Bivariate stats 2

1) Alex and Kris are riding their bicycles together along a bicycle trail and note the following distance markers at the given times.

Time (t hours)	1	2	3	4	5	6	7
Distance (d km)	57	65	72	81	89	97	107

(a) Draw a scatter diagram of the data. Use 1 cm to represent 1 hour and 1 cm to represent 10 km. *[3 marks]*

(b) Write down for this set of data

  (i) the mean time, $\bar{t}$;

  (ii) the mean distance, $\bar{d}$. *[2 marks]*

(c) Mark and label the point $M\left(\bar{t}, \bar{d}\right)$ on your scatter diagram. *[2 marks]*

(d) Draw the line of best fit on your scatter diagram. *[2 marks]*

(e) **Using your graph**, estimate the time when Alex and Kris pass the 85 km distance marker. Give your answer correct to **one decimal place**. *[2 marks]*

(f) Write down the equation of the regression line for the data given. *[2 marks]*

(g) (i) **Using your equation** calculate the distance marker passed by the cyclists at 10.3 hours.

  (ii) Is this estimate of the distance reliable? Give a reason for your answer. *[4 marks]*

2) In a mountain region there appears to be a relationship between the number of trees growing in the region and the depth of snow in winter. A set of 10 areas was chosen, and in each area the number of trees was counted and the depth of snow measured. The results are given in the table below.

Number of trees (x)	Depth of snow in cm (y)
45	30
75	50
66	40
27	25
44	30
28	5
60	35
35	20
73	45
47	25

(a) Use your graphic display calculator to find

(i) the mean number of trees;

(ii) the standard deviation of the number of trees;

(iii) the mean depth of snow;

(iv) the standard deviation of the depth of snow. *[4 marks]*

The covariance, $S_{xy} = 188.5$.

(b) Write down the product-moment correlation coefficient, r. *[2 marks]*

(c) Write down the equation of the regression line of y on x. *[2 marks]*

(d) If the number of trees in an area is 55, estimate the depth of snow. *[2 marks]*

(e) (i) Use the equation of the regression line to estimate the depth of snow in an area with 100 trees.

(ii) Decide whether the answer in (e)(i) is a valid estimate of the depth of snow in the area. Give a reason for your answer. *[3 marks]*

3) Ten students were asked for their average grade at the end of their last year of high school and their average grade at the end of their last year at university. The results were put into a table as follows:

Student	High School grade, x	University grade, y
1	90	3.2
2	75	2.6
3	80	3.0
4	70	1.6
5	95	3.8
6	85	3.1
7	90	3.8
8	70	2.8
9	95	3.0
10	85	3.5
Total	835	30.4

(a) Given that $s_x = 8.96$, $s_y = 0.610$ and $s_{xy} = 4.16$, find the correlation coefficient r, giving your answer to two decimal places. *[2 marks]*

(b) Describe the correlation between the high school grades and the university grades. *[2 marks]*

(c) Find the equation of the regression line for y on x in the form $y = ax + b$. *[2 marks]*

4) The following table shows the cost in AUD of seven paperback books chosen at random, together with the number of pages in each book.

Book	1	2	3	4	5	6	7
Number of pages (x)	50	120	200	330	400	450	630
Cost (y AUD)	6.00	5.40	7.20	4.60	7.60	5.80	5.20

(a) Plot these pairs of values on a scatter diagram. Use a scale of 1 cm to represent 50 pages on the horizontal axis and 1 cm to represent 1 AUD on the vertical axis. *[3 marks]*

(b) Write down the linear correlation coefficient, r, for the data. *[2 marks]*

(c) Stephen wishes to buy a paperback book which has 350 pages in it. He plans to draw a line of best fit to determine the price. State whether or not this is an appropriate method in this case and justify your answer. *[2 marks]*

5)   (ii)   The number of bottles of water sold at a railway station on each day is given in the following table.

Day	0	1	2	3	4	5	6	7	8	9	10	11	12
Temperature $(T°)$	21	20.7	20	19	18	17.3	17	17.3	18	19	20	20.7	21
Number of bottles sold (n)	150	141	126	125	98	101	93	99	116	121	119	134	141

(a)   Write down

    (i)   the mean temperature;

    (ii)   the standard deviation of the temperatures.   *[2 marks]*

(b)   Write down the correlation coefficient, r, for the variables n and T.   *[1 mark]*

(c)   Comment on your value for r.   *[2 marks]*

(d)   The equation of the line of regression for n on T is $n = dT - 100$.

    (i)   Write down the value of d.

    (ii)   Estimate how many bottles of water will be sold when the temperature is 19.6°.   *[2 marks]*

(e)   On a day when the temperature was 36° Peter calculates that 314 bottles would be sold. Give one reason why his answer might be unreliable.   *[1 mark]*