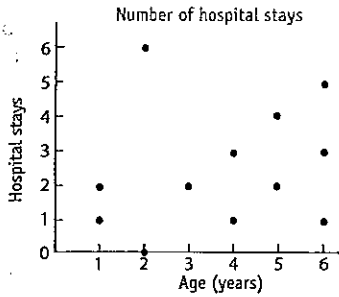


Further Practice: Correlation

Remember: all questions match the numbered examples on pages 95–104.

1 A scatterplot has been drawn showing the age and number of hospital stays of a certain group of children with a particular disease.



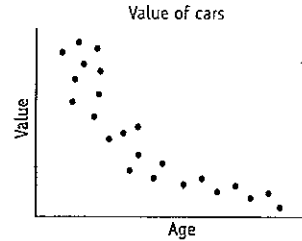
- How many children are shown on the scatterplot?
- How many hospital stays had the child who was 3 years old had?
- What ages are the children who have had three hospital stays?
- What is the difference in the number of hospital stays between the two children who are 5 years old?
- One child has been very ill a number of times for her age. Which child do you think this might be? Justify your answer.

2 The table shows the results of a survey of 18 high school students who were asked how many hours of television they watched each night.

Year	Hours	Year	Hours
7	3	8	4½
10	1	9	3
8	2½	7	2½
12	2	7	5
11	4	12	2½
12	3	9	3½
11	3½	10	4
9	4	12	0
10	2	8	3

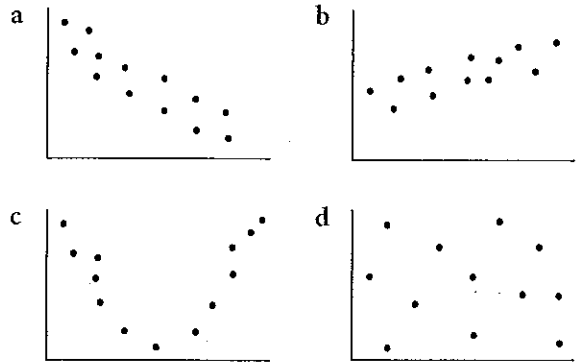
Show the results on a scatterplot, placing the year on the horizontal axis and the hours on the vertical axis.

3 A scatterplot has been drawn showing the value of certain cars as they age.

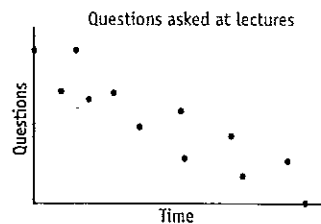


Describe any pattern.

4 Determine whether or not there is a pattern and if so whether it is linear.



5 A scatterplot has been drawn showing the numbers of questions asked at a lecture over time. Draw in a line of fit.

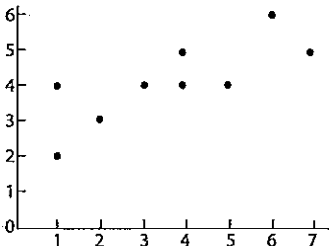


- 6** As an experiment, a group of people were given a different amount of time to memorise a set of objects. The people were then tested to see how many of the objects they could recall. The table below shows the results. Draw a scatterplot of the results and draw in a line of fit.

Seconds studied	Objects remembered
5	3
10	7
15	5
20	8
25	6
30	10
35	9
40	12
45	17
50	15
55	13
60	20

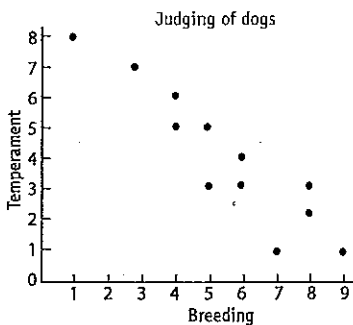
- 7** Find the median point of the points (1, 2), (3, 4) and (6, 3).

- 8** Consider the scatterplot below.



- Find the median of the first three points (moving from left to right across the scatterplot).
- Find the median of the middle three points.
- Find the median of the last three points.
- Draw in the median regression line.

- 9** This scatterplot shows the results when a group of dogs were given a score out of ten for temperament and level of breeding.

