

## Questions

Q1.

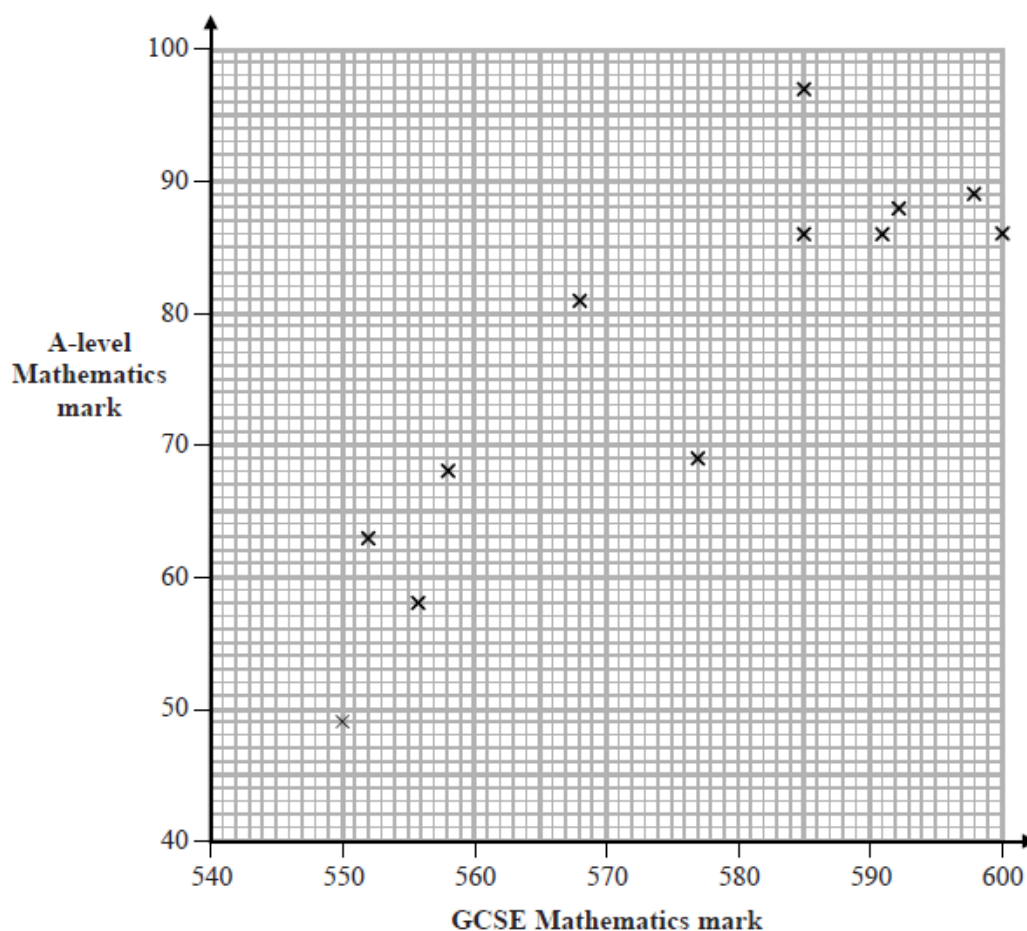
Julie was investigating the relationship between the marks gained by students in their GCSE Mathematics exam and the marks gained by the same students in an A-level Mathematics exam.

(a) Suggest a hypothesis Julie could use.

.....  
.....

(1)

Julie drew a scatter diagram using the marks gained in GCSE Mathematics and the marks gained in A-level Mathematics by each of 12 students.



(b) For this scatter diagram, explain why the GCSE Mathematics mark is the explanatory variable.

.....  
.....  
.....

(1)

(c) Explain, giving a statistical reason, whether or not the scatter diagram supports your hypothesis in part (a).

.....  
.....  
.....

(2)

Using statistical software, Julie obtained the following information for her 12 students.

Mean GCSE Mathematics mark	578
Mean A-level Mathematics mark	78
Gradient of line of best fit	0.7

(d) Using this information, draw a line of best fit on the scatter diagram.

(2)

(e) Interpret the gradient of the line of best fit.

.....  
.....

(1)

Julie wants to use the line of best fit to predict the A-level Mathematics mark for a 13th student. For GCSE Mathematics this student gained a mark of 540

(f) Explain whether or not it would be appropriate to use the line of best fit to make her prediction.

.....  
.....

(1)

**(Total for question = 8 marks)**

Q2.

Audrey is investigating factors that affect life expectancy by using secondary data.

