7.5 Modelling Data Using Logarithmic Functions

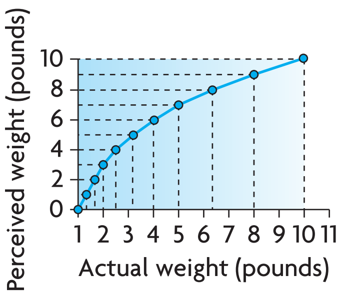
YOU WILL NEED

- graphing technology

EXPLORE

- In 1834, German physiologist E.H. Weber discovered that our perceptions are logarithmically related to the stimuli we receive. For example, it is more difficult for us to distinguish between the weight of two different objects if both objects are heavy than if both objects are light. To detect a change in weight, Weber found that there needs to be a larger increase in the weight of an object as the weight increases. What are the characteristics of the function that models this situation?

Perception vs. Actual Weight



GOAL

Represent data using a logarithmic function and interpret the graph to solve a problem.

INVESTIGATE the Math

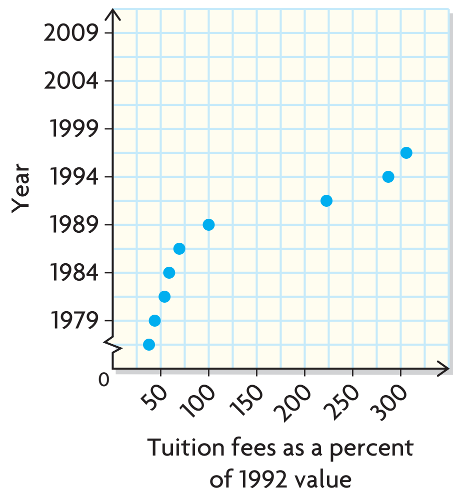
Lydia is researching the rise in tuition fees for post-secondary education in Alberta for her school website. She found some data that uses the tuition fees in 1992 as the benchmark, assigning them a value of 100%. The tuition fees in all other years are compared with the tuition fees in 1992.

Tuition Fees as a Percent of Cost in 1992 (%)	Year
37.8	1979
43.8	1982
54.0	1984
58.8	1986
69.4	1989
100.0	1992
222.7	1999
287.1	2004
305.9	2006

Statistics Canada, Table 326-0002—Consumer Price Index (CPI)

Lydia sketched a scatter plot to help her visualize the trend. She thinks that a logarithmic function may best model the data. She also thinks that writing a summary to describe the rise in tuition fees, in terms of doubling time, would be an interesting feature.

Increase in Tuition Fees



? What is the doubling time for tuition fees in Alberta?

A. Using technology, create a scatter plot to display Lydia's data. Describe the characteristics of the trend in the data. Explain why a logarithmic function is a good model for the data.

B. Estimate the doubling time from the scatter plot. Explain how you determined your estimate.

C. Most graphing calculators and spreadsheets use the natural logarithm when performing a logarithmic regression analysis. Determine the equation of the logarithmic regression function for the data.

D. Graph your regression equation on the scatter plot. Is this regression equation a good model for the data? Explain.

E. Use your graph to estimate the doubling time for tuition fees in Alberta.

Communication *Tip*

The equation of the logarithmic regression function can be written as

$y = (\text{constant}) + (\text{multiplier}) \cdot \ln x$

Most graphing calculators and spreadsheets provide the equation of the logarithmic regression function in the form

$y = a + b \ln x$

Reflecting

F. Choose a different set of values from Lydia's data to calculate the doubling time. Is the doubling time the same as before?

G. Explain why the regression equation does not show 1992 exactly when the tuition fees were 100%.

H. Describe the intervals where the regression equation is a good model for the data and where the regression equation does not model the data very well.

APPLY the Math

EXAMPLE 1 Using logarithmic regression to solve a problem graphically

The flash on most digital cameras requires a charged capacitor in order to operate. The percent charge, Q, remaining on a capacitor was recorded at different times, t, after the flash had gone off.