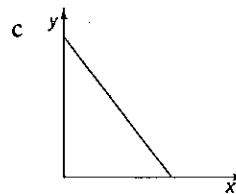
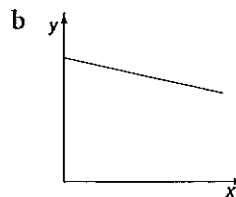
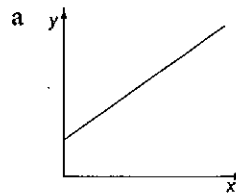


Draw in the median regression line for this scatterplot.

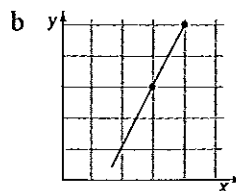
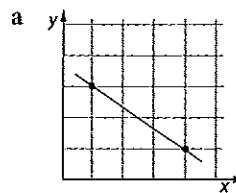
- 10** Fifteen students sat for tests in both spelling and grammar and their results, given as deciles, are shown in the table below. Show the points on a scatterplot and draw in the median regression line.

Spelling	Grammar	Spelling	Grammar
3	7	7	6
2	3	10	9
10	10	5	6
7	7	4	4
9	8	3	5
6	5	6	7
8	9	9	5
4	8		

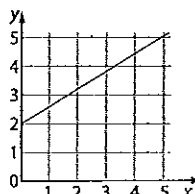
- 11** State whether the gradient of the line is positive or negative.



- 12** Find the gradient of the line joining the two given points:



- 13** Find the vertical intercept for this line.

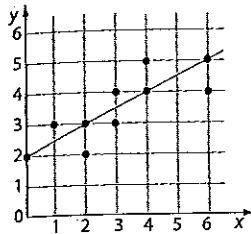


14 A line has gradient 3 and vertical intercept 4. What is the equation of the line?

15 The equation of a line is given by $y = 4x + 7$.

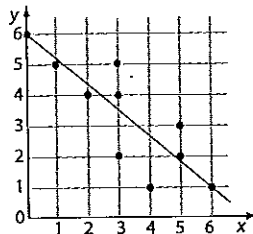
- What is the gradient of the line?
- What is the vertical intercept?

16 A line of fit has been drawn for the points shown on the scatterplot as shown in the diagram.



- What is the gradient?
- What is the vertical intercept?
- What is the equation of the line of fit?

17 What is the equation of the line of fit shown on the scatterplot?



18 Pupils at a school received marks out of 10 for conduct and attendance. The results for one class of twelve pupils are given in the table.

Conduct	Attendance	Conduct	Attendance
1	6	3	3
7	4	7	5
5	7	6	4
4	4	10	7
9	8	2	3
2	5	5	3

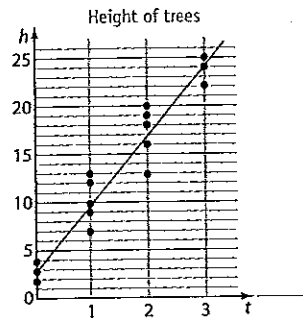
- Show the results on a scatterplot. Place *conduct* on the horizontal axis and *attendance* on the vertical axis.
- Draw in the median regression line.
- Find the equation of the median regression line.

19 A scatterplot has been drawn showing the profit made by a worker for different amounts of output. A line of fit was drawn and its equation was found to be

$y = \frac{1}{20}x + 2.5$ where y is the profit in dollars and x is the number of units produced. Use this equation to answer the follow questions.

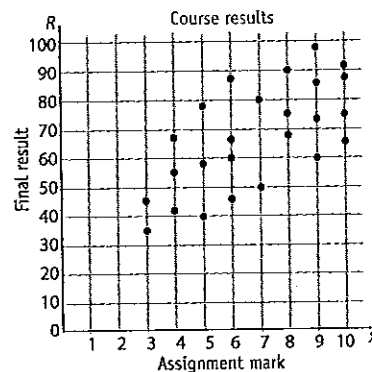
- How much profit would you expect would be made if 100 units are produced?
- One day a profit of \$30.50 was made. How many units would you expect were produced?
- How much more profit would you expect from the production of 2500 units than 1000 units?

20 A number of trees were planted and the height (in metres) recorded at the end of each year for the first three years. A line of fit was drawn as seen in the diagram.



