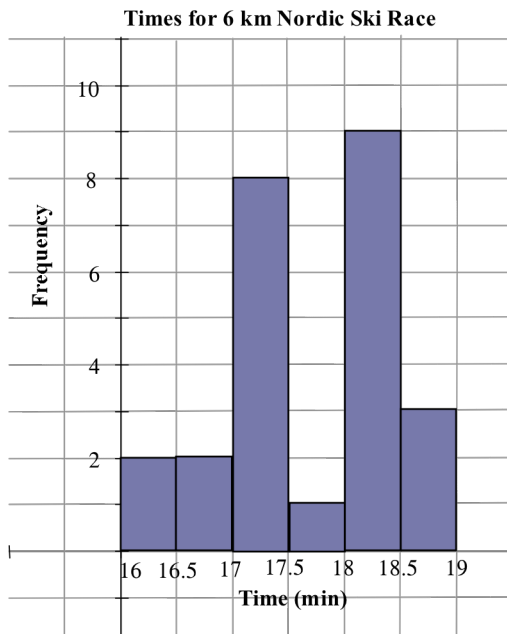


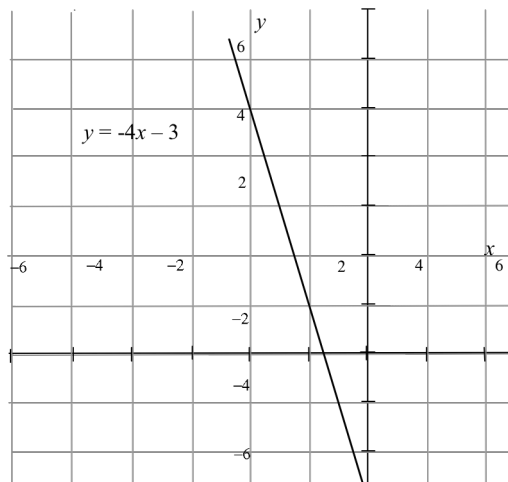
Chapter 3 Answers

Get Set, pages 37–38

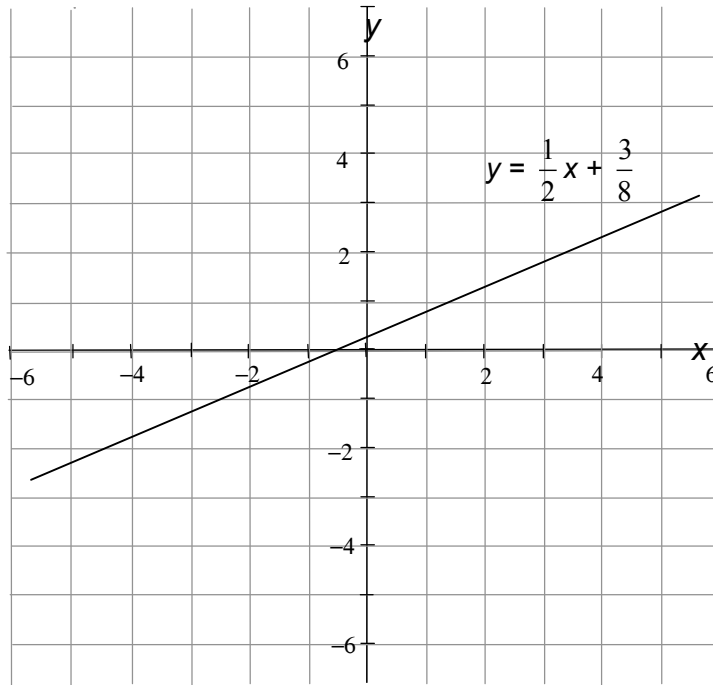
1. a) The sum of a set of data divided by the number of data.
- b) The middle value when a set of data is ordered from least to greatest.
- c) The most frequently occurring value in a set of data.
2. A population is the entire set of people or items being studied, while a sample is a selection of members of a population.
3. A random sample is a sample generated so that each member of the population has an equal chance of being selected, and that it has the same characteristics as the population. A biased sample has an unwanted influence that prevents a study from being representative of the population.
4. a) histogram b) 19 artefacts c) 51 artefacts d) 92 artefacts
5. a) line graph b) after 8 A.M. c) after 1 P.M. d) at 1 P.M.
- 6.