The t.5 flash duration represents the time until a capacitor has only 50% of its initial charge. The t.5 flash duration also represents the length of time that the flash is effective, to ensure that the object being photographed is properly lit.

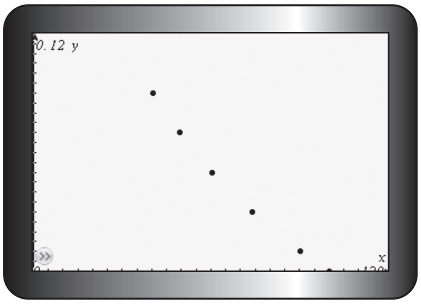
Percent Charge, Q (%)	Time, t (s)
100.00	0
90.26	0.01
73.90	0.03
60.51	0.05
49.54	0.07
40.56	0.09

a) Construct a scatter plot for the given data.

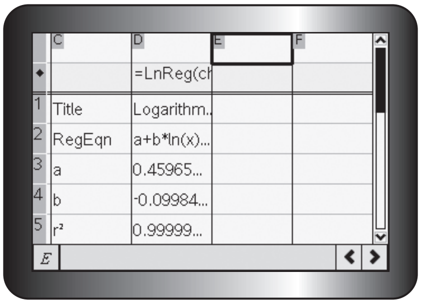
b) Determine a logarithmic model for the data.

c) Use your logarithmic model to determine the t.5 flash duration to the nearest hundredth of a second.

Rico's Solution

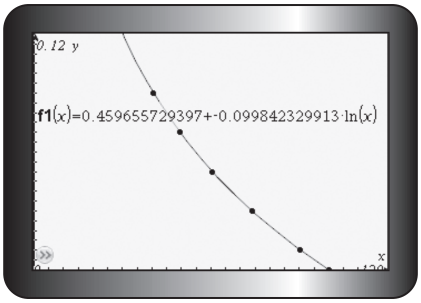
a) 

I entered the data in my graphing calculator and created a scatter plot.

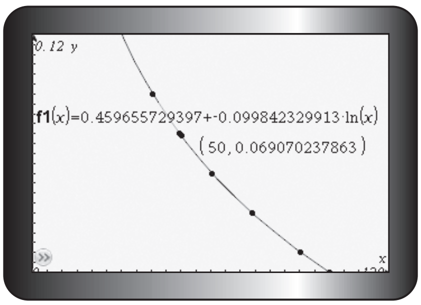
b) 

I used my calculator to determine the equation of the logarithmic regression function for the given data in order to determine the curve of best fit. I knew that my graphing calculator can only provide the equation of the logarithmic regression function using the natural logarithm in the form $y = a + b \ln x$.

The equation is $y = 0.459... - 0.099... (\ln x)$.



I verified the regression equation by entering the equation and graphing it on the same grid as the data points. I noticed that the data points lie on the graph, so the regression equation appears to be a good model for the data.

c) 

To determine the t.5 flash duration for the data, I interpolated the y-value from the graph at $x = 50$.

At about 0.07 s, the t.5 flash duration has been reached.

Your Turn

The t.1 flash duration represents the time until a capacitor has just 10% of its initial charge. Determine the t.1 flash duration for the data above, to the nearest hundredth of a second.

EXAMPLE 2 Using logarithmic regression to solve a problem algebraically

Caffeine is found in coffee, tea, and soft drinks. Many people find that caffeine makes it difficult for them to sleep. The following data was collected in a study to determine how quickly the human body metabolizes caffeine. Each person started with 200 mg of caffeine in her or his bloodstream, and the caffeine level was measured at various times.

