UNIVERSITY OF STIRLING

PDMU9L4
DATA SKILLS

WORKBOOK 1 (of 3)
PATH 3

Computing Science & Mathematics
Faculty of Natural Sciences

Academic Year 16/17

Chloe

UNIVERSITY OF STIRLING

PDMU9L4: DATA SKILLS
PATH 3:
KEY MATHEMATICAL SKILLS

Topic 1: Relationships

Science, whether it be social science or natural science, is about relationships. Here are some examples:

(i) lower CD prices means higher CD sales,
(ii) more smoking means more work for the NHS,
(iii) more CFCs means a thinner ozone layer,
(iv) the larger the animal the greater the body heat produced,
(v) the higher the temperature the greater the pressure,
(vi) the greater the stimulation the greater the response,
(vii) the greater the household income the greater the use of private transport.

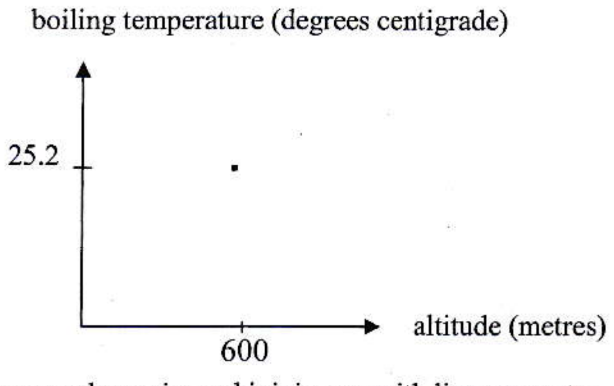
We can make relationships such as these more precise if we can measure numerically the properties being related.

Example 1:
Consider the following data set relating to the phenomenon that occurs when climbing a mountain that water (for a quick brew) boils much more quickly the greater the altitude, i.e. the water boiling temperature decreases with height. Precisely:

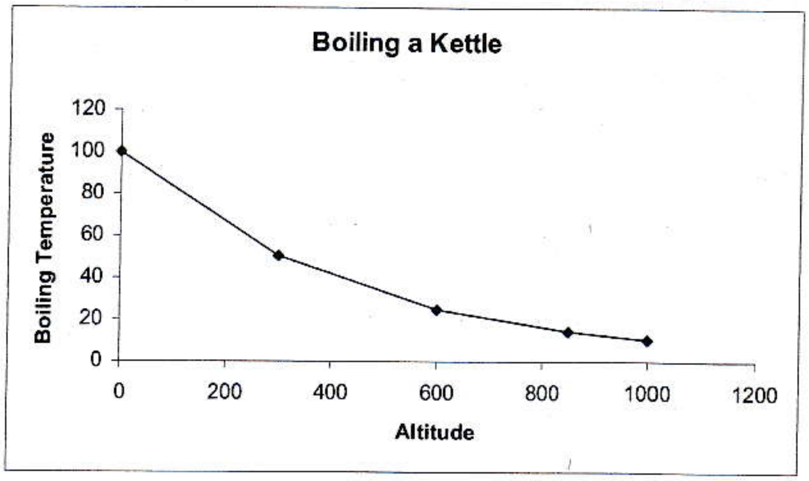
Altitude (metres)	0	300	600	850	1000
Temperature (centigrade)	100	50.2	25.2	14.2	10.1

To "understand" how these properties are related it is helpful to plot the data on graph paper. To do this,

(i) scale the horizontal axis (for altitude) to run from 0 to 1000 metres,
(ii) scalė the vertical axis (for boiling temperature) to run from 0 to 100 (degrees centigrade,
(iii) represent each altitude-temperature number pair (i.e. each column of the table) by a point on the graph paper. This point is located by marking off the altitude value horizontally and the temperature value vertically (Fig.1):



Plotting the four number pairs and joining up with lines we get:

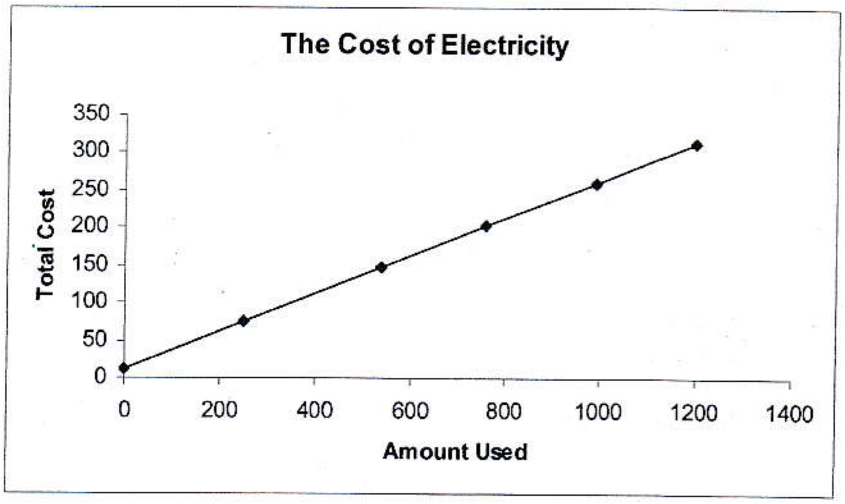


Note: there is quite a dramatic falling off in boiling temperature on the first part of the climb.

Example 2:
This data set relates the cost of electricity in a given quarter (£) and the amount used (unit = kwh = kilowatt hour)

electricity used	0	250	540	760	990	1200
total cost	12.45	74.95	147.45	202.45	259.95	312.45

The plot has the shape:



Comments:
In this second example the plot is a straight line and because of this the relationship is said to be LINEAR.

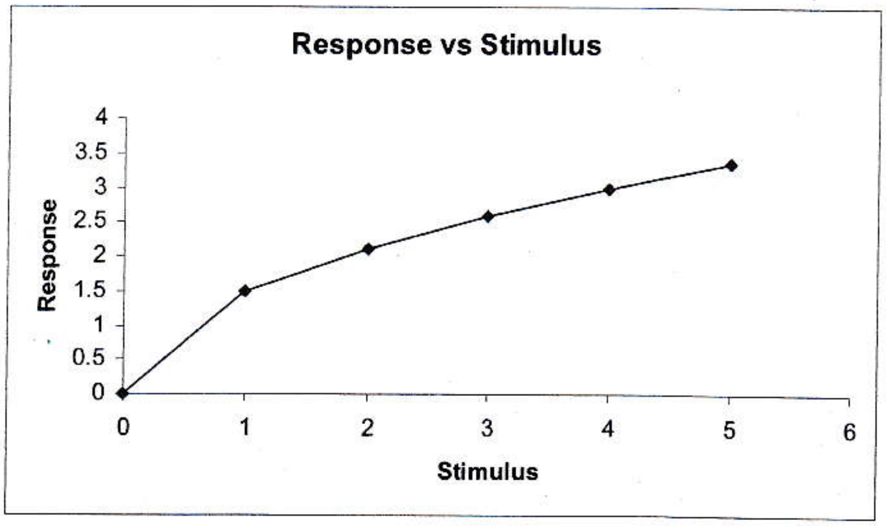
In the first example the plot is not linear and the relationship is said to be NONLINEAR.

There is a second difference between these two plots. In the second example the cost increases with increase in use (the more you use the more you pay!). The relationship is said to be POSITIVE and the two properties are said to be POSITIVELY related. However in the first example the properties are NEGATIVELY (or INVERSELY) related because the *increase* in one property means a *decrease* in the other property.

Example 3: "The Law of Diminishing Returns"
In both Economics and Psychology one is interested in the response to stimuli of various kinds. In Psychology the stimuli will be, for example, visual or aural. In Economics the primary stimulus is 'consumption' and the response 'satisfaction' from that 'consumption'. If the stimulus is increased one would expect the response to increase also but will the increase be proportionate, i.e. will the plot be linear?

Consider the following data with stimulus and response measured relative to a 'benchmark' stimulus and response:

stimulus	0	1	2	3	4	5
response	0	1.5	2.12	2.6	3	3.35



The relationship is clearly not linear – the response increases with the stimulus, but the response is less than proportionate. This makes sense: the more you eat the less you enjoy the extra amount. A loaf when you are starving is of more value to you than a loaf just after you have been out to a good meal at a good restaurant!

Some Definitions

Three key terms we have been using in this Topic Sheet have been:

relationship, properties, plot

From now on we shall, instead, be using the more technical terms:

function, variables, graph (respectively).

A property has different values under different situations, hence the term 'variable'. We will say that the variable is a *function* of the other variable.

Now try the Questions on Worksheet 1/1

PDMU9L4: DATA SKILLS
PATH 3 : KEY MATHEMATICAL SKILLS

Topic 1: Relationships – Worksheet 1/1

In the following questions plot the given data on graph paper and decide whether the relationship is *linear* or *nonlinear* and whether the properties plotted are *negatively* or *positively* related.