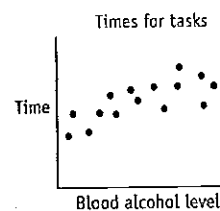
- If h is the height and t the number of years, what is the equation of the line of fit?
- If the trees continue to grow at the same rate, how high would you expect a tree to be after 6 years?
- After how many years would you expect a tree to reach 66 metres?
- Briefly explain why there must be some sort of limit placed on the equation for finding the heights of the trees.

21 A record was kept of the assignment mark achieved by students during a course and their final result in that course. The results are shown in the scatterplot below.



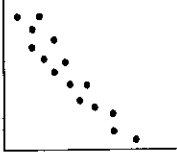
- Draw in a line of fit and find its equation.
- If a student scored 7 on their assignment, what would you expect their final result to be?
- If a student scored 76% as their final result, what was there likely assignment mark?

22 A scatterplot has been drawn showing the relationship between the blood alcohol level of a participant and the time needed to complete a task.

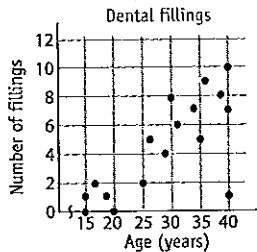


Briefly describe any correlation between the two variables.

23 Describe any correlation seen in the scatterplot?



24 A scatterplot was drawn showing the results of a survey on dental health where the number of fillings was checked on those taking part.

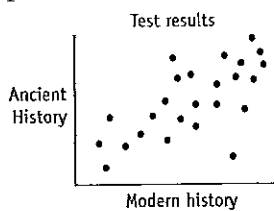


- Briefly describe the correlation.
- Paul commented that everyone who is older has more fillings than those who are younger. Is he correct? Justify your answer.

25 A survey was completed of people with lung disease to see if they had ever smoked. The correlation coefficient between the people with lung disease and those who smoked was found to be 0.8. Briefly describe the relationship between the smokers and lung disease.

26 The correlation coefficient between the age of a car and the distance travelled between breakdowns was found to be -0.5 . What does this mean?

27 The scatterplot shows the results of two different tests taken by history students in year 12 showing their performance in both ancient and modern history.



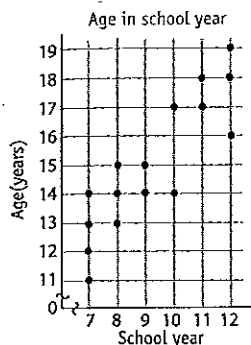
Sue commented that you must be good at modern history to do well in ancient history. Is this correct? Justify your answer.

28 Mandy noticed that there was a strong positive correlation between the number of employees in a business and the return to the business in income. She decided to employ more people in her business so that her income would improve. Briefly explain why this is a bad idea.

Go to p 287 for Quick Answers
or to pp 319–21 for Worked Solutions

Challenge: Correlation

1 A survey of students was taken asking their school year and age. A scatterplot has been drawn showing the results.



- How many students were surveyed?
- Is there any pattern? If so, is it linear?
- Draw in a line of fit.
- What is the equation of the line of fit? *Hint 1*
- Based on these results, what would you predict the age of a student in year 4 to be?

2 Information was collected about the age (in years) and value (in dollars) of certain cars. The results are shown in the table.

Age	Value	Age	Value
4	25 000	3	14 000
8	7 000	2	12 000
5	16 000	6	10 000
3	21 000	10	1 000
7	5 000	8	5 000
9	3 000	5	19 000
5	12 000	3	28 000
0	34 000	0	19 000
2	30 000	2	40 000
4	18 000	1	36 000

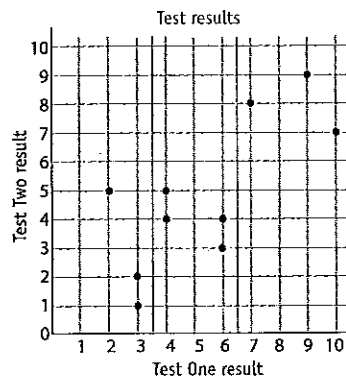
Is there any pattern? *Hint 2*

3 Ten women were surveyed about their height and shoe size. A line of fit was drawn and its equation is approximated by the rule:

$$\text{shoe size} = \frac{2}{5}(\text{height in cm}) - 56$$

- Explain why this rule won't work for people of height less than 140 cm? *Hint 3*
- Explain some other limitations of the rule. *Hint 4*

4 The scatterplot shows the results of two tests given in the same subject, with marks given as deciles. The scatterplot has been divided into three sections with the drawing of two vertical lines.



