Measuring correlation

87 marks

Hints online, answers free with an account: mathswariz.uk/worksheets/a-level-statistics-year-2-measuring-correlation/

Date

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

A Do Now

Quick questions to start: the language of correlation, the last lesson's coding, and your calculator.

A1 Needed today [2 marks]

For a sample of days, the PMCC between the hours of sunshine and the rainfall is -0.71. Describe the correlation, and interpret it in context.

A2 Last lesson [2 marks]

Data are coded with $y = \log P$ and $x = t$, and the coded points lie close to a straight line. State the form of the model for P in terms of t.

Answer .....

A3 Common slip [2 marks]

Rhys says, "The correlation between the age of a car and its value is negative, so as the age of the car goes down, its value goes down." Correct Rhys's statement.

A4 Needed today [2 marks]

There is strong positive correlation between the number of ice creams sold at a beach and the number of people treated for sunburn. Does buying ice cream cause sunburn? Give a reason.

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

A5 Needed today [2 marks]

Use your calculator to find the PMCC of the data, to 3 significant figures.

x	**1**	**2**	**3**	**4**	**5**
y	3	5	4	7	8

Answer .....

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

PART 1 Finding and interpreting the PMCC

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

EXAMPLE 1 · I DO

The age, x years, and the value, y thousand pounds, of eight cars of one model are shown in the table. (a) Find the PMCC between x and y. (b) Describe the correlation. (c) Interpret it in context.

x	1	2	3	4	5	6	7	8
y	18.5	16.2	14.9	12.1	11.4	9.8	8.1	7.6

Hint Give r to 4 decimal places unless told otherwise. A value close to -1 is strong negative.

YOUR TURN · WE DO

B1 The hours of revision, x, and the test mark, y, of eight students are shown in the table.

x	2	4	5	7	8	10	11	12
y	38	45	41	55	60	58	70	72

(a) Find the PMCC between x and y. [2 marks]

Answer .....

(b) Describe the correlation. [1 mark]

Answer .....

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

(c) Interpret the correlation in context. [2 marks]

PRACTICE · YOU DO

C1 A biologist believes that birds with longer wings tend to be heavier. For a sample of 20 birds of one kind, the PMCC between wing length and mass is 0.12. Interpret this value, and state whether it is consistent with the biologist's belief. [4 marks]

~~Total for B1: 5 marks~~

C2 For 12 days, the PMCC between the daily maximum temperature and the number of hours that the heating in a house is on is -0.87. Asha writes "negative correlation" as her interpretation. Ben writes, "As the temperature decreases, the heating is used less."

(a) Explain why Asha's answer does not score the mark for an interpretation. [2 marks]

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

(b) Explain the error in Ben's statement, and write a correct interpretation. [3 marks]

Total for C2: 5 marks

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

PART 2 Interpreting the PMCC

EXAMPLE 2 · I DO

For the data in the table, (a) find the PMCC, (b) explain why it does not show that there is no relationship between x and y.

x	1	2	3	4	5	6	7
y	9.2	4.1	0.8	0.2	1.1	3.9	9.1

Hint Sketch the points. A value of r close to 0 means no LINEAR relationship.

YOUR TURN · WE DO

B2 The speed, x km/h, and the fuel used, y miles per gallon, of a car are recorded on seven tests.

x	30	40	50	60	70	80	90
y	38	45	50	52	49	44	37

(a) Find the PMCC between x and y. [2 marks]

Answer .....

(b) Mia says, "r is close to 0, so the speed of the car does not affect the fuel used." Comment on Mia's statement. [2 marks]

Total for B2: 4 marks

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

EXAMPLE 3 · I DO

A student records the length, x cm, and the mass, y g, of eight plants. (a) Find the PMCC. (b) One value of y was copied wrongly. Identify it, find the PMCC without it, and comment.

x	3	5	6	8	9	11	12	14
y	12	15	17	20	22	25	4	30

Hint Look for the point that does not follow the others. The PMCC is very sensitive to one extreme point.

YOUR TURN · WE DO

B3 The table shows eight pairs of values from an experiment. One value of y was recorded wrongly.

x	3	5	6	8	9	11	13	14
y	17	20	22	25	9	31	34	35

(a) Find the PMCC for all eight points. [2 marks]

Answer .....

(b) Identify the point that was recorded wrongly, and find the PMCC without it. [2 marks]

Answer .....

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

(c) Comment on the effect of the wrong value. [1 mark]

Total for B3: 5 marks

PRACTICE · YOU DO

C3 The height, h cm, and the arm span, s cm, of nine adults are shown in the table. One value of s was recorded in metres by mistake. Correct the error, and find the PMCC between h and s. [3 marks]

h	**158**	**163**	**167**	**170**	**172**	**175**	**178**	**181**	**186**
s	155	164	165	169	1.74	176	177	183	188

Answer .....

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

C4 The values of the PMCC for the three scatter diagrams are 0.15, 0.99 and -0.58, in some order. Match each value to its diagram, and give a reason for each. [4 marks]



Diagram A

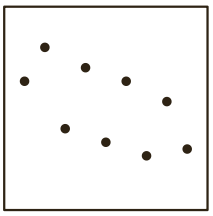

Diagram B

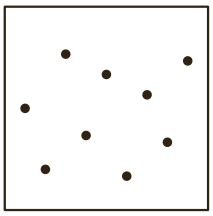

Diagram C

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

PART 3 Coding and choosing a model

EXAMPLE 4 · I DO

For a set of data on x and y, the PMCC is 0.73. The data are coded with $p = \dfrac{x - 20}{5}$ and $q = 100 - y$. Find the PMCC between p and q.