1. The temperature of a dead body brought to the mortuary by the police one night is monitored over time, yielding the following data:

time (hours)	0	0.5	1	2	3	4
temperature (centigrade)	31	27	23	17	13	10

Linear? Yes/No — Nonlinear? Yes/No

Positive? Yes/No — Negative? Yes/No

2. You have been given an increase in salary. You have reduced the choice of how to spend this extra money between spending it on more hours of Pay TV and turning on the heat in your flat for longer periods during the day. The table below gives the 'trade-off' between these possibilities in terms of how many extra hours of heat you can have for less hours of Pay TV.

pay TV (hours)	0.5	1	2	3	4
heat (hours)	3	2.7	2	1.33	0.7

Linear? Yes/No — Nonlinear? Yes/No

Positive? Yes/No — Negative? Yes/No

3. The following data gives the daily heat production of five animals compared to the average body weight for those animals.

	cassowary	sheep	human	pig	cow
weight (kgs)	25	50	100	200	500
body heat (cals)	762	1318	2280	3945	8132

Linear? Yes/No — Nonlinear? Yes/No

Positive? Yes/No — Negative? Yes/No

PDMU9L4: DATA SKILLS
PATH 3: KEY MATHEMATICAL SKILLS

Topic 1: Relationships – Worksheet 1/2

In the following questions plot the given data on graph paper and decide whether the relationship is *linear* or *nonlinear* and whether the properties plotted are *negatively* or *positively* related.

1. Bacteria is being grown in a government laboratory. The number of bacteria over time (in hours) was found to be as follows:

time (hours)	0	1	1.5	2	2.5	3
bacteria	2	16	45	128	362	1024

Linear?	Yes/No	Nonlinear?	Yes/No
Positive?	Yes/No	Negative?	Yes/No

2. Continuing the theme of temperature, there are two scales frequently used – Fahrenheit and Centigrade. Conversion between the two is given in the following table for five chosen Centigrade values:

Centigrade	0	30	50	85	100
Fahrenheit	32	86	122	185	212

Linear?	Yes/No	Nonlinear?	Yes/No
Positive?	Yes/No	Negative?	Yes/No

3. I put money into a bank account at a fixed rate of interest, with income reinvested. This money grows over time in the following way:

years	1	2	3	4	5
money (£)	190.39	213.25	238.84	267.5	299.61

Linear?	Yes/No	Nonlinear?	Yes/No
Positive?	Yes/No	Negative?	Yes/No

PDMU9L4: DATA SKILLS
PATH 3:
KEY MATHEMATICAL SKILLS

UNIVERSITY OF STIRLING

Topic 2: Lines and Logs

Lines:
The simplest relationship is the linear relationship. You can say everything there is to say about a line by specifying just two numbers (i) its gradient and (ii) its intercept – i.e. where it cuts the vertical axis. In a moment we will check whether you can handle these concepts. Before that there is a second point to make.

Logs:
A lot of students don't like logarithms (logs for short). But logarithms are in fact 'magic' and can do for you some pretty amazing things. To convince you of this we are going to reveal the secrets of logarithms bit by bit. Here is the first bit.

Logging the Data:

On your scientific calculator you will find a 'log' button. Key in the number 9 and then press this 'log' button. (***) The calculator should then give you the number 0.954(to 3 decimal places). We write:

$$\log(9) = 0.954$$

Similarly, you should find that:

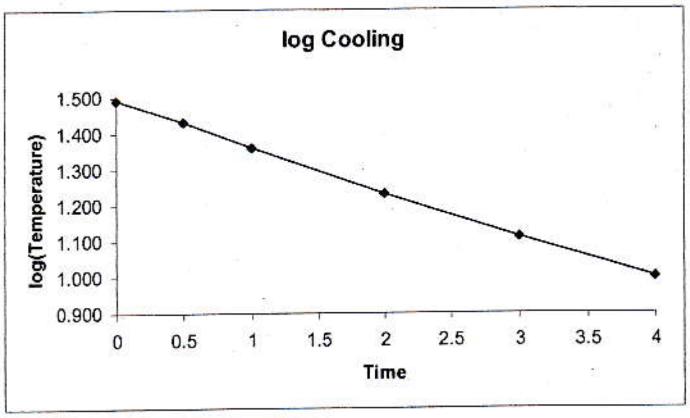
$$\log(0.5) = -0.301$$
$$\log(10.0) = 1.000$$
$$\log(357.1) = 2.553$$

etc etc…

Let us apply this log 'operation' to the temperature values in the first data set in Worksheet 1/1 of Topic 1, the one about the murder victim:

time (hours)	0	0.5	1	2	3	4
temperature (centigrade)	31	27	23	17	13	10
log(temperature)	1.491	1.431	1.362	1.230	1.114	1.000

