4.7 Scatter Graphs and Line of Best Fit: 1-hour Pack

EDEXCEL 1ST0 · Calculator allowed · about 60 minutes

Total Marks

Name: _____ Date: ____ / ____ / _____

Answer ALL questions. Show all your working.

1 A study finds that the product moment correlation coefficient between the price of a house and the number of bedrooms is r = -0.66.

(a) Describe the correlation. Use these descriptions: strong is |r| of 0.7 or more, moderate is from 0.4 up to 0.7, weak is below 0.4.

(b) A student says that a change in the price of a house causes a change in the number of bedrooms. Comment on this claim.

(Total for Question 1 is 4 marks)

2 The data below could be plotted on a scatter graph. A homeowner records the temperature in degrees Celsius (x) and gas used in units (y).

x	1	5	8	12	15	17	19	20
y	50	60	69	77	90	94	99	109

(a) Work out the mean of x and the mean of y, giving the coordinates of the mean point.

(b) Describe the type of correlation between x and y.

(Total for Question 2 is 4 marks)

© Revision Library | revisionlibrary.co.uk

3 The data below could be plotted on a scatter graph. A dealer records the age in years of a car (x) and its value in hundreds of pounds (y).

x	1	2	8	9	17	18	19
y	54	59	70	70	85	87	92

A line of best fit is drawn so that it passes through the points (2, 59) and (18, 87).

(a) Work out the gradient of the line.

(b) Find the equation of the line in the form $y = mx + c$.

(c) Use your equation to estimate y when $x = 15$.

(d) State whether your estimate in (c) is interpolation or extrapolation, and whether it is likely to be reliable.

(Total for Question 3 is 7 marks)

4 A study finds that the product moment correlation coefficient between the number of books in a home and a pupil's exam result is $r = 0.64$.

(a) Describe the correlation. Use these descriptions: strong is $|r|$ of 0.7 or more, moderate is from 0.4 up to 0.7, weak is below 0.4.

(b) A student says that a change in the number of books in a home causes a change in a pupil's exam result. Comment on this claim.

(Total for Question 4 is 4 marks)

© Revision Library | revisionlibrary.co.uk

5 The data below could be plotted on a scatter graph. A taxi firm records distance travelled in kilometres (x) and the fare in pounds (y).

x	6	7	8	15	16	18
y	31	36	37	55	60	66

A line of best fit is drawn so that it passes through the points (7, 36) and (16, 60).

(a) Work out the gradient of the line.

(b) Find the equation of the line in the form $y = mx + c$.

(c) Use your equation to estimate y when $x = 22$.

(d) State whether your estimate in (c) is interpolation or extrapolation, and whether it is likely to be reliable.

(Total for Question 5 is 7 marks)

---

6 The data below could be plotted on a scatter graph. A shop records the number of rainy days in a month (x) and umbrellas sold (y).

x	6	7	10	12	15	18
y	81	80	63	56	45	36

A line of best fit is drawn so that it passes through the points (7, 80) and (15, 45).

(a) Work out the gradient of the line.

(b) Find the equation of the line in the form $y = mx + c$.

(Total for Question 6 is 4 marks)

© Revision Library | revisionlibrary.co.uk

7 A study finds that the product moment correlation coefficient between a child's shoe size and the child's reading score is r = 0.15.

(a) Describe the correlation. Use these descriptions: strong is |r| of 0.7 or more, moderate is from 0.4 up to 0.7, weak is below 0.4.

(b) A student says that a change in a child's shoe size causes a change in the child's reading score. Comment on this claim.

(Total for Question 7 is 4 marks)

---

8 The data below could be plotted on a scatter graph. A shop records the number of rainy days in a month (x) and umbrellas sold (y).

x	3	7	10	12	16	20
y	101	86	69	64	50	36

A line of best fit is drawn so that it passes through the points (7, 86) and (16, 50).

(a) Work out the gradient of the line.

(b) Find the equation of the line in the form y = mx + c.

(Total for Question 8 is 4 marks)

© Revision Library | revisionlibrary.co.uk

9 A study finds that the product moment correlation coefficient between the amount of time spent revising and the test score is r = 0.55.

(a) Describe the correlation. Use these descriptions: strong is |r| of 0.7 or more, moderate is from 0.4 up to 0.7, weak is below 0.4.

(b) A student says that a change in the amount of time spent revising causes a change in the test score. Comment on this claim.

(Total for Question 9 is 4 marks)

---

10 The data below could be plotted on a scatter graph. A homeowner records the temperature in degrees Celsius (x) and gas used in units (y).

x	4	5	6	7	9	10	12
y	50	53	58	63	72	69	81

A line of best fit is drawn so that it passes through the points (5, 53) and (10, 69).

(a) Work out the gradient of the line.

(b) Find the equation of the line in the form y = mx + c.

(c) Use your equation to estimate y when x = 23.

