

6.4

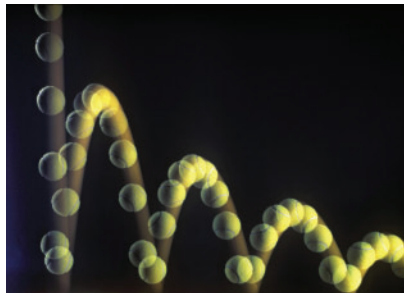
Modelling Data with a Curve of Best Fit

GOAL

Determine the quadratic or cubic function that best fits a set of data, and use the function to solve a problem.

INVESTIGATE the Math

Damien's physics class measured the height of a bouncing ball in the school gymnasium. The ball was repeatedly dropped from the same height. The height was recorded every 0.2 s with a special camera, starting at the first bounce.



Trial 1

Time (s)	0	0.2	0.4	0.6	0.8	1.0
Height (m)	0	1.61	2.43	2.47	1.73	0.20

Trial 2

Time (s)	0	0.2	0.4	0.6	0.8	1.0
Height (m)	0	1.63	2.47	2.53	1.81	0.30

Trial 3

Time (s)	0	0.2	0.4	0.6	0.8	1.0
Height (m)	0	1.60	2.41	2.44	1.69	0.15

? What polynomial function best fits this data?

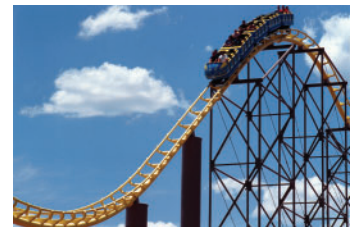
- Decide which variable is the independent variable. Create a scatter plot by plotting the data from all three trials on the same grid.
- Draw a curve that approximates the trend in the data on your scatter plot.
- Describe the shape of the curve. Include the domain and the range in your description. What type of function does the curve represent? Explain.
- Estimate the coordinates of the vertex.
- Determine the equation of the function, in vertex form, that approximates the data.
- Input the data for all three trials into a calculator or spreadsheet. Determine the equation of the quadratic regression function that models the combined data.

YOU WILL NEED

- graphing technology
- ruler
- graph paper

EXPLORE...

- The top of a roller coaster is shown in the photograph below. Describe the characteristics of a polynomial function that might be used to model the shape of this section of the track.



Reflecting

curve of best fit

A curve that best approximates the trend on a scatter plot.

- G. Compare your equation from part E and the equation of the **curve of best fit** from part F. Why are they different?
- H. How could you use one of your functions to determine the time it takes for the ball to reach the second bounce?
- I. How could you use the data or the functions to approximate the maximum height of the ball after the first bounce?
- J. What is the leading coefficient in each function? Why does the coefficient have this value?

APPLY the Math

EXAMPLE 1

Using technology to solve a quadratic problem

Audrey is interested in how speed plays a role in car accidents. She knows that there is a relationship between the speed of a car and the distance needed to stop. She has found the following experimental data on a reputable website, and she would like to write a summary for the graduation class website.

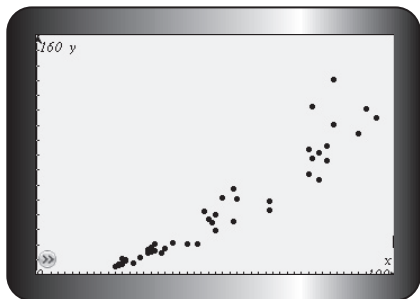
Speed (km/h)	Distance (m)	Speed (km/h)	Distance (m)	Speed (km/h)	Distance (m)
90	94.4	38	21	83	130.4
36	17	92	111	50	29.1
65	49.2	22	5.6	48	37
56	50.3	31	16.8	45	20.7
65	43.1	50	40	81	86
24	10.9	52	51.2	42	20.6
35	14.2	33	15.9	31	14
55	57.3	27	7.4	38	21
81	76.5	33	20.7	29	11
83	100.3	32	17.9	77	112.3
25	9.1	47	41.9	76	84.1
25	10	95	105.2	55	35.3
77	77.8	24	6.7	79	81.8
32	14.9	23	6.9	23	6.2
76	67.3	79	63.6	49	35



- Plot the data on a scatter plot. Determine the equation of a quadratic regression function that models the data.
- Use your equation to compare the stopping distance at 30 km/h with the stopping distance at 50 km/h, to the nearest tenth of a metre.
- Determine the maximum speed that a car should be travelling in order to stop within 4 m, the average length of a car.

