

Scatter Graphs and Correlation

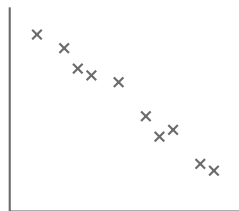
Solutions

For AQA GCSE Mathematics (8300) • Foundation tier

Each question is shown in grey, followed by the working, one step per line. The final answer is in a box. *Marks* shows how the marks are usually given in the exam, and **Tip** points out a common mistake.

1 (2 marks)

Here are two scatter graphs, *A* and *B*.



Graph *A*



Graph *B*

Write down the type of correlation shown by each graph.

(a) Graph *A*

The points go downhill from left to right.

As one variable increases, the other decreases.

Answer: Negative correlation

Tip: Write the full phrase “negative correlation”. Words like “decreasing” or “going down” are not enough.

(b) Graph *B*

The points are scattered all over the graph.

There is no pattern.

Answer: No correlation

2 (1 mark)

Circle the type of correlation you would expect between the age of a car and its value.

Options: positive, negative, no correlation

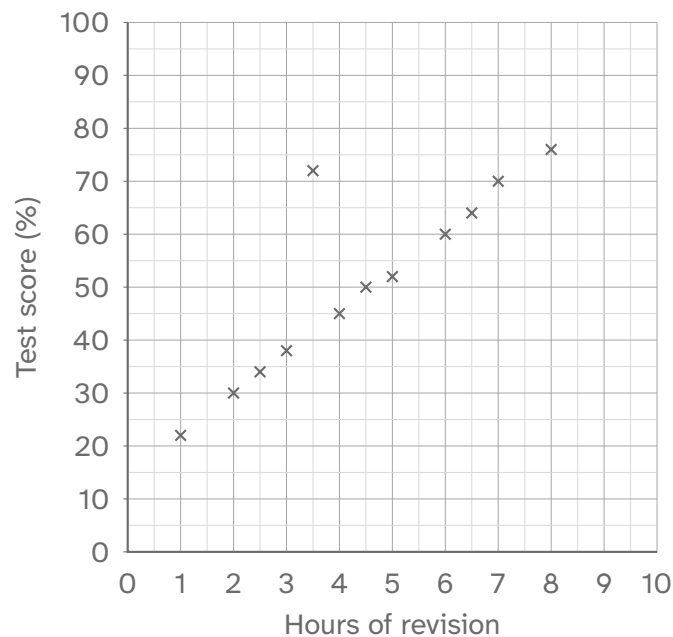
As a car gets older, its value usually goes down.

One variable increases while the other decreases, so this is negative correlation.

Answer: negative

3 (5 marks)

The scatter graph shows the number of hours some students revised and their test scores.



(a) Write down the type of correlation.

The points go uphill from left to right.

Students who revised for more hours tended to get higher scores.

Answer: Positive

- (b) One point is an outlier.
Circle the outlier on the scatter graph.

An outlier is a point that does not fit the pattern of the others.

The point at (3.5, 72) is far above the rest.

It shows a high score for only a little revision.

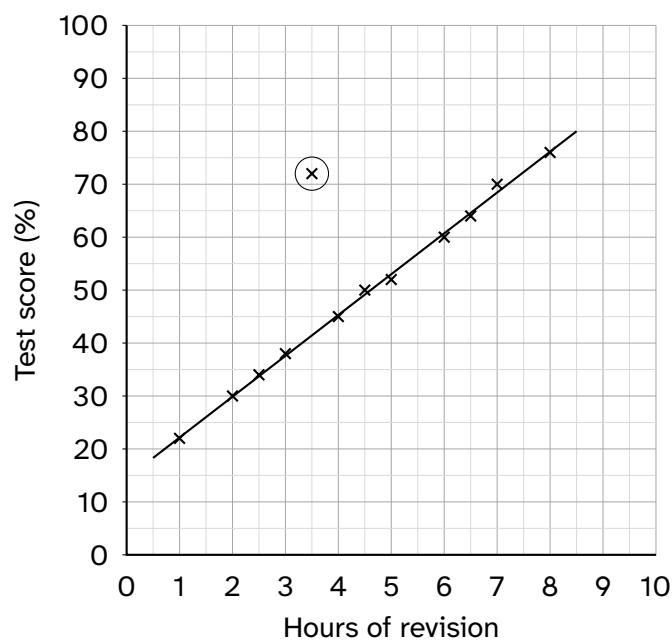
Answer: (3.5, 72)

- (c) Draw a line of best fit on the scatter graph.

