

Name _____

Date _____

Bivariate Statistics

1.

The following table shows the average weights (y kg) for given heights (x cm) in a population of men.

Heights (x cm)	165	170	175	180	185
Weights (y kg)	67.8	70.0	72.7	75.5	77.2

- (a) The relationship between the variables is modelled by the regression equation $y = ax + b$.
- (i) Write down the value of a and of b .
- (ii) Hence, estimate the weight of a man whose height is 172 cm. [4]
- (b) (i) Write down the correlation coefficient.
- (ii) State which **two** of the following describe the correlation between the variables. [3]

strong zero positive
negative no correlation weak

2.

The following table shows the amount of fuel (y litres) used by a car to travel certain distances (x km).

Distance (x km)	40	75	120	150	195
Amount of fuel (y litres)	3.6	6.5	9.9	13.1	16.2

This data can be modelled by the regression line with equation $y = ax + b$.

- (a) (i) Write down the value of a and of b .
- (ii) Explain what the gradient a represents. [3]
- (b) Use the model to estimate the amount of fuel the car would use if it is driven 110 km. [2]

3.

The following table shows the Diploma score x and university entrance mark y for seven IB Diploma students.

Diploma score (x)	28	30	27	31	32	25	27
University entrance mark (y)	73.9	78.1	70.2	82.2	85.5	62.7	69.4

(a) Find the correlation coefficient. [2]

The relationship can be modelled by the regression line with equation $y = ax + b$.

(b) Write down the value of a and of b . [2]

Rita scored a total of 26 in her IB Diploma.

(c) Use your regression line to estimate Rita's university entrance mark. [2]

4.

The following table shows the number of bicycles, x , produced daily by a factory and their total production cost, y , in US dollars (USD). The table shows data recorded over seven days.

	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
Number of bicycles, x	12	15	14	17	20	18	21
Production cost, y	3900	4600	4100	5300	6000	5400	6000

- (a) (i) Write down the Pearson's product-moment correlation coefficient, r , for these data.
- (ii) Hence comment on the result. [4]
- (b) Write down the equation of the regression line y on x for these data, in the form $y = ax + b$. [2]
- (c) Estimate the total cost, **to the nearest USD**, of producing 13 bicycles on a particular day. [3]

All the bicycles that are produced are sold. The bicycles are sold for 304 USD **each**.

- (d) Explain why the factory does **not** make a profit when producing 13 bicycles on a particular day. [2]
- (e) (i) Write down an expression for the total selling price of x bicycles.
- (ii) Write down an expression for the **profit** the factory makes when producing x bicycles on a particular day.
- (iii) Find the least number of bicycles that the factory should produce, on a particular day, in order to make a profit. [5]

5.

As part of his IB Biology field work, Barry was asked to measure the circumference of trees, in centimetres, that were growing at different distances, in metres, from a river bank. His results are summarized in the following table.

Distance, x (metres)	5	12	17	21	24	30	34	44	47
Circumference, y (centimetres)	82	76	70	68	67	60	62	50	50

- (a) State whether *distance from the river bank* is a continuous or discrete variable. [1]
- (b) **On graph paper**, draw a scatter diagram to show Barry's results. Use a scale of 1 cm to represent 5 m on the x -axis and 1 cm to represent 10 cm on the y -axis. [4]
- (c) Write down
- the mean distance, \bar{x} , of the trees from the river bank;
 - the mean circumference, \bar{y} , of the trees. [2]
- (d) Plot and label the point $M(\bar{x}, \bar{y})$ on your graph. [2]
- (e) Write down
- the Pearson's product-moment correlation coefficient, r , for Barry's results;
 - the equation of the regression line y on x , for Barry's results. [4]
- (f) Draw the regression line y on x on your graph. [2]
- (g) Use the equation of the regression line y on x to estimate the circumference of a tree that is 40 m from the river bank. [2]