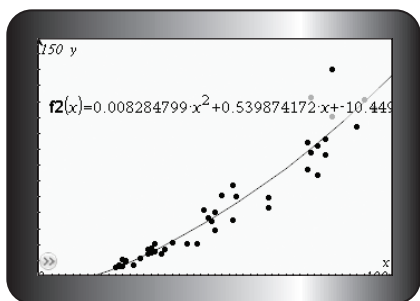
Audrey's Solution

a)



I entered the data into my graphing calculator and created a scatter plot. I decided that speed is the independent variable, because it affects the stopping distance, so I placed speed on the horizontal axis.

The relationship shown on the scatter plot is curved, like half a parabola.



I ran the quadratic regression on my calculator, giving me the curve of best fit and its equation.

The equation of the curve of best fit is
 $y = 0.008...x^2 + 0.539...x - 10.449...$
 where x represents the speed of a car in kilometres per hour and y represents the stopping distance in metres.

b) $y = 0.008...x^2 + 0.539...x - 10.449...$

When $x = 30$ km/h,

$$y = 0.008...(30)^2 + 0.539...(30) - 10.449...$$

$$y = 13.203...$$

When $x = 50$ km/h,

$$y = 0.008...(50)^2 + 0.539...(50) - 10.449...$$

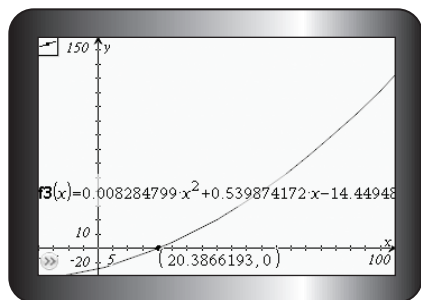
$$y = 37.256... \text{ m}$$

The stopping distance is about 3 times greater for a speed of 50 km/h than it is for a speed of 30 km/h.

I substituted the two different speeds into the regression equation to determine the necessary stopping distance.



c) $y = 0.008...x^2 + 0.539...x - 10.449...$
 $(4) = 0.008...x^2 + 0.539...x - 10.449...$
 $0 = 0.008...x^2 + 0.539...x - 14.449...$



The x -intercept is 20.386...

The car should be travelling at a maximum speed of 20.4 km/h in order to stop within a distance of 4 m, or one car length.

To determine the speed that a car should be travelling in order to stop within 4 m, I substituted 4 for y into the equation and solved for x .

I ended up with a quadratic equation that I could not factor. It is easier to graph the corresponding function than to use the quadratic formula, so I graphed the function.

To determine the roots of the equation, I determined the x -intercepts of the graph of its corresponding function.

There is only one x -intercept in the domain, since $x \geq 0$.

Your Turn

By what factor does the stopping distance increase when the speed is doubled?

EXAMPLE 2 Solving a problem with a cubic regression function

The following table shows the average retail price of gasoline, per litre, for a selection of years in a 30-year period beginning in 1979.

Years after 1979	Price of Gas (¢/L)	Years after 1979	Price of Gas (¢/L)
0	21.98	17	58.52
1	26.18	20	59.43
2	35.63	22	70.56
3	43.26	23	70.00
4	45.92	24	74.48
7	45.78	25	82.32
8	47.95	26	92.82
9	47.53	27	97.86
12	57.05	28	102.27
14	54.18	29	115.29

Statistics Canada

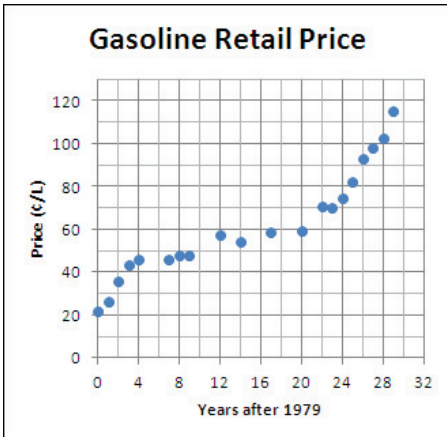
a) Use technology to graph the data as a scatter plot. What polynomial function could be used to model the data? Explain.



