P1 Chapter 14: Logarithms

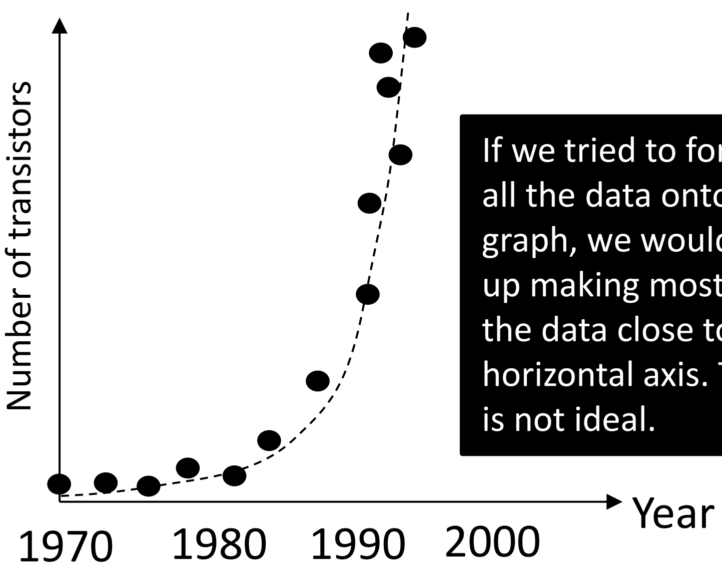
Non-Linear Data

Graphs for Exponential Data

In Science and Economics, **experimental data often has exponential growth**, e.g. bacteria in a sample, rabbit populations, energy produced by earthquakes, my Twitter followers over time, etc.

Because exponential functions increase rapidly, it tends to look a bit rubbish if we tried to draw a suitable graph:

Take for example "Moore's Law", which hypothesised that the processing power of computers would double every 2 years. Suppose we tried to plot this for computers we sampled over time:

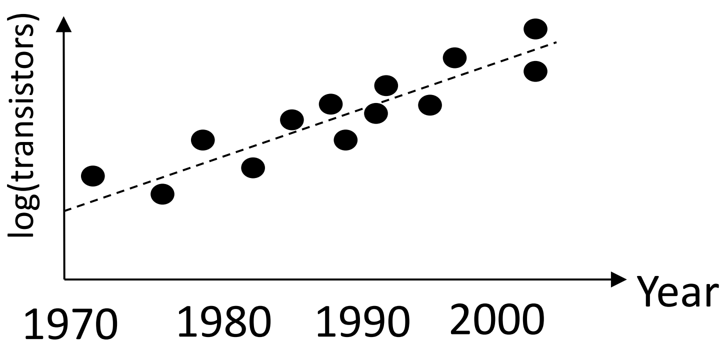


If we tried to force all the data onto the graph, we would end up making most of the data close to the horizontal axis. This is not ideal.

