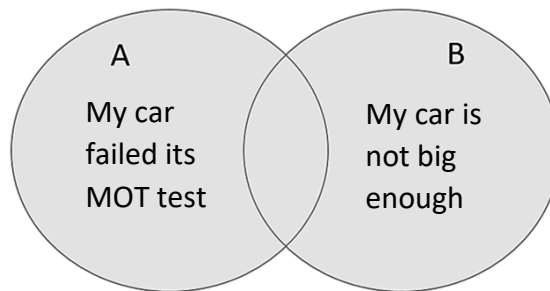Under either of these circumstances, I would need to buy a new car. The extended shaded area of the Venn diagram below now represents purchasing a car:

Figure 291: Venn diagram illustrating (A OR B)

The shaded area represents the total of sets A and B, so either or both of the conditions can be met. In Boolean algebra, this is termed (A or B) and is written with a 'plus' symbol:

A + B

An objective of Boolean algebra is to simplify logic expressions where possible. As an example, we might define the statement:

A: Thursday

This could be represented by a Venn diagram, where Thursday lies within the circle, and all other days of the week lie outside the circle.
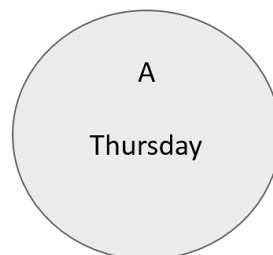
Figure 292: Venn diagram illustrating (A OR A) , (A AND A)

Suppose that we write an AND logic expression as:

A . A

This would imply that it is BOTH Thursday AND Thursday. One of the terms is clearly redundant, and we could simply say 'It is Thursday'.

The expression can be simplified:

$$A \cdot A = A$$

Suppose now that we write an OR logic expression:

$$A + A$$

This would imply that EITHER it is Thursday OR it is Thursday. One of the terms is again redundant and we could simply say 'It is Thursday'. The expression can again be simplified:

$$A + A = A$$

Further laws of Boolean algebra can be derived by examining Venn diagrams. We will define three statements:

P: Mountain climbing holidays

Q: Sailing holidays

T: Holidays in Wales

Suppose that we wish to have a holiday in Wales which involves at least one of the activities of sailing or mountain climbing. This would be represented by the shaded area of the Venn diagram, which includes the overlap of Welsh holidays with either of the other circles.
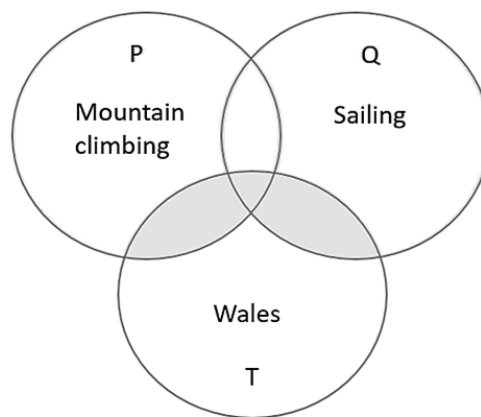

Figure 293: Venn diagram illustrating the Boolean distribution law

If we were to construct a Boolean algebra expression for the shaded area, this could be done in two ways:

Method 1: Combine P and Q to represent all mountain climbing and sailing holidays using a logical OR operation. Specify the overlap of this area with the circle T using an AND operation. This will pick out only the activity holidays which are in Wales. This sequence can be written:

$$(P + Q) \cdot T$$

Method 2: Use a logical AND operation to find the mountain climbing holidays in Wales as the overlap of P and T. Use a similar AND operation to find the sailing holidays in Wales as the overlap of Q and T. Finally, add these two regions using an OR operation to give all activity holidays in Wales.

This sequence can be written:

$$(P \cdot T) + (Q \cdot T)$$

Both methods have led to the same result of the shaded area in the Venn diagram. The logic expressions we deduced are therefore equivalent:

$$(P + Q) \cdot T \; = \; (P \cdot T) + (Q \cdot T)$$

This result is known as the **distribution law** of Boolean algebra.

