

Exercise walk-through

E9.5

Scatter diagrams

IGCSE Mathematics

You must be able to

- describe the type and strength of correlation shown by a scatter diagram
- use a line of best fit to estimate a value, and say when that estimate is unreliable
- explain that correlation does not prove one variable causes the other

1

The method

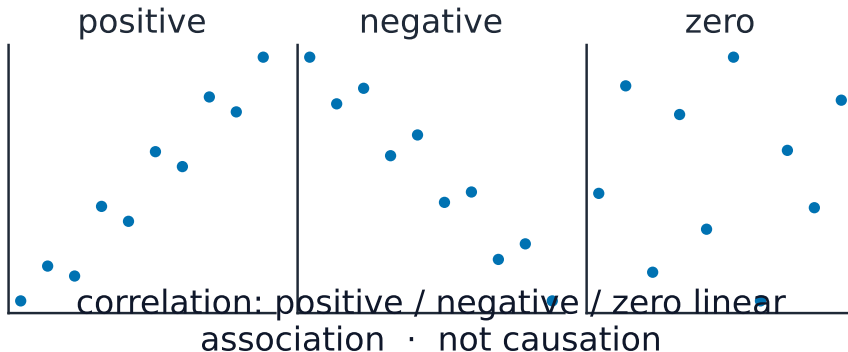
Method — Name the type, then the strength

- **Positive correlation:** as one goes up, so does the other. **Negative:** as one goes up, the other goes down. **No correlation:** no pattern.
- **Strength:** “strong” when the points lie close to a line, “weak” when they are scattered. A full answer gives BOTH words — “strong positive”.
- **Line of best fit:** drawn through the middle of the points, with roughly as many above as below. Use it to read an estimate off the graph.
- **Two warnings:** reading far outside the data (**extrapolation**) is unreliable because the pattern may not continue; and correlation does NOT prove cause — a third factor may explain both.

Answer shapes (marked by **method marks** — M1 for a correct method, A1 for the correct answer): these are B marks for the right words, and a “why” question always wants a reason, not a restatement.

Method — Name the type, then the strength

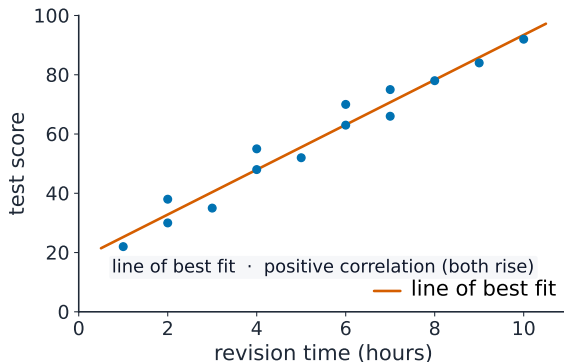
Answer shapes (marked by method marks — M1 for a correct method, A1 for the correct answer): these are B...



Scatter diagrams showing positive, negative and no correlation

Method — Name the type, then the strength

Answer shapes (marked by method marks — M1 for a correct method, A1 for the correct answer): these are B...



A scatter diagram with a line of best fit drawn

2

Practice

Practice 2.1 ■ 1 mark

On a scatter diagram the points rise from the bottom left to the top right. Describe the correlation.

Solution 2.1

Method: Name the type, then the strength — p.4

Worked steps

- as one variable increases so does the other

positive correlation

B1

Practice 2.2 ■ 1 mark

State what a line of best fit is used for.

Solution 2.2

Method: Name the type, then the strength — p.4

Worked steps

- to estimate the value of one variable from a known value of the other **B1**

Practice 2.3 ■ 1 mark

State one reason why an estimate read far outside the range of the data may be unreliable.

Solution 2.3

Method: Name the type, then the strength — p.4

Worked steps

- the pattern shown by the data may not continue outside the range that was measured — this is extrapolation **B1**

3

Exam-style

Exam-style 3.1 ■ 6 marks

A scatter diagram shows the number of hours of revision and the test score for 12 students. The points lie close to a line rising to the right.

Exam-style 3.1 (a) ■ 2 marks

Describe the correlation, giving both its type and its strength.

Solution 3.1 (a)

Method: Name the type, then the strength — p.4

Worked steps

- rising to the right is positive **B1**
- the points lying CLOSE to the line makes it strong, so: **strong positive correlation** **B1**

Exam-style 3.1 (b) ■ 2 marks

Explain how a line of best fit could be used to estimate the score of a student who revised for 6 hours.

Solution 3.1 (b)

Worked steps

- find 6 hours on the horizontal axis and read straight up to the line of best fit
B1
- then read across to the score axis; the estimate is the value there B1

Exam-style 3.1 (c) ■ 2 marks

Explain why this scatter diagram does not prove that revising causes a higher score.

Solution 3.1 (c)

Worked steps

- correlation shows only that the two vary together, not that one produces the other **B1**
- a third factor could explain both — for example a student who is more motivated may both revise more AND pay more attention in class **B1**

Exam-style 3.2 ■ 2 marks

State, with a reason, the type of correlation you would expect between the age of a car and its value.

Solution 3.2

Method: Name the type, then the strength — p.4

Worked steps

- **negative correlation** (B1)
- as a car gets older its value generally falls, so the two move in opposite directions (B1)