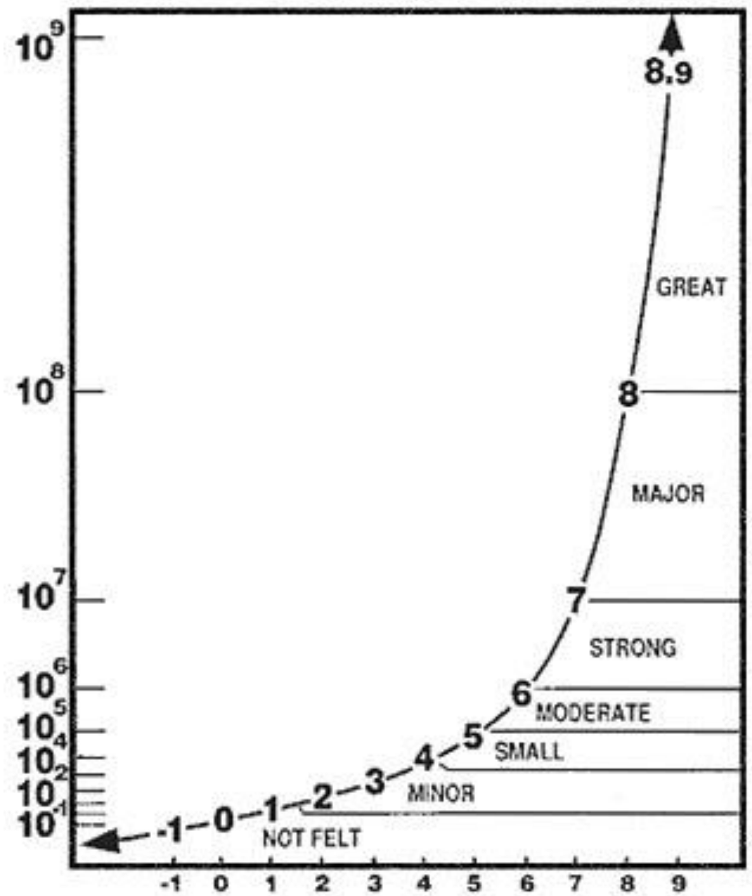
But suppose we **took the log** of the number of transistors for each computer. Suppose the number of transistors one year was y, then doubled 2 years later to get $2y$.
When we log (base 2) these:

$$\begin{aligned} y &\rightarrow \log_2 y \\ 2y &\rightarrow \log_2(2y) = \log_2 2 + \log_2 y \\ &\qquad\qquad\quad = 1 + \log_2 y \end{aligned}$$

The logged value only increased by 1! Thus **taking the log of the values turns <u>exponential growth</u> into <u>linear growth</u> (because each time we would have doubled, we're now just adding 1), and the resulting graph is a straight line.**



Graphs for Exponential Data



Richter Scale

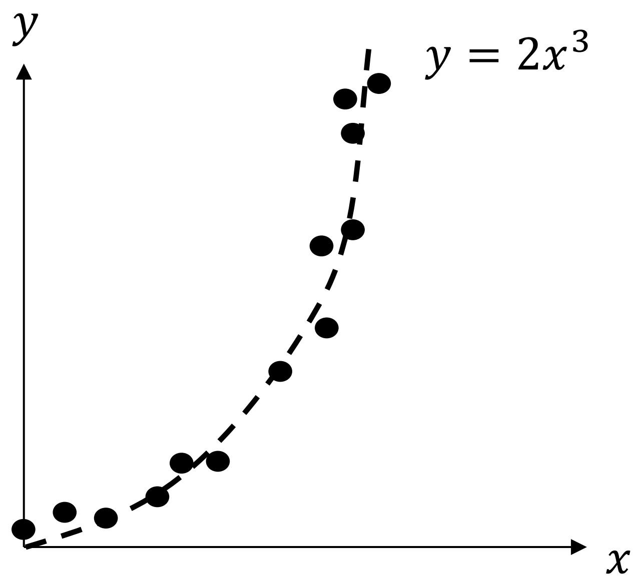
Because the energy involved in **earthquakes** decreases exponentially from the epicentre of the earthquake, such energy values recorded from different earthquakes would **vary wildly**.

The **Richter Scale** is a **logarithmic scale**, and takes the log (base 10) of the amplitude of the waves, giving a more even spread of values in a more sensible range.

(The largest recorded value on the Richter Scale is 9.5 in Chile in 1960, and 15 would destroy the Earth completely – evil scientists take note)

The result is that an earthquake just 1 greater on the Richter scale would in fact be 10 times as powerful.

Other Non-Linear Growth



We would also have similar graphing problems if we tried to plot data that followed some **polynomial function** such as a quadratic or cubic.

