1.4 Trends Using Technology

Visual displays of data taken from a table help not only the identification of trends, but also the drawing of valid conclusions from the information. By examining a scatter plot, for example, you can see whether the relationship between two variables is strong or weak, positive or negative.

In the past, health researchers used treadmills and bicycle ergometers to measure the exercise capacity of patients with cardiac and respiratory illnesses. As this equipment was not always available to everyone, investigators began to use a simpler test more related to day-to-day activity. The simpler test had patients cover as much ground as possible in a specified amount of time by walking in an enclosed corridor. The strength of the relationship between the two tests is important because if the results are strongly related, then one test can be substituted for the other. The strength of the relationship could show how well laboratory tests can predict a patient's ability to undertake physically demanding activities associated with daily living.

You are familiar with scatter plots and finding a line of best fit. Often, a set of data is best represented by a curve of best fit. In this section, you will investigate mathematical tools that will allow you to evaluate the strength of any conclusions drawn from a data set. These tools will give you more confidence in analyzing data and describing trends, which are important mathematical skills.

regression—the process of fitting a line or curve to a set of data

Think about Sanjev's Observations

Sanjev noticed that the turning point for tuition increases was around 1990. What other observations could be added to the list?

Project Connection

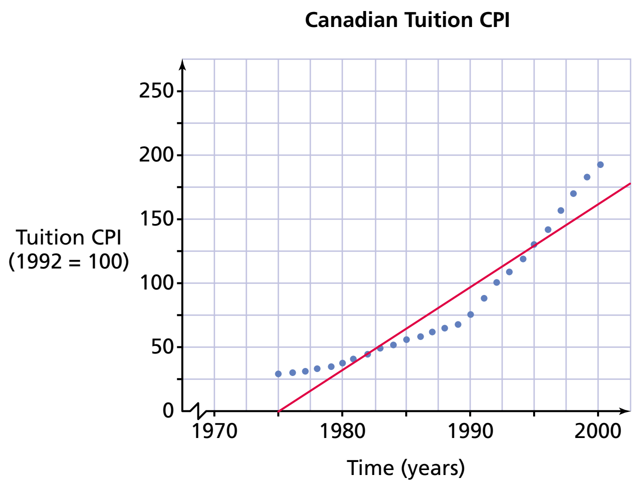
When displaying your data in a graph, consider the type of graph that best displays the data.

Example 1 Finding the Best Model to Fit the Data

Sanjev is planning to enter college or university in 2005. He creates a scatter plot to help him predict the relative costs of tuition fees and must decide which model of **regression** (linear or quadratic) best fits the data.

Solution

The scatter plot with the line of best fit is shown to the right. Extend the line to the year 2005 to predict the relative cost of tuition fees by reading from the graph or by using the equation of the line.



Source: Statistics Canada

The scatter plot shown represents the CPI for tuition fees relative to 1992 levels. For example, the point (1995,130) means the average tuition in the year 1995 was 130% of the 1992 average, or 1.3 times 1992 levels.

Think about The Type of Graph and a Line of Best Fit

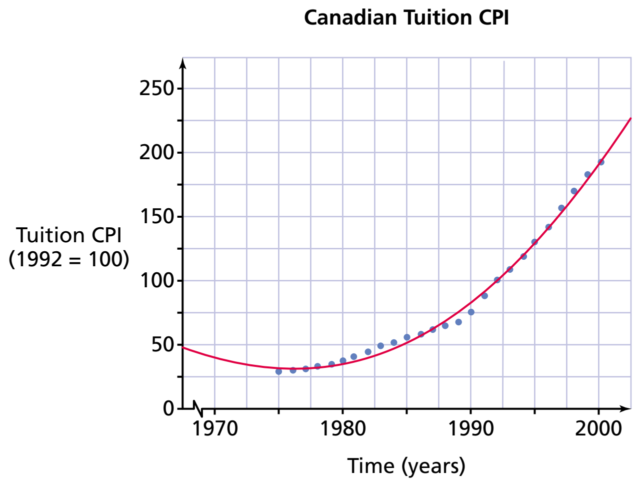
Why is a scatter plot an appropriate way to display the data? How would you decide where to draw a line of best fit?

Rounded to two decimal places, the linear model gives the equation $y = 6.35x - 12\,534.45$. Substituting $x = 2005$ into the equation gives the following:

$$\begin{aligned} y &= 6.35x - 12\,534.45 \\ &= 6.35(2005) - 12\,534.45 \\ &= 197.3 \end{aligned}$$

Sanjev observes that the relative cost of 197.3 in 2005 from his line is about the same as the actual relative cost in the year 2000. Since the cost of tuition is always rising, this value does not seem reasonable to him. Another model is needed.

He notices that there is a curved pattern in the data, so he tries a quadratic regression model.