- b) Determine the cubic regression equation that models the data. Use your equation to estimate the average price of gas in 1984 and 1985.
 c) Estimate the year in which the average price of gas was 56.0¢/L.

Brad's Solution

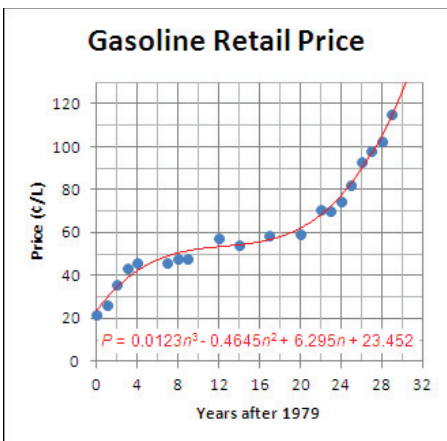
- a) I entered my data into a spreadsheet and created a scatter plot.



----- The price of gas depends on the year, so I placed the price of gas on the vertical axis.

The data is not linear. If the domain were not restricted, it would appear as though the function extends from quadrant III to quadrant I, so a cubic polynomial function may be a good model.

- b) I used the spreadsheet application to draw a cubic curve of best fit for the data.



The equation of the cubic regression function that models this data is

$$P = 0.0123n^3 - 0.4645n^2 + 6.295n + 23.452$$

where P represents the average price of gas per litre and n represents the number of years after 1979.

For average price of gas in 1984, $n = 5$:

$$P = 0.0123(5)^3 - 0.4645(5)^2 + 6.295(5) + 23.452$$

$$P = 44.9¢/L$$

For average price of gas in 1985, $n = 6$:

$$P = 0.0123(6)^3 - 0.4645(6)^2 + 6.295(6) + 23.452$$

$$P = 47.2¢/L$$

I used the spreadsheet application to determine the equation of the cubic regression function.

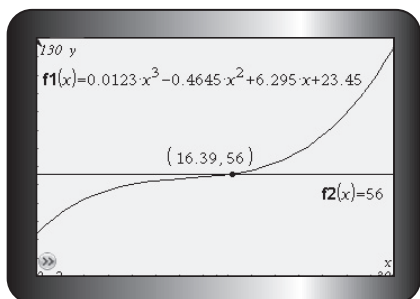
To interpolate the average price of gas in 1984 and 1985, I substituted $n = 5$ and $n = 6$ into the regression equation.

I used 5 and 6 because this is the number of years after 1979.

I estimate that the average price of gas in 1984 was about 44.9¢/L and the average price of gas in 1985 was about 47.2¢/L.

- c) To determine the year in which the average price of gas was 56.0¢/L, I substituted 56.0 for y in the equation of the cubic regression function.

$$56.0 = 0.0123n^3 - 0.4645n^2 + 6.295n + 23.452$$



$$1979 + 16.39 = 1995.39$$

In 1995, the average price of gas was 56.0¢/L.

Since I don't know how to solve a cubic equation algebraically, I used a systems of equations strategy that I have used before with quadratic equations. I graphed each side of the equation as a separate function.

I graphed the functions and determined the n -coordinate of their point of intersection. I know that the n -coordinate of the point of intersection is the solution to the cubic equation.

The n -value for the intersection point is the number of years after 1979 that the price of gas was 56.0¢/L.

I added 1979 to the solution of the equation in order to determine when the price of gas was 56.0 ¢/L.

Your Turn

The actual average prices of gas in 1984, 1989, and 1995 were 69.4¢/L, 72.1¢/L, and 80.1¢/L, respectively. Add these data points to the table, and use interpolation to determine a new average price of gas in 1985.

In Summary

Key Idea

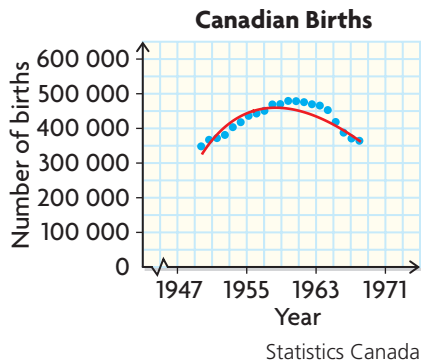
- If the points on a scatter plot seem to follow a predictable curved pattern, then there may be a quadratic or cubic relationship between the independent variable and the dependent variable.