*** If you have a calculator that uses VPAM (Visually Perfect Algebraic Method) then you can enter the expression in the way that you would write it. E.g. log(9) by pressing 'log' button then enter 9 then pressing '='



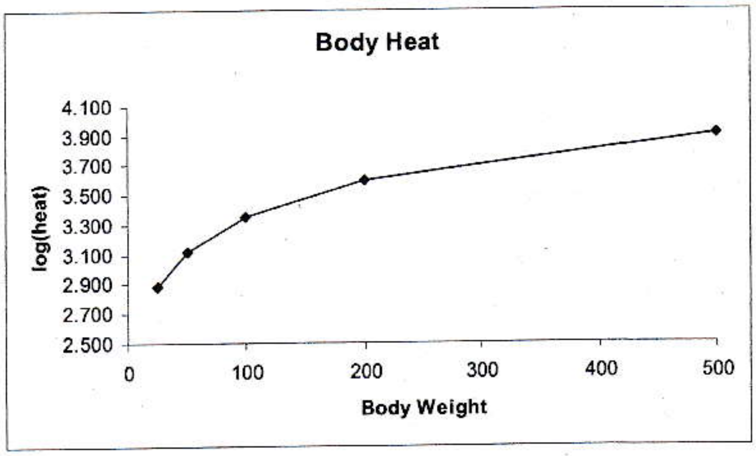
So, if we plot the logs of the temperature values (rather than the temperature values themselves) against time then we obtain a (roughly) linear plot.

Now try Question 1 on Worksheet 2/1.

Let's try the same trick on another of the nonlinear relationships we looked at in the first worksheet of Topic 1 – namely heat generated versus body weight for a range of animals.

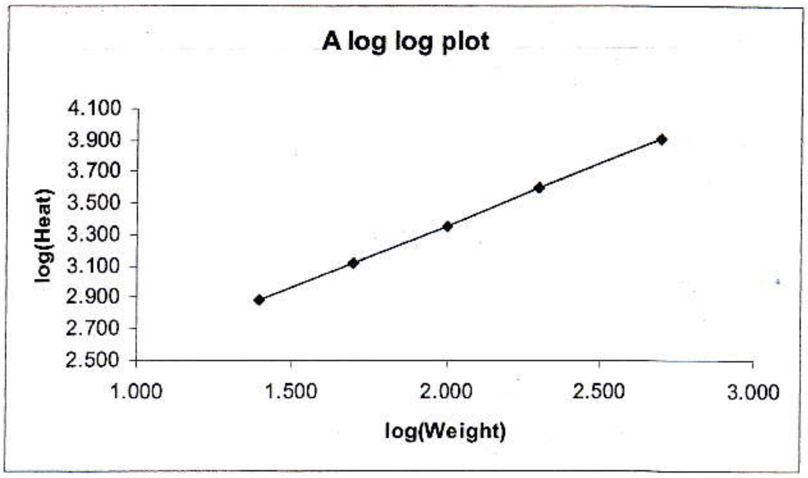
	cassowary	sheep	human	pig	cow
weight (kgs)	25	50	100	200	500
body heat (cals)	762	1318	2280	3945	8132
log (heat)	2.882	3.210	3.358	3.596	3.910

If we plot log(heat) against weight then we don't get a straight line:



However if we plot log(heat) against log(weight) we do get a straight line:

	cassowary	sheep	human	pig	cow
weight (kgs)	25	50	100	200	500
body heat (cals)	762	1318	2280	3945	8132
log (weight)	1.398	1.699	2.000	2.301	2.699
log(heat)	2.882	3.120	3.358	3.596	3.910



Now try Question 2 on Worksheet 2/1.

$y = mx + b$

Message:
So lines are important since many relationships are linear or, if not, then they can be made linear by appropriate log 'transformation'. So we have to learn all about lines.

Properties of Lines:
As we said, we can specify a line completely by two numbers – its *intercept* with the vertical axis and its *gradient*. Let's talk about *gradient*.