Draw one straight line through the middle of the points.

Ignore the outlier.

Have about the same number of points above the line as below it.



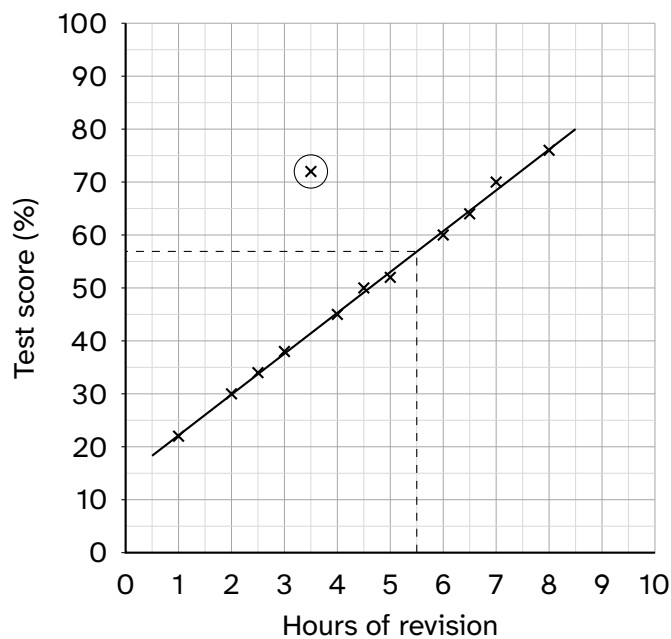
Answer: Line of best fit

Tip: A line of best fit does not have to go through (0, 0) and should not join the points dot-to-dot. It follows the trend of the points.

- (d) Use your line of best fit to estimate the score of a student who revised for 5.5 hours.

Draw a line up from 5.5 hours to your line of best fit.

Then draw a line across to the Test score axis.



This gives about 56%. Any answer from 52 to 60 is fine.

Answer: about 56 (accept 52 to 60)

Tip: Each small square on the vertical axis is 5%. Check the scale before you read off.

- (e) Give a reason why it may not be sensible to use your line to estimate the score for 10 hours of revision.

The data only goes up to 8 hours of revision.

10 hours is outside the range of the data.

We do not know that the pattern carries on, so the estimate may not be reliable.

Answer: Outside the data range

4 (5 marks)

A shop records the temperature and the number of ice creams sold on seven days.

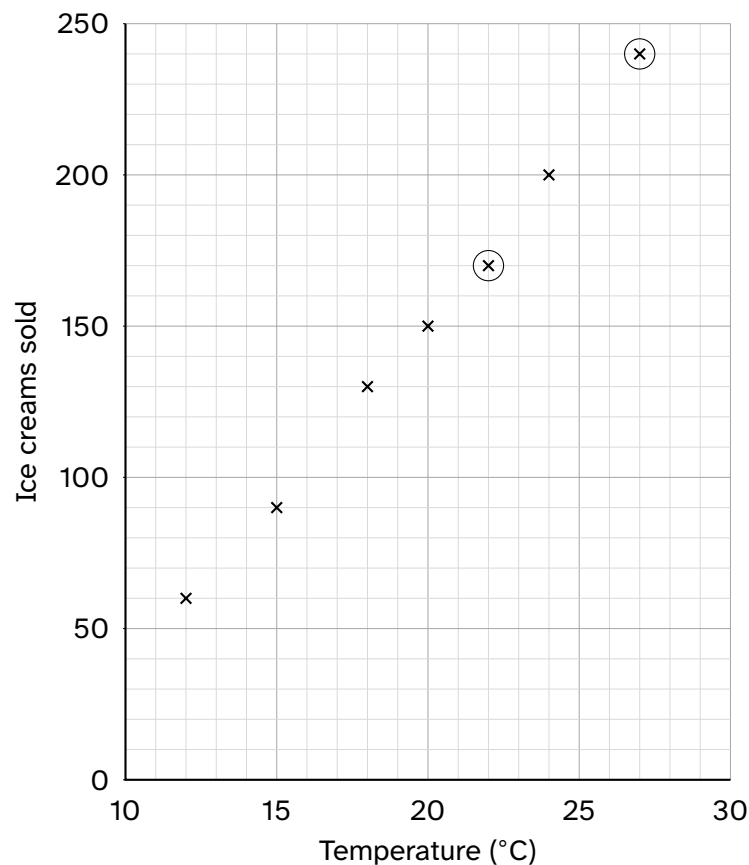
Temperature ($^{\circ}\text{C}$)	12	15	18	20	22	24	27
Ice creams sold	60	90	130	150	170	200	240

