
S1: Correlation

Past Paper Questions
2006 - 2013

Name:

5 [Figure 1, printed on the insert, is provided for use in this question.]

The table shows the times, in seconds, taken by a random sample of 10 boys from a junior swimming club to swim 50 metres freestyle and 50 metres backstroke.

Boy	A	B	C	D	E	F	G	H	I	J
Freestyle (x seconds)	30.2	32.8	25.1	31.8	31.2	35.6	32.4	38.0	36.1	34.1
Backstroke (y seconds)	33.5	35.4	37.4	27.2	34.7	38.2	37.7	41.4	42.3	38.4

(a) On Figure 1, complete the scatter diagram for these data. (2 marks)

(b) Hence:

(i) give **two** distinct comments on what your scatter diagram reveals; (2 marks)

(ii) state, **without calculation**, which of the following 3 values is most likely to be the value of the product moment correlation coefficient for the data in your scatter diagram.

0.912 0.088 0.462 (1 mark)

(c) In the sample of 10 boys, one boy is a junior-champion freestyle swimmer and one boy is a junior-champion backstroke swimmer.

Identify the **two** most likely boys. (2 marks)

(d) **Removing the data for the two boys whom you identified in part (c):**

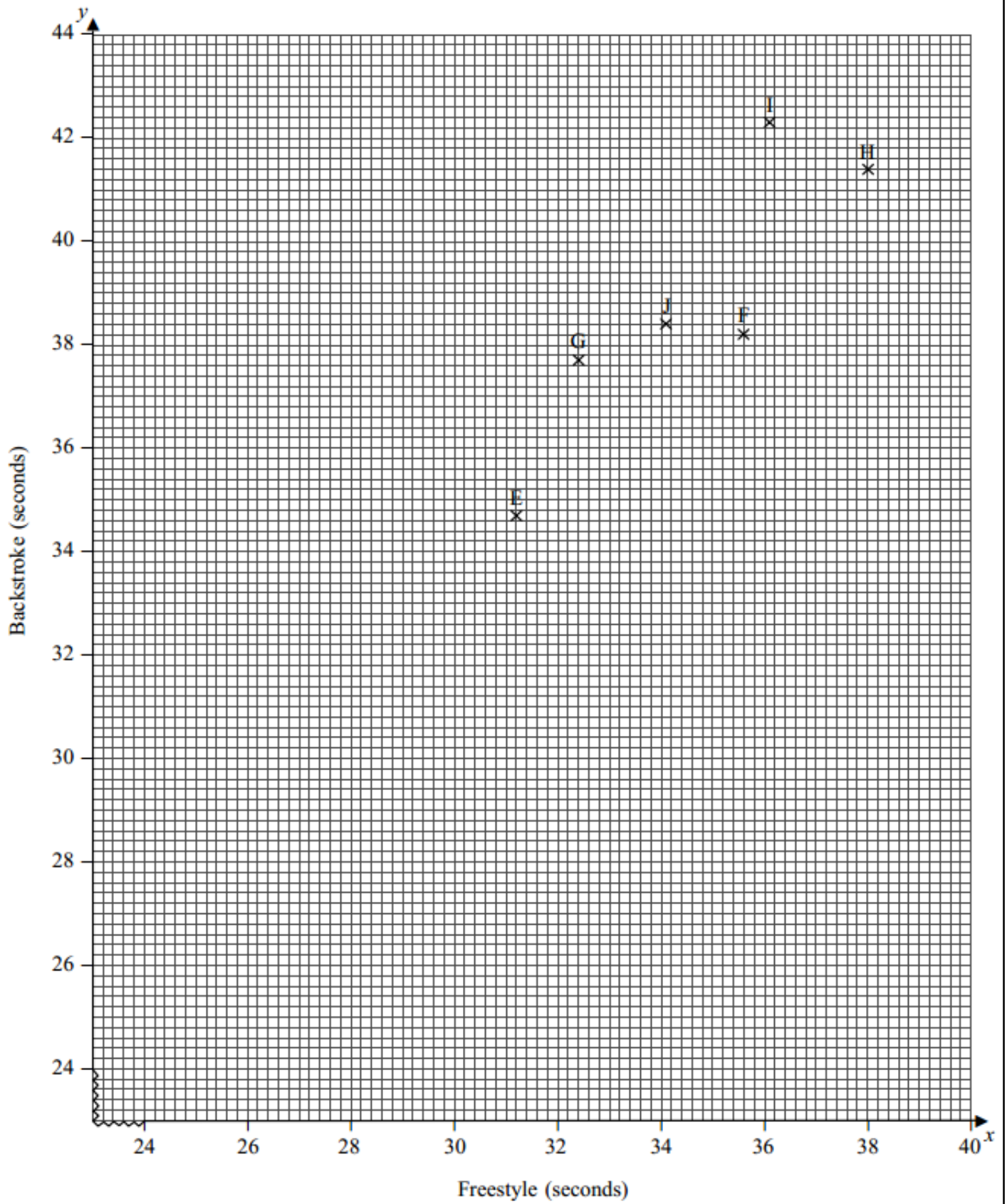
(i) calculate the value of the product moment correlation coefficient for the remaining 8 pairs of values of x and y ; (3 marks)