7. a)



b)



3.1 Two-Variable Data Sets, pages 39-42

Warm-Up

1. $-7.05, -4.02, -0.11, 0.12, 1.95, 16.5$

2. a) $9x - 4y$

b) $s + 11t$

c) $6a + 9b - 1$

3. a) $(0, -14)$

b) $-\frac{4}{5}$

c) $\frac{5}{4}$

4. a) 150.8 cm^3

b) 49.7 m^3

5. a) $\frac{1}{2}$

b) $\frac{1}{2}$

6. $5(s + x) = 30$

7. D

8. 2.135

Practise

1. a) two facts: age and height

b) one fact: age

2. a) Two-variable data. Two attributes are known about each student.

b) One-variable data. One attribute is known about each student.

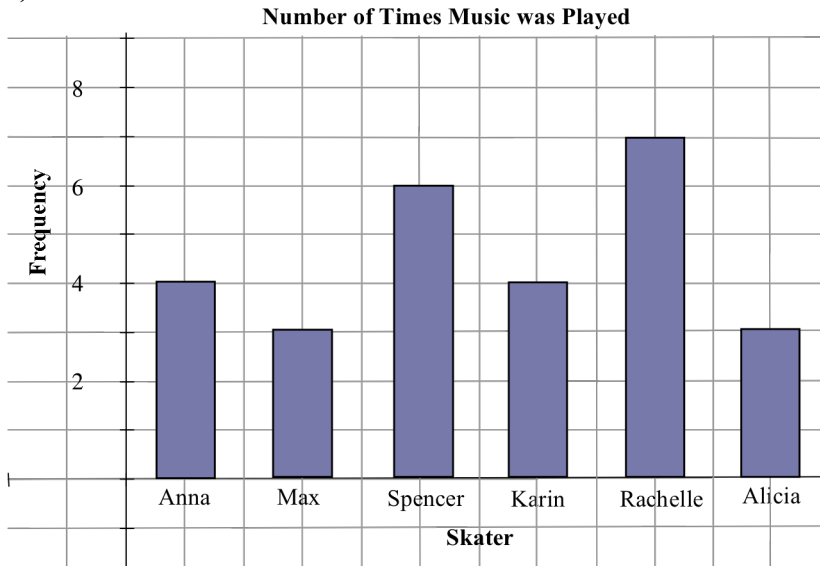
c) One-variable data. One attribute is known about each consumer.

d) Two-variable data. Two attributes are known about each trip.

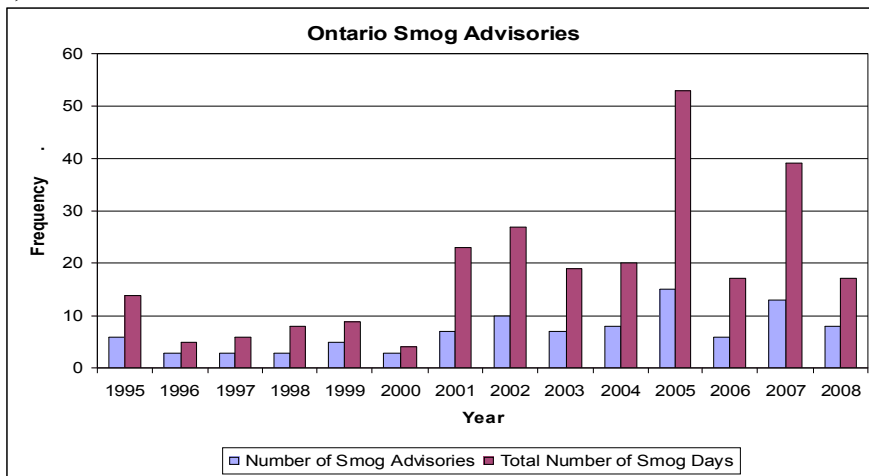
3. a)