Need to Know

- If the points on a scatter plot follow a quadratic or cubic trend, then graphing technology can be used to determine and graph the equation of the curve of best fit.
- To solve an equation, you can graph the corresponding function of each side of the equation. The x -coordinate of the point of intersection is the solution to the equation.
- Technology uses polynomial regression to determine the curve of best fit. Polynomial regression results in an equation of a curve that balances the points on both sides of the curve.
- A 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 curve of best fit.

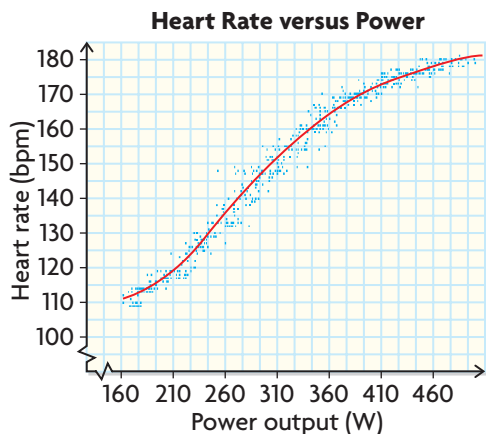
CHECK Your Understanding

- The following graph shows a quadratic curve of best fit (in red) for the number of Canadian births from 1948 to 1968.



- Describe the trend in the data.
 - Based on the graph, in what year did the greatest number of births occur?
 - How many births took place in 1965?
 - In what years did more than 400 000 births occur?
- The data in the following scatter plot represents several trials of an experiment that was conducted by an athlete to compare his heart rate to his power output. The equation of the cubic regression function that models this data is

$$H = -0.000\ 002\ 4x^3 + 0.001\ 930\ 2x^2 - 0.224\ 951\ 6x + 104.820\ 761\ 9$$
 where H represents the athlete's heart rate in beats per minute (bpm) and x represents his power output in watts (W).



- Use the regression equation to estimate the athlete's heart rate when his power output is 310 W.
- Use the curve of best fit to estimate the athlete's power output when his heart rate is 130 bpm.

PRACTISING



3. Josie hit a golf ball from the top of a hill. The height of the ball above the green is given in the table below.

Time (s)	1	2	3	4	5
Height (m)	52.5	73.2	74.6	55.8	16.1

- Describe the characteristics of the data.
 - Determine the equation of the quadratic regression function that models the data.
 - Use your equation to determine the height of the ball at
 - 0 s
 - 2.5 s
 - 4.5 s
 - When did the ball hit the ground?
4. A spherical balloon is being inflated. The volume, V , in cubic centimetres is related to the time, t , in seconds.

Volume, V (cm³)	33.51	113.10	268.08	523.60	904.78
Time, t (s)	0	1	2	3	4

- Use technology to plot the data as a scatter plot. Describe the trend you see.
 - Use cubic regression to create a curve of best fit.
 - Determine the volume of the balloon at 10.5 s.
5. In an experiment, the volume, V , of 1 kg of water is measured as its temperature, T , is increased.

Volume, V (cm³)	999.87	999.75	1000.01	1000.59	1001.44	1002.52	1003.76
Temperature, T (°C)	0	5	10	15	20	25	30

- Use cubic regression to interpolate the temperature at which the water has the minimum volume.
 - Use cubic regression to extrapolate the volume of the water at a temperature of 40 °C.
6. A dolphin jumped out of the water in a tank and then dove back in. The dolphin's height, in metres, above the water is given in the table below. Using quadratic regression, estimate the maximum height of the dolphin during the jump.

Height (m)	0.59	1.22	1.39	0.73	-1.45
Time (s)	0.5	1.0	1.5	2.5	3.5

7. The fertility rate for Canadian women over 40, for several years after 1977, is shown in the table at the right. The rate is the number of babies born, on average, to a group of 1000 women.
- Use quadratic regression to interpolate two missing data values.
 - Compare your interpolated values to the actual values in the table below.