(ii) comment, in context, on the value that you obtain. (1 mark)



Figure 1 (for use in Question 5)

Scatter Diagram for Freestyle and Backstroke Swimming Times



1 The table shows, for each of a random sample of 8 paperback fiction books, the number of pages, x , and the recommended retail price, $\pounds y$, to the nearest 10p.

?

x	223	276	374	433	564	612	704	766
y	6.50	4.00	5.50	8.00	4.50	5.00	8.00	5.50

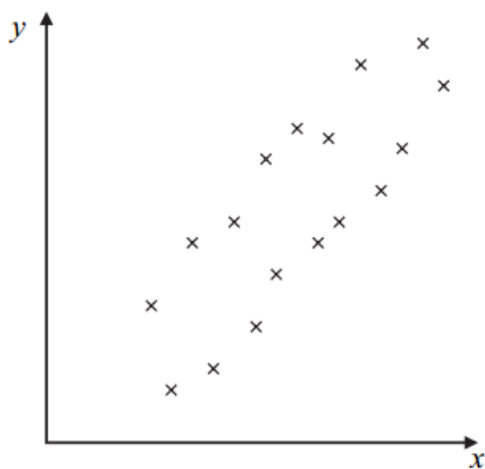
- (a) (i) Calculate the value of the product moment correlation coefficient between x and y . (3 marks)
- (ii) Interpret your value in the context of this question. (2 marks)
- (iii) Suggest one other variable, in addition to the number of pages, which may affect the recommended retail price of a paperback fiction book. (1 mark)
- (b) The same 8 books were later included in a book sale. The value of the product moment correlation coefficient between the number of pages and the sale price was 0.959, correct to three decimal places.

What can be concluded from this value?

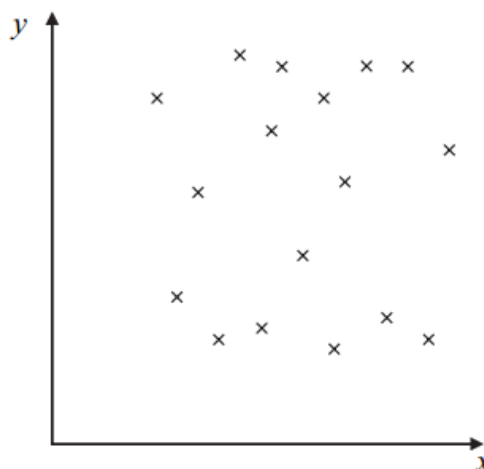
(2 marks)

3 Estimate, **without undertaking any calculations**, the value of the product moment correlation coefficient between the variables x and y in each of the three scatter diagrams.

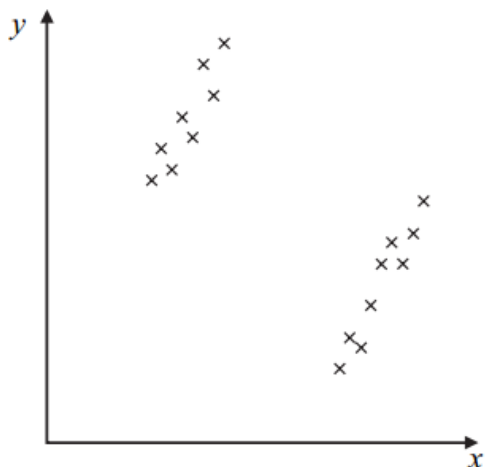
(a)



(b)



(c)



(5 marks)

- 1 The table shows the length, in centimetres, and maximum diameter, in centimetres, of each of 10 honeydew melons selected at random from those on display at a market stall.

Length	24	25	19	28	27	21	35	23	32	26
Maximum diameter	18	14	16	11	13	14	12	16	15	14

- (a) Calculate the value of the product moment correlation coefficient. *(3 marks)*
- (b) Interpret your value in the context of this question. *(2 marks)*

- 2 The head and body length, x millimetres, and tail length, y millimetres, of each of a sample of 20 adult dormice were measured. The following statistics are derived from the results.

$$S_{xx} = 1280.55 \quad S_{yy} = 281.8 \quad S_{xy} = 416.3$$

- (a) Calculate the value of the product moment correlation coefficient between x and y . *(2 marks)*
- (b) Interpret your value in the context of this question. *(2 marks)*
- (c) Write down the value of the product moment correlation coefficient if the measurements had been recorded in centimetres. *(1 mark)*
- (d) Give a reason why it is not generally advisable to calculate the value of the product moment correlation coefficient without first viewing a scatter diagram of the data. Illustrate your answer with a sketch. *(2 marks)*

3 [Figure 1, printed on the insert, is provided for use in this question.]

