



F02-30min Correlation and Causation (Foundation): 30-minute Pack

EDEXCEL 1ST0 · Calculator allowed · about 30 minutes

Total Marks

Name: _____ Date: ____ / ____ / ____

Answer ALL questions. Show all your working.

- 1** State whether this pair of variables would most likely show positive correlation, negative correlation or no correlation.

The number of hours a candle has been burning and the height of the candle remaining.

(Total for Question 1 is 1 mark)

- 2** For each pair of variables, state whether you would expect positive correlation, negative correlation or no correlation between them.

(a) The distance a delivery driver travels in a day and the amount of fuel used. (1)

(b) The age of a car and its value in pounds. (1)

(c) A person's height and the day of the week they were born on. (1)

(d) The number of hours spent revising for a test and the number of mistakes made in the test. (1)

(Total for Question 2 is 4 marks)

3 Complete the sentence with the correct word.

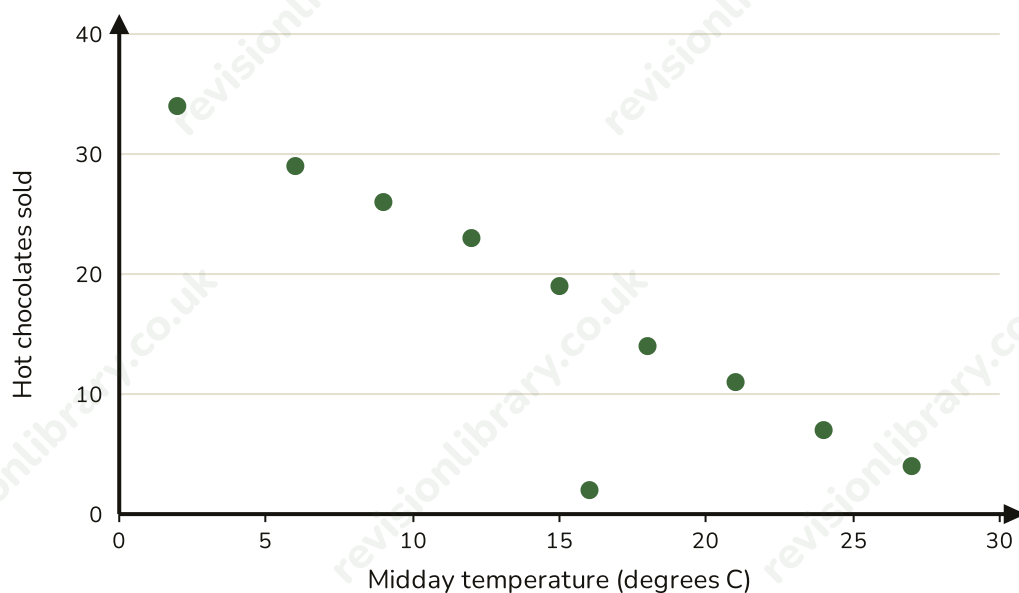
A sample of 8 towns shows a correlation between the number of umbrella shops and the number of raincoat shops, but neither business causes the other to exist; both are affected by how often it _____ in that area.

(Total for Question 3 is 2 marks)

4 A gardener records the amount of fertiliser used, in grams, and the height of 8 sunflowers, in cm, at the end of the season. The data shows a clear upward trend, with the points lying close to a straight line rising from (10, 40) to (80, 190). Describe the correlation shown by this data.

(Total for Question 4 is 2 marks)

- 5 A cafe manager records the midday temperature, in degrees C, and the number of hot chocolates sold on 10 days. The scatter graph shows the results.



- (a) Describe the correlation shown on the scatter graph, ignoring the unusual point. (2)
- (b) What name is given to a point, such as the one at 16 degrees C, that does not fit the pattern of the rest of the data? (1)

(Total for Question 5 is 3 marks)

6 For each statement about lines of best fit, write down True or False.

(i) A line of best fit must always pass through the origin (0, 0). **(1)**

(ii) A line of best fit should be drawn so that it passes through the data with roughly the same number of points above the line as below it. **(1)**

(iii) Using a line of best fit to estimate a value inside the range of the given data (interpolation) is generally more reliable than using it to estimate a value far outside that range (extrapolation). **(1)**

(Total for Question 6 is 3 marks)

7 For each statement about lines of best fit, write down True or False.

(i) A line of best fit is always drawn so it touches every point on the scatter graph.

(ii) Estimating a value that lies inside the range of the collected data is called interpolation.

(Total for Question 7 is 2 marks)

8 For each statement about lines of best fit, write down True or False.

(i) Estimating a value well beyond the range of the collected data (extrapolation) is generally less reliable than interpolation.

(ii) A line of best fit must have a positive gradient.

