1.   Gary compared the total attendance, x, at home matches and the total number of goals, y, scored at home during a season for each of 12 football teams playing in a league. He correctly calculated:

$$S_{xx} = 1022500 \qquad S_{yy} = 130.9 \qquad S_{xy} = 8825$$

(a)   Calculate the product moment correlation coefficient for these data.

(2)

(b)   Interpret the value of the correlation coefficient.

(1)

Helen was given the same data to analyse. In view of the large numbers involved she decided to divide the attendance figures by 100. She then calculated the product moment   correlation coefficient between $\dfrac{x}{100}$ and y.

(c)   Write down the value Helen should have obtained.

(1)
(Total 4 marks)

2.   The blood pressures, p mmHg, and the ages, t years, of 7 hospital patients are shown in the table below.

Patient	A	B	C	D	E	F	G
t	42	74	48	35	56	26	60
p	98	130	120	88	182	80	135

$$\left[\sum t = 341, \ \sum p = 833, \ \sum t^2 = 18181, \ \sum p^2 = 106397, \ \sum tp = 42948 \right]$$

(a)   Find S_{pp}, S_{tp} and S_{tt} for these data.
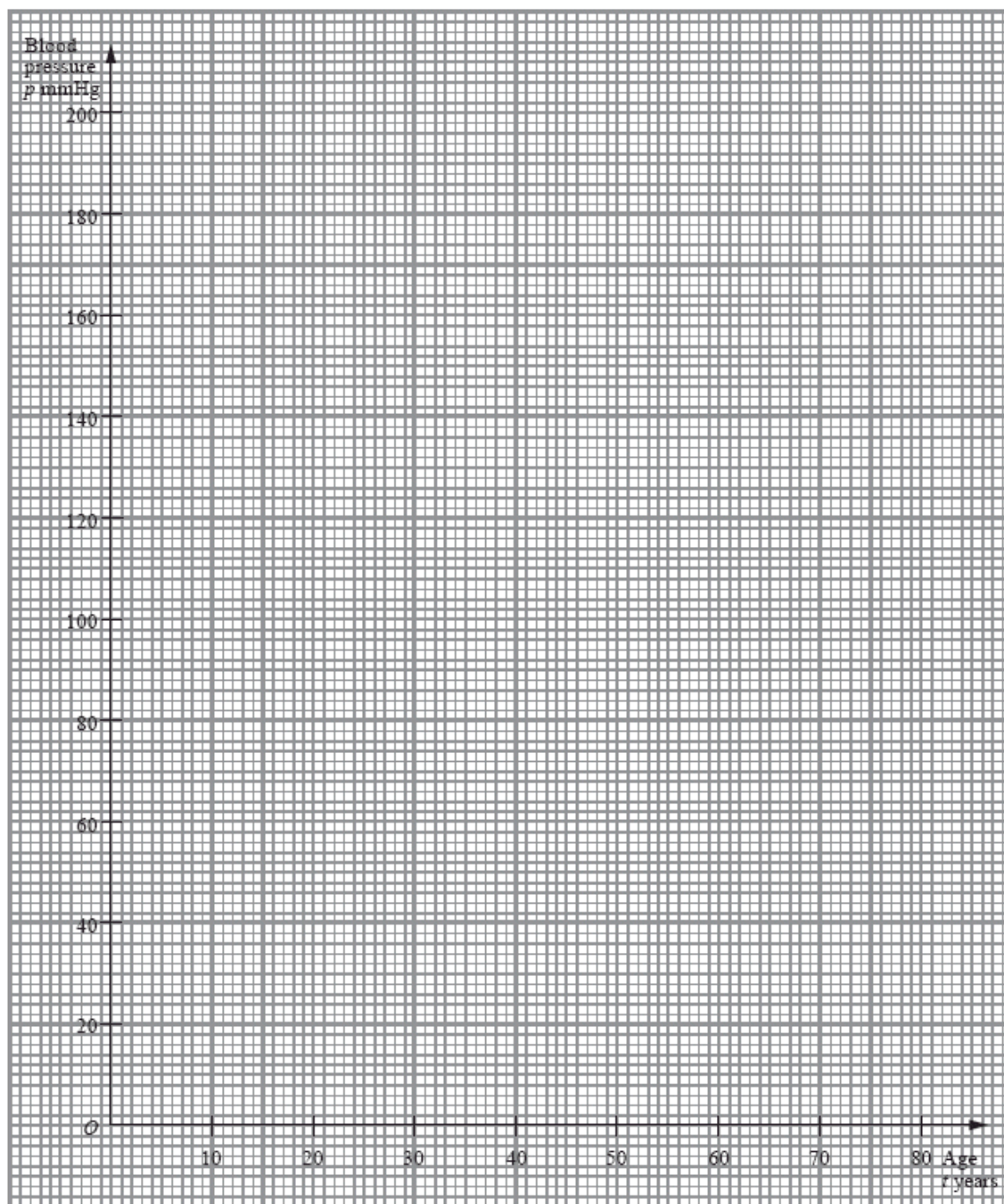
(4)