Skater	Frequency
Anna	4
Max	3
Spencer	6
Karin	4
Rachelle	7
Alicia	3

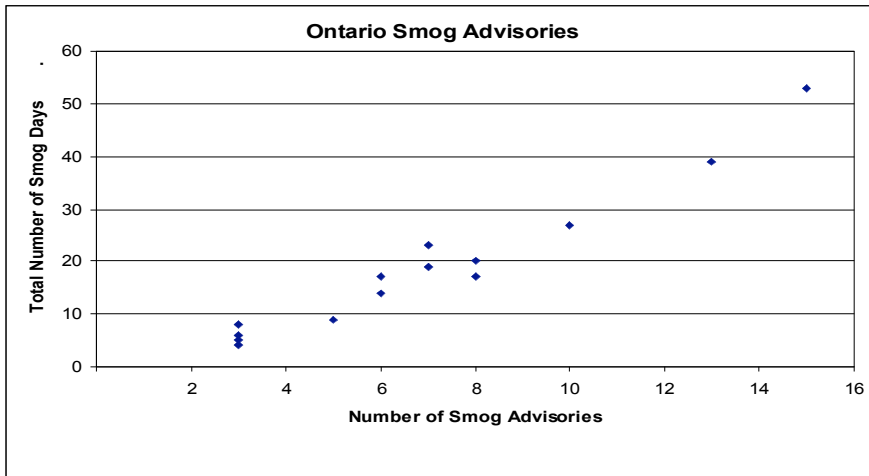
- b) The number of times each skater's music was played.
- c) One-variable data set. One attribute is known about each skater.
- d) Number of times each skater's music was played.
- e)



- 4. a) There are two facts known about each skater, so this is a two-variable data set.
- b) number of spins and number of jumps
- c) Answers may vary. For example: Karin has the highest number of spins and jumps combined. Alicia has the fewest number of spins and jumps.
- 5. a) the number of smog advisories and the total number of smog days
- b) Answers may vary. For example: Which year had the greatest number of smog advisories?
- c) Answers may vary. For example: As the number of smog advisories increase, does the total number of smog days increase?
- d)



e)



f) The scatter plot. It displays the data as a two-variable data set and shows the relationship between the two variables clearly.

3.2 Effective Surveys, pages 43-45

Warm-Up

- a) 2.0 b) 2.7 c) 40.5
- a) $4xy - 5x + 5y$ b) $-2a + 10b - 3$ c) $x + y$
- 2 roots
- a) 20 b) 8.7
- a) $\frac{1}{26}$ b) $\frac{1}{2}$
- 23, 37
- a) octagon

b) isosceles
- Two-variable data. Two attributes are recorded for each sale.

Practise

- ethics in surveys, design for honest responses, eliminate bias
- a) Design for honest responses. Options such as “Do not know” or “Other” should be included.

b) Eliminate bias. Change the order of options in questions so as not to influence responses.

c) Ethics in surveys. The introduction should inform respondents of the intentions of the survey.
- Answers may vary. For example:
 - Have you ever hired a landscaping company for your property? Yes/No
 - What is the area of property that you own? a) I do not own any property b) less than 200 m² c) 200 m² to 400 m² d) 401 m² to 600 m² e) more than 600 m²
 - On a scale of 1 to 5, how satisfied are you with your current landscaping? 1—dissatisfied 2—a little dissatisfied 3—neutral 4—a little satisfied 5—satisfied
 - Which month do you think is best for updating the landscaping of your home? _____
 - What landscaping feature would you like to add to the front of your home? Explain.
- a) The survey question is leading the respondent by suggesting a factor that would be most important, and this is a form of bias.

b) Answers may vary. For example: “What is the most important factor to consider when choosing a post-secondary school?”
- Answers may vary. For example:

a) Advantage: This is a way to meet people who cannot be reached in other ways. Disadvantage: Respondents may not be comfortable providing answers in public.

b) Advantage: Graphics and videos can be included as part of the survey. Disadvantage: Some people in a sample may not have access to the Internet, especially older people.

c) Advantage: Many people can be contacted in a short period of time. Disadvantage: Some people in a sample might not have telephone service.

d) Advantage: It allows respondents to take their time while answering the questionnaire. Disadvantage: Some respondents might ignore the questionnaire.