For a sample of 12 countries, Audrey recorded the mean age, in years, of women when they had their first child.

She also recorded the mean life expectancy, in years, for these 12 countries.

Audrey drew a scatter diagram for this information.



(a) Explain why a scatter diagram is appropriate for the type of data Audrey collected.

.....

(1)

For these 12 countries, the double mean point of the data is (25.2, 72).

(b) Using this information, draw a line of best fit on the scatter diagram.

(2)

Audrey includes the following two statements in the conclusion to her investigation.

A) There is strong positive correlation between the mean age at which women in a country have their first child and the mean life expectancy in that country.

B) Life expectancy can be increased if women have their first child later in life.

(c) Discuss whether or not Audrey's conclusions are appropriate.

You should consider her graph and how she carried out her investigation.

.....

.....

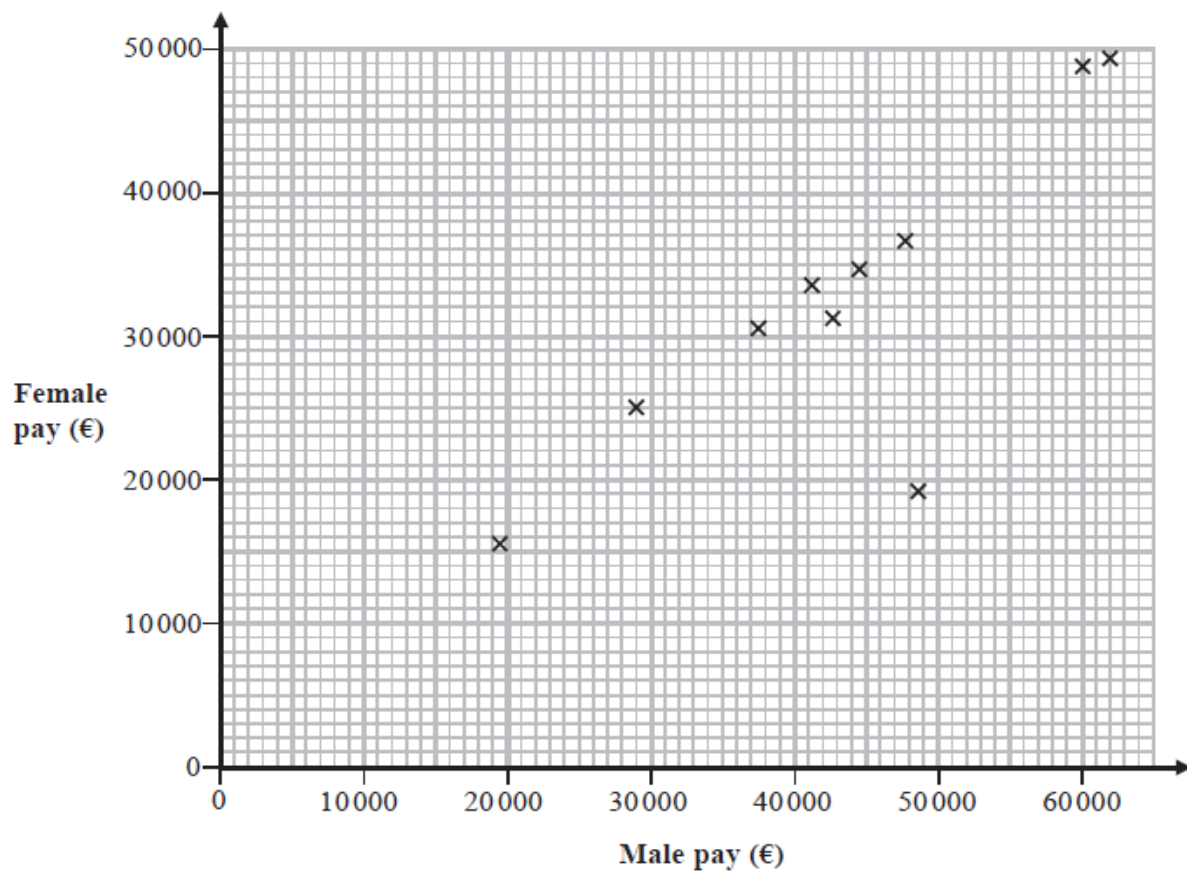
.....



Country	Average gross pay (€)	
	Male	Female
Denmark	62 120	49 254
Germany	44 465	34 740
Ireland	48 459	19 177
Spain	29 009	25 101
France	37 627	30 406
Netherlands	47 373	36 696
Portugal	19 424	15 299
Sweden	41 311	33 305
United Kingdom	42 710	31 115
Switzerland	60 135	48 972

Source: Eurostat

The data have been plotted on the scatter diagram.