Source: Statistics Canada

The quadratic model gives the equation

$$y = 0.306\,916\,971\,9x^2 - 1213.646\,792x + 1\,199\,818.279$$

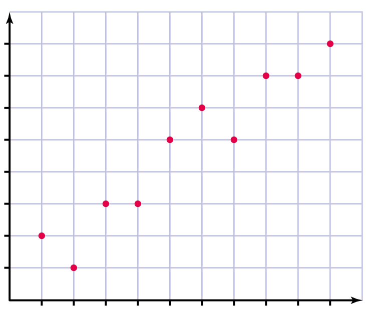
Substituting $x = 2005$ into the equation gives the following:

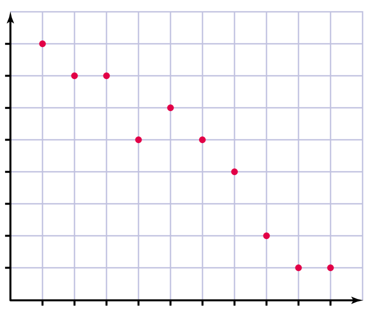
$$\begin{aligned} y &= 0.306\,916\,971\,9x^2 - 1213.646\,792x + 1\,199\,818.279 \\ &= 0.306\,916\,971\,9(2005)^2 - 1213.646\,792(2005) + 1\,199\,818.279 \\ &= 270.36 \end{aligned}$$

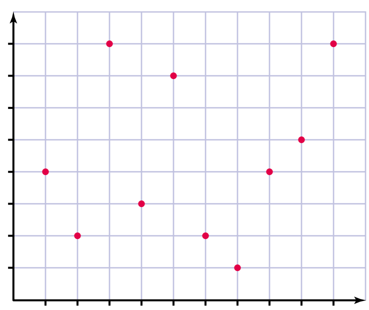
Sanjev concludes that the quadratic model is a better fit than is the linear model since the curve passes through or near the majority of data points. The model suggests that, in 2005, tuition fees will be 2.7 times the 1992 levels. Based on the data, this seems reasonable.

SCATTER PLOTS AND CORRELATION

The following three scatter plots show different types and strengths of correlation.


Strong Positive Correlation


Strong Negative Correlation


No Correlation

When two variables increase in the same proportion and simultaneously, they have a positive correlation. If one variable increases in the same proportion as the other decreases, they have a negative correlation.

INVESTIGATION: THE CORRELATION COEFFICIENT

Purpose

To investigate the correlation coefficient between two variables using technology.

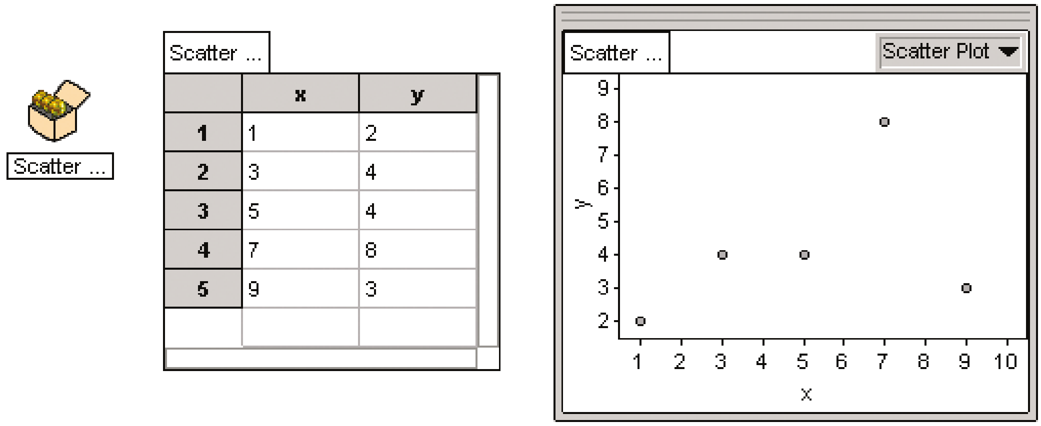
Procedure—Using Fathom™

Technolink For more information about creating scatter plots in Fathom™, see Appendix D.6 on page 421.

A. Enter the data shown below into a new case table, labelling the attributes *x* and *y*.

x	1	3	5	7	9
y	2	4	4	8	3

B. Drop a new graph into the workspace, and drag *x* to the horizontal axis and *y* to the vertical axis to create a scatter plot.



C. Drag the points so that the value of r^2 is as large as possible. Record the greatest value of r.

D. Adjust the points so that the value of r^2 is as small as possible. Record the smallest value of r.

E. Adjust the points so that r^2 is 0. Do this in as many ways as possible.

F. Drag the points so that they are on top of each other.

Discussion Questions

1. What is the largest value of r^2?
2. Describe the relationship between the points when r^2 is a maximum.
3. What is the smallest value of r^2?
4. Describe the relationship between the points when r^2 is a minimum.
5. What happens when you try to drag the points on top of each other? Zoom in on the points to help explain the results.