Caffeine Level in Bloodstream, *m* (mg)	Time after Ingesting, *t* (h)	Caffeine Level in Bloodstream, *m* (mg)	Time after Ingesting, *t* (h)
168	1.0	33	14.0
167	1.5	80	7.5
113	5.0	145	3.0
145	3.0	100	6.0
90	6.5	71	8.5
125	4.0	156	2.0
138	3.5	153	2.5
77	8.0	130	4.0
83	7.0	90	6.5
50	12.0	112	5.0
150	2.5	32	16.0
55	12.0	23	18.0
112	5.0	25	17.5
84	7.0	45	13.0
136	3.5	27	18.5
180	1.0	18	20.0
110	5.0	29	15.0
75	8.0	43	12.0
76	9.0	25	17.5
49	12.5	21	19.0

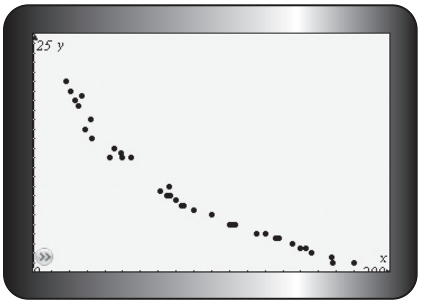
a) Determine the equation of the logarithmic regression function for the data representing time as a function of caffeine level.

b) Determine the time it takes for an average person to metabolize 50% of the caffeine in her or his bloodstream. Round your answer to the nearest tenth of an hour.

c) Paula drank a cup of coffee that contained 200 mg of caffeine at 10:00 a.m. How much caffeine will be in her bloodstream at 9:00 p.m. that evening? Round your answer to the nearest milligram.

Kourosh's Solution

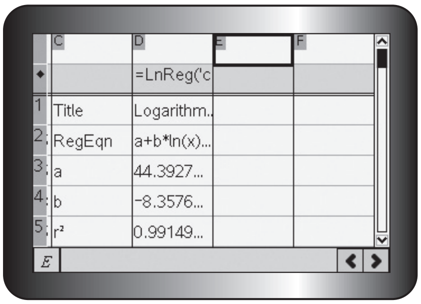
a) Let x represent the caffeine level in the bloodstream in milligrams.
Let y represent the time in hours.



I used x and y as variables since I knew that I was going to be using technology.

Since time is a function of caffeine level, caffeine level is the independent variable and time is the dependent variable.

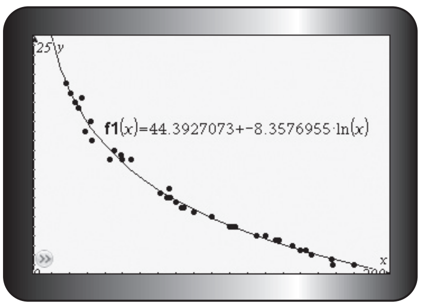
I entered the data in my graphing calculator and created a scatter plot.



	C	D	E	F
◆		=LnReg('c		
1	Title	Logarithm..		
2	RegEqn	a+b*ln(x)...		
3	a	44.3927...		
4	b	-8.3576...		
5	r²	0.99149...		

I used my calculator to determine the equation of the logarithmic regression function for the given data in order to determine the curve of best fit.

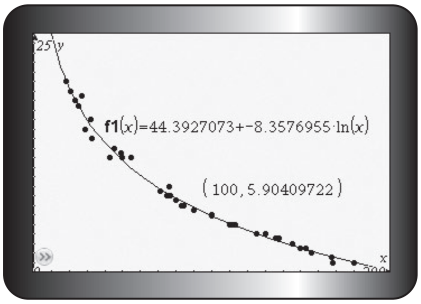
The equation is $y = 44.392... - 8.357... (\ln x)$.



I verified the regression equation by entering the equation and graphing it on the same grid as the data points. I noticed that, although some data points lie above or below the graph, the regression equation appears to be a good model for the data.

b) To determine the time it takes for an average person to metabolize 50% of the caffeine in her or his bloodstream, I need to determine 50% of the initial amount of caffeine.

$$0.5 \cdot 200 \text{ mg} = 100 \text{ mg}$$

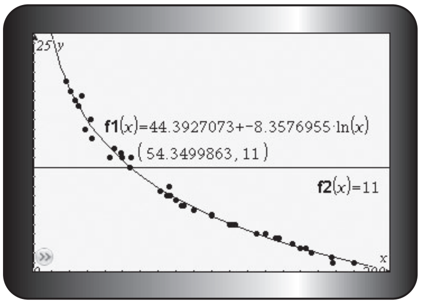


50% of the initial amount of caffeine is 100 mg, so I interpolated the *y*-value from the graph at $x = 100$.