(d) State whether your estimate in (c) is interpolation or extrapolation, and whether it is likely to be reliable.

(Total for Question 10 is 7 marks)

© Revision Library | revisionlibrary.co.uk

11 The data below could be plotted on a scatter graph. A homeowner records the temperature in degrees Celsius (x) and gas used in units (y).

x	3	4	6	9	11	14	16	19
y	98	93	91	85	73	67	63	60

A line of best fit is drawn so that it passes through the points (4, 93) and (16, 63).

(a) Work out the gradient of the line.

(b) Find the equation of the line in the form $y = mx + c$.

(c) Use your equation to estimate y when $x = 10$.

(d) State whether your estimate in (c) is interpolation or extrapolation, and whether it is likely to be reliable.

(Total for Question 11 is 7 marks)

12 The data below could be plotted on a scatter graph. A shop records the number of rainy days in a month (x) and umbrellas sold (y).

x	1	3	6	9	12	13	14	20
y	66	63	61	52	41	40	37	31

A line of best fit is drawn so that it passes through the points (3, 63) and (14, 37).

(a) Work out the gradient of the line.

(b) Find the equation of the line in the form $y = mx + c$.

(Total for Question 12 is 4 marks)

© Revision Library | revisionlibrary.co.uk

Mark scheme · 4.7 Scatter Graphs and Line of Best Fit: 1-hour Pack

Question 1

negative	B1
moderate	B1
correlation does not prove causation	B1
a third factor, such as the size of the house, may affect both	B1

Answer: (a) moderate negative correlation; (b) correlation does not imply causation; a third factor such as the size of the house may be the cause

Question 2

sum of x divided by n or sum of y divided by n	M1
(12.1, 81) cao	A1
positive correlation	B1
as x increases y increases oe	B1

Answer: (a) (12.1, 81); (b) positive correlation

Question 3

change in y divided by change in x	M1
1.75 cao	A1
substitutes a point to find c	M1
y = 1.75x + 55.5 cao	A1
substitutes x = 15	M1
81.75 cao	A1
interpolation and more reliable as within the data range	B1

Answer: (a) 1.75; (b) y = 1.75x + 55.5; (c) 81.75; (d) interpolation

Question 4

positive	B1
moderate	B1
correlation does not prove causation	B1
a third factor, such as family circumstances, may affect both	B1

Answer: (a) moderate positive correlation; (b) correlation does not imply causation; a third factor such as family circumstances may be the cause

Question 5

change in y divided by change in x	M1
2.6666666667 cao	A1
substitutes a point to find c	M1
y = 2.6666666667x + 17.3333333333 cao	A1
substitutes x = 22	M1
76 cao	A1
extrapolation and less reliable as outside the data range	B1

Answer: (a) 2.6666666667; (b) y = 2.6666666667x + 17.3333333333; (c) 76; (d) extrapolation

Question 6

change in y divided by change in x	M1
gradient -4.375 cao	A1
substitutes a point into y = mx + c	M1
y = -4.375x + 110.625 cao	A1

Answer: (a) -4.375; (b) y = -4.375x + 110.625

Question 7

positive	B1
weak	B1
correlation does not prove causation	B1
a third factor, such as age, may affect both	B1

Answer: (a) weak positive correlation; (b) correlation does not imply causation; a third factor such as age may be the cause

Question 8

change in y divided by change in x	M1
gradient -4 cao	A1
substitutes a point into y = mx + c	M1
y = -4x + 114 cao	A1

Answer: (a) -4; (b) y = -4x + 114

Question 9

positive	B1
moderate	B1
correlation does not prove causation	B1
a third factor, such as natural ability or prior learning, may affect both	B1

Answer: (a) moderate positive correlation; (b) correlation does not imply causation; a third factor such as natural ability or prior learning may be the cause

Question 10

change in y divided by change in x	M1
3.2 cao	A1
substitutes a point to find c	M1
y = 3.2x + 37 cao	A1
substitutes x = 23	M1
110.6 cao	A1
extrapolation and less reliable as outside the data range	B1

Answer: (a) 3.2; (b) y = 3.2x + 37; (c) 110.6; (d) extrapolation

Question 11

change in y divided by change in x	M1
-2.5 cao	A1
substitutes a point to find c	M1
y = -2.5x + 103 cao	A1
substitutes x = 10	M1
78 cao	A1
interpolation and more reliable as within the data range	B1

Answer: (a) -2.5; (b) y = -2.5x + 103; (c) 78; (d) interpolation

Question 12

change in y divided by change in x	M1
gradient -2.3636363636 cao	A1
substitutes a point into y = mx + c	M1
y = -2.3636363636x + 70.0909090909 cao	A1

Answer: (a) -2.3636363636; (b) y = -2.3636363636x + 70.0909090909

© Revision Library | revisionlibrary.co.uk