



eClassroom

GCSE MATHEMATICS • DATA & STATISTICS

Scatter Graphs

Worked Solutions

Pearson Edexcel GCSE (1MA1) & International GCSE (4MA1) • Calculator allowed

Full worked solutions for the free Scatter Graphs worksheet, GCSE and iGCSE Maths



Tap a section to jump straight to it

Full worked solutions for every question on the Scatter Graphs worksheet, in the same order and with the same question numbers.

How to mark your work. Tick each step you got right, not just the final answer. In the exam, method marks are given for correct working even when the final answer is wrong. If you lost marks, find the *first* line where your working is different from ours. That is the thing to fix.

In Section E: **M** = method mark, **P** = process mark, **A** = accuracy mark, **B** = independent mark, **C** = communication mark.

Question sheet

Data & Statistics revision engine

Free courses

More worksheets



Join students already acing their exams.

Book your FREE lesson at www.eclassroom.org.uk

A**Your turn answers**

Compare each line of your working with the solution. Estimates from a line of best fit can differ slightly: any value in the stated range is fine.

FOUNDATION

A1 **Positive** correlation: taller people tend to have a longer arm span.

A2 Go across from $y = 30$ to the line, then down: $x = 10$

A3 $(4, 70)$: every other point follows the downward trend.

HIGHER

A4 Gradient = $\frac{40 - 90}{10 - 0} = -5$, and the line crosses the y -axis at 90 $y = -5x + 90$

A5 12 more lollies are sold for each extra degree. At 0°C the model gives -50 , which is impossible: the model should not be used far outside the data.

A6 x : $25 \div 5 = 5$ y : $150 \div 5 = 30$ $(5, 30)$

A7 **No**: 30 is far outside the ages 4 to 12 used for the line (extrapolation), and children stop growing, so the trend does not continue.

B**Fluency solutions**

One mark for each correct part. The notes in italics explain why each sequence was chosen.

FOUNDATION

B1 (a) **positive** (b) **negative** (c) **none** (d) **negative**

B2 (a) **8** (b) $(8, 9)$ (c) $(9, 2)$ (d) **positive**

B3 (a) **40** (b) **55** (c) **6** (d) **No**: the data only goes up to $x = 10$ *Up then across to read y ; across then down to read x .*



- B4** (a) **No**: bigger towns have more people, so both more doctors and more deaths (b) **Yes**: it describes a correlation, not a cause
(c) **No**: hot sunny weather causes both (d) **No**: they may be related in a way that is not a straight-line trend

- B5** (a) **positive** (b) about 16.5 (accept 15 to 18) (c) about 8.8 (accept 8 to 10)

- B6** (a) **strong positive** (b) **weak negative** (c) **no correlation** (d) **strong negative**

Strong: the points lie close to a line. Weak: the trend is there but the points are spread out.

HIGHER

- B7** (a) $y = 5x + 10$ (b) gradient 3, $8 = 6 + c$: $y = 3x + 2$ (c) $y = -3x + 50$ (d) gradient -2.5 , $30 = -10 + c$:
 $y = -2.5x + 40$

- B8** (a) each extra km adds 2 minutes; 20 minutes for starting and stopping (b) each hour of gaming lowers the mark by 4; about 80 with no gaming
(c) a child grows about 6 cm a year; about 50 cm at birth (d) the car loses about £1500 a year; worth about £20 000 when new

- B9** (a) $(3, 9)$ (b) $(5, 15)$ (c) $(30, 9)$ (d) $(3, 30)$

- B10** (a) **23** (b) **35** (c) $3x + 8 = 29$: **7** (d) 40 is far outside the data, so the prediction is not reliable

C

Reasoning solutions

Your wording can be different. Check that your reason makes the same point.

FOUNDATION

- C1** Positive correlation means there is an upward trend: as one variable increases the other tends to increase. The points do not have to lie on a line.
- C2** The first and last points may be unusual. A line of best fit should follow the trend of all the points, with about as many above it as below.
- C3** No. Correlation does not show cause. Hot weather brings more people to the beach, which means more ice-cream sales and more people to see sharks.



C4 A line of best fit should follow the trend of the points. Forcing it through (0, 0) can pull it away from the data.

HIGHER

C5 $x = 30$ is far outside the data (extrapolation). The trend may not continue beyond the data.

C6 The mean point is the centre of the data. A line through it balances the points above and below it.

C7 Measure how far each point is from each line (vertically). The better line has points closer to it overall, with points balanced above and below.