THE LINE OF BEST FIT: LINEAR REGRESSION

If a data set shows a linear correlation, a line of best fit is drawn to model the data.

Example 2 Using a TI-83 Plus Calculator to Draw a Line of Best Fit

One of the science classes is growing a bean plant. The height of the plant is measured every few days. The results are randomly collected in the table below.

Day	0	10	8	13	9	11	14
Height (cm)	1	12	7	14	10	11	13

(a) Find the line of best fit.

(b) How confident are you in your answer to part (a)?

(c) Use the equation to predict the height of the bean plant at 20 days.

Technolink See Appendix C starting on page 401 for instructions on entering data into lists, creating a scatter plot, and graphing a line of best fit.

If your calculator does not display r and r^2, refer to page 402 of Appendix C and activate **Diagnostic On**.

coefficient of correlation—a number from +1 to −1 that gives the strength and direction of the relationship between two variables

coefficient of determination—a number from 0 to +1 that gives the relative strength of the relationship between two variables. (If r^2=0.44, this means that 44% of the variation of the dependent variable is due to variation in the independent variable.)

Solution

(a) Enter the data into $\mathbf{L_1}$ and $\mathbf{L_2}$ using Stat Editor.

To construct a line of best fit, you must choose the type of line/curve you think will best fit the data, where the data are located, and where you would like the equation of the line to be written.

The slope and y-intercept of the line are given, as well as two measures of correlation: r, the **coefficient of correlation**, and r^2, the **coefficient of determination**. Plot the points and the line of best fit.

(b) You can be confident in the conclusion that the relationship is linear because the r and r^2 values are very close to 1. Also, the line of best fit appears to fit the data closely.

(c) Using the **CALC** feature on a TI-83 Plus calculator, select **1:value**, hit [ENTER], and enter 20 for the x-value. The y-value returned is 19.9. The plant would be approximately 19.9 cm in height, provided that it continued to grow at the rate demonstrated by the data.

CORRELATION COEFFICIENT

The correlation coefficient, r, is an indicator of both the strength and direction of a linear relationship. A value of $r = 0$ indicates no correlation, while $r = \pm 1$ indicates perfect positive or negative correlation.

The coefficient of determination, r^2, does not give the direction of correlation, but does make the scale constant. A value of $r^2 = 0.4$ indicates that 40% of the variation in y is due to the variation in x.

Example 3 Using Fathom™ to Draw the Line of Best Fit

The manager of a service station changes the price of unleaded gasoline and records the amount of gas sold per hour at each price. The results are shown in the table below.

Technolink
For instructions on creating a collection, a case table, a scatter plot, and the least-squares line using Fathom™, refer to pages 416, 417, and 421 of Appendix D.

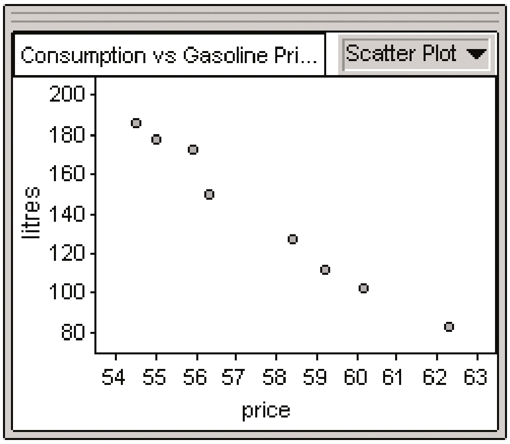
Price (¢/L)	54.5	55.0	55.9	56.3	58.4	59.2	60.2	62.3
Amount Sold (L/h)	186	178	172	150	127	112	102	83

(a) Is the relationship linear? Give reasons for your answer.

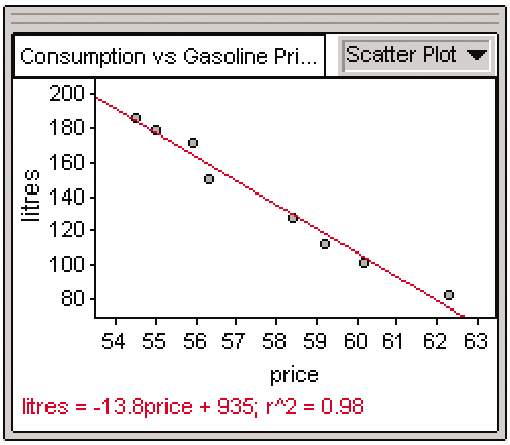
(b) How many litres of gasoline would be sold if the price were \$0.57/litre?

Solution

(a) Using Fathom™, start a new worksheet and create a collection with two attributes: **price** and **litres**. Enter the data from the table. Drag a new graph onto the worksheet and drag the attribute names onto the appropriate axes.