Here is a list of statistical words.

rank	quantitative	dependent	primary	bivariate	categorical
------	--------------	-----------	---------	-----------	-------------

(a) Circle the **two** words that describe the data on pay for the scatter diagram.

(2)

\*(b) Using information from the table and the scatter diagram, discuss the relationship between male pay and female pay in these ten countries.

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

(3)

**(Total for question = 5 marks)**

Q4.

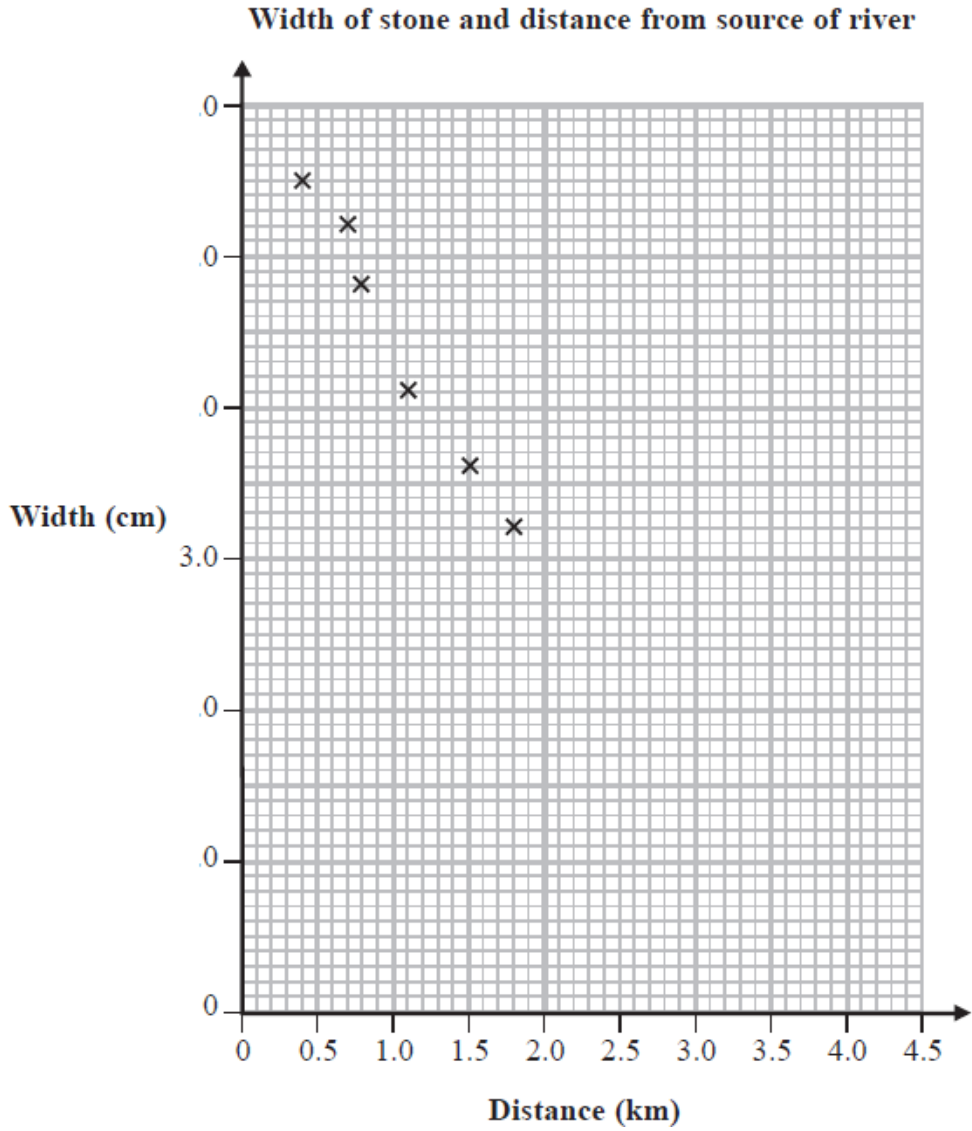
Nadia is investigating the width of stones in a river.

She collects stones at different distances from the source of the river.

The table gives information about her results.

Stone	A	B	C	D	E	F	G	H
Distance from source of river (km)	0.4	0.7	0.8	1.1	1.5	1.8	2.0	2.5
Width of stone (cm)	5.5	5.2	4.8	4.1	3.6	3.2	3.0	2.3

Some of this information is shown on the scatter diagram.



