

# Scatter Graphs and Correlation

## Solutions

For Pearson Edexcel GCSE Mathematics (1MA1) • Foundation tier

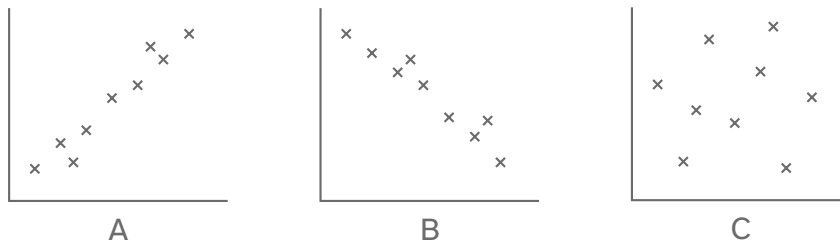
---

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 three scatter graphs.



Write down the letter of the graph that shows

(a) negative correlation

Graph *A*: the points go uphill, so it shows positive correlation.

Graph *B*: the points go downhill, so it shows negative correlation.

**Answer: B**

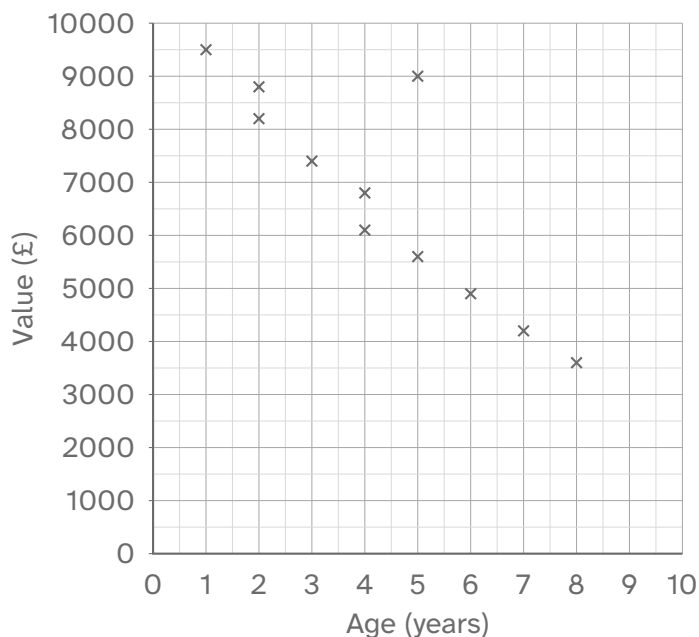
(b) no correlation

Graph *C*: the points are scattered with no pattern, so it shows no correlation.

**Answer: C**

**2** (4 marks)

The scatter graph shows information about the age and the value of some cars of the same model.



- (a) Describe the relationship between the age of a car and its value.

The points go downhill, so there is negative correlation.

In words: as the age of the car increases, its value decreases.

**Answer:** As age increases, value decreases

**Tip:** “Describe the relationship” means say it in words using the context, e.g. “older cars are worth less”. Just “negative” is not enough.

- (b) One of the points is an outlier.

Write down the age and value of this car.

Look for the point that does not fit the pattern.

The point at age 5 years has a value of £9000, which is much higher than other cars of that age.

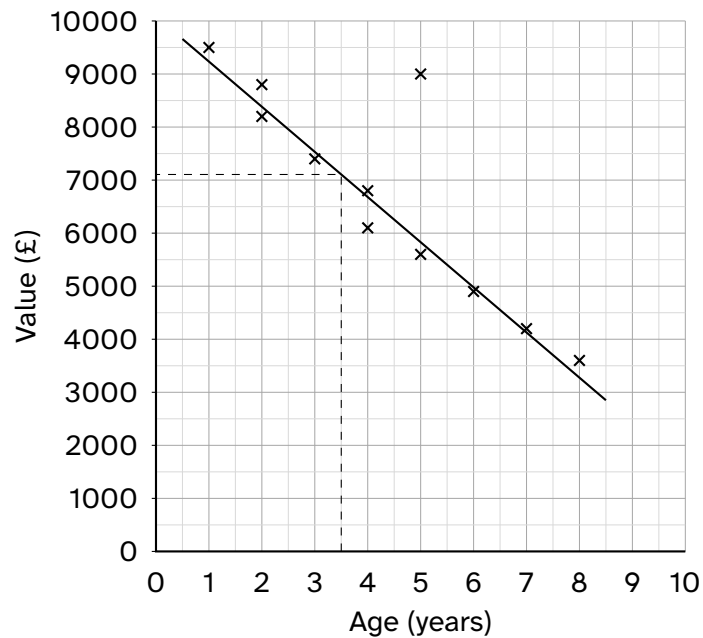
**Answer:** 5 years, £9000

- (c) A car of this model is 3.5 years old.