C8 (a) The mean point lies on the line, so $\bar{y} = m\bar{x} + c$

(b) The sum is $\sum y_i - m \sum x_i - nc = n\bar{y} - mn\bar{x} - nc = n(\bar{y} - m\bar{x} - c) = 0$, using part (a)

D**Problem solving solutions**

There is often more than one correct method. Any complete, correct method scores full marks.

FOUNDATION

D1 From the line: about **61** marks. Reliable, because 6 hours is within the data and the points are close to the line.

D2 (a) **positive** correlation: students with higher maths marks tend to have higher science marks

(b) about 56 (accept 51 to 60)

(c) It is a trend, not a rule; some students may not follow it, and correlation does not guarantee every case

D3 For example: it is a rare or classic model; it has very low mileage or is in excellent condition.

D4 **Graph P**: its points lie much closer to a straight line.



HIGHER

D5 (a) The line passes through (0, 12) and (10, 52): gradient $\frac{40}{10} = 4$ $y = 4x + 12$

(b) $4 \times 7 + 12 = 40$

D6 (a) The line gives $4 \times 5 + 15 = 35$; $35 - 30 = 5$, not more than 8, so **not an outlier**

(b) The line gives $4 \times 8 + 15 = 47$; $60 - 47 = 13$, so it **is an outlier**

D7 $2w + 10 = 3w + 4$, so $w = 6$. After that B is taller, because its gradient (3 cm a week) is larger.

D8 (a) $\bar{x} = \frac{2+4+6+8+10}{5} = 6$. The line passes through $(\bar{x}, \bar{y})$, so $\bar{y} = 2 \times 6 + 5 = 17$

(b) The y-values add up to $5 \times 17 = 85$: $9 + 13 + k + 21 + 27 = 85$ $k = 15$

(c) The line gives 17 at $x = 6$, and $17 - 15 = 2$ units

E

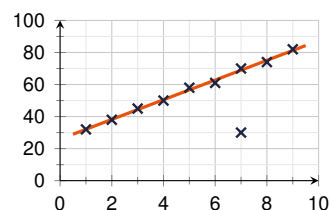
Exam practice solutions

Marks are shown on the right of each line. Any value in the stated range scores the mark for an estimate.

FOUNDATION

Foundation and Higher crossover

- E1 (a) **positive**
 (b) **(7, 30)**
 (c) a straight line following the trend, ignoring the outlier
 (d) about 60 (accept 55 to 64)



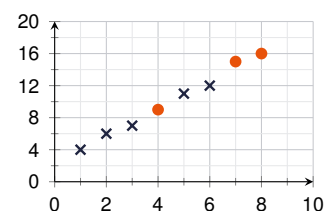
B1 B1 B1 B1

- E2 20 hours is far outside the data, which only goes up to 9 hours
 The trend may not continue: the mark cannot go above 100

C1

C1

- E3 (a) Points at (4, 9), (7, 15), (8, 16)
 (b) **positive** correlation: as x increases, y increases



B2 B1

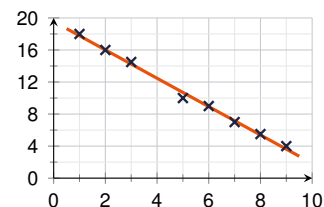


Join students already acing their exams.

Book your FREE lesson at www.eclassroom.org.uk

- E4** (a) **positive** B1
 (b) Correlation does not show cause C1
 Bigger towns have more people, so more coffee shops and more cars C1

- E5** (a) a line following the downward trend
 (b) about £12 500 (accept £11 000 to £14 000)
 (c) No: 15 years is outside the data, and the value cannot go below 0



B1 B1 C1

HIGHER

- E6** (a) Gradient = $\frac{70 - 20}{10} = 5$ M1
 $M = 5h + 20$ A1
 (b) Each extra hour of revision is linked to about 5 more marks C1

- E7** (a) $38 = 4 \times 6 + c$, so $c = 14$ M1
 $y = 4x + 14$ A1
 (b) $4 \times 9 + 14 = 50$ B1

- E8** (a) $(9, 10)$ B1
 (b) $x: 35 \div 7 = 5$; $y: 126 \div 7 = 18$ M1
 $(5, 18)$ A1
 (c) $x = 5.5$: it is inside the data (interpolation), while 20 is far outside it C1

- E9** (a) Gradient = $\frac{54 - 14}{10 - 2} = 5$; $14 = 10 + c$ M1
 $y = 5x + 4$ A1
 (b) $35 + 4 = 39$ B1
 (c) When $x = 2$, $y = 14$; when $x = 4$, $y = 24$, which is not $2 \times 14 = 28$ M1
 The +4 means y is not proportional to x , so doubling x does not double y C1