6. Answers may vary. For example:

a) Survey the student population at my school because it is a very large group that is growing fast in terms of purchasing power.

b) The trends in cellular telephone use among teenagers.

c) What type of cellular telephone plan do you currently have? a) None b) Pay as you go. c)

Contract d) Other (specify). _____. How much is your monthly bill? \$0–\$14 \$15–\$29 \$30–\$45 over \$45

7. Answers may vary. For example:

a) 1. What grade are you in? _____

2. What would you like the student council to do with the surplus money? a) Invest it in next year’s budget. b) Spend it on more activities for this year. c) Donate it to a local charity. d) Other (specify). _____

3. Rate how well you feel the student council has managed the budget for this year. 1—very well 2—well 3—no opinion 4—poorly 5—very poorly.

b) 1. Enrolment figures show that the student population is going to increase by 10% next year, and this means increased fees for student council. Should this year’s student council invest the surplus money in order to reduce the amount of money the fees will increase by next year?

Yes/No

2. Student council events take about four weeks to plan in order to be successful. With only three weeks left in this school year, what should student council do with the surplus money? a) Invest it for next year. b) Plan another event. c) Other (specify). _____

3. Costs of planning new events are high at this time of year. Would student council be best to invest the money and get better value for it later on? Yes/No

3.3 Collect and Organize Data, pages 46-48

Warm-Up

1. a) -4 **b)** 1

2. a) $7x^2 + 7x + 13$ **b)** $5x^2 + 3x$ **c)** $-5x - 2y$

3. $(-8, -12)$

4. 7755.5 mL

5. a) $\frac{11}{20}$ **b)** $\frac{1}{2^{40}}$

6. $A = \pi \left(\frac{d}{2} \right)^2$

7. C

8. Eliminate bias. The respondent is being led to a particular answer.

Practise

1. Answers may vary (except for part b)). For example:

a) In front of a basketball net, use masking tape to mark several toss lines on the ground, each 5 m apart. Recruit at least five participants to volunteer for the experiment. Do not tell the

participants what you are testing, and make sure none of them are watching the others while performing the experiment. One at a time, have the participants toss a basketball from the closest line and record the participant, the distance and whether the ball goes in the basket or not. After all participants have had a turn, repeat the procedure from the next closest line. Repeat until all participants have tossed the ball from each line.

b) participant, distance, whether the ball goes in the basket or not

c) basketball, basketball net, masking tape, paper and pencil

d) At least five participants, who are not highly skilled basketball players.

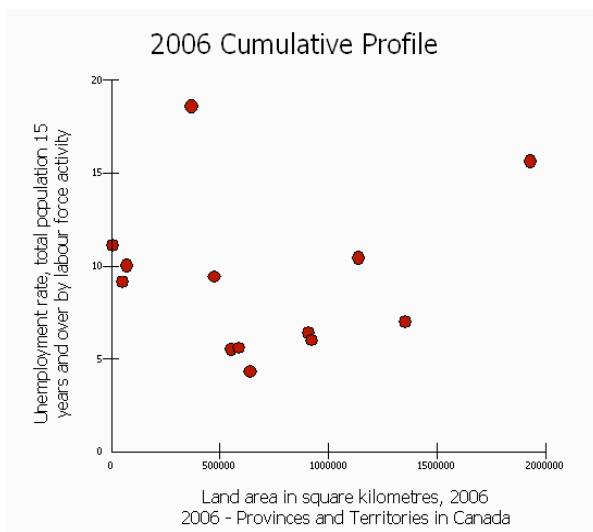
e) Make sure each participant tosses the ball from each line before moving to the next distance to avoid fatigue. Do not allow the participants to watch the performance of the others to avoid influencing their tossing strategy. Use the same ball and the same net for all participants.

f)

Participant	Distance from Basket (m)	Did the ball go in the basket? (Y/N)	Accuracy (%)

g) I would analyse the data by examining how many baskets were scored at each distance. I would present the data using a bar graph to determine how distance affects the accuracy of all participants combined. To analyse the data for the accuracy of each player, I would examine the distance versus the number of baskets at each distance and present the data using a histogram.