Estimate its value.

Draw a line of best fit, ignoring the outlier.

Read up from 3.5 years to the line, then across to the Value axis.



This gives about £7100 to £7200. Any answer from £6800 to £7600 is fine.

**Answer:** about £7200 (accept £6800 to £7600)

**Tip:** Always draw the line of best fit first, then read from it. Estimates made without a line are often outside the accepted range.

**Marks:** 1 mark for a line of best fit or a reading from 3.5 years. 1 mark for an answer in the range.

**3** (3 marks)

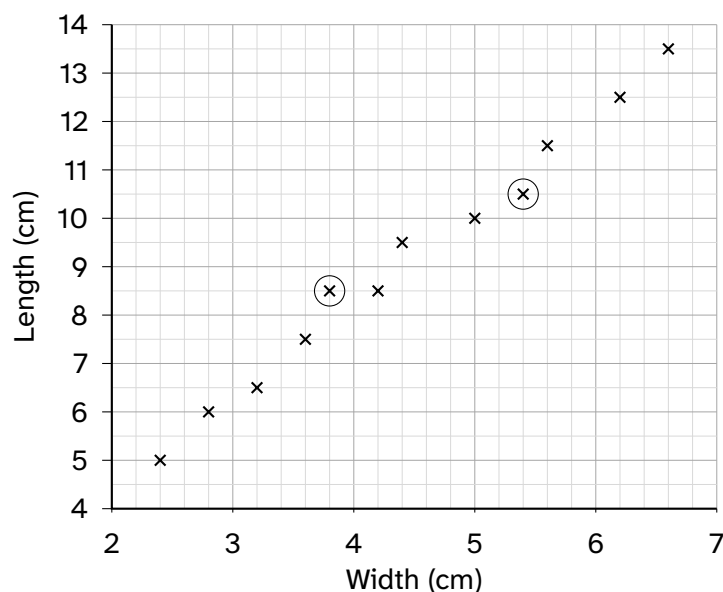
Emma measures the width and the length of some leaves from a tree. The scatter graph below shows information about her results. The table shows the width and the length of two more leaves.

|             | Leaf <i>P</i> | Leaf <i>Q</i> |
|-------------|---------------|---------------|
| Width (cm)  | 3.8           | 5.4           |
| Length (cm) | 8.5           | 10.5          |

- (a) On the scatter graph, plot the information for leaf *P* and leaf *Q*.

Leaf *P*: go across to 3.8 on the Width axis, then up to 8.5 on the Length axis.

Plot leaf *Q* at (5.4, 10.5) in the same way. Both new points are circled here.



**Answer:** Both points plotted

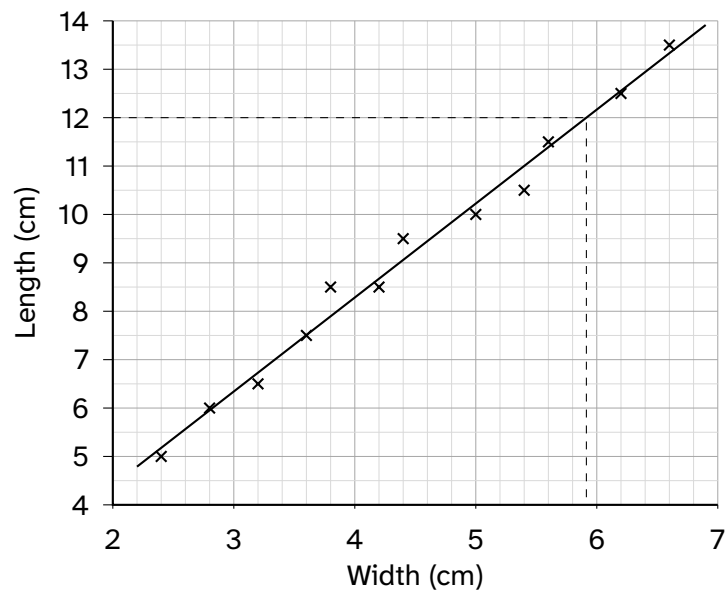
**Tip:** Each small square on the Width axis is 0.2 cm, so 3.8 is one small square before 4.