Two other useful results in Boolean algebra are **de Morgan's laws**. To demonstrate these, suppose that an employment agency is recruiting staff. A vacancy exists for a French language tutor. We can define two statements:

> P: the applicant is a qualified teacher

> Q: the applicant is a fluent French speaker

In order to be considered for the post, both of these conditions must be met. We could therefore rule out candidates falling in the shaded area of the Venn diagram.
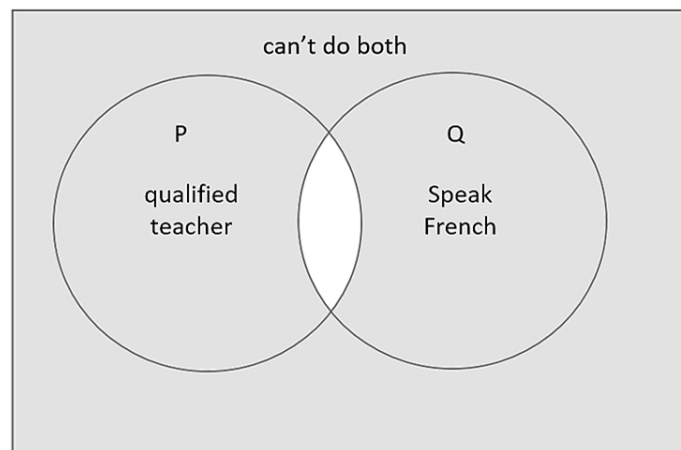

Figure 294: Venn diagram illustrating de Morgan's conjunction law

We could again find this shaded area in two different ways:

Method 1: Use a logical AND operation to find the overlap of P and Q. This would represent the ***suitable*** candidates who are both qualified teachers and speak French. We then take all the remaining area to represent the ***unsuitable*** candidates. This could be written as:

$$\overline{(P \cdot Q)}$$

The line above the expression indicates the logical NOT operation. In this case, it is any part of the Venn diagram which is not in the area (P . Q)

Method 2: We could use a NOT operation on P to find all candidates who are not qualified teachers. We could similarly use a NOT operation on Q to find all candidates who do not speak French. A candidate in either of these groups would be ***unsuitable***.

This could be written as:

$$\overline{P} + \overline{Q}$$

which means 'either (not in P) or (not in Q) '. This will include candidates who are not qualified teachers and also do not speak French.

Both methods have led to the same result of the shaded area in the Venn diagram. The logic expressions are therefore equivalent:

$$\overline{(P \cdot Q)} \; = \; \overline{P} + \overline{Q}$$

This result is known as **de Morgan's conjunction law**.

Suppose now that the employment agency has a vacancy for a forestry worker. We can define the statements:

P: the applicant is qualified to drive a tractor

Q: the applicant is qualified to operate a chain saw

Candidates can be considered for the post if they have one or both of these qualifications. We would therefore rule out candidates with neither qualification, who fall in the shaded area of the Venn diagram.
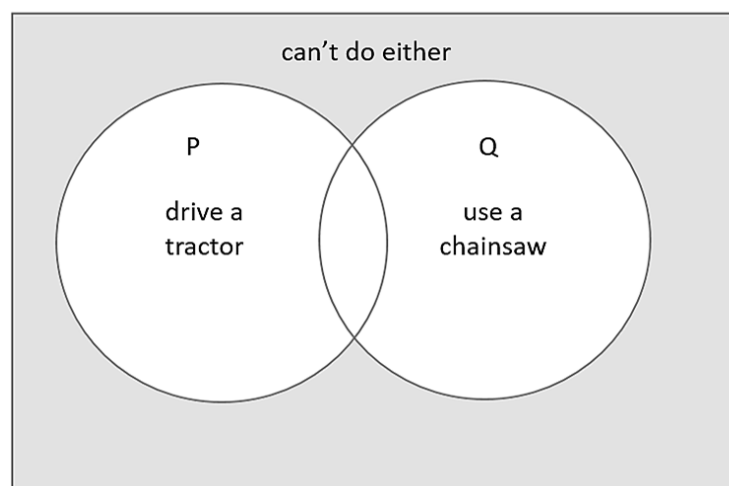

Figure 295: Venn diagram illustrating de Morgan's disjunction law

As before, we could find this shaded area in two different ways:

Method 1: Use a logical OR operation to find the overlap of P and Q. This would represent the *suitable* candidates who have qualification in at least one of the skills: tractor driving and chain saw operation. We then take all the remaining area to represent the *unsuitable* candidates. This could be written as:

$$\overline{(P + Q)}$$

The line above the expression indicates any part of the Venn diagram which is not in the area (P + Q)