2. a)



Source: Statistics Canada, E-STAT Table, Cumulative Profile, 2006 - Provinces and Territories in Canada, 2006 Census of Population (Provinces, Census Divisions, Municipalities) (database), Using E-STAT (distributor).

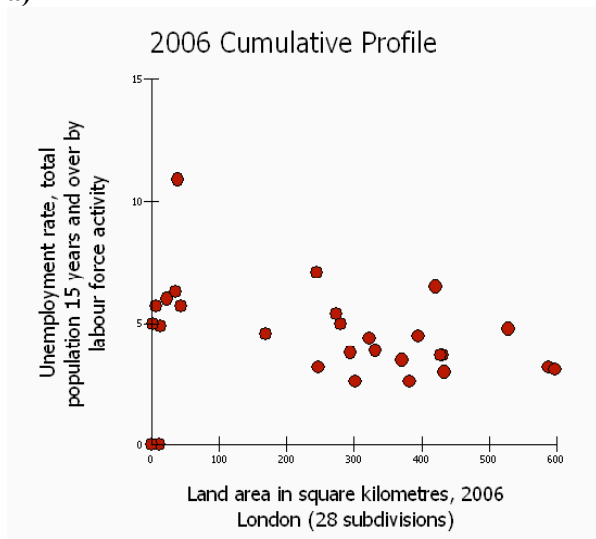
b) This is secondary data.

c) Outlier for Nunavut, at point approximately (2 000 000, 15.6).

d) For provinces or territories with a small land area, the unemployment rate is approximately 10%. But overall, there appears to be no relationship between the land area of a province or territory and its unemployment rate. The unemployment rate appears to be randomly spread between approximately 5% and 10%.

3. Answers may vary (except for part b)). For example:

a)



Source: Statistics Canada, E-STAT Table, Cumulative Profile, 2006 - Provinces and Territories in Canada, 2006 Census of Population (Provinces, Census Divisions, Municipalities) (database), Using E-STAT (distributor).

b) This is secondary data.

c) There are no outliers.

d) There is no trend between land area of regions in London, Ontario, and unemployment rate. The unemployment rate is randomly scattered between approximately 3% and 7% for all land areas.

e) This is the same pattern as in question 2 for the provinces and territories.

3.4 The Line of Best Fit, pages 49-52

Warm-Up

1. a) 14

b) $\frac{1}{4}$

2. a) $\frac{3}{2}$

b) 3

c) $-\frac{1}{2}$

3. vertical translation down by 3 units

4. 4.3 m

5. mean: 64.8; median: 70; no mode

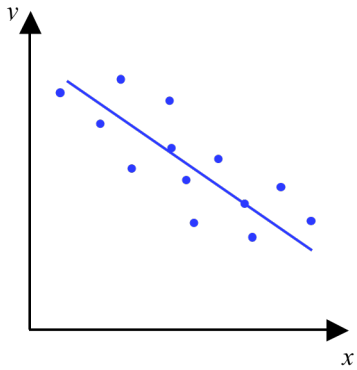
6. Let x be Alan's age and let y be his sister's age. Then, the equations to model this situation are: $2x = y + 10$ and $x + y = 40$.

7. B

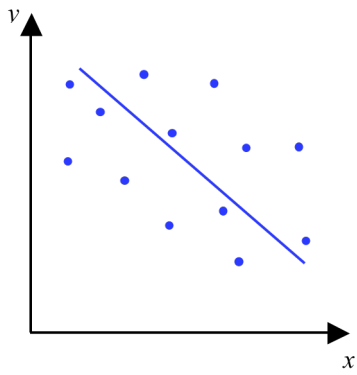
8. An outlier is an extreme value in a set of data that is separated from the main body of the data and does not follow the pattern of the other values in the set of data.