We will therefore look at the process to convert a **polynomial graph into a linear one**, as well as a **exponential graph into a linear one…**

Turning non-linear graphs into linear ones

Case 1: Polynomial → Linear

Suppose our original model was a polynomial one*:

$$y = ax^n$$

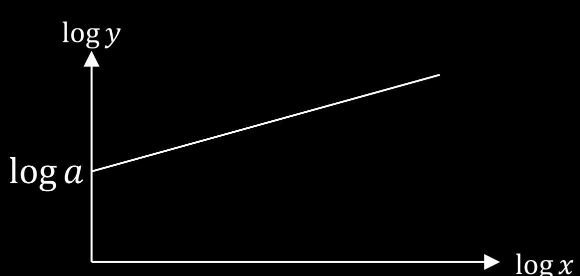
Then taking logs of both sides:

$$\log y = \log ax^n$$
$$\log y = \log a + n \log x$$

We can compare this against a straight line:

$$Y = mX + c$$

✎ If $y = ax^n$, then the graph of $\log y$ against $\log x$ will be a straight line wih gradient n and vertical intercept $\log a$.



* We could also allow non-integer n; the term would then not strictly be polynomial, but we'd still say the function had "polynomial growth".

Case 2: Exponential → Linear

Suppose our original model was an exponential one:

$$y = ab^x$$

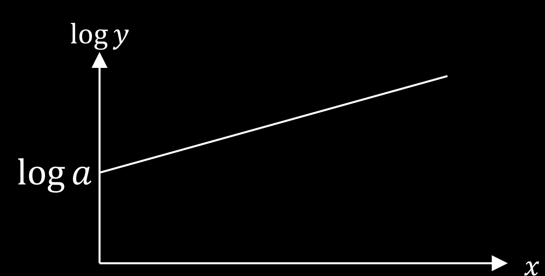
Then taking logs of both sides:

$$\log y = \log ab^x$$
$$\log y = \log a + x \log b$$

Again we can compare this against a straight line:

$$Y = mX + c$$

✎ If $y = ab^x$, then the graph of $\log y$ against x will be a straight line with gradient $\log b$ and vertical intercept $\log a$.



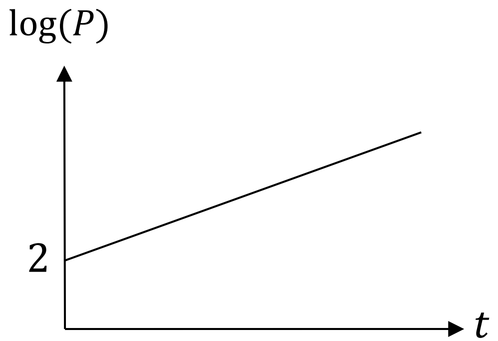
The key difference compared to Case 1 is that we're **only logging the y values** (e.g. number of transistors), not the x values (e.g. years elapsed). **Note that you <u>do not need to memorise</u> the contents of these boxes and we will work out from scratch each time…**

In summary, logging the y-axis **turns an exponential graph into a linear one**. Logging **both** the x and y-axis turns a polynomial graph into a linear one.

Example

[Textbook] The graph represents the growth of a population of bacteria, P, over t hours. The graph has a gradient of 0.6 and meets the vertical axis at $(0,2)$ as shown.
A scientist suggest that this growth can be modelled by the equation $P = ab^t$, where a and b are constants to be found.

a. Write down an equation for the line.
b. Using your answer to part (a) or otherwise, find the values of a and b, giving them to 3 sf where necessary.
c. Interpret the meaning of the constant a in this model.



a ?

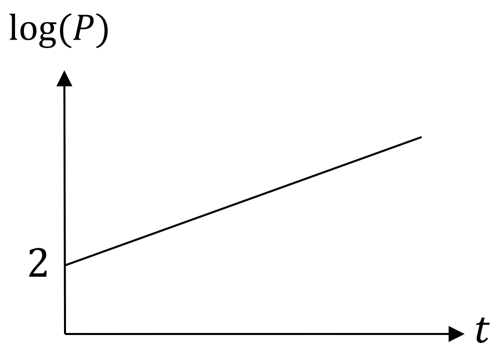
b ?

c ?