(a) Complete the scatter diagram by plotting the points for stones **G** and **H**.

(2)

(b) Describe and interpret the correlation.

.....

.....

.....

.....

.....

(2)

(c) Draw a line of best fit on the scatter diagram.

(1)

(d) Use your line of best fit to estimate the width of a stone at a distance of 1.3 km from the source of the river.

..... cm

(1)

Nadia wants to predict the width of a stone 3.2 km from the source of the river.

She uses the line of best fit to make this prediction.

This may not be reliable.

(e) Explain why.

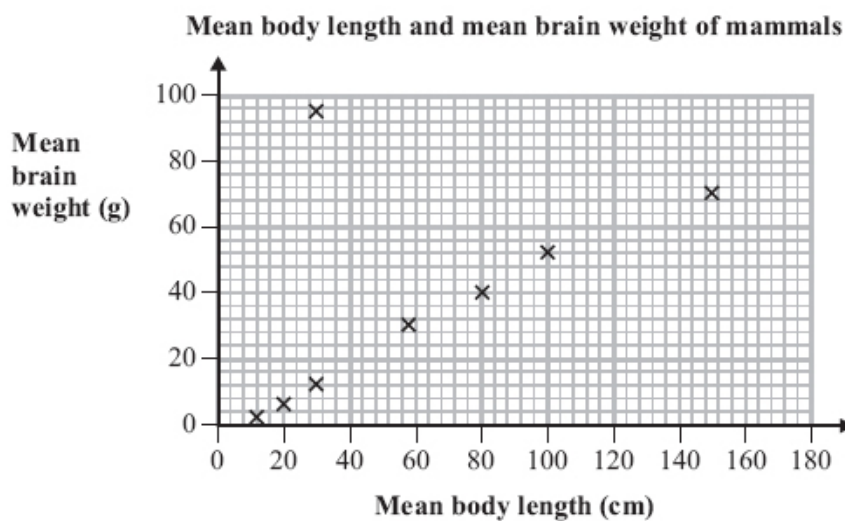
.....  
.....

(1)

**(Total for Question = 7 marks)**

Q5.

The scatter diagram shows the mean body length and the mean brain weight for eight mammals.



*Data Source: Adapted from serendip.brynmawr.edu/*

One point on the scatter diagram represents a mammal that does **not** fit the general pattern.

(a) Draw a circle round this point.



(1)

(b) Using the other seven points, draw a line of best fit.

(1)

(c) Describe and interpret the correlation.

.....

.....

.....

.....

.....

(2)

A mammal has a body length of 120 cm.

(d) (i) Estimate the brain weight of this mammal.

.....g

(ii) Is this estimate likely to be reliable?  
Give a reason for your answer.

.....

.....

.....

.....

.....

(3)

A mammal has a body length of 180 cm.

(e) (i) Estimate the brain weight of this mammal.

.....g

(ii) Is this estimate likely to be reliable?  
Give a reason for your answer.

.....  
.....

.....  
.....

.....  
.....

.....  
.....

(4)

**(Total for Question is 11 marks)**

Q6.

Irina is investigating whether the percentage of the population of a country living in urban areas has an effect on the life expectancy in that country.

.....

.....

(1)

Irina collected the following information about 10 countries.

The Urban population (%) refers to the percentage of the population of the country who live in urban areas.

Country	Urban population (%)	Life expectancy (years)
Australia	89	82
Bangladesh	34	73
China	56	76
Egypt	43	73
France	80	82
Macedonia	57	76
Norway	81	82
Panama	67	79
Thailand	50	75
USA	82	80

(Source: *CIA World Factbook*)

Irina used statistical software to draw a scatter diagram for the information in the table.

(b) Give a reason why a scatter diagram is an appropriate diagram to use.

.....

.....

(1)

(c) For this investigation, which variable is the explanatory variable?  
Give a reason for your answer.