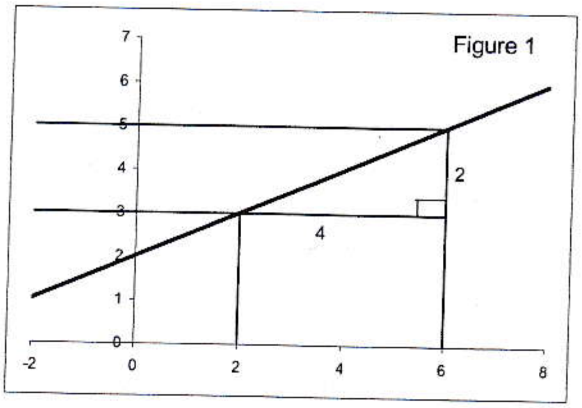
We have an intuitive notion of gradient from everyday life since we associate 'steepness' with going up or down hill. Let's now be absolutely precise about what 'gradient' means and how we calculate it:

$$\text{gradient} = \frac{\text{amount 'travelled' vertically upwards}}{\text{amount 'travelled' horizontally to the right}}$$

In the example of Fig 1 below the gradient is given by:

$$\text{gradient} = \frac{(5.0 - 3.0)}{(6.0 - 2.0)} = \frac{2.0}{4.0} = 0.5$$

(The intercept is given by 2.0 since this is the value at which the line cuts the vertical axis.)



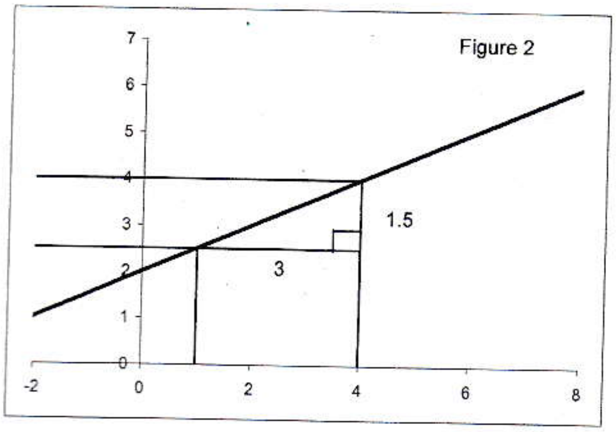
Note that for a line, it does not matter over what interval we undertake our horizontal travel, the gradient values at all points on the line are equal, unlike actual 'hills' where the steepness varies.

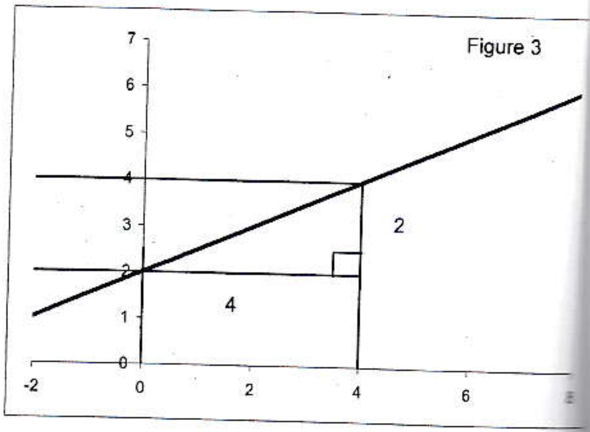
In Fig 1 we travelled horizontally from 2 to 6 and as a result moved vertically from 3 to 5 to get back to the line. However, equally we could have travelled horizontally from 1 to 4 and would then have risen from 2.5 to 4.0 giving us a gradient of (Fig 2):

$$\text{gradient} = \frac{(4.0-2.5)}{(4.0-1.0)} = \frac{1.5}{3.0} = 0.5 \quad \text{(as before)}$$

Or perhaps most simply, we could have travelled from 0.0 to 4.0 horizontally rising from 2.0 to 4.0 to give a gradient of (Fig 3):

$$\text{gradient} = \frac{(4.0-2.0)}{(4.0-0.0)} = \frac{2.0}{4.0} = 0.5 \quad \text{(as before)}$$



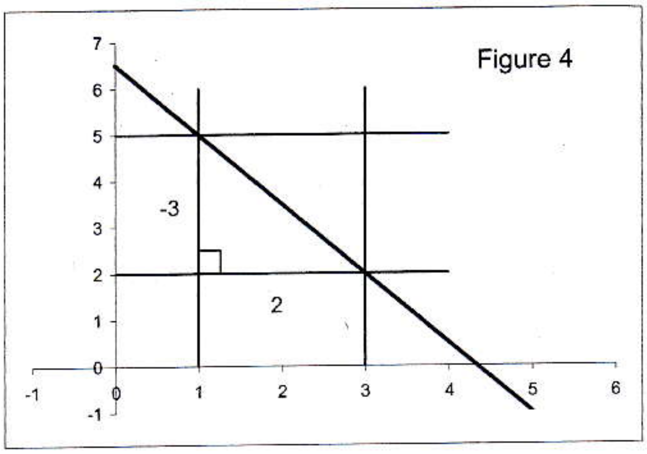


The formula for gradient gives us a positive gradient for the common line of Figs 1, 2, 3.