Years after 1977	1	2	7	8	12	13	18	19	20	24	25
Rate per 1000 Females	3.9	3.5	3.1	3.1	3.8	3.9	5.0	5.3	5.4	6.3	6.4

8. A 225 L hot-water tank sprung a leak at $t = 0$ min. The remaining volume was measured every 5 min for the first 40 min.

Volume, V (L)	225	188	155	124	100	72	55	36	23
Time, t (min)	0	5	10	15	20	25	30	35	40

- Plot the data as a scatter plot. Describe the trend.
 - Determine the equation of the quadratic regression function that models the data.
 - Determine when the tank was half full.
 - Determine when the tank was empty.
9. Statistics Canada recorded the following data for the incidence of lung cancer per 100 000 Canadian males in the years after 1976.
- Use a cubic regression function to model the data and extrapolate the incidence of lung cancer in 2010.
 - According to your polynomial model, when should Canada expect to see the incidence of lung cancer drop below 65 per 100 000 males?

Years after 1976	Cancer Incidence	Years after 1976	Cancer Incidence	Years after 1976	Cancer Incidence
0	75.7	10	96.4	20	82.2
1	78.6	11	95.0	21	79.3
2	85.1	12	95.5	22	80.5
3	83.9	13	93.6	23	79.5
4	83.2	14	92.7	24	77.1
5	91.2	15	90.7	25	75.8
6	92.6	16	90.5	26	73.0
7	95.2	17	91.5	27	70.0
8	97.1	18	86.7	28	69.6
9	93.2	19	84.6		

Years after 1977	Rate per 1000 Females
0	3.9
3	3.3
4	3.4
5	3.3
6	3.1
9	3.3
10	3.5
11	3.7
14	4.0
15	4.3
16	4.6
17	4.8
20	5.4
21	5.4
22	5.8
23	6.1
26	6.8
27	7.2
28	7.4

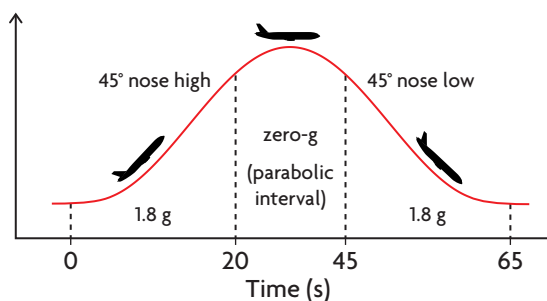
x	y	x	y
0	1	12	69
5	49	15	85
10	75	18	65
15	80	21	82
20	78	24	80
25	83	27	100
30	100	30	105
35	150	33	110
40	200	36	150
50	400	39	190
3	15	42	220
6	55	48	400
9	70	51	450

Closing

10. a) Create a scatter plot for the data in the table at the left.
- b) Determine the equation of the cubic regression function that models the data.
- c) Use your equation to determine the y -value when x is 25. How close is your calculated value to the y -value in the table?

Extending

11. To prepare astronauts for the experience of weightlessness, NASA uses versions of an airplane officially known as the Weightless Wonder (and unofficially known as the Vomit Comet). To produce the effect of weightlessness, the aircraft ascends at a 45° angle, transitions into a parabolic flight path, and finishes by descending at a 45° angle. The feeling of weightlessness occurs only during the parabolic interval. For example, during this flight, the feeling of weightlessness lasted for 25 s.



The table below shows the time, t , and the altitude, h , from a different Weightless Wonder flight.

t (s)	h (m)	t (s)	h (m)	t (s)	h (m)	t (s)	h (m)
0	7000	8	7926	16	8501	24	7907
1	7116	9	8043	17	8490	25	7791
2	7232	10	8159	18	8459	26	7675
3	7347	11	8265	19	8408	27	7559
4	7463	12	8351	20	8338	28	7443
5	7579	13	8418	21	8248	29	7327
6	7695	14	8465	22	8139	30	7212
7	7811	15	8493	23	8023		

- a) Model the data with a combination of linear and quadratic functions using the appropriate types of regression.
- b) Use your regression models to determine the length of time that the feeling of weightlessness occurred during this flight.