Practise

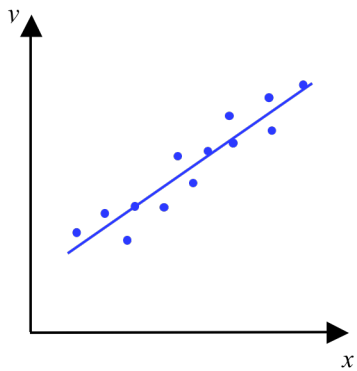
1. a)



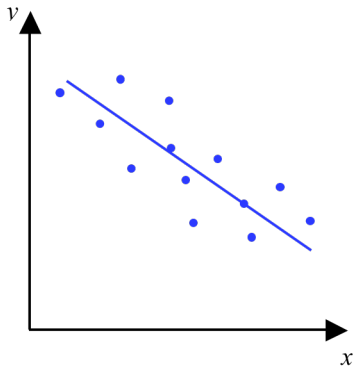
b)



c)

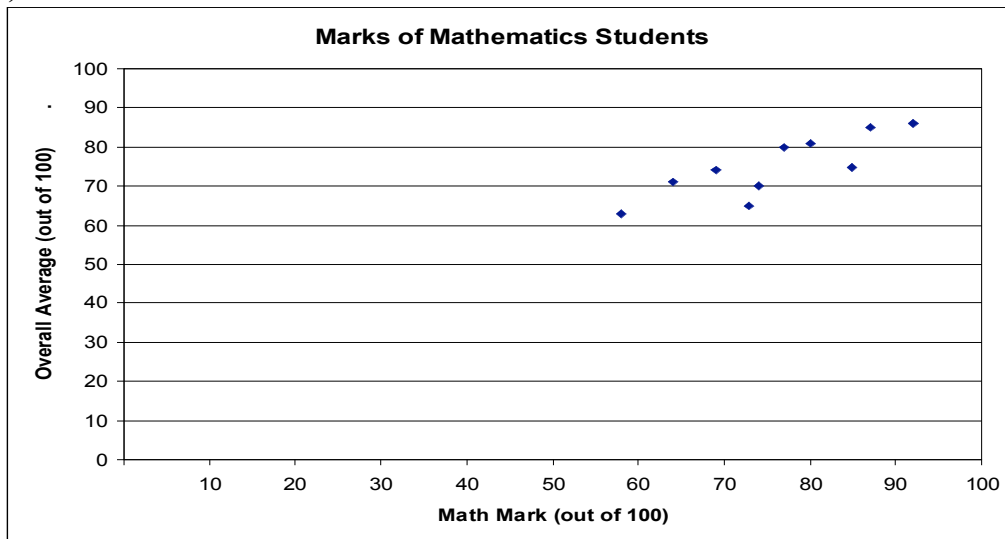


d)



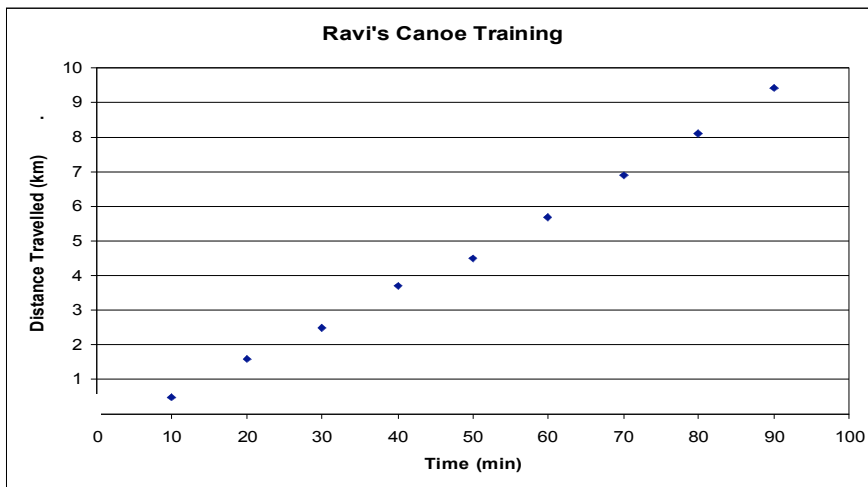
2. a) Positive. With a higher income there will be more money to give to charitable organizations.
b) Zero. There is no relationship between IQ and muscle flexibility.
c) Positive. More time spent preparing and writing the essay will lead to a better mark.
d) Negative. The higher the price of gasoline the fewer the litres will be sold.