Example

[Textbook] The graph represents the growth of a population of bacteria, P, over t hours. The graph has a gradient of 0.6 and meets the vertical axis at $(0,2)$ as shown.
A scientist suggest that this growth can be modelled by the equation $P = ab^t$, where a and b are constants to be found.

a. Write down an equation for the line.
b. Using your answer to part (a) or otherwise, find the values of a and b, giving them to 3 sf where necessary.
c. Interpret the meaning of the constant a in this model.



a

$$\log P = 0.6t + 2$$

Equation of straight line is $y = mx + c$ where here:
$y = \log P, m = 0.6, c = 2, x = t$

b

Just like on previous slide, start with the model then log it:
$P = ab^t$
$\log P = \log a + t \log b$
<u>**Comparing with our straight line equation in (a)**</u>:
$\log a = 2 \quad \rightarrow \quad a = 10^2 = 100$
$\log b = 0.6 \quad \rightarrow \quad b = 10^{0.6} = 3.98\ (3sf)$

Recall that $\log a$ means $\log_{10} a$

c

a gives the initial size of the bacteria population.

Recall that the coefficient of an exponential term gives the 'initial value'.

Example

[Textbook] The table below gives the rank (by size) and population of the UK's largest cities and districts (London is number 1 but has been excluded as an outlier).

City	B'ham	Leeds	Glasgow	Sheffield	Bradford
Rank, R	2	3	4	5	6
Population, P	1 000 000	730 000	620 000	530 000	480 000

The relationship between the rank and population can be modelled by the formula:
$P = aR^n$ where a and n are constants.

a) Draw a table giving values of $\log R$ and $\log P$ to 2dp.

b) Plot a graph of $\log R$ against $\log P$ using the values from your table and draw the line of best fit.

c) Use your graph to estimate the values of a and n to two significant figures.

Textbook Error: They use $R = aP^n$ but then plot $\log P$ against $\log R$.

a

?

b

?

c

?

Example

[Textbook] The table below gives the rank (by size) and population of the UK's largest cities and districts (London is number 1 but has been excluded as an outlier).

City	B'ham	Leeds	Glasgow	Sheffield	Bradford
Rank, *R*	2	3	4	5	6
Population, *P*	1 000 000	730 000	620 000	530 000	480 000

The relationship between the rank and population can be modelled by the formula:
$P = aR^n$ where a and n are constants.

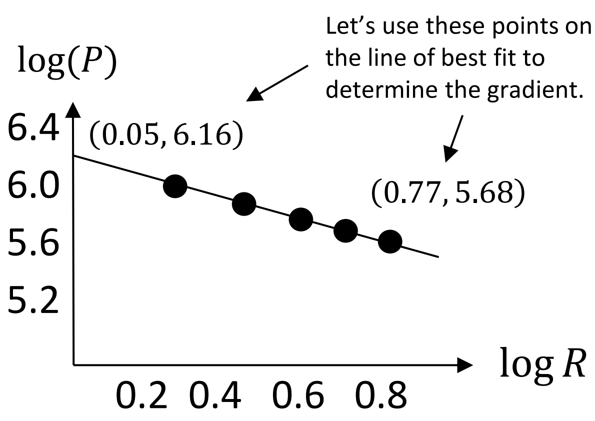
a) Draw a table giving values of $\log R$ and $\log P$ to 2dp.
b) Plot a graph of $\log R$ against $\log P$ using the values from your table and draw the line of best fit.
c) Use your graph to estimate the values of a and n to two significant figures.