- Find the median of the three points in the first section.
- Find the median of the four points in the second section.
- Find the median of the three points in the third section.
- Draw a median regression line. *Hint 5*
- Find the equation of the median regression line.

Go to p 287 for **Quick Answers**
or to pp 321–2 for **Worked Solutions**

Hint 1: The vertical intercept of the line of fit is not 12 because there is a gap in the scale on the horizontal axis, the line of fit doesn't intercept the vertical axis at 12.

Hint 2: Plot the points on a scatterplot and look for a pattern.

Hint 3: Substitute a height of 140 cm into the rule.

Hint 4: How reliable was the original survey? What features did the original survey have?

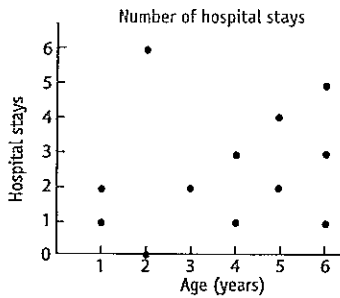
Hint 5: If the line joining the medians of the first and third sections also passes through the second median, then that line is the median regression line.

SOLUTIONS

Ch 6: Correlation

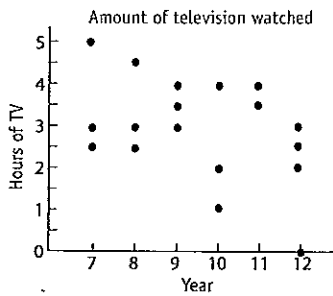
Further Practice p105

- 1** a There are 12 children shown.
 b The three-year-old child has had 2 hospital stays.
 c There are two children who have had three hospital stays. One is 4 years old and the other is 6 years old.
 d Of the 5-year-old children, one has had two stays and the other 4 stays. Difference = 2 stays
 e The two-year-old child who has had 6 hospital stays. The number of hospital stays is more than any other child and yet she is younger than most.

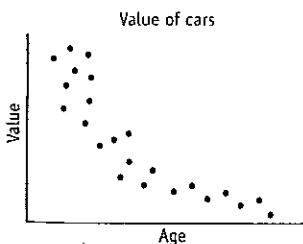


2

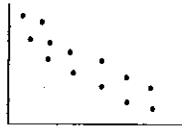
Year	Hours	Year	Hours
7	3	8	4½
10	1	9	3
8	2½	7	2½
12	2	7	5
11	4	12	2½
12	3	9	3½
11	3½	10	4
9	4	12	0
10	2	8	3



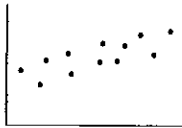
- 3** The dots follow a pattern. The value decreases quickly with age at first, then much more slowly. The pattern resembles an example of exponential decay [see page 251].



- 4** a linear pattern



- b linear pattern



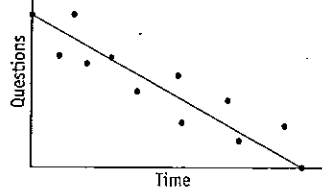
- c pattern, but not linear



- d no pattern



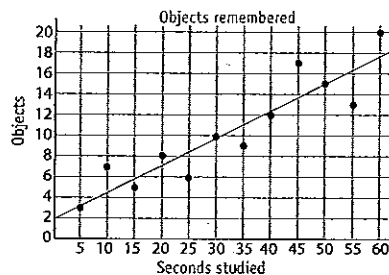
- 5** Questions asked at lectures



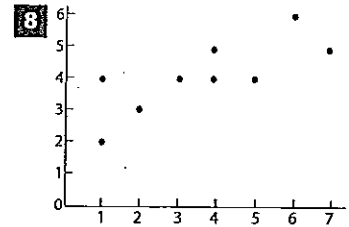
[There is no exact answer. Slight variations are correct.]