- (b) Another leaf from the same tree has a length of 12 cm. Estimate the width of this leaf.

Draw a line of best fit.

You know the length, so start on the vertical (Length) axis.

Read across from 12 cm to the line, then down to the Width axis.



This gives about 5.9 cm. Any answer from 5.6 to 6.2 is fine.

**Answer:** about 5.9 cm (accept 5.6 to 6.2)

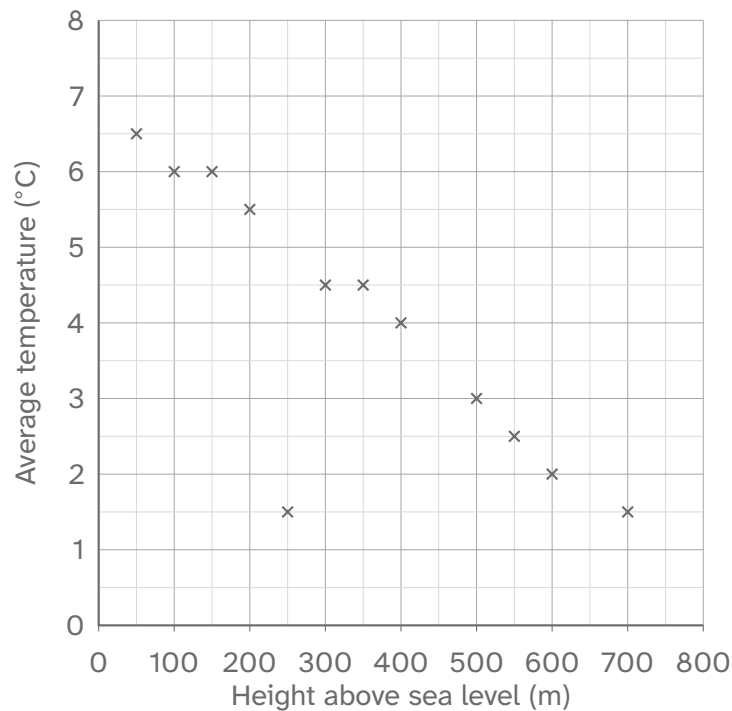
**Tip:** Start from 12 on the **Length** axis. Starting from 12 on the wrong axis is a common mistake.

*Marks:* 1 mark for a line of best fit or reading across from 12. 1 mark for an answer from 5.6 to 6.2.

**4** (5 marks)

The scatter graph shows information about 12 villages.

It shows the height above sea level of each village and the average temperature in the village in January.



- (a) One of the points is an outlier.

Write down the coordinates of this point.

The point at 250 m and 1.5 °C does not fit the pattern.

Other villages at about 250 m have an average temperature of about 5 °C.

**Answer:** (250, 1.5)

**Tip:** Give the coordinates the right way round: horizontal value first, (250, 1.5).

- (b) Ignoring the outlier, describe the relationship between the height above sea level and the average temperature in January.

The other points go downhill from left to right (negative correlation).

In words: the higher a village is above sea level, the lower its average temperature in January.

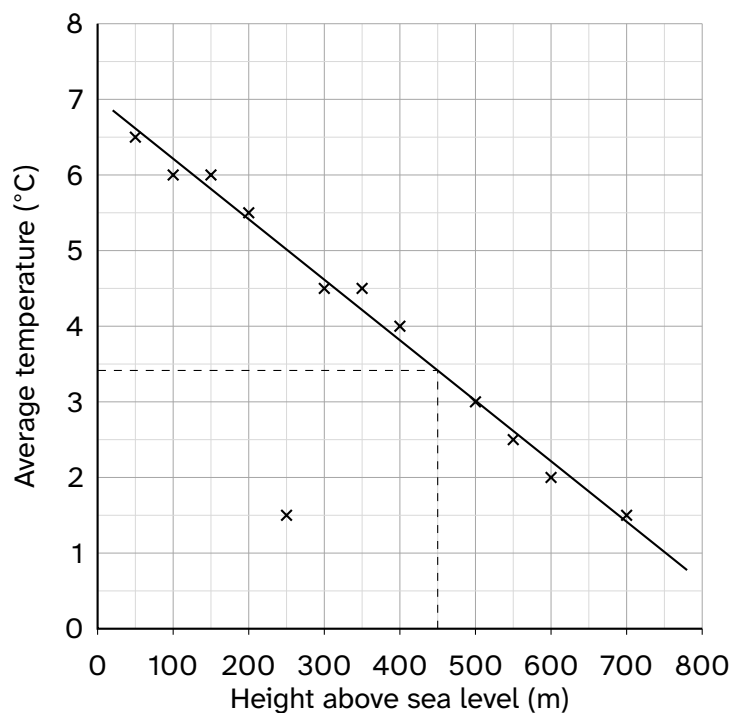
**Answer:** The higher the village, the lower the average January temperature

**Tip:** “Describe the relationship” means say it in words using the context. Just “negative” is not enough.

- (c) Another village is 450 m above sea level.  
Estimate the average temperature in January in this village.