Select the graph and choose **Least-Squares Line** under the **Graph** menu.



Based on the r^2 value, the data are linear. The line shown on the graph appears accurate.

(b) Substitute 57 for the *price* in the equation

$$\begin{aligned} litres &= -13.8price + 935 \\ &= -13.8(57) + 935 \\ &= 148.4 \text{ L} \end{aligned}$$

Example 4 Using a Spreadsheet to Draw a Line of Best Fit

For each player, the basketball coach keeps track of the amount of time played (in minutes) and the number of points scored. The results are shown below.

Name	Minutes Played	Points Scored
Fred	151	50
Tony	19	11
Samir	164	38
Jared	87	28
Anuj	135	39
Matt	111	39
Steve	54	8
Ali	163	61
Darcy	192	52
Ryan	98	33
Travis	71	26

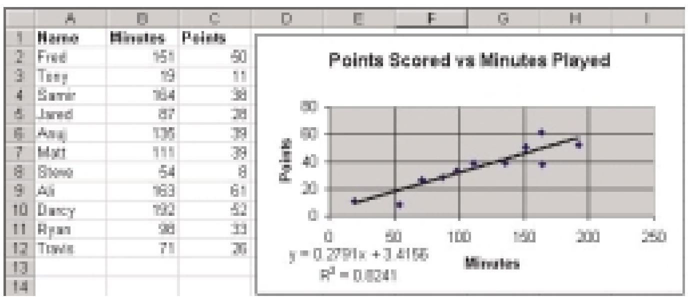
Technolink
See Appendix E.I on page 425 for instructions on finding lines of best fit.

(a) Create a scatter plot and find a line of best fit.

(b) If Jude scored 45 points, how many minutes did he play?

Solution

(a) Open a new spreadsheet. Enter the titles and data in columns A, B, and C.



Select the entries in columns B and C, select **insert chart**, select **XY(Scatter)**. Column B contains the independent data and column C contains the dependent data.

To add the line of best fit, the equation of the line of best fit, and the value of r^2, select **add trendline** from the **chart** menu. Be sure that a checkmark appears in the box before **Display equation on chart** and **Display R squared value on chart**.

(b) Substitute $y = 45$ and solve for x. This gives a value of 149, so Jude played 45 min.

RESIDUALS

Although the *r*-value indicates the strength and direction of a linear relationship, a lower *r*-value does not necessarily mean that the linear model should be rejected. Another method of analyzing data is also useful. This involves analyzing the distance the data points are from the line of best fit.

residual value (residual)—the vertical distance between a data point and the line of best fit

The vertical distance between a data point and the line of best fit is called the **residual value** (or **residual**). It may be calculated for a single point (x_1, y_1) by subtracting the calculated value from the actual value

$$R_1 = y_1 - [a(x_1) + b]$$

where a and b are the slope and intercept of the line of best fit, respectively.

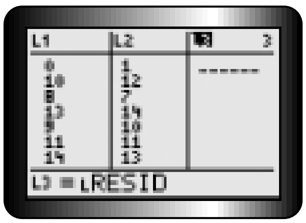
The residuals should be graphed. If the model is a good fit, the residuals should be fairly small, and there should be no noticeable pattern. Large residuals or a noticeable pattern are indicators that another model may be more appropriate. If only a few pieces of data cause large residuals, you may wish to disregard them.

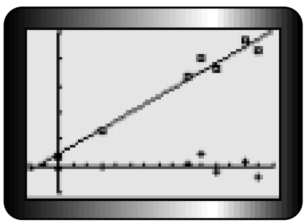
Example 5 Using a TI-83 Plus Calculator to Determine the Residuals

Using the data in Example 2 and the line of best fit you created, graph the residuals.

Solution

Technolink
For instructions on creating a residual plot using a TI-83 Plus calculator, refer to Appendix C.10 on page 407.





When the calculator computes a line of best fit, it also computes the residuals and stores these in a list called **RESID**. Copy these values into $\mathbf{L_3}$ for easy comparison. Press [STAT] **Edit** and move the cursor onto list $\mathbf{L_3}$. When you press [ENTER], the blinking cursor moves to the command line at the bottom of the screen. You can then insert the list name **RESID** by pressing [2nd] [LIST] and selecting the list name **RESID**.

When you press [ENTER], the residual values are copied into $\mathbf{L_3}$. Set up **STAT PLOT 2** to plot the values of the residuals for each point.