3. a)



- b) There is a positive relationship between the math mark and the overall average.
c) Overall Average = $0.6229 \times (\text{Math Mark}) + 27.718$
d) 84
e) The correlation coefficient is $r = 0.83$, so the value of r^2 is approximately 0.69. This means that predictions made with this model are accurate approximately 69% of the time.

4. a)



b) $d = 0.1098t - 0.725$; slope: 0.1098; vertical intercept: -0.725

c) The value of the correlation coefficient is $r = 0.9986$, so the regression equation fits the data very well.

d) 6.4 km

e) No. That would be extrapolating too far beyond the data set. Other factors such as fatigue would cause the relationship to become non-linear.

3.5 Analysis and Conclusions, pages 53-55

Warm-Up

1. a) 34

b) 6

2. a) -2

b) -8

c) 0

3. -5 , maximum

4. 198 m^2

5. 12 values

6. $32 = l(l - 4)$ or $l^2 - 4l = 32$

7. a) right triangle

b) obtuse triangle

8. The data would be quite scattered about the line of best fit but would have a slight trend with a negative slope.

Practise

1. a) Using linear regression for a non-linear relation.

b) Too little data.

c) Using linear regression when the correlation is weak.

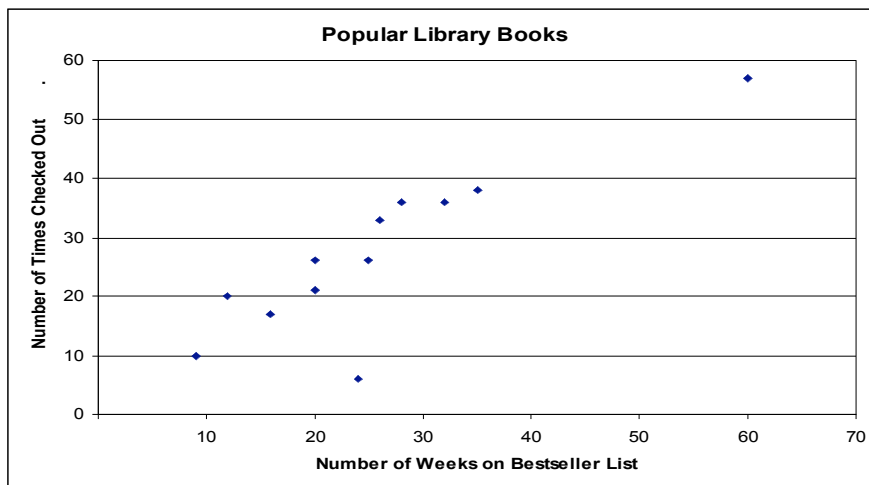
d) Not considering the effects of outliers or influential points.

2. a) Too little data.

b) The cause and effect relationship is reversed.

c) Extrapolating outside the range of the data set.

3. a)



b) Outlier at (60, 57) and influential point at (24, 6).

c) Yes. As the number of weeks on the bestseller list increases the number of times the book is checked out increases.

d) $y = 0.9106x + 3.8695$; slope: 0.9106; vertical intercept: 3.8695

e) Yes. There is sufficient data, the relation is linear, and the correlation is strong.

Chapter 3 Review, pages 56-58

1. a) i) one fact; number of community service hours completed

ii) two facts; number of field hockey goals scored this year and last year

b) Answers may vary. For example: Is there a relationship between the number of goals scored last year and the number of goals scored this year?

2. a) Eliminate bias. Consider the mode of delivery for the types of questions being asked.

b) Design for honest responses. Do not force the respondent to think hard about their responses.

c) Ethics in surveys and design for honest responses. Names of respondents must be kept confidential and answers kept anonymous and confidential.