6

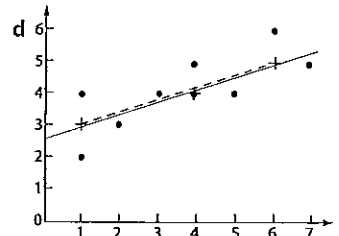
Seconds studied	Objects remembered
5	3
10	7
15	5
20	8
25	6
30	10
35	9
40	12
45	17
50	15
55	13
60	20



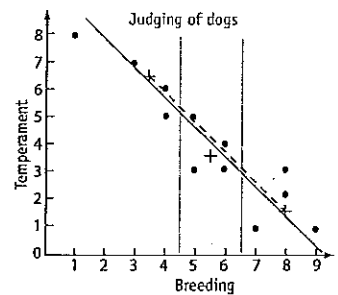
- 7** Horizontal values are 1, 3 and 6.
 The median is 3.
 Vertical values are 2, 3 and 4. [In order]
 The median is 3.
 Median point is at (3, 3).



- a Horizontal values: 1, 1, 2
 Median = 1
 Vertical values: 2, 3, 4
 Median = 3
 Median point is (1, 3).
 b Horizontal values: 3, 4, 4
 Median = 4
 Vertical values: 4, 4, 5
 Median = 4
 Median point is (4, 4).
 c Horizontal values: 5, 6, 7
 Median is 6.
 Vertical values: 4, 5, 6
 Median is 5.
 Median point is (6, 5).

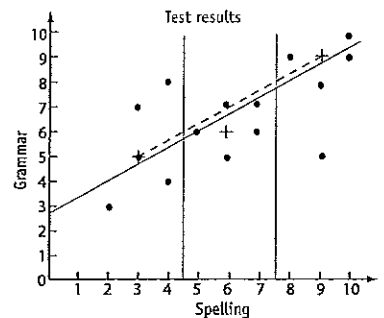


- 9** median 1st group = (3.5, 6.5)
 median 2nd group = (5.5, 3.5)
 median 3rd group = (8, 1.5)

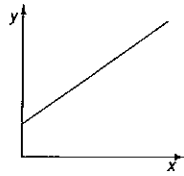


10

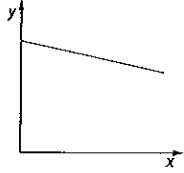
Spelling	Grammar	Spelling	Grammar
3	7	7	6
2	3	10	9
10	10	5	6
7	7	4	4
9	8	3	5
6	5	6	7
8	9	9	5
4	8		



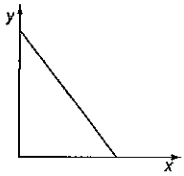
11 a positive



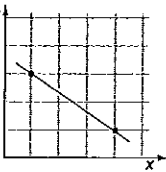
b negative



c negative



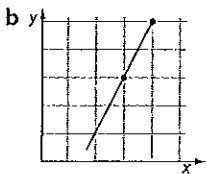
12 a



$$m = \frac{\text{vertical change in position}}{\text{horizontal change in position}}$$

$$= \frac{-2}{3}$$

$$= -\frac{2}{3}$$

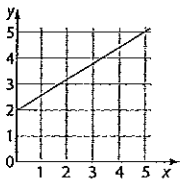


$$m = \frac{\text{vertical change in position}}{\text{horizontal change in position}}$$

$$= \frac{2}{1}$$

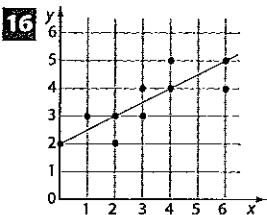
$$= 2$$

13 vertical intercept = 2



14 $m = 3$, $b = 4$
 $y = 3x + 4$

15 $y = 4x + 7$
a $m = 4$
b $b = 7$



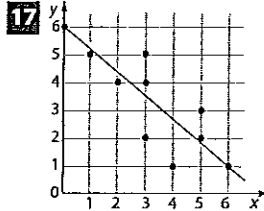
a $m = \frac{\text{vertical change in position}}{\text{horizontal change in position}}$

$$= \frac{2}{4}$$

$$= \frac{1}{2}$$

b $b = 2$

c $y = mx + b$
 $y = \frac{1}{2}x + 2$



17 $m = \frac{\text{vertical change in position}}{\text{horizontal change in position}}$