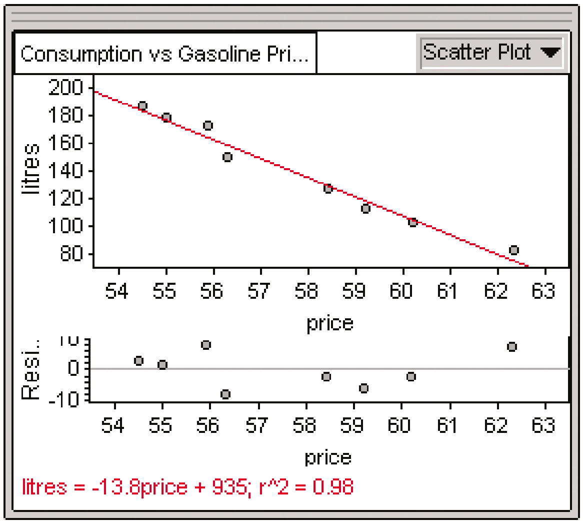
The residuals are small and show both positive and negative values. This means that the line of best fit chosen is a good fit.

Example 6 Using Fathom™ to Determine the Residuals

Use the gasoline data given in Example 3 and create a residual plot.

Solution

Select the graph and choose **Make Residual Plot** under the **Graph** menu. The residuals are small and show both positive and negative values. This means that the line of best fit chosen is a good fit.



Technolink
For detailed instructions on creating a residual plot using Fathom™, refer to Appendix D.7 on page 422.

Example 7 Ontario Youth Crime on the Rise?

When you read the newspaper or watch the news on television, you sometimes hear something like:

> *…the Young Offenders Act is not tough enough. Crimes committed by today's youth are out of control and something needs to be done.*

What is the current rate and general trend of youth crime in Ontario? What model might be appropriate to display trends and allow you to make reasonable predictions for the future?

Solution

The data table to the right displays the number of young males and females from Ontario who have committed crimes between the years 1983 and 1998.

L1	L2	L3	L4
1983	13146	2690	15836
1984	12156	2671	14827
1985	24233	5227	29460
1986	30406	6816	37222
1987	32026	7352	39378
1988	31762	7471	39233
1989	34376	8331	42707
1990	36662	9214	45876
1991	41567	10980	52547
1992	39812	11698	51510
1993	37796	11018	48814
1994	36762	10138	46900
1995	37608	10849	48457
1996	36221	11045	47266
1997	33482	10801	44283
1998	33026	10276	43302

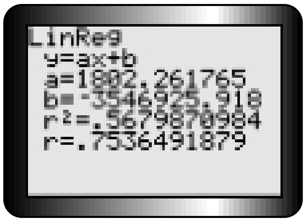
Source: Statistics Canada

Note
The headings L1, L2, L3, and L4 represent Year, Number of Males, Number of Females, and Total Number of Males and Females who committed crimes between 1983 and 1998, respectively.

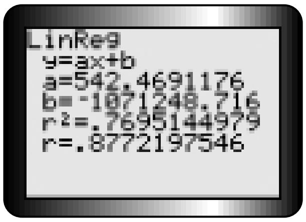
Technolink
The electronic version of the data is on the textbook CD and can be loaded into a TI-83 Plus calculator.

Technolink
For additional help creating scatter plots and using regression to determine an algebraic model, refer to page 404 of Appendix C.

Totals



Males

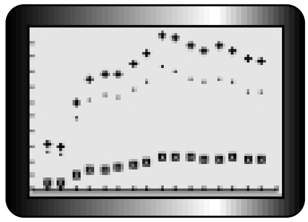


Females

Technolink
To perform a quadratic regression on a TI-83 Plus calculator, refer to Appendix C.3 on page 403.

The following scatter plots were created on a TI-83 Plus calculator.

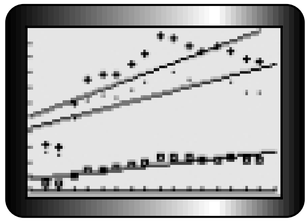
\+ STAT PLOT 1 is L1 and L4.
• STAT PLOT 2 is L1 and L2.
□ STAT PLOT 3 is L1 and L3.



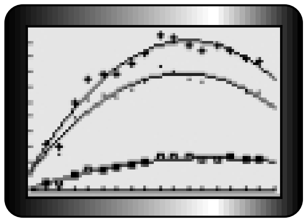
Linear regression formulas are shown in the margin for the three series of data. The table and the scatter plots show that

- except for 1988, there was a steady increase in the number of youth crimes from 1983 until around 1991–92
- except for 1995, the number of crimes has been slowly decreasing since 1991

A linear model for the data is shown to the right. The graphing calculator gives the linear regression information for each plot. Notice that the correlation coefficient, r, is closer to 1 for Males than for the other two groups. The graphs show that the linear model is reasonable for the relationship pertaining to Males. The linear models for Females and Totals do not capture the decreasing trend since 1992 and are, therefore, not appropriate.



A quadratic model for the same data is shown here. This models the data better than does the linear model. It captures the decreasing trend in the number of youth crimes in Ontario since 1992, and allows you to predict the number of crimes for each gender, provided that the trend continues.