Hint Look at the number each variable is MULTIPLIED by. Only a negative multiplier changes r, and only its sign.

YOUR TURN · WE DO

B4 For a set of data on s and t, the PMCC is -0.58. The data are coded with $u = 3s + 40$ and $v = \dfrac{80 - t}{4}$.

(a) Find the PMCC between u and v. [2 marks]

Answer .....

(b) Find the PMCC between u and t. [2 marks]

Answer .....

Total for B4: 4 marks

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

EXAMPLE 5 · I DO

For the data in the table, find the PMCC between x and $\log y$, and between $\log x$ and $\log y$. Hence choose a model of the form $y = kb^x$ or $y = ax^n$, and find it.

x	1	2	3	4	5	6
y	2.1	3.6	5.6	9.9	14.8	25.6

Hint The better model is the coding whose PMCC is closer to ± 1.

YOUR TURN · WE DO

B5 The table shows six pairs of values of x and y.

x	2	4	6	8	10	12
y	8.1	24.9	43	69.8	93.1	127.5

(a) Find the PMCC between x and $\log y$, and between $\log x$ and $\log y$. [2 marks]

Answer .....

(b) State, with a reason, which model is more suitable. [2 marks]

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

(c) Write down the straight line that this model gives, in terms of $\log y$, $\log x$, a and n. [1 mark]

PRACTICE · YOU DO

C5 For a set of data, the PMCC between x and y is 0.64. Find the PMCC between (a) x and $\frac{y}{10}$, (b) $50 - x$ and $\frac{y}{10}$, (c) $50 - x$ and $20 - 2y$. Give a reason for (b). [4 marks]

Answer .....

Total for B5: 5 marks

Answer .....

C6 For a set of data, the PMCC between x and y is 0.962, between x and $\log y$ is 0.998, and between $\log x$ and $\log y$ is 0.981.

(a) State, with a reason, which of $y = kb^x$ and $y = ax^n$ is the better model. [2 marks]

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

(b) The regression line of $\log y$ on x is $\log y = 0.42 + 0.115x$. Find the model, giving constants to 3 significant figures. [3 marks]

Answer .....

Total for C6: 5 marks

D Exam-style questions

Edexcel exam style

D1 A plant that spreads quickly is studied. The area, A m², that it covers is measured at the end of week w.

w	1	2	3	4	5	6	7
A	3.0	4.8	6.2	10.1	13.2	20.9	28.4

(a) Find the PMCC between w and A, and between w and $\log A$. [2 marks]

Answer .....

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

(b) Explain why a model of the form $A = kb^w$ is suitable. [2 marks]

(c) Find the equation of the regression line of $\log A$ on w. [2 marks]

Answer .....

(d) Find the model, giving k and b to 3 significant figures. [3 marks]

Answer .....

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

(e) Interpret b in context. [2 marks]

Total for D1: 11 marks

D2

For 15 days, the PMCC between the daily mean temperature, T °C, and the daily mean pressure, P hPa, is -0.82, and the regression line of P on T is $P = 1031 - 1.4T$. Ella writes three statements. (1) "Higher temperatures cause the pressure to fall." (2) "The pressure falls by 0.82 hPa for every 1 °C rise in temperature." (3) "If I change the temperatures to °F, the PMCC will change."

Each of Ella's statements is wrong.

(a) Explain the error in statement (1), and write a correct statement. [3 marks]

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

(b) Explain the error in statement (2), and write a correct statement. [3 marks]

(c) Explain the error in statement (3). [3 marks]

Total for D2: 9 marks

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

E Extension

Stretch

No route is given. The hints are at the end of the sheet. Use one at a time.

E1 Two classes sit different tests. For each student, x is the number of hours of revision and y is the number of marks lost. Class P sat an easy test; class Q sat a much harder one. A teacher puts all ten students together and says, "The more students revise, the more marks they lose." Find the PMCC for class P, for class Q, and for all ten students. Explain how the teacher's data can give this result, and say what the data really show. [9 marks]

Class	x	y
P	2	14
P	3	13
P	4	12
P	5	11.5
P	6	10
Q	8	26
Q	9	25
Q	10	23
Q	11	22
Q	12	21

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

Extension hints

Cover these with paper. Uncover one at a time.

E1

Hint 1 Find the PMCC for each class on its own first. What sign does each one have?

Hint 2 Now look at where the two classes sit: class Q revised more AND lost more marks, because its test was harder.

Hint 3 The test (easy or hard) is a third factor. Within each class, what happens to the marks lost as revision goes up?

How did it go?

Colour one circle on each line: not yet, nearly, yes.

I can find the PMCC, r, on my calculator, and describe and interpret it in context. ○ ○ ○

I can explain why r close to 0 does not mean there is no relationship. ○ ○ ○

I can use the PMCC of coded data to choose between $y = kb^x$ and $y = ax^n$. ○ ○ ○

I can say how coding and a recording error change r. ○ ○ ○

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