The table shows, for each of a sample of 12 handmade decorative ceramic plaques, the length, x millimetres, and the width, y millimetres.

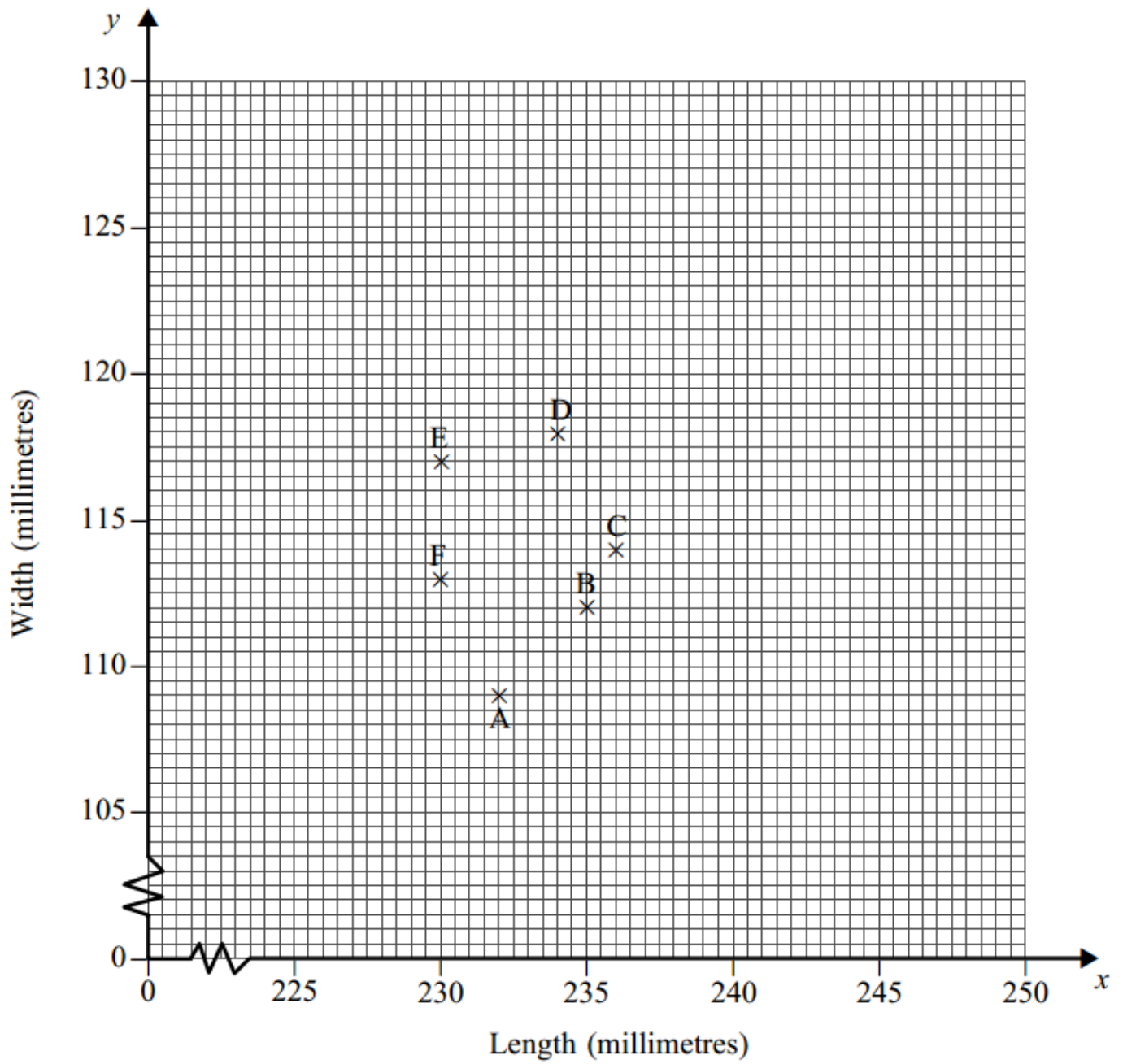
Plaque	x	y
A	232	109
B	235	112
C	236	114
D	234	118
E	230	117
F	230	113
G	246	121
H	240	125
I	244	128
J	241	122
K	246	126
L	245	123

- (a) Calculate the value of the product moment correlation coefficient between x and y .
(3 marks)
- (b) Interpret your value in the context of this question.
(2 marks)
- (c) On **Figure 1**, complete the scatter diagram for these data.
(3 marks)
- (d) In fact, the 6 plaques A, B, ..., F are from a different source to the 6 plaques G, H, ..., L.