KEY IDEAS

predictive models—linear and quadratic—a linear model is based on the relationship $y = mx + b$, and a quadratic model is based on the relationship $y = ax^2 + bx + c$, where x is the independent variable and y is the dependent variable

coefficient of correlation, r—a number from $+1$ to -1 that gives the strength and direction of the relationship between two variables

coefficient of determination, r^2—a number from 0 to $+1$ that gives the relative strength of the relationship between two variables. (If $r^2=0.44$, this means that 44% of the variation of the dependent variable is due to variation in the independent variable.)

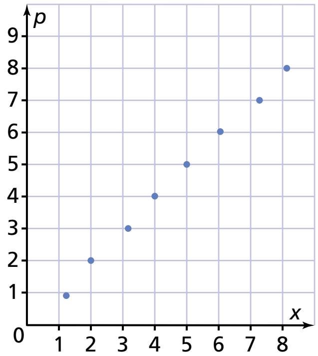
residual value (residual)—the vertical distance between a data point and the line of best fit

1.4 Exercises

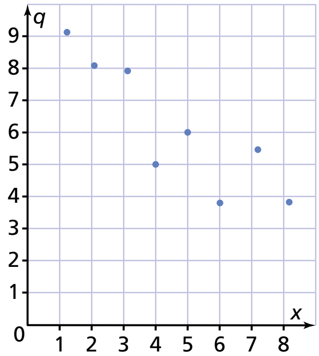
A

1. **Knowledge and Understanding** Classify these scatter plots as having a correlation that is positive or negative, and describe the strength.

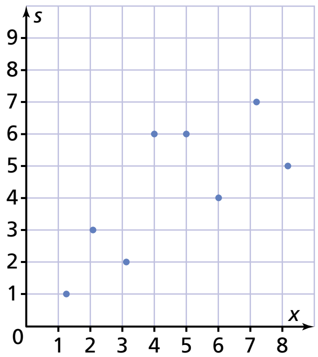
(a)



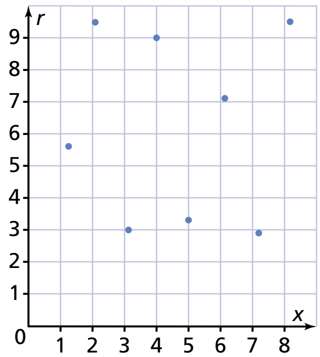
(b)



(c)



(d)



2. **Application** Do the following situations describe relationships with positive or negative correlations? Give reasons for your answers.
   - **(a)** As you get older, the number of years until retirement changes.
   - **(b)** The taller you are, the longer your arms.
   - **(c)** The farther you drive, the less gas you have.
   - **(d)** The more you study, the higher your marks.

B

3. Use technology to create scatter plots and lines of best fit.

(a)

Age (years)	Height (cm)
0	20
1	40
2	65
3	80
4	92
5	108

(b)

Speed (km/h)	Time (min)
10	60
20	30
30	20
40	15
60	10

(c)

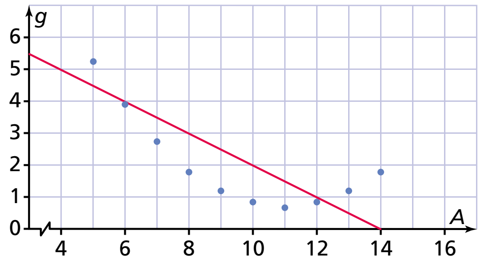
Study Time (h)	Mark
0	50
1	62
2	68
3	72
4	74
5	75

4. Use technology to calculate the correlation coefficient for each of the data sets in Question 3. Tell how each describes the strength of the relationship.

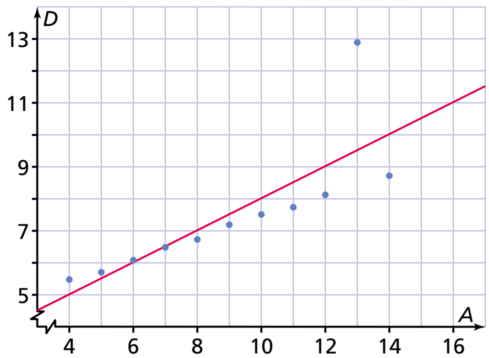
5. Create a residual plot for each of the data sets in Question 3. How do the residual plots enable you to determine the strength of each relationship?