Method 2: We could use a NOT operation on P to find all candidates who are not qualified tractor drivers. We could similarly use a NOT operation on Q to find all

candidates who are not qualified to use a chain saw. A candidate would have to be in **both** of these groups to be **unsuitable**. The unsuitability condition is written as:

$$\overline{P} . \overline{Q}$$

The two methods have led to the same result of the shaded area in the Venn diagram. The logic expressions are therefore equivalent:

$$\overline{(P + Q)} \ = \ \overline{P} . \overline{Q}$$

This result is known as **de Morgan's disjunction law**.

Boolean algebra is widely used during the design of digital electronic circuits. Two components commonly used in electronic devices are AND gates and OR gates.

An AND gate, represented by the symbol:
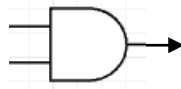
takes two input signals which may be at logic state 0 or 1. The output of the gate is at logic state 0, except in the case when both inputs are at logic 1.

An OR gate, represented by the symbol:
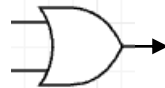
takes two input signals which again may be at logic state 0 or 1. The output of the gate is at logic state 1 if either or both of the inputs are at logic 1. The output is 0 only if both input signals are at logic state 0.

To illustrate how the laws of Boolean algebra can be used to simplify logic expressions, consider the electronic circuit represented by the Boolean expression:

$$(A . B) + ((B . C) . (B + C))$$

A circuit diagram can be constructed using AND and OR logic gates:
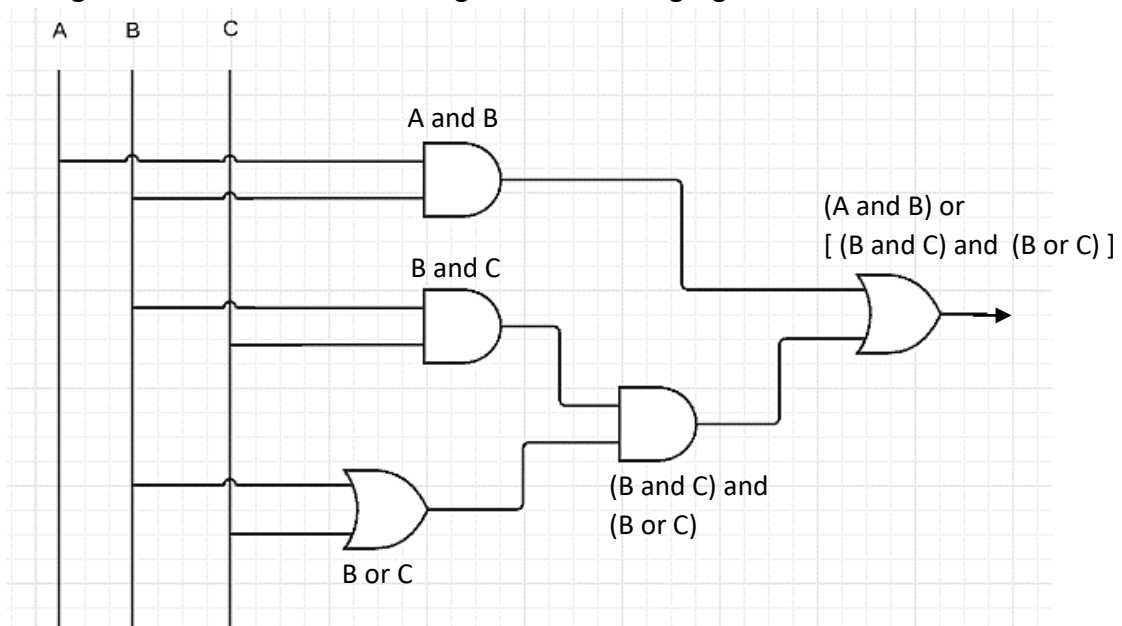

Figure 296: Digital circuit composed of AND and OR logic gates

Inputs A, B and C might represent signals from sensors in a machine, and an output of logic 1 represents an error condition when a warning light or buzzer should operate.