With reference to your scatter diagram, **but without further calculations**, estimate the value of the product moment correlation coefficient between x and y for **each** source of plaque.
(2 marks)

Figure 1 (for use in Question 3)

Decorative Plaques



2 A greengrocer sells bunches of 9 carrots at his Saturday market stall. Tom and Geri are two Statistics students who work on the stall. Each selects a bunch of carrots at random.

- (a) At home, Tom measures the length, x centimetres, and the maximum diameter, y centimetres, of each carrot in his selected bunch with the following results.

x	16.2	13.1	10.4	12.1	14.6	9.7	11.8	13.6	17.3
y	4.2	3.9	4.7	3.3	3.7	2.4	3.1	3.5	2.7

- (i) Calculate the value of the product moment correlation coefficient. (3 marks)

- (ii) Interpret your value in context. (2 marks)

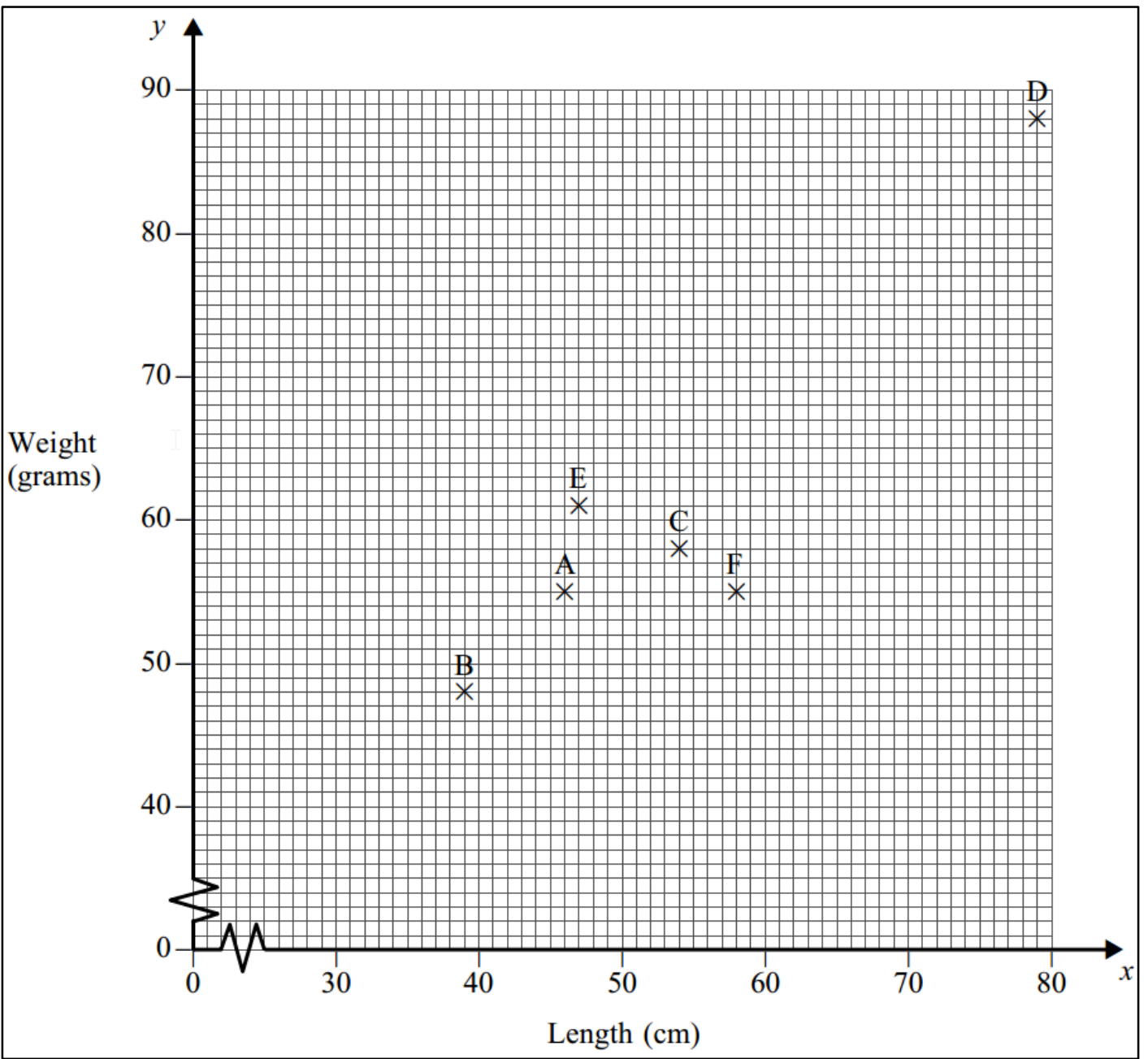
- (b) At her home, Geri measures the length, in centimetres, and the **weight**, in grams, of each carrot in her selected bunch and then obtains a value of -0.986 for the product moment correlation coefficient.