The scatter graph shows some of the information from the table.

- (a) Complete the scatter graph by plotting the two missing points.

The two missing days are 22°C with 170 ice creams and 27°C with 240 ice creams.

Plot (22,170) and (27,240). They are circled here.



Answer: Both points plotted

Tip: On the vertical axis each small square is 10 ice creams. 170 is between 150 and 200, so count up carefully.

Marks: 1 mark for each point plotted correctly.

- (b) Write down the type of correlation shown in the scatter graph.

As the temperature increases, the number of ice creams sold increases.

Answer: Positive

- (c) Write down the strength of the correlation.

The points lie close to a straight line.

Answer: Strong

Tip: The answer to a “strength” question is “strong” or “weak”. Do not write “positive” again.

- (d) The shop says, *“Hot weather causes people to buy ice creams.”*
Does the scatter graph prove this? Give a reason.

No.

The graph shows a link between temperature and ice cream sales.

But correlation does not prove that one thing causes the other.

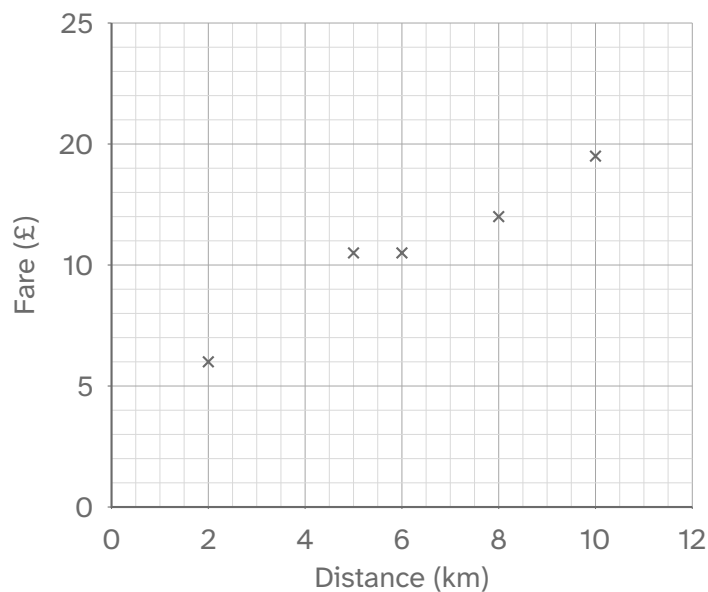
Answer: No, correlation does not prove causation

5 (2 marks)

Nadia records the distance and the fare for six taxi journeys.

Journey	A	B	C	D	E	F
Distance (km)	2	5	8	3	10	6
Fare (£)	6	11	14	7	19	11

Nadia draws this scatter graph.



Write down **two** mistakes she has made.

Check each point against the table.

Mistake 1: journey D (3 km, £7) has not been plotted. There are only five points, but there are six journeys.

Check the scales.

Mistake 2: the scale on the Fare axis is not even.

The numbers go 0, 5, 10, 20, 25, so the gap from 10 to 20 is the same size as the gap from 0 to 5.

