4.2 Correlation & Regression
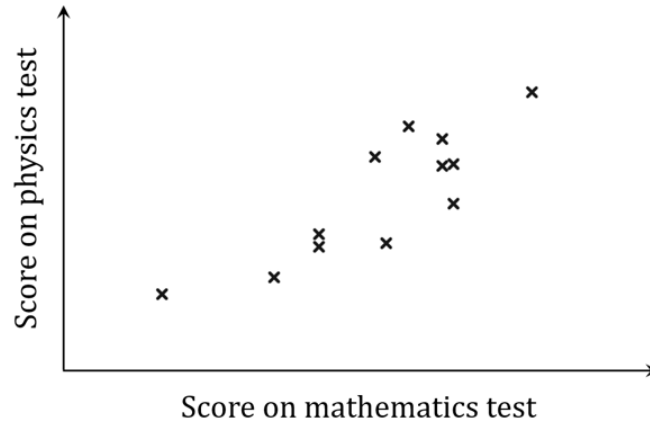
Question Paper

Course	DP IB Maths
Section	4. Statistics & Probability
Topic	4.2 Correlation & Regression
Difficulty	Medium

Time allowed: 80

Score: /65

Percentage: /100

Question 1a

A teacher collected the maths and physics test scores of a number of students and drew a scatter diagram to represent this data.
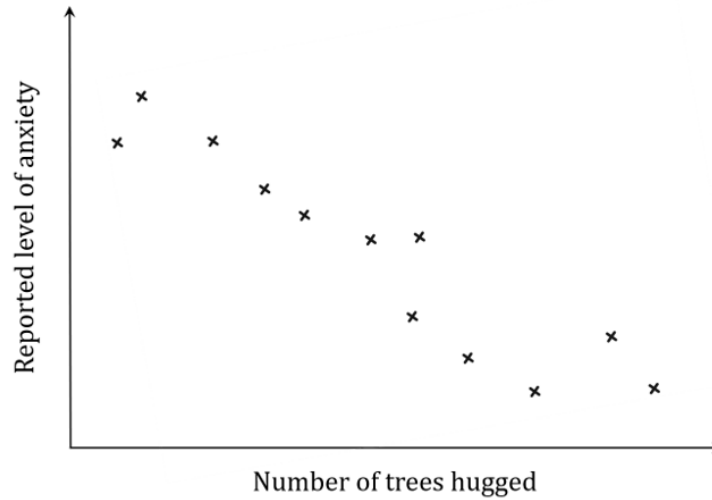

(a) Describe the correlation shown by the scatter diagram, and interpret the correlation in context.

[2 marks]

Question 1b

An alternative therapist collected data on his clients' reported levels of anxiety as well as the number of trees they had hugged in the course of therapy. He drew a scatter diagram to represent this data.

(b) Describe the correlation shown by the scatter diagram, and interpret the correlation in context.

[2 marks]

Question 2a

Jennifer sells cups of tea at her shop and has noticed that she sells more tea on cooler days. On five different days, she records the maximum daily temperature, T, measured in degrees Celsius, and the number of cups of teas sold, C. The results are shown in the following table.

Maximum Daily Temperature, T.	3	5	8	9	12
Cups of tea sold, C.	37	34	33	26	21

(a)   (i)   Write down the equation of the regression line of C on T.

      (ii)   Write down the value of the Pearson's product-moment correlation coefficient, r.

[4 marks]

Question 2b

(b) Use your regression equation from part (a)(i) to estimate the number of teas that Jennifer will sell on a day when the maximum temperature is 11°C.

[2 marks]

Question 2c

(c) Being sure to consider the result from part (a)(ii) in your answer, state how confident you would be in your estimate from part (b).

[2 marks]

Question 3a

The following table shows the mean height, y cm, of primary school children who are age x years old.

Age, x years	6.25	7.35	8.5	9.25	10.75
Mean Height, y cm	115	121	129	136	140

(a) (i) Write down the equation of the regression line of y on x.

(ii) Write down the value of the Pearson's product-moment correlation coefficient, r.

[4 marks]

Question 3b

(b) Use your regression equation from part (a)(i) to estimate the height of a child aged 9 years old.

[2 marks]

Question 3c

(c) Explain why it is not appropriate to use the regression equation to estimate the age of a child who is 133 cm tall.

[1 mark]

Question 4a

Rebecca, a regular jogger, ran the "Thao Dien Loop" on 7 consecutive days. The following table shows the distance, x km, that she ran and the corresponding number of calories, y, that she was able to burn during the run.

Distance (x)	2	5	6	7	10	12	14
Calories (y)	180	315	365	435	619	830	871

The number of calories burnt during a run is dependent upon on the length of the run.

(a) (i) Write down the equation of the regression line of y on x, giving your answer in the form $y = ax + b$ where a and b are constants to be found.

(ii) Write down the value of the Pearson's product-moment correlation coefficient, r.

[4 marks]

Question 4b

(b) Interpret, in the context of the question, the value of a found in part (a)(i).

[1 mark]

Question 4c

On the 8th day, Rebecca is only able to run for 8 kilometres.

(c) Use the result from part (a)(i) to estimate the number of calories Rebecca will lose.

[2 marks]

Question 4d

(d) Comment on the validity of using the result from part (a)(i) answer part (c).

[1 mark]

Question 5a

The percentage of people who are willing to get a particular vaccine is dependent on their age. The following table shows the age, A years old, and the corresponding percentage of people, V, that are willing to receive a vaccine for 6 different ages.