3. Answers may vary. For example:

a) Have you ever purchased organic produce? Yes/No

b) What types of organic produce do you grow? Select all that apply. a) apples b) oranges c) bananas d) potatoes e) tomatoes d) other (specify) _____

c) How often do you eat organic produce? 1—Never 2—A few times a year, 3—A few times a month. 4—A few times a week. 5—Daily

d) Which type of organic produce is your greatest revenue? _____

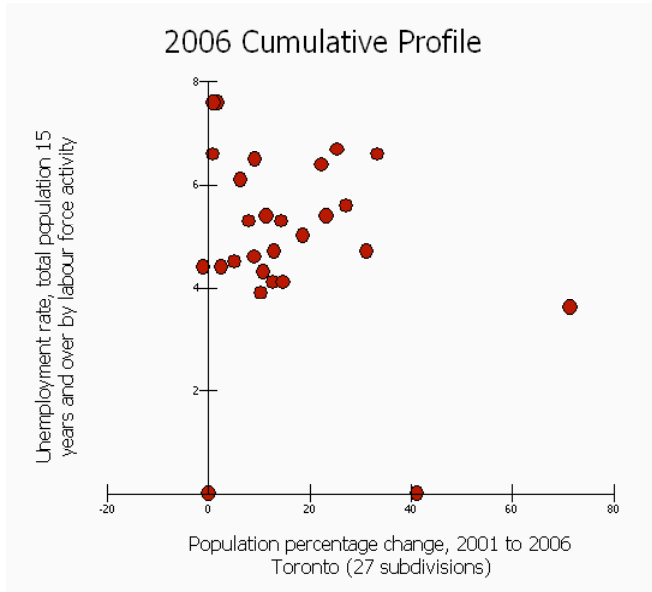
e) Why did you decide to start growing organic produce?

4. a) Primary data. The agency collected the data themselves.

b) Secondary. The data was collected by another source.

c) Primary. The data was collected by Heather herself.

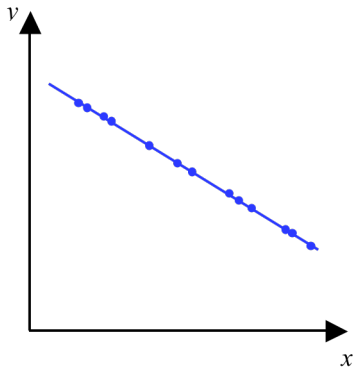
5. a) Answers may vary. For example:



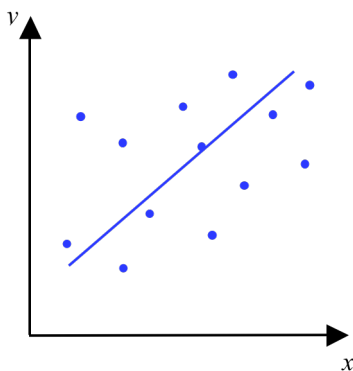
Source: Statistics Canada. 2006 Cumulative Profile, Toronto (27 subdivisions) (graph), 2006 Census of Population (Provinces, Census Divisions, Municipalities) (database), Using E-STAT (distributor).

b) There is a slight trend to indicate that as the population of a major city increases the unemployment rate also increases.

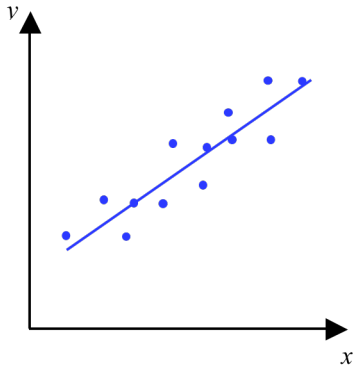
6. a)



b)



c)



7. Answers may vary. For example:

a) $y = -0.0376x + 5.743$; slope: -0.0376 ; vertical intercept: 5.743

b) The correlation coefficient is $r = -0.3844$, so $r^2 = 0.1478$ and the line of best fit does not approximate the data very well.

8. a) Not considering the effects of outliers or influential points.

b) Using linear regression when the correlation is weak.

c) Too little data.

9. a) $y = 76.501x - 119.78$; $r = 0.9612$

b) Yes, a cause and effect relationship exists. An increase in the walking speed causes an increase in the number of calories burned.