It takes about 5.9 h for an average person to metabolize 50% of the caffeine in her or his bloodstream.

c) The difference between 10 a.m. and 9 p.m. is 11 h. I need to solve the equation

$$11 = 44.392... - 8.357... (\ln x).$$



I graphed the horizontal line $y = 11$ on my calculator and determined the point of intersection of the linear-logarithmic system of equations. The intersection point gives the value of the independent variable, caffeine level, when the dependent variable, time, is 11.

At 9:00 p.m., Paula will have about 54 mg of caffeine in her bloodstream.

Your Turn

Paula is only able to sleep if her caffeine levels are below 40 mg. If Paula usually goes to bed at 10 p.m., what is the latest time that she can enjoy a cup of coffee? State your answer to the nearest quarter hour.

In Summary

Key Idea

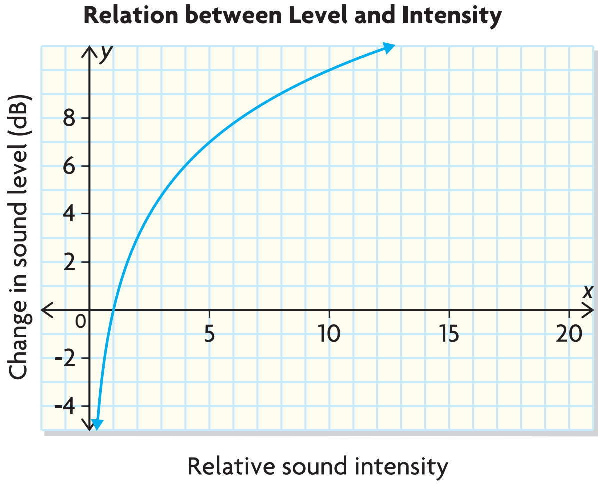
- A logarithmic function may be a good model for a set of data if the points on a scatter plot form an increasing or decreasing curve, where the domain is restricted to the set of positive real numbers.

Need to Know

- The general form of the logarithmic regression model is
$$y = (\text{constant}) + (\text{multiplier}) \cdot \ln x$$
- Most graphing calculators and spreadsheets provide the equation of the logarithmic regression function in the form
$$y = a + b \ln x$$
- A logarithmic curve of best fit can be used to predict values that are not recorded or plotted. Predictions can be made by reading values from the curve of best fit on a scatter plot or by using the equation of the logarithmic regression function.

CHECK Your Understanding

1. This logarithmic graph shows changes in the decibel level, y, as a function of the relative sound intensity, x.



a) Use the following characteristics to describe the curve, and explain why it is logarithmic:
- the location of any intercepts
- the end behaviour
- the domain and range
- whether the function is increasing or decreasing

b) By what factor does the relative sound intensity change when the decibel level increases by 10?

c) When the relative sound intensity is doubled, how does the decibel level change?

2. Determine the equation of the logarithmic regression function that models the given data, and describe the following characteristics of the graph:
- the location of any intercepts
- the end behaviour
- the domain and range
- whether the function is increasing or decreasing

x	*y*	*x*	*y*	*x*	*y*
5	165	18	305	11	255
12	265	25	340	10	243
16	295	15	285	5	170
3	112	17	300	3	115
17	300	5	165	18	310
12	263	22	330	21	325

PRACTISING

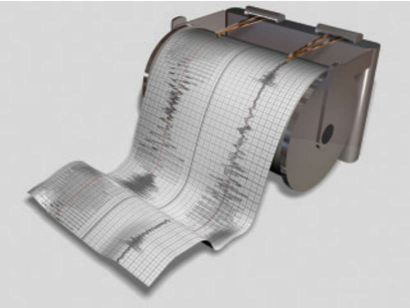
3. Time as a function of population can be approximated by a logarithmic function. The table below shows the population of Alberta in 10-year intervals, from 1901.