Textbook Error: They use $R = aP^n$ but then plot $\log P$ against $\log R$.

a

log R	0.30	0.48	0.60	0.70	0.78
log P	6	5.86	5.79	5.72	5.68

b



c

First get equation of straight line:

$$m = \frac{\Delta y}{\Delta x} = \frac{5.68 - 6.16}{0.77 - 0.05} = -0.67$$

$c = 6.2$ (reading from the graph)

$$\therefore \log(P) = -0.67\log(R) + 6.2$$

As with previous example, let's log the original model so we can compare against our straight line:

$$P = aR^n$$
$$\log P = \log a + n \log R$$

Comparing this with our straight line:

$\log a = 6.2 \quad \rightarrow \quad a = 10^{6.2} = 1600000\ (2sf)$

$n = -0.67$

Test Your Understanding

TheRealSanta wants to predict his number of Twitter followers P (@ TheRealSanta) t years from the start 2015. He predicts that his followers will increase exponentially according to the model $P = ab^t$, where a, b are constants that he wishes to find.

He records his followers at certain times. Here is the data:

Years t after 2015:	0.7	1.3	2.2
Followers P:	2353	3673	7162

a) Draw a table giving values of t and $\log P$ (to 3dp).

b) A line of best fit is drawn for the data in your new table, and it happens to go through the first data point above (where $t = 0.7$) and last (where $t = 2.2$). Determine the equation of this line of best fit. (The y-intercept is 3.147)

c) Hence, determine the values of a and b in the model.

d) Estimate how many followers TheRealSanta will have at the start of 2020 (when $t = 5$).

a

t	?
$\log P$	

b

?

c

?

d

?

Reflections: Consider what we're doing in this whole process in case you don't understand why we're doing all of this:

1. We want to find the **parameters of a model**, e.g. $P = ab^t$ that **best fits the data** (in this case the parameters we want to find are a and b).
2. If the data had a linear trend, then this would be easy! We know from KS3 that we'd just plot the data, find the line of best fit, then use the gradient and y-intercept to work out the m and c in our linear model.
3. But the original data wasn't linear, and it would be harder to draw an 'exponential curve of best fit'.
4. We therefore log the model so that the plotted data then roughly forms a straight line, and then we can then draw a (straight) line of best fit.
5. The gradient and y-intercept of this line then allows us to estimate the parameters a and b in the original model that best fit the data.

The process of finding parameters in a model, that best fits the data, is known as **regression**.

Test Your Understanding

TheRealSanta wants to predict his number of Twitter followers P (@ TheRealSanta) t years from the start 2015. He predicts that his followers will increase exponentially according to the model $P = ab^t$, where a, b are constants that he wishes to find.

He records his followers at certain times. Here is the data:

Years t after 2015:	0.7	1.3	2.2
Followers P:	2353	3673	7162

a) Draw a table giving values of t and $\log P$ (to 3dp).
b) A line of best fit is drawn for the data in your new table, and it happens to go through the first data point above (where $t = 0.7$) and last (where $t = 2.2$). Determine the equation of this line of best fit. (The y-intercept is 3.147)
c) Hence, determine the values of a and b in the model.
d) Estimate how many followers TheRealSanta will have at the start of 2020 (when $t = 5$).

a

t	0.7	1.3	2.2
$\log P$	3.372	3.565	3.855

b

$$m = \frac{3.855 - 3.372}{2.2 - 0.7} = 0.322$$
$$c = 3.147$$
$$\therefore \log P = 0.322t + 3.147$$

c

$$P = ab^t$$
$$\log P = \log a + t \log b$$
$$\therefore \log a = 3.147 \quad \rightarrow \quad a = 1403$$
$$\log b = 0.322 \quad \rightarrow b = 2.099$$

d

$$P = 1403(2.099)^t$$
When $t = 5, P = 57164$

Reflections: Consider what we're doing in this whole process in case you don't understand why we're doing all of this:

1. We want to find the **parameters of a model**, e.g. $P = ab^t$ that **best fits the data** (in this case the parameters we want to find are a and b).
2. If the data had a linear trend, then this would be easy! We know from KS3 that we'd just plot the data, find the line of best fit, then use the gradient and y-intercept to work out the m and c in our linear model.
3. But the original data wasn't linear, and it would be harder to draw an 'exponential curve of best fit'.
4. We therefore log the model so that the plotted data then roughly forms a straight line, and then we can then draw a (straight) line of best fit.
5. The gradient and y-intercept of this line then allows us to estimate the parameters a and b in the original model that best fit the data.