$$= \frac{-5}{6}$$

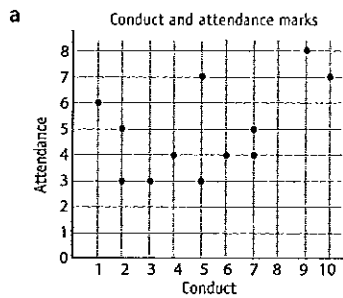
$$= -\frac{5}{6}$$

b $b = 6$

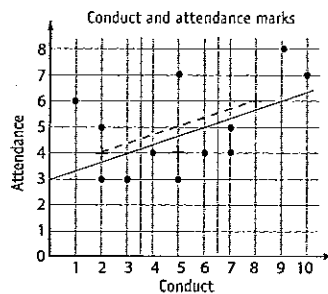
$y = mx + b$
 $y = -\frac{5}{6}x + 6$

18

Conduct	Attendance	Conduct	Attendance
1	6	3	3
7	4	7	5
5	7	6	4
4	4	10	7
9	8	2	3
2	5	5	3



b Median of first four points is (2, 4).
Median of middle four points is (5, 4).
Median of last four points is (8, 6).



c $m = \frac{\text{vertical change in position}}{\text{horizontal change in position}}$

$$= \frac{2}{6}$$

$$= \frac{1}{3}$$

Vertical intercept = 3

Equation is $y = \frac{1}{3}x + 3$

19 a $y = \frac{1}{20}x + 2.5$

When $x = 100$,

$$y = \frac{1}{20} \times 100 + 2.5$$

$$= 5 + 2.5$$

$$= 7.5$$

Expected profit is \$7.50

b $y = \frac{1}{20}x + 2.5$

When $y = 30.5$,

$$30.5 = \frac{1}{20}x + 2.5$$

$$28 = \frac{1}{20}x$$

$$x = 560$$

560 units are likely to be produced.

c $y = \frac{1}{20}x + 2.5$

When $x = 1000$,

$$y = \frac{1}{20} \times 1000 + 2.5$$

$$= 52.5$$

When $x = 2500$,

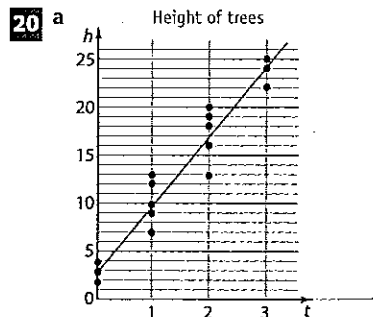
$$y = \frac{1}{20} \times 2500 + 2.5$$

$$= 127.5$$

$$\text{Extra profit} = \$127.50 - \$52.50$$

$$= \$75$$

You would expect the profit to be \$75 more.



vertical change = $24 - 10 = 14$

horizontal change = $3 - 1 = 2$

$$m = \frac{\text{vertical change in position}}{\text{horizontal change in position}}$$

$$= \frac{14}{2}$$

$$= 7$$

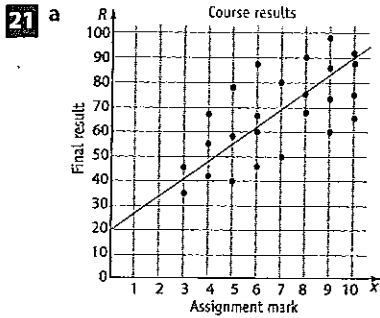
$$b = 3, \quad h = mt + b$$

$$h = 7t + 3$$

b $h = 7t + 3$
 When $t = 6$,
 $h = 7 \times 6 + 3$
 $= 45$
 You would expect the tree to be 45 m high.

c $h = 7t + 3$
 When $h = 66$,
 $7t + 3 = 66$
 $7t = 63$
 $t = 9$
 You would expect the tree to be 9 years old.

d Trees won't continue to grow at the rate they do to begin with. They eventually stop growing.



