PRACTICE 6.2 – Linear Regression

* Full, worked solutions can be found in the folder linked on the Course Website ☺

Exercise 6E

1 For each set of data:

- **i** Plot the points on a scatter diagram.
- **ii** Describe the type of correlation.
- **iii** Find the mean of x and the mean of y.
- **iv** Plot and label the mean point on your diagram and draw a line of best fit by eye through the mean point.

a

x	10	11	12	13	14	15	16	17	18	20
y	18	20	19	21	24	23	24	26	27	32

b

x	−3	−2	−1	0	1	2	3	4	5	6
y	15	13	10	7	6	4	5	2	3	0

2 The table shows the area (a, in millions of square feet) of ten shopping malls, along with the annual number of visitors (v, in millions).

a	2.7	2.2	1.8	2.6	1.8	2.2	2.7	2	1.4	1
v	28	27	26	25	23	22	22	21	20	18

- **a** Plot a scatter diagram for this data.
- **b** Describe the correlation.
- **c** Find the mean area of the shopping malls.
- **d** Find the mean annual number of visitors.
- **e** Plot and label the mean point on the scatter diagram.
- **f** Draw a line of best fit by eye through the mean point.

3 The table shows the gross domestic product (GDP) (x) and the average number of books read per person in a year (y) for 15 countries.

x	7904	10326	7616	11387	12645	14528	20122	22122	28166	29861	54304	44189	42651	46678	50169
y	2.2	2.9	3.0	4.0	4.6	5.4	6.0	8.5	9.7	10.3	12.0	15.0	16.3	16.8	17.0

- **a** Plot a scatter diagram for this data.
- **b** Describe the correlation.
- **c** Find the mean GDP.
- **d** Find the mean number of books read per person in a year.
- **e** Plot and label the mean point on the scatter diagram.
- **f** Draw a line of best fit by eye through the mean point.

continued on back □

Exercise 6F

1 The travel times in minutes (x) and the prices in euros (y) of ten different train journeys between various places in Spain are shown in the table.

x	128	150	102	140	140	98	75	130	80	132
y	25.95	40	24.85	31.8	30.2	28.95	21.85	34.5	23.25	26

a Calculate the value of r. Comment on the correlation.

b Write down the equation of the regression line of y on x.

c Find the mean travel time.

d Find the mean price.

e Plot the data points on a scatter diagram together with the regression line of y on x, clearly marking and labelling the mean point.

2 A memory test was given to a group of ten people. Each of them was shown images of 20 different objects, then after 5 minutes they were asked to name the objects they remembered. The times they took to remember those names (in minutes) were recorded. The results are shown in the table.

Objects remembered (x)	12	16	9	10	14	18	12	15	17	15
Time taken (y)	1.9	2.2	1.7	2	2	2.5	2.3	2.2	2.4	2.4

a Calculate the value of r. Comment on the correlation.

b Write down the equation of the regression line of y on x.

c Find the mean number of objects remembered.

d Find the mean time taken to remember the objects.

e Plot the data points on a scatter diagram along with the regression line of y on x, clearly marking and labelling the mean point.

3 The heights in metres (x) and weights in kilograms (y) of ten male gorillas are shown in the table.

x	1.9	1.83	1.81	1.79	1.74	1.91	1.93	1.86	1.81	1.95
y	275	267	260	257	258	272	273	268	261	273

a Calculate the value of r. Comment on the correlation.

b Write down the equation of the least squares regression line for this data.

c Find the mean height for this group of male gorillas.

d Find the mean weight for this group of male gorillas.

e Plot the data points on a scatter diagram together with the regression line of y on x, clearly marking and labelling the mean point.

4 The headmaster of a secondary school is investigating whether there is any relationship between a student's age (x) and the monthly average number of absences for that age group (y). The data collected is shown in the table.

x	12	13	14	15	16	17	18
y	4.2	4	3.9	3.5	3.4	3.4	3.2

a Calculate Pearson's product moment correlation coefficient for this data. Comment on the correlation.

b Write down the equation of the regression line of y on x.

c Find the mean point of this data.

d Plot the data points on a diagram along with the regression line of y on x, clearly marking and labelling the mean point.

e Hence determine how many absences per month you might expect from an 11-year-old, and comment on the reliability of the prediction.