Answer: Journey D (3, 7) is missing; the scale on the Fare axis is not even (10 to 20 takes the same space as 0 to 5)

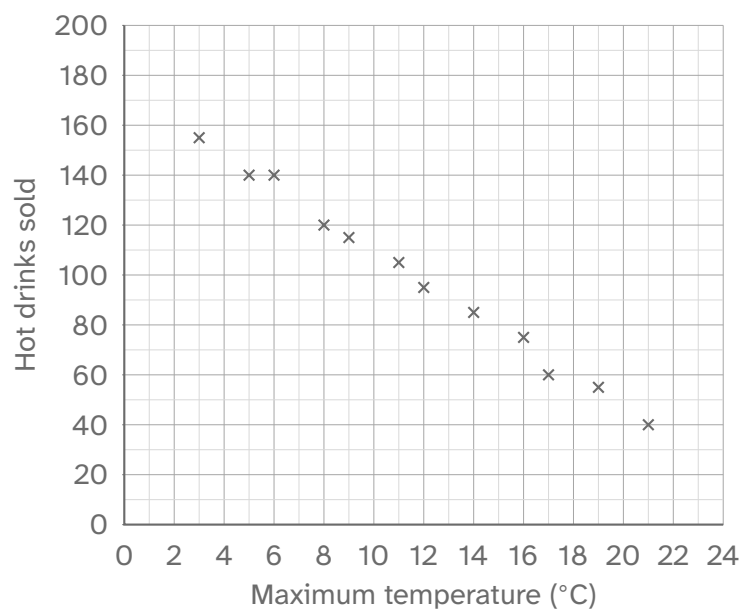
Tip: Be specific. Say which point is missing, or exactly what is wrong with the scale, not just “the graph is wrong”.

Marks: 1 mark for each mistake.

6 (4 marks)

A café records the maximum temperature and the number of hot drinks sold on 12 days.

The scatter graph shows this information.



(a) Circle the strength and type of correlation shown.

Options: weak positive, strong positive, weak negative, strong negative

The points go downhill, so the correlation is negative.

They lie close to a straight line, so it is strong.

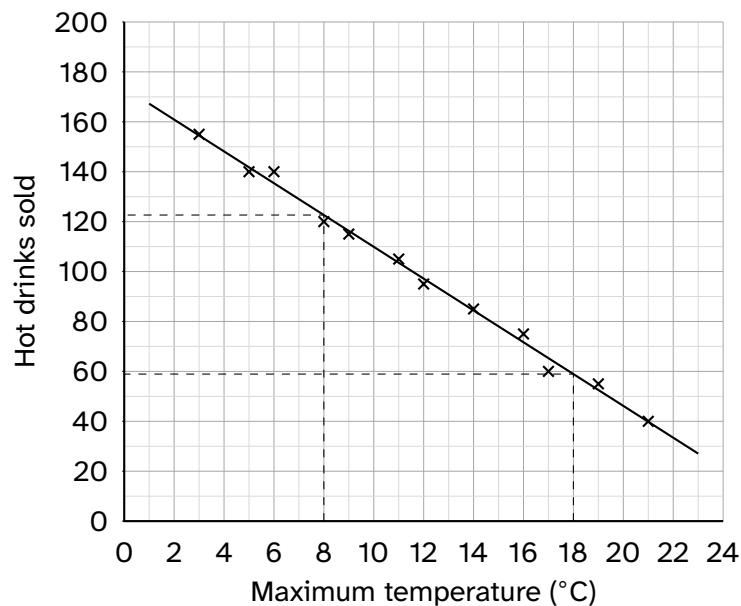
Answer: strong negative

- (b) The weather forecast for next weekend is
- | | |
|----------|---------------------------|
| Saturday | maximum temperature 8 °C |
| Sunday | maximum temperature 18 °C |

Use a line of best fit to estimate the total number of hot drinks the café will sell next weekend.

Draw a line of best fit.

Read up from 8 °C to the line, then across. Do the same from 18 °C.



Saturday: about 123 hot drinks.

Sunday: about 59 hot drinks.