$$m = \frac{\text{vertical change in position}}{\text{horizontal change in position}}$$

$$= \frac{70}{10}$$

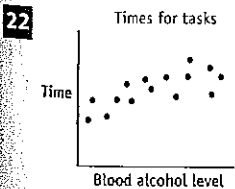
$$= 7$$

Vertical intercept = 20
 Equation is $R = 7x + 20$

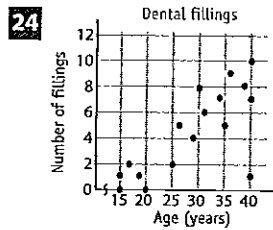
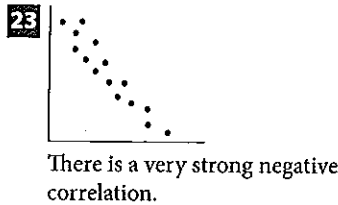
[The line of fit is not unique. A slightly different line will have a different equation.]

b When $x = 7$,
 $R = 7 \times 7 + 20$
 $= 69$
 The final mark is likely to be 69.

c When $R = 76$,
 $76 = 7x + 20$
 $56 = 7x$
 $8 = x$
 The assignment mark was likely to have been 8.



There is a strong positive correlation between the blood alcohol level and the time taken to complete the task.



a There is a strong positive correlation between the age and the number of fillings.

b Paul is not correct. The strong positive correlation tells us that in general those who are older have more fillings, but it is not true for every person. For example, one 40-year-old person had just one filling.

25 There is a very strong positive correlation between smoking and lung disease.

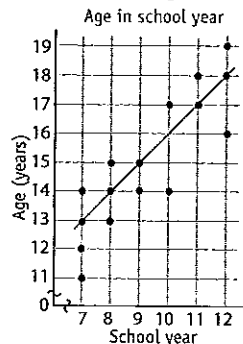
26 There is a moderate negative correlation between the age of a car and the distance travelled between breakdowns.

27 It is not correct. Most of those who did well in ancient history also did well in modern history, but there is no proof that being good at modern history causes someone to do well in ancient history.

28 There is no reason to believe that it is the number of employees who cause the income to be higher. In fact Mandy would probably find that her business would suffer from having to pay the extra employees.

Challenge p109

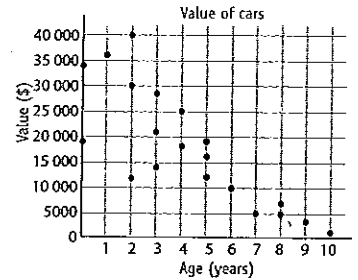
- 1** a 16 students were surveyed.
 b Yes there is a pattern and it is linear.
 c [Other lines are possible.]



d gradient = $\frac{1}{1}$
 $= 1$
 Equation is $y = x + 6$.
 [$y = x + b$. When $x = 7$, $y = 13$ so $b = 6$.]

e When $x = 4$,
 $y = 4 + 6$
 $= 10$
 The child is likely to be 10 years old.

2 There is a pattern. It is almost linear. [In general, the older the car the less its value. This is an example of strong negative correlation.]



3 a shoe size = $\frac{2}{5}(\text{height in cm}) - 56$

When height is 140 cm:

$$\text{shoe size} = \frac{2}{5} \times 140 - 56$$

$$= 0$$

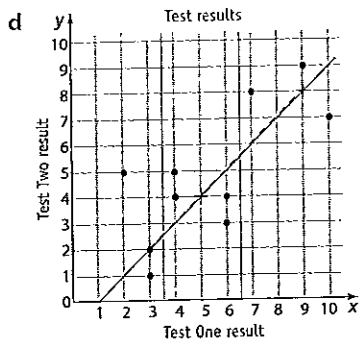
If the height was less than 140 cm the rule would give a negative shoe size which is not possible.

b The original survey was of women so the rule may not be able to be applied to men or children. Only ten women were surveyed so the rule might not be very reliable. If the height gets too large the rule may not be reliable.

4 a x-values: 2, 3, 3
 Median is 3
 y-values: 1, 2, 5
 Median is 2
 Median point is (3, 2).

b x-values: 4, 4, 6, 6
 Median is 5
 y-values: 3, 4, 4, 5
 Median is 4
 Median point is (5, 4).

c x-values: 7, 9, 10
 Median is 9
 y-values: 7, 8, 9
 Median is 8
 Median point is (9, 8).



e Gradient = $\frac{1}{1}$
 $= 1$
 Vertical intercept = -1
 Equation is $y = x - 1$