Draw a line of best fit, ignoring the outlier.

Read up from 450 m to the line, then across to the Temperature axis.



This gives about 3.4 °C. Any answer from 3 to 3.8 is fine.

**Answer:** about 3.4 °C (accept 3 to 3.8)

**Tip:** Each small square on the Height axis is 50 m, so 450 is halfway between 400 and 500. Each small square on the Temperature axis is 0.5 °C.

**Marks:** 1 mark for a line of best fit or reading up from 450. 1 mark for an answer in the range.

(d) Karim says,

*“The outlier must be a mistake, so it should be crossed out.”*

Is Karim right?

Give a reason for your answer.

No.

An outlier does not fit the pattern, but it can still be a real value.

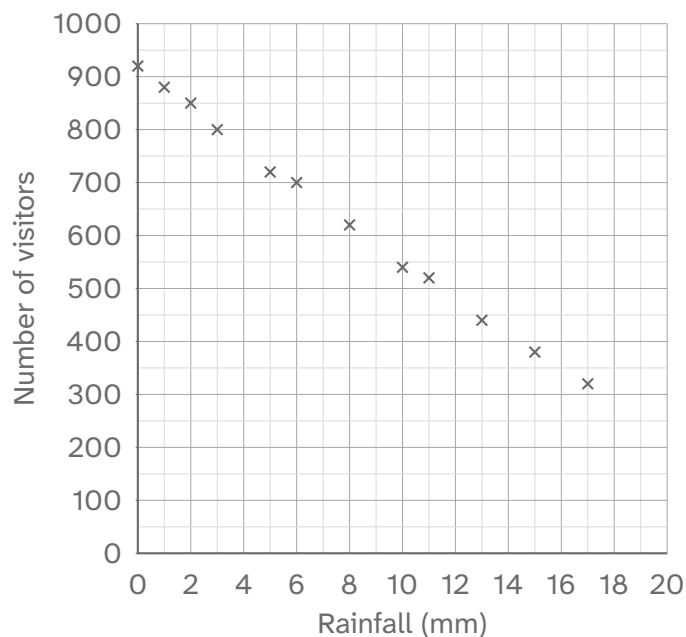
For example, this village might be in a cold valley.

You should only remove it if you know it was recorded wrongly.

**Answer:** No, an outlier can be a genuine value (it is not necessarily a mistake)

**5** (3 marks)

The scatter graph shows the rainfall and the number of visitors to a zoo on each of 12 days.



(a) Write down the type of correlation.

The points go downhill from left to right.

On days with more rain there were fewer visitors.

**Answer:** Negative

- (b) The zoo manager wants to use the scatter graph to estimate the number of visitors on a day with 40 mm of rain.  
Explain why this estimate may not be reliable.

The data only goes up to 17 mm of rain.

40 mm is well outside the range of the data.

The trend may not carry on, so the estimate may not be reliable.

**Answer:** 40 mm is outside the range of the data

**Tip:** Using a line of best fit outside the range of the data is called extrapolation. It can give unreliable results.

- (c) The manager says,  
*“The scatter graph proves that rain makes people stay away from the zoo.”*  
Is the manager right?  
Give a reason for your answer.

No.

The graph shows a link (correlation) between rainfall and the number of visitors.

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

Other things, such as the time of year, could also affect the number of visitors.

**Answer:** No, correlation does not prove causation

**Tip:** A scatter graph can show that two things are linked. It cannot prove that one causes the other.

**6** (2 marks)

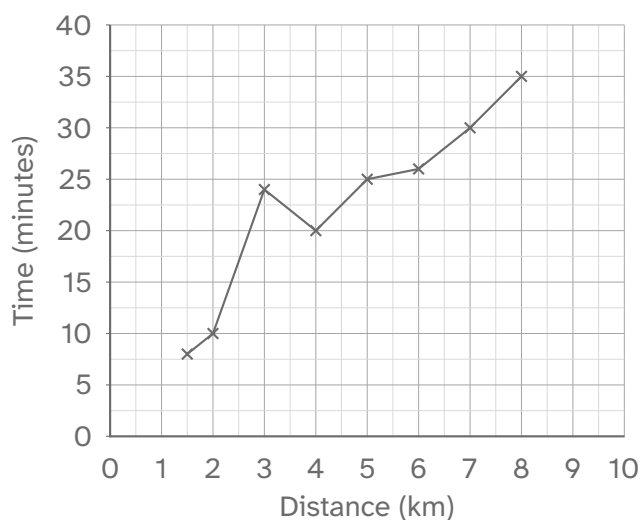
Mia asks eight students how far they live from school and how long it takes them to get to school.