Age, (A)	25	30	35	40	45	50
Percentage of willing people, (V)	57	59	61	62	68	75

(a) (i)   Write down the equation of the regression line of V on A, giving your answer in the form $V = aA + b$ where a and b are constants to be found.

(ii)   Write down the value of the Pearson's product-moment correlation coefficient, r.

[4 marks]

Question 5b

(b) Interpret, in the context of the question, the value of a found in part (a)(i).

[1 mark]

Question 5c

(c) Use the result from part (a)(i) to estimate the percentage of people aged 95 years old who are in willing to receive a vaccine.

[2 marks]

Question 5d

(d) Comment on the validity of using the result from part (a)(i) to answer part (c).

[1 mark]

Question 6a

The price, P, of an airline ticket is dependent on the distance, d km, between two cities. The table below shows the airfares in US dollars from Prague in the Czech Republic, to eight different destinations in Europe.

Distance (d)	885	340	835	330	1270	295	650	1930
Price (P)	2	50	90	45	119	42.5	59	139

The relationship between d and P can be modelled by the regression line of P on d with equation $P = ad + b$.

(a) (i) Write down the equation of the regression line of P on d

(ii) Write down the value of the Pearson's product-moment correlation coefficient, r.

[4 marks]

Question 6b

(b) Use the result from part (a) to estimate the price of an airline ticket for a flight from Prague to a destination that is 2635 km away.

[2 marks]

Question 6c

Madlenka buys a ticket for a flight from Prague to Cairo, a distance of 2635 km. The airfare in US dollars is $245.

(c) Compare this price to your result from part (c), suggesting possible reasons for any discrepancies.

[2 marks]

Question 7a

A snack company is trialling a new series of crisp flavours. The eight flavours have been assigned the letters A through H, and the company has conducted taste tests in which volunteers assign each flavour a score on a scale from 1 to 10 (where '1' indicates 'I hate it!', and '10' indicates 'This is my new favourite!'). The following table collates the scores assigned to the flavours by each of three volunteers – Idris, Jameel and Kevin.

Flavours	A	B	C	D	E	F	G	H
Idris' score	1	9	4	8	10	3	7	6
Jameel's score	6	4	2	10	7	9	3	8
Kevin's score	9	4	7	2	1	7	5	6

The company would like to find the Spearman's rank correlation coefficients for these taste testers' rankings.

(a) Complete the information in the following table.

Flavours	A	B	C	D	E	F	G	H
Idris' rank	1				8			
Jameel's rank			1	8				
Kevin's rank	8				1	6.5		

[3 marks]

Question 7b

(b) Find the value of the Spearman's rank correlation coefficient, r_s, for:

     (i)   Idris' and Jameel's ranks

     (ii)   Idris' and Kevin's ranks

     (iii) Jameel's and Kevin's ranks.

[3 marks]

Question 7c

(c) Comment, in the context of the question, on the results obtained for r_s in part (b)(i), (ii) and (iii).

[3 marks]

Question 7d

It is noticed that a greasy fingerprint has caused Kevin's score of 9 for flavour A to be misread. The score actually assigned by Kevin to flavour A was 8.

(d) Explain, with a reason, whether this will change any of the values for the Spearman's rank correlation coefficient r_s that were calculated in part (b).
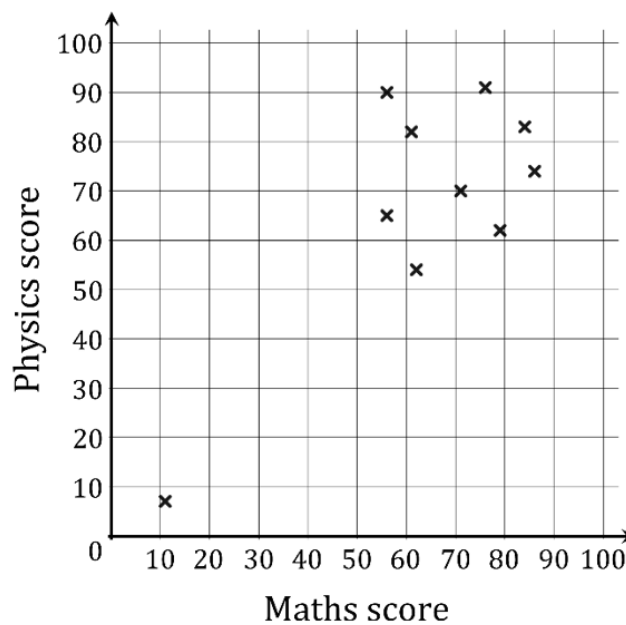
[2 marks]

Question 8a

The following table collates the scores achieved on a recent maths test and a recent physics test by a group of 10 students.

Student	A	B	C	D	E	F	G	H	I	J
Maths score	62	79	71	56	61	84	86	76	56	11
Physics score	54	62	70	65	82	83	74	91	90	7

The scatter diagram for these scores is shown in the diagram below:

(a) Write down the value of the Pearson's product-moment correlation coefficient, r,

    (i)   with student J included

    (ii)   without student J included.

[3 marks]

Question 8b

(b) Complete the information in the following table, ranking the scores from highest to lowest:

Student	A	B	C	D	E	F	G	H	I	J
Maths rank			5	8.5			1			10
Physics rank							5	1		10

[2 marks]

Question 8c

(c) Find the value of the Spearman's rank correlation coefficient, r_s,

    (i)   with student J included

    (ii)   without student J included.

[3 marks]

Question 8d

(d) Comment on the results of parts (a) and (c).

[3 marks]