Years since 1900	Population of Alberta	Years since 1900	Population of Alberta
101	2 974 807	41	796 169
91	2 545 553	31	731 605
81	2 237 724	21	588 454
71	1 627 874	11	374 295
61	1 331 944	1	73 022
51	939 501		

a) Create a scatter plot to show how time, *t*, is related to population, *P*.

b) Determine the equation of the logarithmic regression function that models the data, and describe the following characteristics of the graph:
- the location of any intercepts
- the end behaviour
- the domain and range
- whether the function is increasing or decreasing

c) Interpolate the year in which the population exceeded 2 000 000.



4. A seismograph records the amplitude of the vibrations during an earthquake by recording the size of the needle deflection in microns $\left(\frac{1}{1000}\text{ mm}\right)$. This data for an earthquake and its many aftershocks was recorded by a seismograph located 100 km from the earthquake.

Seismographic Reading (microns), *r*	Richter Scale Magnitude, *M*
75 023 200	7.88
2 500 010	6.40
500 320	5.70
400 250	5.60
35 400	4.55
5 005	3.70
1 053	3.02
206	2.31

a) Identify the independent and dependent variables.
b) Create a scatter plot to compare the magnitude of the earthquake to its seismographic reading.
c) Determine the equation of the logarithmic regression function that models the data in the table.
d) How many times more intense is a magnitude 5.7 earthquake than a magnitude 4.5 earthquake?

Pressure (kPa)	Altitude (m)
101.30	0
95.40	400
87.14	1000
64.50	3000
74.90	2000
39.90	6187

5. Altitude above sea level is a logarithmic function of atmospheric pressure. Michael just purchased an altimeter watch, which measures the current altitude in metres above sea level. The watch also measures the current atmospheric pressure in kilopascals (kPa). Michael recorded the atmospheric pressure at six different altitudes, as shown in the table to the left.
a) Identify the independent and dependent variables.
b) Use Michael's data to determine the equation of the logarithmic regression function for the altitude, h, as a function of the pressure, P.
c) Describe the following characteristics of the function:
- the intercepts
- the end behaviour
- the domain and range
- whether the function is increasing or decreasing

d) Michael lives at an altitude of 139 m. Determine the pressure setting that he needs to use to calibrate his watch. Round your answer to the nearest tenth.
e) Determine the atmospheric pressure at the summit of Mt. Everest, which is 8848 m above sea level, to the nearest tenth of a kilopascal.

6. Global warming is causing glaciers to melt in the Mackenzie Mountains along the border between Yukon and the Northwest Territories. This is allowing Canadian scientists to discover hunting tools used by ancestral peoples who inhabited this region thousands of years ago. The age of some ancient objects can be estimated using carbon dating. The process involves measuring the percent of carbon-14 in the ancient object and comparing it with current values. By knowing the rate at which carbon-14 decays, scientists can then estimate the age of the object.

The age, t, in years, of an object is a logarithmic function of the percent, P, of carbon-14 remaining in the object. Data for six objects is recorded in the table.

Percent of Carbon-14, P (%)	Age, t (years)
11	18 247
18	14 176
23	12 149
28	10 523
44	6 787
63	3 819

a) Create a scatter plot to display the data, and determine the equation of the logarithmic regression function that models the data.

b) In 2000, Canadian archaeologist Tom Andrews, with members of the Shutaot'ine of the Mountain Dene Nation, discovered fragments of a birch arrow with 96.8% of its original carbon-14 still present. Estimate the age of this arrow. Round your answer to the nearest year.

c) Determine how old an object would be if only 50% of the original carbon-14 remains. Round your answer to the nearest year.