Comment, with a reason, on the likely validity of Geri's value. (2 marks)

2 Hermione, who is studying reptiles, measures the length, x cm, and the weight, y grams, of a sample of 11 adult snakes of the same type. Her results are shown in the table.

Snake	A	B	C	D	E	F	G	H	I	J	K
x	46	39	54	79	47	58	73	35	43	51	36
y	55	48	58	88	61	55	82	51	50	66	57

- (a) Calculate the value of the product moment correlation coefficient, r , between x and y . (3 marks)
- (b) Interpret your value in context. (2 marks)
- (c) Complete the scatter diagram, opposite, for these data. (2 marks)
- (d) Subsequently it is found that, of the 11 adult snakes, 9 are male and 2 are female.
- (i) Given that female adult snakes are generally larger than male adult snakes, identify the 2 snakes which are most likely to be female. (1 mark)
- (ii) Hence, **without further calculation**, estimate the value of r for the 9 male snakes and revise, as necessary, your interpretation in part (b). (2 marks)



7 [Figure 1, printed on the insert, is provided for use in this question.]

Harold considers himself to be an expert in assessing the auction value of antiques. He regularly visits car boot sales to buy items that he then sells at his local auction rooms.

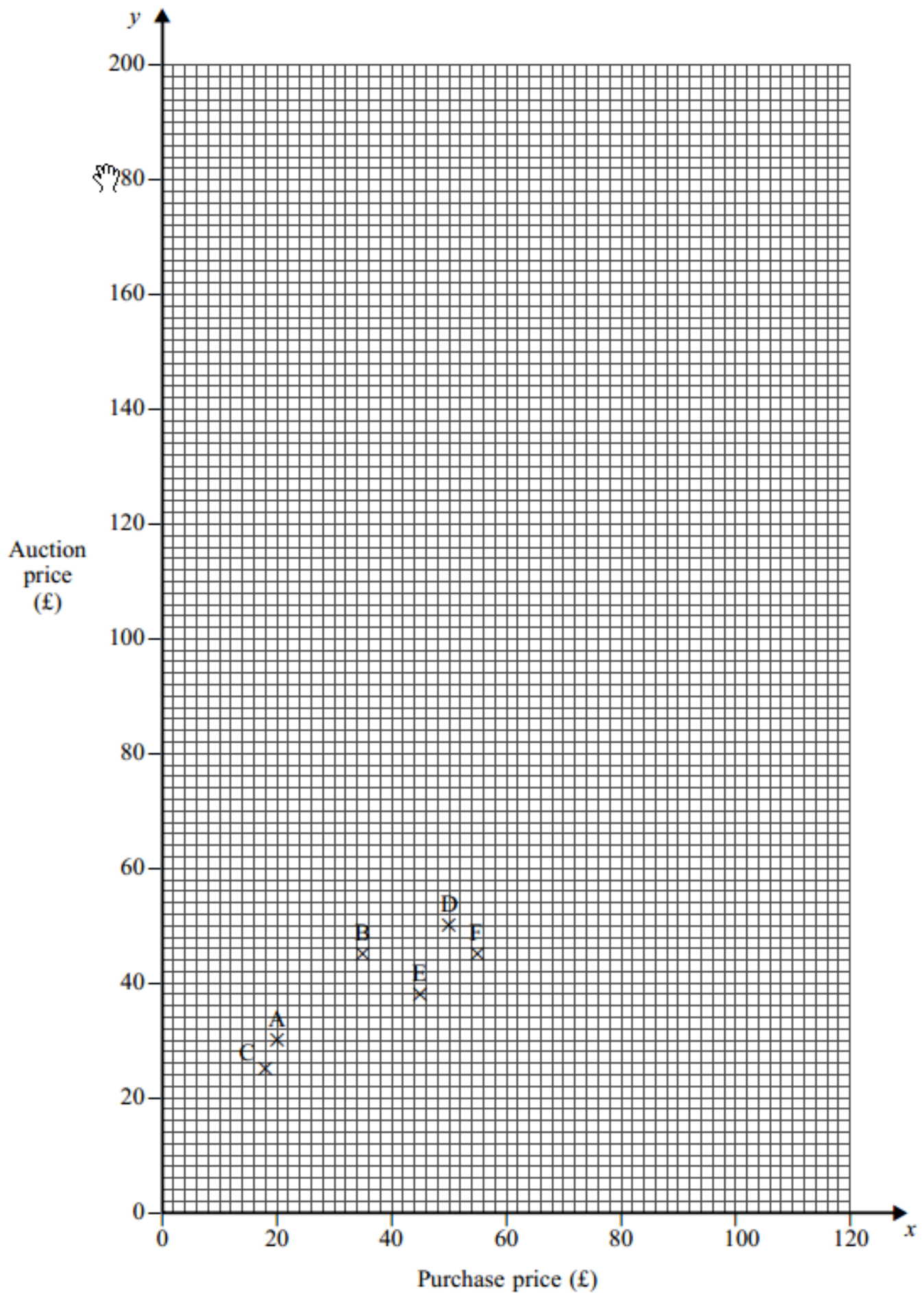
Harold's father, Albert, who is not convinced of his son's expertise, collects the following data from a random sample of 12 items bought by Harold.