The table shows her results.

| Student        | A   | B  | C  | D  | E  | F  | G  | H  |
|----------------|-----|----|----|----|----|----|----|----|
| Distance (km)  | 1.5 | 2  | 3  | 4  | 5  | 6  | 7  | 8  |
| Time (minutes) | 8   | 10 | 14 | 20 | 25 | 26 | 30 | 35 |

Mia draws a scatter graph for this information.

Here is her answer.



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

Check each point against the table.

1. Student C (3 km, 14 minutes) has been plotted at (3,24) instead of (3,14).

Check the line.

2. Mia has joined the points with straight lines, dot to dot.

She should draw a single straight line of best fit through the middle of the points.

**Answer:** Student C plotted at (3, 24) instead of (3, 14); points joined dot to dot instead of a line of best fit

**Tip:** Be precise. Say which point is wrong and where it should be, e.g. “(3, 14) has been plotted at (3, 24)”.

*Marks:* 1 mark for each correct mistake.

---

**7** (4 marks)

The scatter graph (shown in part (a) with a line of best fit) shows the age and the height of 10 trees of the same type in a park.

(a) Josh says,

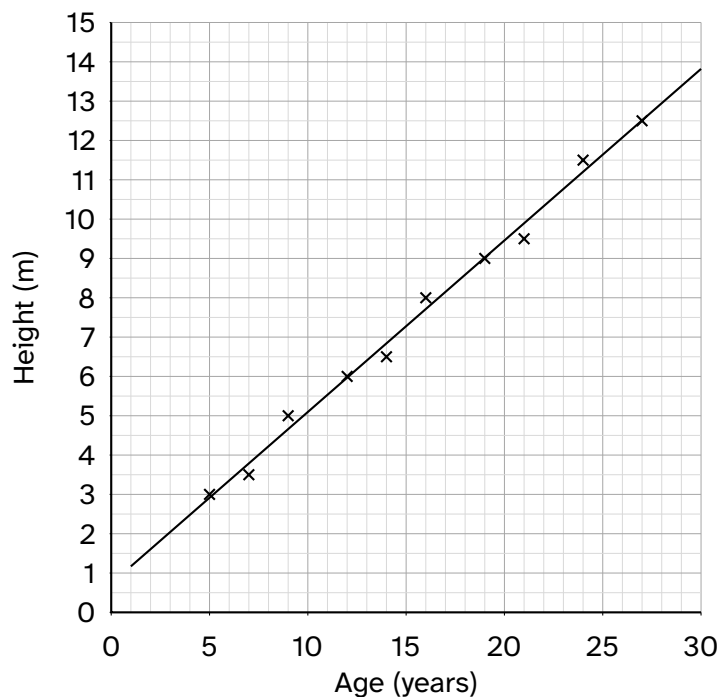
*“On average, these trees grow 1 metre taller each year.”*

Is Josh right?

Show how you get your answer.

Draw a line of best fit.

Pick two points on the line that are far apart, e.g. (10, 5.1) and (30, 13.8).



Change in height:  $13.8 - 5.1 = 8.7$  m

Change in age:  $30 - 10 = 20$  years

Increase per year:  $8.7 \div 20 = 0.435$  m

This is about 0.4 or 0.5 m each year, not 1 m. So Josh is not right.

**Answer:** No: the trees grow about 0.44 m each year (accept 0.35 to 0.55)

**Tip:** Use two points on the line that are far apart. Do not just compare two crosses next to each other.

*Marks:* 1 mark for a line of best fit or two readings from it. 1 mark for change in height  $\div$  change in age. 1 mark for saying No, with a value from 0.35 to 0.55.

- (b) Another tree of the same type is 12 years old.  
Use a line of best fit to estimate its height.

Read up from 12 years to the line of best fit, then across to the Height axis.

This gives about 6 m.

**Answer:** about 6 m (accept 5.5 to 6.5)