Total: $123 + 59 = 182$

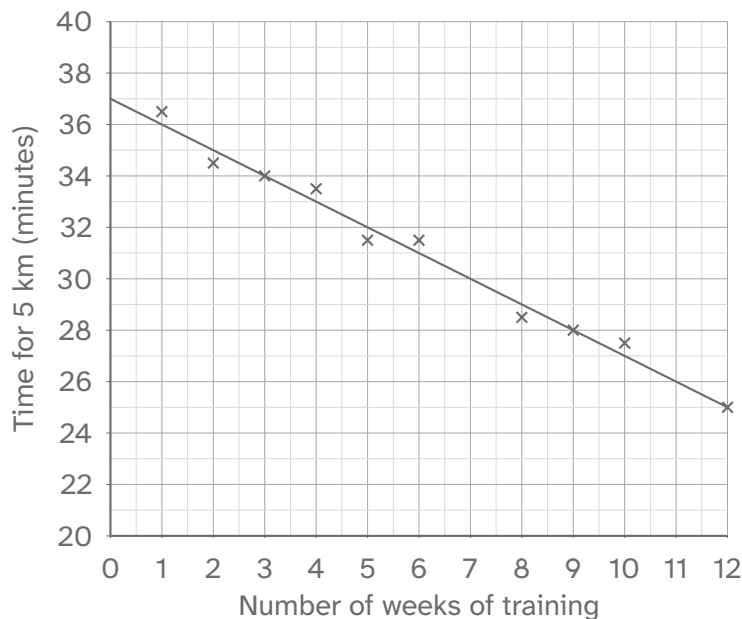
Answer: about 182 (accept 170 to 195)

Tip: Each small square on the vertical axis is 10 drinks. Read both days from the line, not from the nearest cross, then add.

Marks: 1 mark for a line of best fit. 1 mark for reading at 8 °C and at 18 °C. 1 mark for adding your two readings.

7 (4 marks)

The scatter graph shows the number of weeks that 10 members of a running club have been training, and the time each of them takes to run 5 km.
A line of best fit has been drawn.



- (a) Describe the relationship between the number of weeks of training and the time to run 5 km.

Describe the relationship in words, using the context.

The points go downhill: members who had trained for more weeks ran 5 km in less time.

Answer: As the number of weeks of training increases, the time to run 5 km decreases

Tip: Do not just write “negative correlation”. Say what happens: “the more weeks of training, the less time it takes to run 5 km”.

- (b) Zara joins the club.

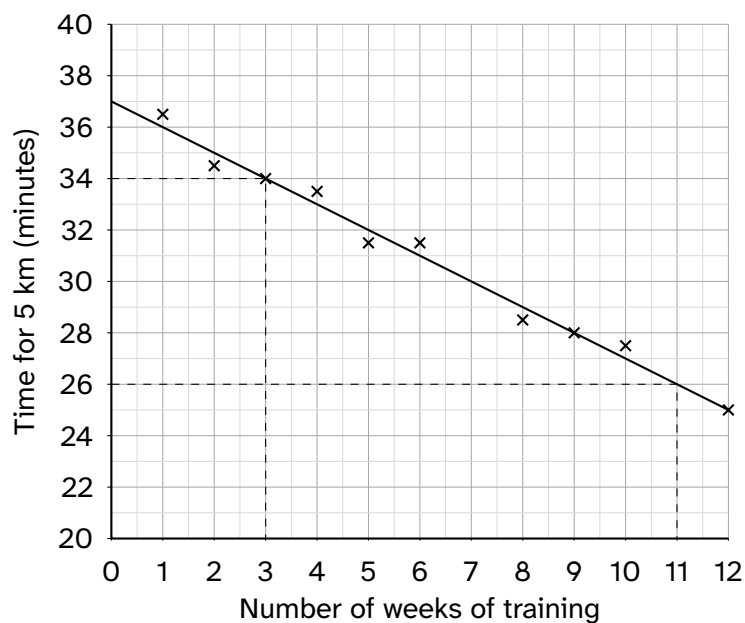
At the start she can run 5 km in 34 minutes.

She wants to get her time down to 26 minutes.

Use the line of best fit to estimate how many weeks of training this will take.

Read across from 34 minutes to the line, then down: about 3 weeks.

Read across from 26 minutes to the line, then down: about 11 weeks.



Zara is already at the 3-week level, so she needs $11 - 3 = 8$ more weeks.

Answer: about 8 weeks (accept 7 to 9)

Tip: A common mistake is to give 11 weeks. Zara can already run as fast as a member with about 3 weeks of training, so find the **difference**.

Marks: 1 mark for reading about 3 weeks at 34 minutes. 1 mark for reading about 11 weeks at 26 minutes. 1 mark for the difference, about 8.