(b)   Calculate the product moment correlation coefficient for these data.

(3)

(c)   Interpret the correlation coefficient.

(1)

(d)   On the graph paper below, draw the scatter diagram of blood pressure against age for these 7 patients.

(2)

(e)   Find the equation of the regression line of p on t.

(4)

(f)   Plot your regression line on your scatter diagram.

(2)

(g)   Use your regression line to estimate the blood pressure of a 40 year old patient.

(2)
(Total 18 marks)

3.   The volume of a sample of gas is kept constant. The gas is heated and the pressure, p, is measured at 10 different temperatures, t. The results are summarised below.

$\Sigma p = 445$ $\Sigma p^2 = 38\ 125$ $\Sigma t = 240$ $\Sigma t^2 = 27\ 520$ $\Sigma pt = 26\ 830$

(a)   Find S_{pp} and S_{pt}.

(3)

Given that $S_{tt} = 21\ 760$,

(b)   calculate the product moment correlation coefficient.

(2)

(c)   Give an interpretation of your answer to part (b).

(1)
(Total 6 marks)

4.   In a study of how students use their mobile telephones, the phone usage of a random sample of 11 students was examined for a particular week.

The total length of calls, y minutes, for the 11 students were

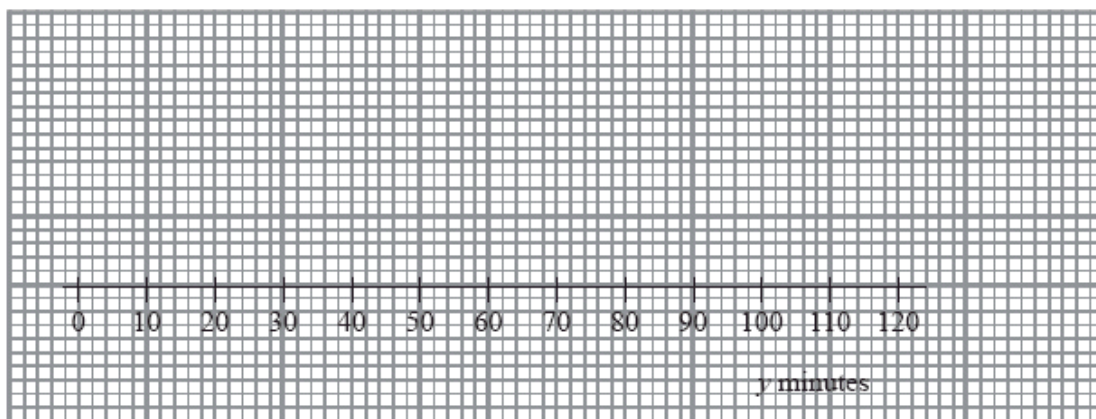$$17,\ 23,\ 35,\ 36,\ 51,\ 53,\ 54,\ 55,\ 60,\ 77,\ 110$$

(a)   Find the median and quartiles for these data.

(3)

A value that is greater than $Q_3 + 1.5 \times (Q_3 - Q_1)$ or smaller than $Q_1 - 1.5 \times (Q_3 - Q_1)$ is defined as an outlier.

(b)   Show that 110 is the only outlier.

(2)

(c)   Using the graph below draw a box plot for these data indicating clearly the position of the outlier.

(3)

The value of 110 is omitted.

(d)   Show that S_{yy} for the remaining 10 students is 2966.9

(3)

These 10 students were each asked how many text messages, x, they sent in the same week.

The values of S_{xx} and S_{xy} for these 10 students are $S_{xx} = 3463.6$ and $S_{xy} = -18.3$.

(e)   Calculate the product moment correlation coefficient between the number of text messages sent and the total length of calls for these 10 students.

(2)

A parent believes that a student who sends a large number of text messages will spend fewer minutes on calls.