(Total for Question 8 is 2 marks)

9 A line of best fit for a set of data is $y = 4x + 6$, where x is the number of hours a heater is used and y is the electricity used, in units. Work out the estimated electricity used when the heater is used for 5 hours.

(Total for Question 9 is 2 marks)

10 A researcher collected data from 12 primary school classes and found a strong positive correlation between the average shoe size of pupils in a class and their average score on a reading test. A newspaper headline claims: 'Bigger feet make children better readers.'

(a) Explain why the newspaper's claim is not a valid conclusion to draw from this data. (2)

(b) Suggest one factor, other than age, that could affect a pupil's reading score. (1)

(Total for Question 10 is 3 marks)

11 A shop owner notices that in months when sunglasses sales are higher, ice cream sales are also higher. Suggest a factor, other than one causing the other, that could explain this correlation, and explain how it affects both variables.

(Total for Question 11 is 2 marks)

12 Fire brigade records show that fires attended by a larger number of fire engines tend to result in more damage, in pounds. A local councillor claims: 'Sending fewer fire engines to fires would reduce the amount of damage caused.' Explain why the councillor's claim is not supported by this data.

(Total for Question 12 is 2 marks)

13 An airline notices that in months when sales of sun cream are higher, the number of long-haul flight bookings is also higher. Suggest a factor, other than one directly causing the other, that could explain this correlation, and explain how it affects both variables.

(Total for Question 13 is 3 marks)

Mark scheme · F02-30min Correlation and Causation (Foundation): 30-minute Pack

Question 1

- **B1** negative correlation cao
- **Answer: Negative correlation**

Question 2

- (a) **B1** Positive correlation cao
- (a) **Answer: Positive correlation**
- (b) **B1** Negative correlation cao
- (b) **Answer: Negative correlation**
- (c) **B1** No correlation cao
- (c) **Answer: No correlation**
- (d) **B1** Negative correlation cao
- (d) **Answer: Negative correlation**

Question 3

- **M1** recognises the missing word describes local weather/climate oe
- **A1** rains cao
- **Answer: rains**

Question 4

- **B1** positive correlation
- **B1** strong (correlation), since the points lie close to a straight line
- **Answer: Strong positive correlation.**

Question 5

- (a) **B1** negative correlation
- (a) **B1** strong (correlation)
- (a) **Answer: Strong negative correlation**
- (b) **B1** outlier cao
- (b) **Answer: Outlier**

Question 6

- (i) **B1** False cao
- (i) **Answer: False**
- (ii) **B1** True cao
- (ii) **Answer: True**
- (iii) **B1** True cao
- (iii) **Answer: True**

Question 7

- **B1** (i) False cao
- **B1** (ii) True cao
- **Answer: (i) False (ii) True**

Question 8

- **B1** (i) True cao
- **B1** (ii) False cao
- **Answer: (i) True (ii) False**

Question 9

- **M1** $4(5) + 6$ oe
- **A1** 26 cao
- **Answer: 26 units**

Question 10

- (a) **B1** correlation does not prove causation
- (a) **B1** identifies age as a plausible confounding variable: older pupils tend to have both bigger feet and higher reading scores (from more years of practice), so age affects both variables rather than foot size causing better reading
- (a) **Answer: The correlation does not prove causation. A likely explanation is a confounding variable, age: older pupils tend to have both bigger feet (as they grow) and higher reading scores (from more years of practice), so age affects both variables rather than foot size causing better reading.**
- (b) **B1** any sensible factor, e.g. the amount of reading practice at home, the quality of teaching, or whether English is the pupil's first language, oe
- (b) **Answer: For example, the amount of time spent reading at home (or any other sensible factor affecting reading ability).**

Question 11

- **B1** identifies a plausible common cause, e.g. warm or sunny weather
- **B1** explains that this factor increases both variables, e.g. on hot sunny days people are more likely to buy both sunglasses and ice cream
- **Answer: Warm, sunny weather is a likely confounding factor: on hot sunny days people are more likely to buy both sunglasses and ice cream, so the weather affects both variables rather than one causing the other.**

Question 12

- **B1** correlation does not prove causation
- **B1** identifies the size or severity of the fire as the confounding variable: a larger, more serious fire causes both more fire engines to be sent and more damage, rather than the fire engines causing the damage
- **Answer: The claim is not supported. The confounding variable is the size (severity) of the fire: a larger fire causes both more fire engines to be sent and more damage, so the number of fire engines does not cause the damage. Sending fewer engines to a serious fire could make the damage worse, not better.**

Question 13

- **B1** identifies a plausible common cause, e.g. the summer holiday season oe
- **B1** explains that this factor increases both variables, e.g. more people go on holiday and buy sun cream during the summer months oe
- **B1** states that neither variable directly causes the other oe
- **Answer: The summer holiday season is a likely confounding factor: more people go abroad and buy sun cream during the summer months, so the season affects both variables rather than one causing the other.**