6. Sketch the residual plot for each of the following.

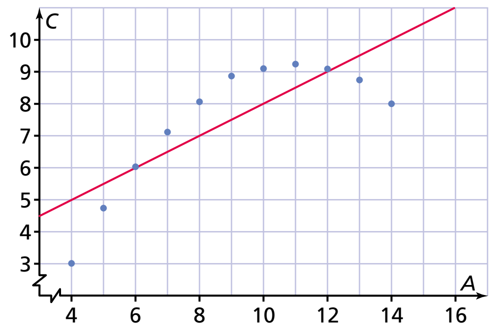
(a)



(b)

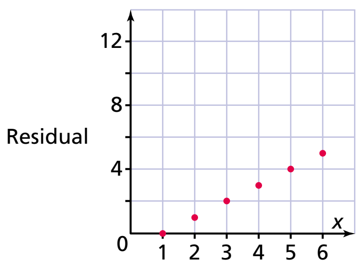


(c)

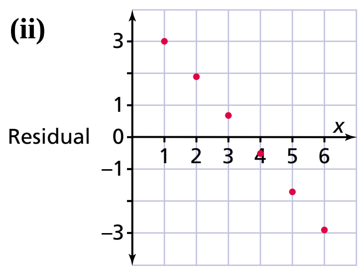


7. Consider the following residual plots.

(i)



(ii)



(a) Sketch a graph for each of the residual plots.
(b) What do the residual plots tell you about the line of best fit?
(c) Describe the residual plot for a perfect line of best fit.
(d) Based on your responses in parts (a) and (b), describe the general characteristics of a residual plot for a line of best fit that accurately models the data.

8. Given the following coefficients of correlation, state what percent of variation in y is due to the variation in x.

(a) $r = 0.5$	(b) $r = 0.85$	(c) $r = 0.66$
(d) $r = 0.9$	(e) $r = -0.1$	(f) $r = 0.32$

9. Roll a pair of dice 10 times and record the values shown on the first die, x, and on the second die, y. It is easier to keep track of the values if the dice are different colours.

(a) Make a scatter plot of the data. The data consists of random pairs of values (x, y). What would you expect the correlation coefficient to be? Give reasons for your answer.

(b) Use technology to calculate the correlation coefficient. Compare your value to those generated by other students in your class. What percent of the students had a correlation coefficient greater than 0.5?

10. **Communication** Correlation does not necessarily prove causation. For example, amount of study time and marks may have a positive correlation, but other factors besides study time may cause higher marks. Before doing an experiment and evaluating the correlation coefficient, you must examine the possibility of cause and effect. Examine the scenarios below and decide whether a cause-and-effect relationship exists. Give reasons for your answers.

(a) A medical study tracks the use of aspirin and the incidence of heart attacks. A weak negative correlation is found between the amount taken and the likelihood of having an attack. Does it appear that aspirin is preventing heart attacks?

(b) Consider the hypothesis that severe illness is caused by depression and anger. It has been observed that people who are ill are very often depressed and angry. Thus, it follows that the cause of severe illness is depression and anger. So, a good and cheerful attitude is the key to staying healthy.

(c) Over the course of several weeks, the needles from the pine trees along the Wombat River fell into the water. Shortly thereafter, many dead fish washed up on the riverbank. When environmental officials investigated, the owners of Wombat River Chemical Company claimed that the pine needles had killed the fish.

11. Complete the activities that follow using the data on the textbook CD that show the changes in Canada's population.

(a) Describe the trends in the data.

(b) Find the model that best represents the data.

(c) Use your model to predict what the population of Canada would have been in 1998, provided that the population growth is consistent. Check your answers with data from the Statistics Canada Web site.

(d) Use your model to predict what the population would have been for the most recent year for which data is available, provided that the population growth had continued in the same manner. Check your answers with data from the Statistics Canada Web site.

12. The winning distance for the Olympic men's discus event is provided below.

Year	Distance (m)	Year	Distance (m)	Year	Distance (m)
1896	29.15	1932	49.50	1972	64.39
1900	36.04	1936	50.48	1976	67.50
1904	39.28	1948	52.78	1980	66.65
1908	40.89	1952	55.04	1984	66.60
1912	45.21	1956	56.35	1988	68.82
1920	44.68	1960	59.18	1992	65.12
1924	46.15	1964	61.00	1996	69.40
1928	47.32	1968	64.78	2000	69.30

Source: British Broadcasting Corporation (BBC)

(a) Describe any trends in the data. Give reasons for your answer.
(b) Create a graph to display the data.

13. The population, births, deaths, and infant deaths for Canada for the years 1948 to 1972 are listed in the table below.

Year	Population	Births	Deaths	Infant Deaths
1948	13 167 000	359 860	122 974	15 965
1949	13 475 000	367 092	124 567	15 935
1950	13 737 000	372 009	124 220	15 441
1951	14 050 000	381 092	125 823	14 673
1952	14 496 000	403 559	126 385	15 408
1953	14 886 000	417 884	127 791	14 859
1954	15 330 000	436 198	124 855	13 934
1955	15 736 000	442 937	128 476	13 884
1956	16 123 000	450 739	131 961	14 399
1957	16 677 000	469 093	136 579	14 517
1958	17 120 000	470 118	135 201	14 178
1959	17 522 000	479 275	139 913	13 595
1960	17 909 000	478 551	139 693	13 077

Year	Population	Births	Deaths	Infant Deaths
1961	18 238 000	475 700	140 985	12 940
1962	18 614 000	469 693	143 699	12 941
1963	18 964 000	465 767	147 367	12 270
1964	19 325 000	452 915	145 850	11 169
1965	19 678 000	418 595	148 939	9 862
1966	20 048 000	387 710	149 863	8 960
1967	20 412 000	370 894	150 283	8 151
1968	20 729 000	364 310	153 196	7 583
1969	21 028 000	369 647	154 477	7 149
1970	21 297 100	371 988	155 961	7 001
1971	22 962 082	362 187	157 272	6 356
1972	22 219 560	347 319	162 413	5 938

Source: Data have been extracted from Fathom Dynamic Statistics™, Key Curriculum Press.