The process of finding parameters in a model, that best fits the data, is known as **regression**.

Exercise 14.8

Pearson Pure Mathematics Year 1/AS
Page 120

Homework Exercise

1 Two variables, S and x satisfy the formula $S = 4 \times 7^x$.

a Show that $\log S = \log 4 + x \log 7$.

b The straight line graph of $\log S$ against x is plotted. Write down the gradient and the value of the intercept on the vertical axis.

2 Two variables A and x satisfy the formula $A = 6x^4$.

a Show that $\log A = \log 6 + 4 \log x$.

b The straight line graph of $\log A$ against $\log x$ is plotted. Write down the gradient and the value of the intercept on the vertical axis.

3 The data below follows a trend of the form $y = ax^n$, where a and n are constants.

x	3	5	8	10	15
y	16.3	33.3	64.3	87.9	155.1

a Copy and complete the table of values of $\log x$ and $\log y$, giving your answers to 2 decimal places.

log x	0.48	0.70	0.90	1	1.18
log y	1.21				2.19

b Plot a graph of $\log y$ against $\log x$ and draw in a line of best fit.

c Use your graph to estimate the values of a and n to one decimal place.

Homework Exercise

4 The data below follows a trend of the form $y = ab^x$, where a and b are constants.

x	2	3	5	6.5	9
y	124.8	424.4	4097.0	30 763.6	655 743.5

a Copy and complete the table of values of x and $\log y$, giving your answers to 2 decimal places.

x	2	3	5	6.5	9
$\log y$	2.10				

b Plot a graph of $\log y$ against x and draw in a line of best fit.

c Use your graph to estimate the values of a and b to one decimal place.

5 Kleiber's law is an empirical law in biology which connects the mass of an animal, m, to its resting metabolic rate, R. The law follows the form $R = am^b$, where a and b are constants.

Animal	Mouse	Guinea pig	Rabbit	Goat	Cow
Mass, m (kg)	0.030	0.408	4.19	34.6	650
Metabolic rate R (kcal per day)	4.2	32.3	195	760	7637

a Copy and complete this table giving values of $\log R$ and $\log m$ to 2 decimal places. **(1 mark)**

$\log m$	−1.52				
$\log R$	0.62	1.51	2.29	2.88	3.88

b Plot a graph of $\log R$ against $\log m$ using the values from your table and draw in a line of best fit. **(2 marks)**

c Use your graph to estimate the values of a and b to two significant figures. **(4 marks)**

d Using your values of a and b, estimate the resting metabolic rate of a human male with a mass of 80 kg. **(1 mark)**

Homework Exercise

6 Zipf's law is an empirical law which relates how frequently a word is used, f, to its ranking in a list of the most common words of a language, R. The law follows the form $f = AR^b$, where A and b are constants to be found.

The table below contains data on four words.

Word	'the'	'it'	'well'	'detail'
Rank, R	1	10	100	1000
Frequency per 100 000 words, f	4897	861	92	9

a Copy and complete this table giving values of $\log f$ to 2 decimal places.

$\log R$	0	1	2	3
$\log f$	3.69			

b Plot a graph of $\log f$ against $\log R$ using the values from your table and draw in a line of best fit.

c Use your graph to estimate the value of A to two significant figures and the value of b to one significant figure.

d The word 'when' is the 57th most commonly used word in the English language. A trilogy of novels contains 455 125 words. Use your values of A and b to estimate the number of times the word 'when' appears in the trilogy.