We can calculate the outputs from the circuit for different combinations of logic states for the inputs A, B and C:

A	B	C	A.B	B.C	B + C	(B . C) . (B + C)	(A . B) + ((B . C) . (B + C))
0	0	0	0	0	0	0	0
0	0	1	0	0	1	0	0
0	1	0	0	0	1	0	0
0	1	1	0	1	1	1	1
1	0	0	0	0	0	0	0
1	0	1	0	0	1	0	0
1	1	0	1	0	1	0	1
1	1	1	1	1	1	1	1

If a circuit can be constructed which has fewer logic gates but still produces the same final outputs, this has a number of benefits:

- Manufacturing costs are reduced
- Less space is required, which might be important in small electronic devices
- Less power would be needed
- Less heat would be generated, which reduces problems of heat dissipation

We will attempt to simplify the logic expression

$$(A . B) + ((B . C) . (B + C))$$

using Boolean algebra. We begin by using the **distribution law**. We showed earlier that

$$(P + Q) . T = (P . T) + (Q . T)$$

Equating P = B, Q = C, and T = (B.C) in the logic expression for the circuit gives:

$$(A . B) + (B.B.C) + (B.C.C)$$

We know that (B . B) can be simplified to just B, and (C . C) can be simplified to C, giving:

$$(A . B) + (B . C) + (B . C)$$

One of the (B . C) terms is redundant, so can be omitted:

$$(A . B) + (B . C)$$

We can finally use the distribution law to produce:

$$B . (A + C)$$

This simplified expression is used to produce the revised circuit design shown in figure 297. The number of logic gates has been reduced from five to two.
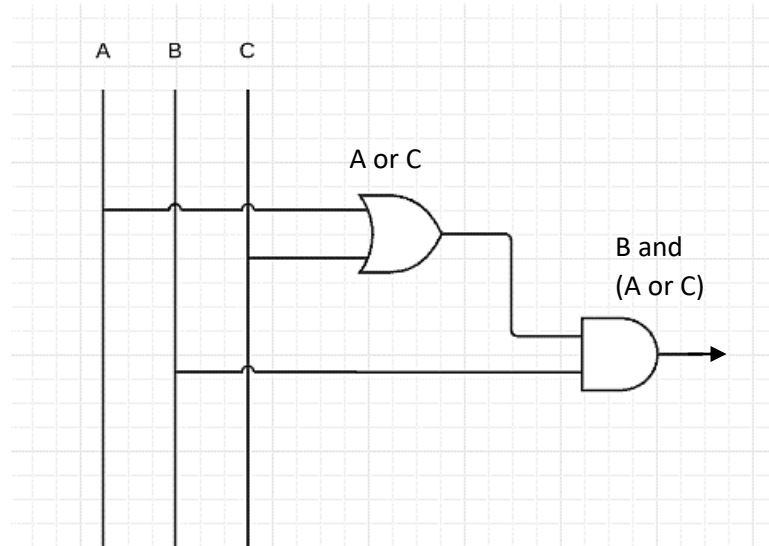
Figure 297: Simplified digital circuit

We can verify that the outputs from the simplified circuit are the same as in the original design. The simplified circuit does indeed produce the required pattern of output.

A	B	C	A + C	B . (A + C)
0	0	0	0	0
0	0	1	1	0
0	1	0	0	0
0	1	1	1	1
1	0	0	1	0
1	0	1	1	0
1	1	0	1	1
1	1	1	1	1

Boolean algebra is also important in the design of computer programs, particularly in algorithms for expert systems or artificial intelligence applications. Examples are:

- Medical diagnosis, combining a range of factors such as diseases which the patient has or has not had, and symptoms the patient does or does not exhibit.
- Careers guidance, based on the subjects which a student did or did not enjoy at school, and skills which they do or do not possess.
- Holiday choice, based on the countries the client does or does not want to visit, and the activities they do or do not want to include during their holiday.

In this chapter we have examined a variety of situations in which algebra can be used to identify patterns, simplify or display data. An important skill in numeracy is to convert raw data into information in a format suitable for decision making. This might, for example, take the form of graphs, maps, diagrams or general equations. Algebra may be essential to the production of these materials.

Algebra can be a difficult subject. However, algebraic techniques have become so important in modern data processing that a study of algebra will be well worth the effort required.