

C9.5 Scatter diagrams

Name: _____ Class: _____ Date: _____

Total: 20 marks

KEY WORDS correlation 相关性 • positive correlation 正相关 • scatter diagram 散点图
 • line of best fit 最佳拟合线 • negative correlation 负相关

Objective

Use scatter diagrams to solve Core IGCSE Mathematics problems.

You must be able to:

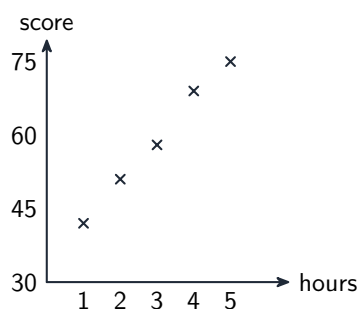
- Plot paired observations and identify correlation
- Draw a balanced by-eye best-fit line
- Distinguish interpolation, extrapolation and a causal claim

1 Key knowledge and worked examples

■ Key knowledge: Scatter diagrams

A scatter diagram plots paired observations with small crosses; do not join consecutive points. Positive correlation rises overall, negative falls, zero has no clear pattern. Draw a single ruled line of best fit by eye across the whole data range, with roughly balanced points on either side. It need not pass through the origin or a data point. Use it to estimate within the observed range; extrapolation outside is less reliable. Correlation does not establish cause.

■ Worked example: Scatter diagrams



Paired study hours and scores are (1, 42), (2, 51), (3, 58), (4, 69), (5, 75).

- Plot each pair as a cross. Scores generally increase: positive correlation.
- A reasonable best-fit line is $s = 8.5h + 34$, giving $s = 8.5(3.5) + 34 = 63.75$, about 64 marks.

- This is an estimate for these data, not a guarantee for a student. Extending to 12 hours goes far outside the sample.
- A line of best fit need not go through a data point. For five observations, judge the balance across the full range rather than forcing two selected points to define it.

■ Method check

Keep each pair together when plotting. A fitted line is an estimate for the observed relationship, not proof that one variable causes the other.

2 Practice

2.1 Name the correlation when points tend to fall from left to right. [2]

2.2 Explain why scatter points should not be joined in order. [2]

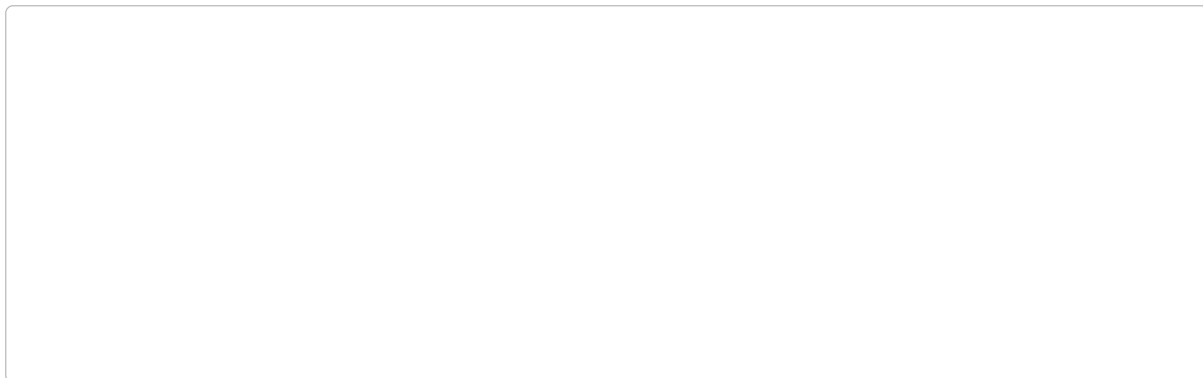
2.3 A best-fit line is $y = 2x + 10$. Estimate y at $x = 7$. [2]

2.4 Observations cover $2 \leq x \leq 8$. Which estimate is usually safer: $x = 5$ or $x = 20$? Explain. [2]

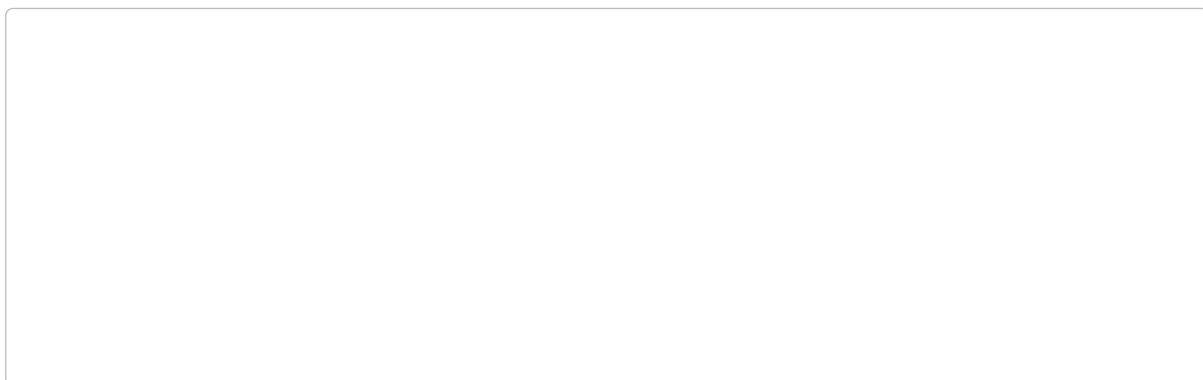
3 Exam-style questions

3.1 Apply the scatter diagrams methods below. Show working and include units where appropriate.

- (a) Plot crosses for $(1, 12)$, $(2, 15)$, $(3, 19)$, $(4, 20)$, $(5, 24)$ and state the correlation. [2]



- (b) Draw a line of best fit and estimate y at $x = 3.5$. [2]



- (c) Explain why this diagram does not show that increasing x causes y to increase. [2]

3.2 A scatter diagram records room temperatures from 15°C to 25°C and drink sales. A by-eye best-fit line is $s = 4T - 20$, where T is temperature and s is sales.

(a) Use the line to estimate sales at 22°C. Find the temperature corresponding to an estimate of 60 sales. [3]

(b) Evaluate a prediction at 35°C and the claim that raising the room temperature must increase sales. Give reasons using the data range and meaning of correlation. [3]

Solutions

Award method and answer marks as indicated; accept equivalent correct methods.

2.1

- The variables move oppositely: **negative correlation**.

2.2

- **Each point is a separate paired observation**; connecting them suggests a sequence the data may not have.

2.3

- $y = 2(7) + 10 = \mathbf{24}$, a line-based estimate.

2.4

- $x = 5$ is within the data; $x = 20$ extrapolates and the trend may change.

3.1

(a)

- **Five correctly plotted crosses; positive correlation**.

(b)

- A reasonable line is $y = 3x + 9$, so $y = 19.5$. Accept **about 19 to 21** from a sensible balanced line.

(c)

- **Another variable may affect both quantities; the diagram shows association only**.

3.2

(a)

- Substitute $T = 22$: $s = 4(22) - 20 = \mathbf{68}$.
- For 60 sales, $60 = 4T - 20$, so $T = \mathbf{20^\circ C}$.

(b)

- The line predicts $s = 4(35) - 20 = \mathbf{120}$, but 35°C is outside the observed range, so this is extrapolation.
- A relationship can change outside that range.
- Correlation does not prove causation; other factors, such as visitor numbers, may affect both variables.