Homework Exercise

7 The table below shows the population of Mozambique between 1960 and 2010.

Year	1960	1970	1980	1990	2000	2010
Population, P (millions)	7.6	9.5	12.1	13.6	18.3	23.4

This data can be modelled using an exponential function of the form $P = ab^t$, where t is the time in years since 1960 and a and b are constants.

a Copy and complete the table below.

Time in years since 1960, t	0	10	20	30	40	50
$\log P$	0.88					

b Show that $P = ab^t$ can be rearranged into the form $\log P = \log a + t \log b$.

c Plot a graph of $\log P$ against t using the values from your table and draw in a line of best fit.

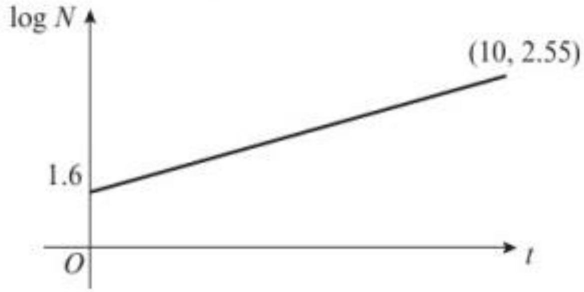
d Use your graph to estimate the values of a and b.

e Explain why an exponential model is often appropriate for modelling population growth.

Hint For part **e**, think about the relationship between P and $\frac{dP}{dt}$.

Homework Exercise

8 A scientist is modelling the number of people, N, who have fallen sick with a virus after t days.



From looking at this graph, the scientist suggests that the number of sick people can be modelled by the equation $N = ab^t$, where a and b are constants to be found.

The graph passes through the points (0, 1.6) and (10, 2.55).

a Write down the equation of the line. **(2 marks)**

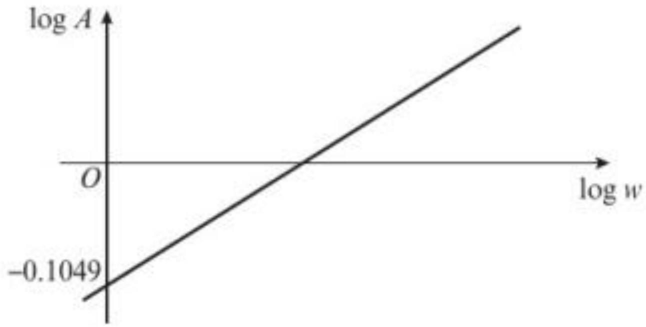
b Using your answer to part **a** or otherwise, find the values of a and b, giving them to 2 significant figures. **(4 marks)**

c Interpret the meaning of the constant a in this model. **(1 mark)**

d Use your model to predict the number of sick people after 30 days. Give one reason why this might be an overestimate. **(2 marks)**

Homework Exercise

9 A student is investigating a family of similar shapes. She measures the width, w, and the area, A, of each shape. She suspects there is a formula of the form $A = pw^q$, so she plots the logarithms of her results.



The graph has a gradient of 2 and passes through -0.1049 on the vertical axis.

a Write down an equation for the line.

b Starting with your answer to part **a**, or otherwise, find the exact value of q and the value of p to 4 decimal places.

c Suggest the name of the family of shapes that the student is investigating, and justify your answer.

Hint Multiply p by 4 and think about another name for 'half the width'.

Challenge

Find a formula to describe the relationship between the data in this table.

x	1	2	3	4
y	5.22	4.698	4.2282	3.805 38

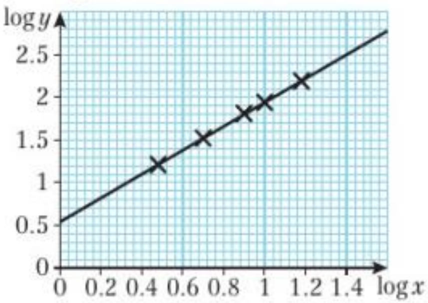
Hint Sketch the graphs of log y against log x and log y against x. This will help you determine whether the relationship is of the form $y = ax^n$ or $y = ab^x$.

