A-LEVEL MATHS (EDEXCEL 9MA0) · STATISTICS YEAR 1 · CHAPTER 4: CORRELATION · A* EXAM PRACTICE

Hints online, answers free with an account:
mathswariz.uk/worksheets/a-level-statistics-year-1-chapter-4-a-star/

Correlation: A* questions

74 marks

Date

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

X A* exam questions

A to A*

Extended Edexcel-style questions. Show every step. The hints are at the end of the sheet. Use one at a time.

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

X1 Two varieties of sunflower, A and B, are grown in a trial. For variety A the height, h cm, was recorded each week from week 2 to week 12, and the regression line of h on the number of weeks, w, is $h = 4.0 + 2.6w$. For variety B the height was recorded from week 0 to week 8, and the regression line of h on w is $h = 12.5 + 1.6w$. Interpret the gradient of each line. Use the lines to estimate the week in which the two varieties have the same height, and that height, and comment on the reliability of your estimate. Ravi wants to estimate the week in which variety A reaches 30 cm by rearranging the line for A. Comment on his method. [10 marks]

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

X2 A stone is thrown upwards from the end of a jetty. Its height above the water, h metres, is recorded every 0.5 seconds for 2.5 seconds. The results are in the table.

t (s)	0	0.5	1	1.5	2	2.5
h (m)	1.8	7.4	10.6	11.4	9.6	5.5

(a) Use your calculator to find the regression line of h on t. [2 marks]

Answer .....

(b) With reference to the data, explain why this linear model is not suitable. [3 marks]

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

(c) Jane suggests the model $h = c - 4.9(t - k)^2$. Using the points $(0, 1.8)$ and $(2, 9.6)$, find the values of k and c, giving each to 3 significant figures. [4 marks]

Answer .....

(d) Use Jane's model to estimate the time at which the stone reaches the water, and comment on the reliability of your estimate. [3 marks]

Total for X2: 12 marks

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

X3 Ava uses the large data set to study Camborne from May to October 2015 (184 days). She takes a sample of 12 days and records the daily total sunshine, s hours, and the daily mean temperature, T °C. (The sample values below are for this question only; they are not taken from the large data set.)

(a) Describe how Ava could take a simple random sample of 12 days. [2 marks]

(b) For her sample, $\sum s = 66.0$ and $\sum s^2 = 513.6$. The greatest value of s is 13.4. An outlier is a value more than 2 standard deviations from the mean. Show that 13.4 is an outlier. [4 marks]

Answer .....

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

(c) Ava says she will remove the value 13.4 before drawing her scatter diagram. Explain whether she should. [2 marks]

(d) The regression line of T on s is $T = 13.2 + 0.46s$. Interpret the value 0.46, and estimate T on a day with 8 hours of sunshine. [2 marks]

Answer .....

(e) Ava concludes: "More sunshine causes higher temperatures at every weather station in the large data set." Give two reasons why this conclusion is not valid. [2 marks]

Total for X3: 12 marks

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

X4 Five pairs of data are $(1, 3)$, $(2, 5)$, $(3, 4)$, $(4, 7)$ and $(5, 8)$. Bea draws a line by eye: $y = 1.5 + 1.3x$. A residual is the observed y minus the y given by a line.

(a) Find the sum of the squares of the residuals for Bea's line. [3 marks]

Answer .....

(b) Consider lines through $(\bar{x}, \bar{y})$ with gradient m. Show that the sum of the squares of the residuals is $S = 10m^2 - 24m + 17.2$. [3 marks]

Answer .....

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

(c) Find the value of m that makes S smallest, and hence the least squares regression line of y on x. Check it with your calculator. [3 marks]

Answer .....

(d) Explain, using the least squares idea, which line fits the data better. [2 marks]

Total for X4: 11 marks

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

X5 A kiosk records the noon temperature, x °C, and the number of drinks sold, y, on 9 days. The results are in the table.

x (°C)	14	16	17	19	20	22	23	25	26
y	112	128	131	146	15.4	165	174	190	197

(a) One value of y does not follow the pattern of the others. Identify it and suggest a reason for it. [2 marks]

(b) Use your calculator to find the regression line of y on x for all 9 days, and for the 8 days without the value in part (a). [2 marks]

Answer .....