(f)   Comment on this belief in the light of your calculation in part (e).

(1)
(Total 14 marks)

5. As part of a statistics project, Gill collected data relating to the length of time, to the nearest minute, spent by shoppers in a supermarket and the amount of money they spent.
Her data for a random sample of 10 shoppers are summarised in the table below, where t represents time and £m the amount spent over £20.

t (minutes)	£m
15	–3
23	17
5	–19
16	4
30	12
6	–9
32	27
23	6
35	20
27	6

(a) Write down the actual amount spent by the shopper who was in the supermarket for 15 minutes.

(1)

(b) Calculate S_{tt}, S_{mm} and S_{tm}.

(You may use $\Sigma t^2 = 5478$ $\Sigma m^2 = 2101$ $\Sigma tm = 2485$)

(6)

(c) Calculate the value of the product moment correlation coefficient between t and m.

(3)

(d) Write down the value of the product moment correlation coefficient between t and the actual amount spent. Give a reason to justify your value.

(2)

On another day Gill collected similar data. For these data the product moment correlation coefficient was 0.178

(e)   Give an interpretation to both of these coefficients.

(2)

(f)   Suggest a practical reason why these two values are so different.

(1)
(Total 15 marks)

6.   Students in Mr Brawn's exercise class have to do press-ups and sit-ups. The number of press-ups x and the number of sit-ups y done by a random sample of 8 students are summarised below.

$$\Sigma x = 272, \quad \Sigma x^2 = 10\ 164, \quad\quad \Sigma xy = 11\ 222,$$

$$\Sigma y = 320, \quad \Sigma y^2 = 13\ 464.$$

(a)   Evaluate S_{xx}, S_{yy} and S_{xy}.

(4)

(b)   Calculate, to 3 decimal places, the product moment correlation coefficient between x and y.

(3)

(c)   Give an interpretation of your coefficient.

(2)

(d)   Calculate the mean and the standard deviation of the number of press-ups done by these students.

(4)

Mr Brawn assumes that the number of press-ups that can be done by any student can be modelled by a normal distribution with mean μ and standard deviation σ. Assuming that μ and σ take the same values as those calculated in part (d),

(e)   find the value of a such that $P(\mu - a < X < \mu + a) = 0.95$.

(3)

(f)   Comment on Mr Brawn's assumption of normality.

(2)
(Total 18 marks)

7.   A researcher thinks there is a link between a person's height and level of confidence. She measured the height h, to the nearest cm, of a random sample of 9 people. She also devised a test to measure the level of confidence c of each person. The data are shown in the table below.

h	179	169	187	166	162	193	161	177	168
c	569	561	579	561	540	598	542	565	573

[You may use $\sum h^2 = 272\ 094$, $\sum c^2 = 2\ 878\ 966$, $\sum hc = 884\ 484$]

(a)   Draw a scatter diagram to illustrate these data.

(4)

(b)   Find exact values of S_{hc} S_{hh} and S_{cc}.

(4)

(c)   Calculate the value of the product moment correlation coefficient for these data.

(3)

(d)   Give an interpretation of your correlation coefficient.

(1)

(e)   Calculate the equation of the regression line of c on h in the form $c = a + bh$.

(3)

(f)   Estimate the level of confidence of a person of height 180 cm.

(2)

(g)   State the range of values of h for which estimates of c are reliable.

(1)
(Total 18 marks)

8.   A company owns two petrol stations P and Q along a main road. Total daily sales in the same week for P (£p) and for Q (£q) are summarised in the table below.

	p	q
Monday	4760	5380
Tuesday	5395	4460
Wednesday	5840	4640
Thursday	4650	5450
Friday	5365	4340
Saturday	4990	5550
Sunday	4365	5840

When these data are coded using $x = \dfrac{p-4365}{100}$ and $y = \dfrac{q-4340}{100}$,

$$\Sigma x = 48.1, \quad \Sigma y = 52.8, \quad \Sigma x^2 = 486.44, \quad \Sigma y^2 = 613.22 \text{ and } \Sigma xy = 204.95.$$

(a)   Calculate S_{xy}, S_{xx} and S_{yy}.

(4)

(b)   Calculate, to 3 significant figures, the value of the product moment correlation coefficient between x and y.

(3)

(c)   (i)   Write down the value of the product moment correlation coefficient between p and q.

(ii)   Give an interpretation of this value.

(2)
(Total 9 marks)