However for the line Fig 4, the formula gives us a negative gradient:

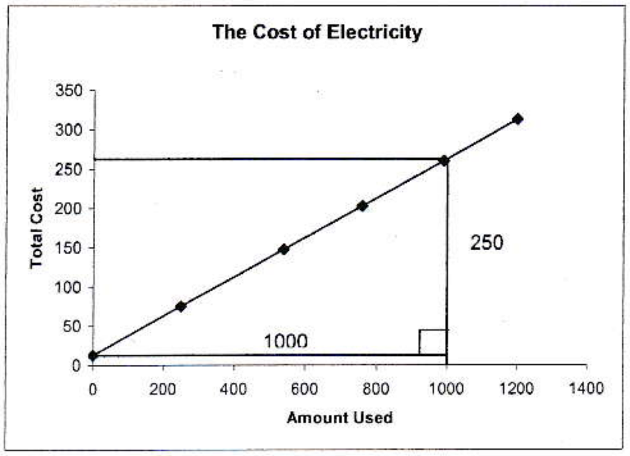
$$\text{gradient} = \frac{(2.0 - 5.0)}{(3.0 - 1.0)} = \frac{-3.0}{2.0} = -1.5.$$

The amount travelled vertically *upward* turns out to be negative since, in fact, we are travelling *downwards* in moving from left to right.



Remember:
Going uphill in travelling from left to right gives a positive gradient.
Going downhill in travelling from left to right gives a negative gradient.

An application:
Consider again the linear function relating total cost to electricity used with graph shown below. The intercept is read off from where the line cuts the axis. This is 12.45 (checking with the data set). The gradient is found to be roughly 250/1000 = 0.25 (i.e. 25 pence).



The intercept is the 'standing charge', i.e. how much you have to pay even before consuming any electricity. The gradient is the price of electricity per unit (of electricity) used, i.e. in this case 25 pence.

Now try the other questions on Worksheet 2/1.

PDMU9L4: DATA SKILLS
PATH 3: KEY MATHEMATICAL SKILLS

Topic2: Lines & Logs – Worksheet 2/1

1. *The boiling kettle problem again.*
   Find the logarithms of the temperature values and plot against altitude.

Altitude (metres)	0	300	600	850	1000
Temperature (centigrade)	100	50.2	25.2	14.2	10.1

2. *Satisfaction revisited.*
   Plot the logarithms of the response values against the logarithms of the stimulus values.

stimulus	0	1	2	3	4	5
response	0	1.5	2.12	2.6	3	3.35

3. *Intercept and gradient.*
   Find the gradients and intercepts of the lines you obtained in questions 1 and 2.

Question 1:

Gradient $\boxed{-\frac{1}{1000}}$ Intercept $\boxed{2}$

(2, 0)

Question 2:

Gradient $\boxed{0,49}$ Intercept $\boxed{0}$

4. *Pay TV.*
   Find the gradient and intercept of the line you obtained in Worksheet 1/1 when plotting extra heating time against extra Pay TV time.

Gradient $\boxed{}$ Intercept $\boxed{}$

−0,67 3.25

PDMU9L4: DATA SKILLS
PATH 3: KEY MATHEMATICAL SKILLS

Topic2: Lines & Logs – Worksheet 2/2

1. *The Bacteria problem*
   Find the logarithms of the population values and plot against time. Find the gradient and intercept of the line you obtain.

Time (hrs)	0	1	1.5	2	2.5	3
Bacteria	2	16	45	128	362	1024

Gradient [ ] Intercept [1,204]

2. *Money in the bank.*
   Plot the logarithms of the amount of the money in the bank against time. Find the gradient and intercept of the line so obtained.

Years	1	2	3	4	5
Money (£)	190.39	213.25	238.84	267.5	299.61

3. *Temperature conversion*
   The following data set relates the temperatures as measured in Centigrade and Fahrenheit.

Centigrade	0	30	50	85	100
Fahrenheit	32	86	122	185	212

Plot this data and find the gradient and intercept of the line obtained.