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

(c) Estimate the number of drinks sold on a day when the noon temperature is 21 °C using each line, and say which estimate is more reliable. [3 marks]

(d) The two gradients are close, but the estimates differ by about 15 drinks. Explain why. [2 marks]

Total for X5: 9 marks

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

X6 A researcher records the speed, x mph, and the fuel economy, y miles per gallon, of 5 test runs of a small car (S) and 5 test runs of a large car (L). The results are in the table. Use your calculator to find the regression line of y on x for the small car, for the large car, and for all 10 runs together. Compare the three gradients, explain why they differ, and say which line or lines should be used to describe how fuel economy changes with speed. [10 marks]

Small car: x (mph)	y (mpg)	Large car: x (mph)	y (mpg)
30	41	50	22
32	42	52	24
34	44	54	25
36	45	56	27
38	47	58	28

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

X7 A café owner finds, over 50 days with 40 to 120 customers a day, that the regression line of daily takings, £R, on the number of customers, c, is $R = 140 + 6.8c$. Her daily costs, £K, are modelled by $K = 180 + 1.2c$.

(a) Find a model for the daily profit, £Q, in terms of c, and interpret its gradient. [3 marks]

Answer .....

(b) Estimate the profit on a day with 80 customers. [2 marks]

Answer .....

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

(c) Find the increase in daily profit when the number of customers rises from 60 to 90. [2 marks]

Answer .....

(d) The owner says: "On a quiet day with 20 customers, I still make a profit." Use the model to test this, and comment on the reliability of the model here. [3 marks]

Total for X7: 10 marks

F How did it go?

Colour one circle on each line.

	Not yet	Nearly	Yes
I can compare two regression lines, decide when an estimate is extrapolation, and say which variable a line can estimate.	○	○	○
I can criticise a linear model using the data, and fit a better model through given points.	○	○	○
I can explain what least squares means and how an anomalous point changes a regression line.	○	○	○
I can explain when a combined line is misleading and how to combine regression models.	○	○	○

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

Hints

Use one hint at a time. Cover the ones below it with a sheet of paper.

X1	**Hint 1**	A gradient is the change in h for each extra week. Give the units.
	Hint 2	The two heights are equal when $4.0 + 2.6w = 12.5 + 1.6w$. Then check w against the weeks each variety was measured.
	Hint 3	Which variable did the researcher choose, and which was measured? A regression line of h on w is for estimating h.
X2	**Hint 1**	Enter t in the first list and h in the second. For (b) compare the data with the line at the ends and in the middle.
	Hint 2	Put each point into $h = c - 4.9(t - k)^2$ and subtract to remove c. Expand $(2 - k)^2$.
	Hint 3	The stone reaches the water when $h = 0$. Take the root greater than k, and compare it with the times in the data.
X3	**Hint 1**	For (a), every day must have an equal chance. For (b), use $\sigma^2 = \dfrac{\sum s^2}{n} - \bar{s}^2$.
	Hint 2	For (c), think about how long the days are at Camborne in summer.
	Hint 3	For (e), think about cause, and about how much of the large data set the sample covers.
X4	**Hint 1**	Work out the value of $1.5 + 1.3x$ at each x, then subtract it from the observed y.
	Hint 2	Write each residual as $(y_i - \bar{y}) - m(x_i - \bar{x})$, square, and add the three kinds of term separately.
	Hint 3	S is a quadratic in m. Differentiate with respect to m and set the result equal to 0.
X5	**Hint 1**	Look at the size of each y value compared with the others near it.
	Hint 2	Enter all 9 pairs first, then delete the odd pair and find the line again.
	Hint 3	Find $\bar{x}$. Where is the odd point compared with the mean of x? Every regression line passes through $(\bar{x}, \bar{y})$.

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.

X6

Hint 1 Enter the five small-car pairs, find the line, then do the same for the large car.

Hint 2 Now enter all ten pairs. Is the gradient still positive?

Hint 3 Sketch the ten points: where is each group? A variable that is not x or y may explain the combined pattern.

X7

Hint 1 Profit is takings minus costs. Subtract the two models term by term.

Hint 2 The change in Q over 30 customers is 30 times the gradient, not the value of Q at 30.

Hint 3 Work out Q at $c = 20$, then compare 20 with the range of the data.

© 2026 Mathswariz. All rights reserved. Free for personal and classroom use. Not to be sold or republished.