Item	Purchase price (£ x)	Auction price (£ y)
A	20	30
B	35	45
C	18	25
D	50	50
E	45	38
F	55	45
G	43	50
H	81	90
I	90	85
J	30	190
K	57	65
L	112	25

- (a) Calculate the value of the product moment correlation coefficient between x and y . *(3 marks)*
- (b) Interpret your value in the context of this question. *(2 marks)*
- (c) (i) On **Figure 1**, complete the scatter diagram for these data. *(3 marks)*
- (ii) Comment on what this reveals. *(2 marks)*
- (d) When items J and L are omitted from the data, it is found that
- $$S_{xx} = 4854.4 \quad S_{yy} = 4216.1 \quad S_{xy} = 4268.8$$
- (i) Calculate the value of the product moment correlation coefficient between x and y for the remaining 10 items. *(2 marks)*
- (ii) Hence revise as necessary your interpretation in part (b). *(1 mark)*

Figure 1 (for use in Question 7)

Prices of Antiques



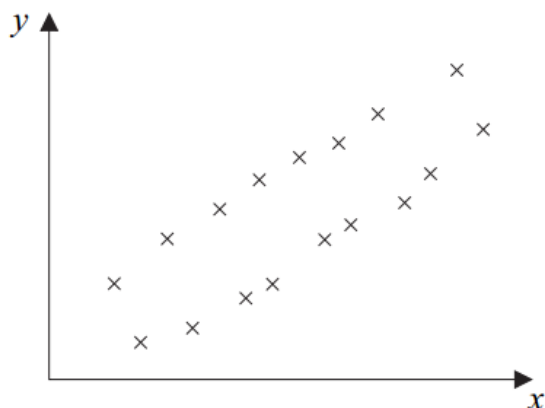
1 The weight, x kg, and the engine power, y bhp, of each car in a random sample of 10 hatchback cars are shown in the table.

x	1196	1062	1335	1429	1012	1355	1145	1417	1275	1284
y	123	88	150	158	69	120	94	143	107	128

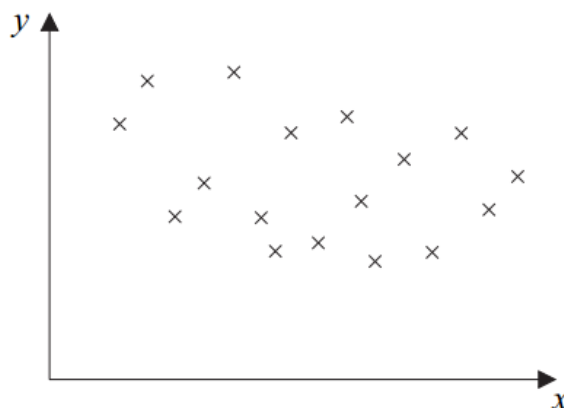
- (a) Calculate the value of the product moment correlation coefficient between x and y . (3 marks)
- (b) Interpret your value in the context of the question. (2 marks)

1 (a) Estimate, **without undertaking any calculations**, the value of the product moment correlation coefficient between the variables x and y for each of the two scatter diagrams.

(i)



(ii)



(2 marks)

(b) The table gives the circumference, x centimetres, and the weight, y grams, of each of 12 new cricket balls.

x	22.5	22.7	22.6	22.4	22.5	22.8	22.6	22.7	22.8	22.4	22.9	22.6
y	160.3	159.4	157.8	158.0	157.3	159.8	158.3	159.6	161.3	156.4	162.5	161.2

- (i) Calculate the value of the product moment correlation coefficient between x and y . (3 marks)
- (ii) Assuming that the 12 balls may be considered to be a random sample, interpret your value in context. (2 marks)

7 (a) Three airport management trainees, Ryan, Sunil and Tim, were each instructed to select a random sample of 12 suitcases from those waiting to be loaded onto aircraft.