(a) Find a model that best fits the data given. Use it to predict the population, live births, deaths, and infant deaths for 1973.

(b) Consider the following data from 1974 to 1997. If these data were included in the table, how would your model change?

Year	Population	Births	Deaths	Infant Deaths
1974	22 808 446	345 645	166 794	5192
1975	23 142 275	348 110	167 404	5130
1976	23 449 793	348 857	167 009	4847
1977	23 726 345	361 400	167 498	4475
1978	23 963 967	358 852	168 179	4289
1979	24 202 205	366 064	168 183	3994
1980	24 516 278	370 709	171 473	3868
1981	24 820 382	371 346	171 029	3562
1982	25 117 424	373 082	174 413	3385
1983	25 366 965	373 689	174 484	3182
1984	25 607 555	377 031	175 727	3058
1985	25 842 590	375 727	181 323	2982

Year	Population	Births	Deaths	Infant Deaths
1986	26 100 587	372 431	184 224	2938
1987	26 449 888	369 441	184 953	2706
1988	26 798 303	375 743	190 011	2705
1989	27 286 239	391 925	190 965	2795
1990	27 700 856	404 669	191 973	2766
1991	28 030 864	411 910	196 050	2677
1992	28 376 550	398 642	196 535	2432
1993	28 703 142	388 394	204 912	2448
1994	29 035 981	385 112	207 077	2418
1995	29 353 854	378 011	210 733	2321
1996	29 671 892	364 732	213 649	2042
1997	30 003 955	361 785	216 970	1925

Source: Data have been extracted from Fathom Dynamic Statistics™, Key Curriculum Press.

(c) Using the data from 1974 to 1997, find the equation(s) you would use to predict the live births in the future, provided the trend continues.

C **14.** **Thinking, Inquiry, Problem Solving**

(a) Find some data containing two or three sets that show trends relating to the general direction of growth or decline.

(b) Describe the rate of the trend and state whether or not the trend shows a change that is steady or erratic. Give reasons for your answer.

(c) Provide a model that could be used for predictions.

(d) Explain why your model is accurate.

ADDITIONAL ACHIEVEMENT CHART QUESTIONS

15. **Knowledge and Understanding** Use the graphing technology of your choice to create a scatter plot, the line of best fit, and the residual plots for men's and women's data shown below.

Retail Merchandising in Canada

Year	1993	1994	1995	1996	1997	1998
Men's clothing stores ($ millions)	1730	1687	1623	1516	1570	1582
Women's clothing stores ($ millions)	4303	4812	5377	5522	5830	6259

Source: Statistics Canada

16. **Application** The table below shows the value of RRSP holdings of Canadians (in millions of dollars) from 1983 to 1997. Determine whether a linear or a quadratic model more accurately fits the data, and then use your model to predict the RRSP holdings in 2010.

Year	1983	1984	1985	1986	1987
RRSP ($ millions)	70 736	76 292	85 084	92 916	102 660

Year	1988	1989	1990	1991	1992
RRSP ($ millions)	115 884	116 356	132 316	143 704	158 212

Year	1993	1994	1995	1996	1997
RRSP ($ millions)	163 548	171 468	179 032	177 220	183 832

17. **Thinking, Inquiry, Problem Solving**

(a) Search the Statistics Canada Web site to find a set of data that shows a strong negative correlation.

(b) Verify this using appropriate models and technology.

(c) Describe the behaviour shown between the variables and investigate the reasons for the behaviour.

18. **Communication**

(a) What is a trend in data?

(b) What type of graph should be used when looking for trends in data? Give reasons for your answer.

(c) Describe how you can use the correlation coefficient produced by regression analysis to determine the accuracy of fit of a particular algebraic model.

Chapter Problem

Trends in Canada's Population

Use the technology of your choice to answer the following questions.

CP12. For each age class since 1951, determine whether a linear or a quadratic model best represents the data. Justify your decision with the appropriate equations and correlation coefficients.

CP13. Use your models to predict the percent of the population within each age class in the year 2000.

CP14. Use the 2000 Census data from the Statistics Canada Web site (**www.statcan.ca**) to check the accuracy of your models. Comment on any differences that you see.