7. Jussi received an inheritance of $15 000. He invested it in a GIC that earns 4.5% per year, compounded annually. He would like his investment to accumulate to $25 000, which he can put toward a down payment on a house. The table below shows Jussi's balance over the first 5 years.

Time, t (years)	Amount, A ($)
0	15 000
1	15 675
2	16 380
3	17 117
4	17 888
5	18 693

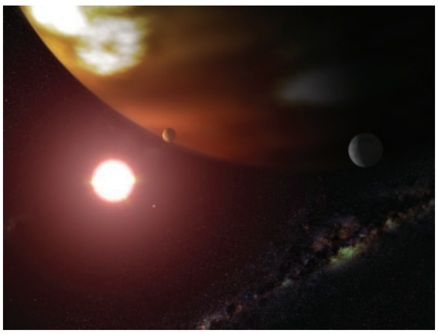
a) Determine the equation of the exponential regression function that models this growth.

b) Determine the equation of the logarithmic regression function for time as a function of amount.

c) Use each equation to determine the time it takes for the balance to equal $25 000. Which equation do you prefer to use? Explain.

8. In astronomy, the distance to a star is determined by comparing its brightness as seen from Earth, called its apparent brightness, to its actual brightness. The distance modulus, m, is defined as the difference between the apparent brightness of a star and its brightness when measured from a distance of 10 pc (parsecs), where 1 pc is 3.09×10^{13} km. The distance modulus is related to the distance, d, of the star from Earth.

Star	Distance from Earth, d (pc)	Distance Modulus, m
Sun	4.85×10^{-6}	−31.57
Proxima Centauri	1.29	−4.44
Alpha Centauri	1.30	−4.40
Barnard's Star	1.82	−4.70
Sirius	2.64	−2.89
Vega	7.76	−0.55
Arcturus	11.25	0.26
Canopus	30.10	2.38
Spica	80.39	4.53
Polaris	110.00	5.19
Betelgeuse	160.00	6.01
Rigel	276.10	7.24
Deneb	490.80	8.36



a) Identify the independent and dependent variables.
b) Create a scatter plot to show the relationship between the distance modulus, m, and the distance from Earth, d.
c) Use the data to determine the equation of the logarithmic regression function for the distance modulus, m, as a function of the distance from Earth, d.
d) Wolf 359 is considered to be the fifth-closest star to Earth, at 2.39 pc. Determine its distance modulus to the nearest hundredth.
e) Gliese 876 had four known planets in 2011 and a distance modulus of −1.66. Determine the distance of this star from Earth. Round your answer to the nearest tenth.

9. In February 2004, Mark Zuckerberg launched Facebook from his Harvard dorm room. Facebook had 1 000 000 registered users by December of that year, and it has been growing rapidly ever since.

Number of Months since Feb., 2004	Number of Registered Users (millions), *n*	Number of Months since Feb., 2004	Number of Registered Users (millions), *n*
10	1.0	62	200.0
22	5.5	65	250.0
34	12.0	67	300.0
38	20.0	70	350.0
44	50.0	72	450.0
54	100.0	77	500.0
59	150.0	79	550.0
60	175.0		

a) Create a scatter plot that relates the date, in number of months, to the number of registered users of Facebook.

b) Use the data to determine the equation of the logarithmic regression function for time, t, in months, as a function of the number of registered users, n.

c) Interpolate to determine when Facebook first surpassed 275 million registered users.

Closing

10. Describe the process you would use to interpolate a value for a set of data that can be modelled by a logarithmic function.

Extending

11. Lara invested $5000 at an annual interest rate of 5%. The compounding period, n, is related to the balance, A, in dollars.

a) Determine the equation of the logarithmic regression function that relates n to A for each compounding period.

i) annually
ii) semi-annually
iii) quarterly
iv) daily

b) Describe how the equation changes as the number of compounding periods per year increases.