Each trainee also had to measure the volume, x , and the weight, y , of each of the 12 suitcases in his sample, and then calculate the value of the product moment correlation coefficient, r , between x and y .

- Ryan obtained a value of -0.843 .
- Sunil obtained a value of $+0.007$.

Explain why neither of these two values is likely to be correct. *(2 marks)*

(b) Peggy, a supervisor with many years' experience, measured the volume, x cubic feet, and the weight, y pounds, of each suitcase in a random sample of 6 suitcases, and then obtained a value of 0.612 for r .

- Ryan and Sunil each claimed that Peggy's value was different from their values because she had measured the volumes in cubic feet and the weights in pounds, whereas they had measured the volumes in cubic metres and the weights in kilograms.
- Tim claimed that Peggy's value was almost exactly half his calculated value because she had used a sample of size 6 whereas he had used one of size 12.

Explain why neither of these two claims is valid. *(2 marks)*

(c) Quentin, a manager, recorded the volumes, v , and the weights, w , of a random sample of 8 suitcases as follows.

v	28.1	19.7	46.4	23.6	31.1	17.5	35.8	13.8
w	14.9	12.1	21.1	18.0	19.8	19.2	16.2	14.7

(i) Calculate the value of r between v and w . *(3 marks)*

(ii) Interpret your value in the context of this question. *(2 marks)*

- 2** Dr Hanna has a special clinic for her older patients. She asked a medical student, Lenny, to select a random sample of 25 of her male patients, aged between 55 and 65 years, and, from their clinical records, to list their heights, weights and waist measurements.
- Lenny was then asked to calculate three values of the product moment correlation coefficient based upon his collected data. His results were:
- (a) 0.365 between height and waist measurement;
- (b) 1.16 between height and weight;
- (c) -0.583 between weight and waist measurement.
- For **each** of Lenny's three calculated values, state whether the value is definitely correct, probably correct, probably incorrect or definitely incorrect. *(3 marks)*

- 1** A production line in a rolling mill produces lengths of steel.
- A random sample of 20 lengths of steel from the production line was selected. The minimum width, x centimetres, and the minimum thickness, y millimetres, of each selected length was recorded.
- The following summarised information was then calculated from these records.
- $$S_{xx} = 2.030 \quad S_{yy} = 1.498 \quad S_{xy} = -0.410$$
- (a) Calculate the value of the product moment correlation coefficient between x and y . *(2 marks)*
- (b) Interpret your value in the context of the question. *(2 marks)*

4 Ashok is a work-experience student with an organisation that offers two separate professional examination papers, I and II.

For each of a random sample of 12 students, A to L, he records the mark, x per cent, achieved on Paper I, and the mark, y per cent, achieved on Paper II.

	A	B	C	D	E	F	G	H	I	J	K	L
x	34	46	53	62	67	72	60	54	70	71	82	85
y	61	66	72	78	88	81	49	60	54	44	49	36

- (a) (i) Calculate the value of the product moment correlation coefficient, r , between x and y .
(3 marks)
- (ii) Interpret your value of r in the context of this question.
(2 marks)

4 (b) (i) Give **two** possible advantages of plotting data on a graph before calculating the value of a product moment correlation coefficient.
(2 marks)

(ii) Complete the plotting of Ashok's data on the scatter diagram on the page opposite.
(2 marks)