Gradient [1.8] Intercept [32]

UNIVERSITY OF
STIRLING

PDMU9L4: DATA SKILLS
PATH 3:
KEY MATHEMATICAL SKILLS

Topic 3: Formulae

Formulae:
Most people are aware of Einstein's formula:

$$E = mc^2$$

telling us how much energy (E) we get if we turn mass (m) into energy.
[Note: c denotes the velocity of light which is roughly 300,000 kms per second (pretty fast!) while c^2 (pronounced c squared) means c multiplied by itself: $c \times c = c^2$.]

Formulae are to be found throughout the social and natural sciences. A formula defines precisely a function (relationship), enabling us to find the value of one variable (the dependent variable) for any given value of the other variable: the independent variable.

For example:

(a) (Degrees Fahrenheit) $= 32 + \frac{9}{5} \times$ (degrees Centigrade). (1)

The dependent variable is degrees Fahrenheit;
the independent variable is degrees Centigrade.

If the temperature measured in degrees Centigrade is 20 then in degrees Fahrenheit that temperature would be:

(degrees Fahrenheit) $= 32 + \left(\frac{9}{5} \times 20\right) = 32 + 36 = 68.$

(b) (Cost of electricity) = (standing charge) + (price) $\times$ (amount used) (2)

In the example discussed in Topic 1 the standing charge was £12.45 and the price £0.25 per unit of electricity used. Inserting these numbers in the formula we get:

(cost of electricity) = $12.45 + 0.25 \times$ (amount used).

If I use 627 units then the cost will be:

(cost of electricity) = $12.45 + 0.25 \times 627 = 12.45 + 156.75 =$ £169.20.

A comment and a definition:

First the comment. One problem that sometimes arises in the use of formulae such as (1) and (2) is deciding in which order to carry out the several arithmetic operations involved. The order matters a lot, the difference between getting the answer right and getting the answer wrong! In (1) and (2) we have to add once and multiply once. We multiplied first and then added. What if we had done it in the reverse order?

$$\text{Degrees Fahrenheit} = 32 + \frac{9}{5} \times 20 = \left(32 + \frac{9}{5}\right) \times 20 = (32 + 1.8) \times 20 = 33.8 \times 20 = 676.0$$

Mighty hot compared to the 68 degrees obtained previously! So how do we know which is the correct order? There are two ways to indicate the correct order,

(i) insert brackets to indicate the order;
(ii) have an 'understanding' as to which operations have priority.

The agreed priority listing is as follows:

Highest priority	Evaluate within brackets
Next priority	taking powers
Next priority	division
Next priority	multiplication
Next priority	addition
Next priority	subtraction

If this ordering is not to be adhered to then this will be indicated in the formula with the occurrence of brackets to steer you to the correct ordering.

Examples:

(i) $3/4 + 2 = 0.75 + 2 = 2.75$ (*/ means divide*) (*do division first*)
while $3/(4 + 2) = 3/6 = 0.5$

(ii) $3 \times 6/2 = 3 \times 3 = 9$ (*do division first*)
while $(3 \times 6)/2 = 18/2 = 9$

One way to indicate when to do division is to use the numerator, denominator format. For example:

$$\frac{4 + 7 \times 3}{4 \times 2 - 3} = \frac{25}{5} = 5$$

We evaluate the numerator (top) and the denominator (bottom) separately and then divide.

Now try question 1 on Worksheet 3/1.

Definition:

The highest priority but one is given to 'taking powers'. What precisely does this mean? We have seen one example already from Einstein's famous formula: $E = mc^2$.

A power of a number is that number multiplied by itself a given number of times. For example:

$$3 \times 3 = 3^2$$
$$7 \times 7 \times 7 \times 7 = 7^4$$
$$10 \times 10 \times 10 \times 10 \times 10 \times 10 = 10^6 (= 1 \text{ million})$$

and:

$$a^n = a \times a \times a .... \times a.$$

n denotes the number of factors of a in the product and is called the *exponent.* a^n is said to be the *nth power* of a. Calculating a^n is called taking the nth power of a.

For example:

(i) 3^2 is the second power of 3 (with exponent2). It is also called the square of 3 or 3 squared.

(ii) 8^3 is the third power of 8 (with exponent 3). It is also called the cube of 8 or 8 cubed.

We have a much richer family of formulae if we include powers of a variable.

Formulae (continued):

(c) (volume of sphere) = $\frac{4}{3} \times \pi \times$ (radius of sphere)^3

where π is the special number 'pi' equal to 3.1416 (to 4 decimal places).