Homework Answers

1 **a** $\log S = \log(4 \times 7^x)$
$\log S = \log 4 + \log 7^x$
$\log S = \log 4 + x \log 7$
b gradient $\log 7$, intercept $\log 4$

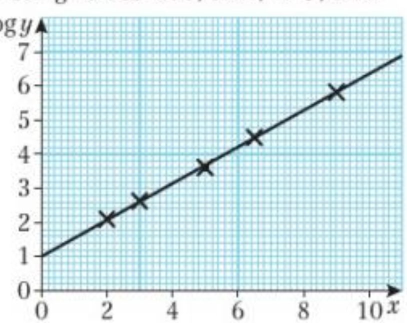
2 **a** $\log A = \log(6x^4)$
$\log A = \log 6 + \log x^4$
$\log A = \log 6 + 4 \log x$
b gradient 4, intercept $\log 6$

3 **a** Missing values 1.52, 1.81, 1.94
b



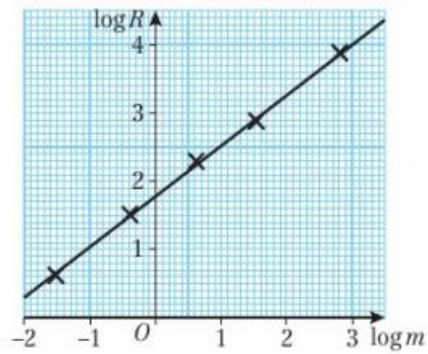
c $a = 3.5$, $n = 1.4$

4 **a** Missing values 2.63, 3.61, 4.49, 5.82
b



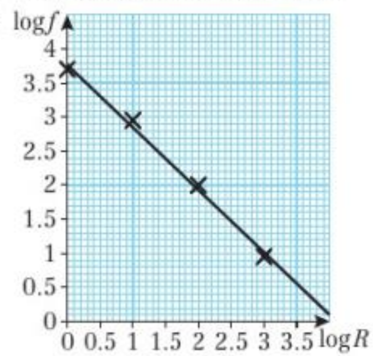
c $b = 3.4$, $a = 10$

5 **a** Missing values −0.39, 0.62, 1.54, 2.81
b



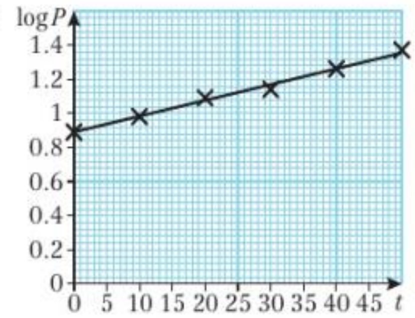
c $a = 60$, $b = 0.75$
d 1,600 kcal per day (2 s.f.)

6 **a** Missing values 2.94, 1.96, 0.95
b



c $A = 5800$, $b = -0.9$
d 694 times

7 **a** Missing values 0.98, 1.08, 1.13, 1.26, 1.37
b $P = ab^t$
$\log P = \log(ab^t)$
$\log P = \log a + \log b^t$
$\log P = \log a + t \log b$
c



d $a = 7.6$, $b = 1.0$
e The rate of growth is often proportional to the size of the population

8 **a** $\log N = 0.095t + 1.6$
b $a = 40$, $b = 1.2$
c The initial number of sick people
d 9500 people. After 30 days people may start to recover, or the disease may stop spreading as quickly.

9 **a** $\log A = 2 \log w - 0.1049$
b $q = 2$, $p = 0.7854$
c Circles: p is approximately one quarter π, and the width is twice the radius, so $A = \frac{\pi}{4}w^2 = \frac{\pi}{4}(2r)^2 = \pi r^2$.

Challenge
$y = 5.8 \times 0.9^x$