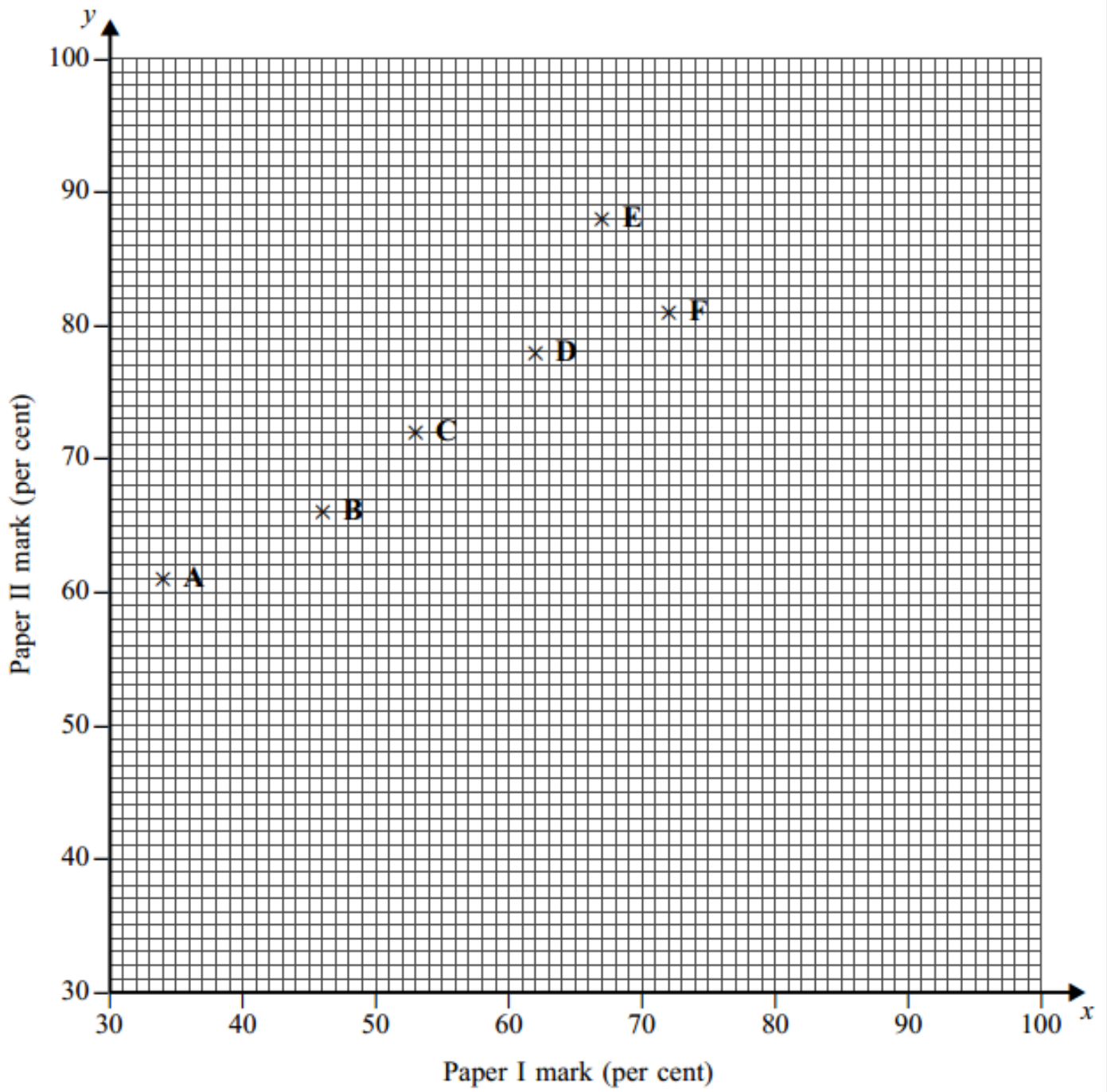
(iii) State what is now revealed by the scatter diagram.
(1 mark)

(c) Ashok subsequently discovers that students A to F have a more scientific background than students G to L.

With reference to your scatter diagram, estimate the value of the product moment correlation coefficient for **each** of the two groups of students. You are **not** expected to calculate the two values.
(2 marks)

	G	H	I	J	K	L
x	60	54	70	71	82	85
y	49	60	54	44	49	36

Examination Marks



- 4 The girth, g metres, the length, l metres, and the weight, y kilograms, of each of a sample of 20 pigs were measured.

The data collected is summarised as follows.

$$S_{gg} = 0.1196 \quad S_{ll} = 0.0436 \quad S_{yy} = 5880 \quad S_{gy} = 24.15 \quad S_{ly} = 10.25$$

- (a) Calculate the value of the product moment correlation coefficient between:
- girth and weight;
 - length and weight. (3 marks)
- (b) Interpret, in context, **each** of the values that you obtained in part (a). (3 marks)
- (c) Weighing pigs requires expensive equipment, whereas measuring their girths and lengths simply requires a tape measure. With this in mind, the following formula is proposed to make an estimate of a pig's weight, x kilograms, from its girth and length.

$$x = 69.3 \times g^2 \times l$$

Applying this formula to the relevant data on the 20 pigs resulted in

$$S_{xx} = 5656.15 \quad S_{xy} = 5662.97$$

- By calculating a third value of the product moment correlation coefficient, state which of g , l or x is the most strongly correlated with y , the weight. (2 marks)
- Estimate the weight of a pig that has a girth of 1.25 metres and a length of 1.15 metres. (2 marks)
- Given the additional information that $\bar{x} = 115.4$ and $\bar{y} = 116.0$, calculate the equation of the least squares regression line of y on x , in the form $y = a + bx$. (3 marks)
- Comment on the likely accuracy of the estimated weight found in part (c)(ii). Your answer should make reference to the value of the product moment correlation coefficient found in part (c)(i) and to the values of b and a found in part (c)(iii). (4 marks)