If the radius of the sphere is 1.23 then the volume is equal to:

(volume of sphere) =
$\frac{4}{3} \times 3.1416 \times (1.23)^3 = \frac{4}{3} \times 3.1416 \times 1.23 \times 1.23 \times 1.23 = 7.795$.

(d) (money in bank) = (initial investment) $\times \left(1 + \frac{\text{rate of interest}}{100}\right)^n$

where n denotes the number of years the money has been invested. If the initial investment is £2500 and the rate of interest is 8.5% then the formula yields:

(money in bank) = $2500 \times (1 + 0.085)^n = 2500 \times (1.085)^n$

For example after four years ($n = 4$) the money in the account has risen to:

(money in bank) =
$2500 \times (1.085)^4 = 2500 \times 1.085 \times 1.085 \times 1.085 \times 1.085 =$ £3464.65.

Now try Question 2 on Worksheet 3/1!

Calculating Powers:

If we want to find out how much money is in the account after 10 years we have to multiply by 1.085 ten times! Fortunately we can void this pain. Your calculator most probably has a power button labelled by:

$$x^y$$

accessed by using the shift button (or equivalent). To calculate $(1.085)^{10}$ involves the following four steps:

(i) enter 1.085
(ii) press the power button
(iii) enter 10
(iv) evaluate by pressing the '=' button.

and similarly for any other power.

For our example, after 10 years:

$$(\text{money in bank}) = 2500 \times (1.085)^{10} = 2500 \times 2.261 = £5652.46$$

The money has more than doubled.

Now try Question 3 on Worksheet 3/1!

PDMU9L4: DATA SKILLS
PATH 3 : KEY MATHEMATICAL SKILLS

Topic3: Formulae – Worksheet 3/1

1. Consider again the dilemma of deciding whether to spend the salary increase on 'inessentials' such as Pay TV or 'essentials' like heating. The amount left over after deciding on inessentials is given by:

   (Money left over) = (Salary Increase) – (hourly fee charged) × (extra hours of Pay TV)

   If the (net) salary increase is £10.25 (per week) and the Pay TV fee is £1.75 per hour then:

   (Money left over) = 10.25 – 1.75 × (extra hours Pay TV)

   Use this formula to complete the following table:

Extra Pay TV (hrs)	0	1.35	2.74	3.91	6.01
Money left	10.25	7.89	5.46	3.41	-0,27

2. The volume of a cone (as in traffic jam) is given by:

   $$(\text{volume of cone}) = \frac{1}{3} \times \pi \times \text{ height } \times \text{ (base radius)}^2.$$

   Use this formula to complete the following table when the height is 0.91 metres:

Radius (m)	0.1	0.75	1.37	2.14	3.44
Volume (m)	1/300 π	3/16 π	1,97	4.80	12.39

3. The world's population is growing rapidly, according to the 'Law':

   $$(\text{future population}) = (\text{present population}) \times (1+0.20)^n$$

   where n is the number of decades from now and the present population equals 5.6 billion. The formula tells us that the world's population is growing by 20% each decade. Use your calculator power button to complete the following table:

n(decades)	1	3	5	7	10
population	6,72 B	9,68 B	13.93 B	20.06 B	34.67 B

PDMU9L4: DATA SKILLS
PATH 3 : KEY MATHEMATICAL SKILLS

Topic3: Formulae – Worksheet 3/2

1. The amount of blood flowing through a main artery is given by:

   (blood flow) = 0.18 x (radius of artery)^4

   Use this formula to complete the following table:

radius (mm)	0.5	1.5	2.5	3.5	5
flow (cc/sec)	0.01	0.91	7.03	27.01	112
Flow with half (mm)	0.0007	0.056	0.49	1.69	7.03

   Plot this data n graph paper (i) as is, and (ii) after taking lots of both variables.

   What is the gradient and intercept of the log log graph?

   gradient $\boxed{0,71}$ intercept $\boxed{-1,38}$

   If the radius is halved (through disease or otherwise) by what factor is the blood flow reduced: It reduced by 16 times

2. The number of cars parking in Stirling Railway Station depends on the parking charge. Supposing the formula relating 'number of cars' and 'charge' is given by

   (number of cars parking) = 132 – 24 × (parking charge in £).

   Fill in the following table and plot on graph paper:

charge (£)	1	2	3	4	5
cars	108	84	60	36	12

   Find the gradient and intercept of the line so obtained.

gradient $\boxed{-24}$ intercept $\boxed{132}$

How many cars are likely to park if the charge is £7? $\boxed{0}$

Explain: Because there are no **-36** cars.