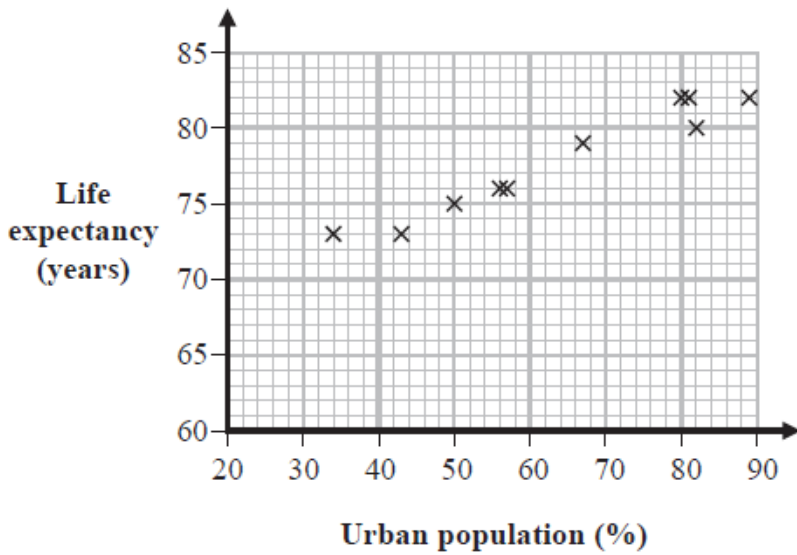
.....

.....

.....

(2)

The scatter diagram from the statistical software is shown below.



(d) Explain, giving a statistical reason, whether or not this scatter diagram supports your hypothesis in part (a).

.....

.....

(2)

For these 10 countries, the double mean point of the data is (63.9, 77.8).

(e) Using this information, draw a line of best fit on the scatter diagram.

(2)

Using statistical software, Irina finds that the gradient of the line of best fit should be 0.19

(f) Interpret the gradient of the line of best fit.

.....

.....

(1)

Irina now finds that South Africa has Urban population 65% and Life expectancy 63 years.

(g) Determine how this information for South Africa fits with the relationship shown in the scatter diagram for the other countries.

.....

.....

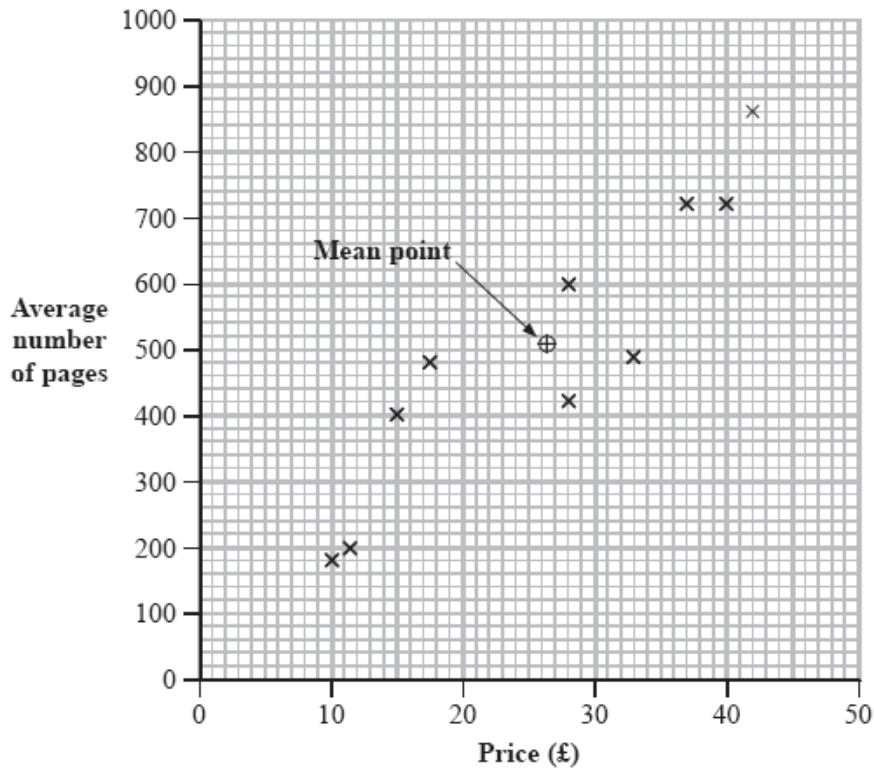
.....

(2)

**(Total for question = 11 marks)**

Q7.

The scatter diagram shows the price of ten printer cartridges and the average number of pages each cartridge prints.



\*(a) Describe and interpret the relationship between the price of a cartridge and the average number of pages it prints.

.....

.....

.....

(2)

(b) Draw a line of best fit through the mean point on the scatter diagram.

(1)

The table shows information about two more cartridges.

	Price	Average number of pages
Cartridge A	£18	200
Cartridge B	£25	680

(c) (i) Plot this information on the scatter diagram and label the points A and B.

\*(ii) Using your line of best fit, explain which of the cartridges in the table is the better value for money.

.....

.....

.....

.....

(4)

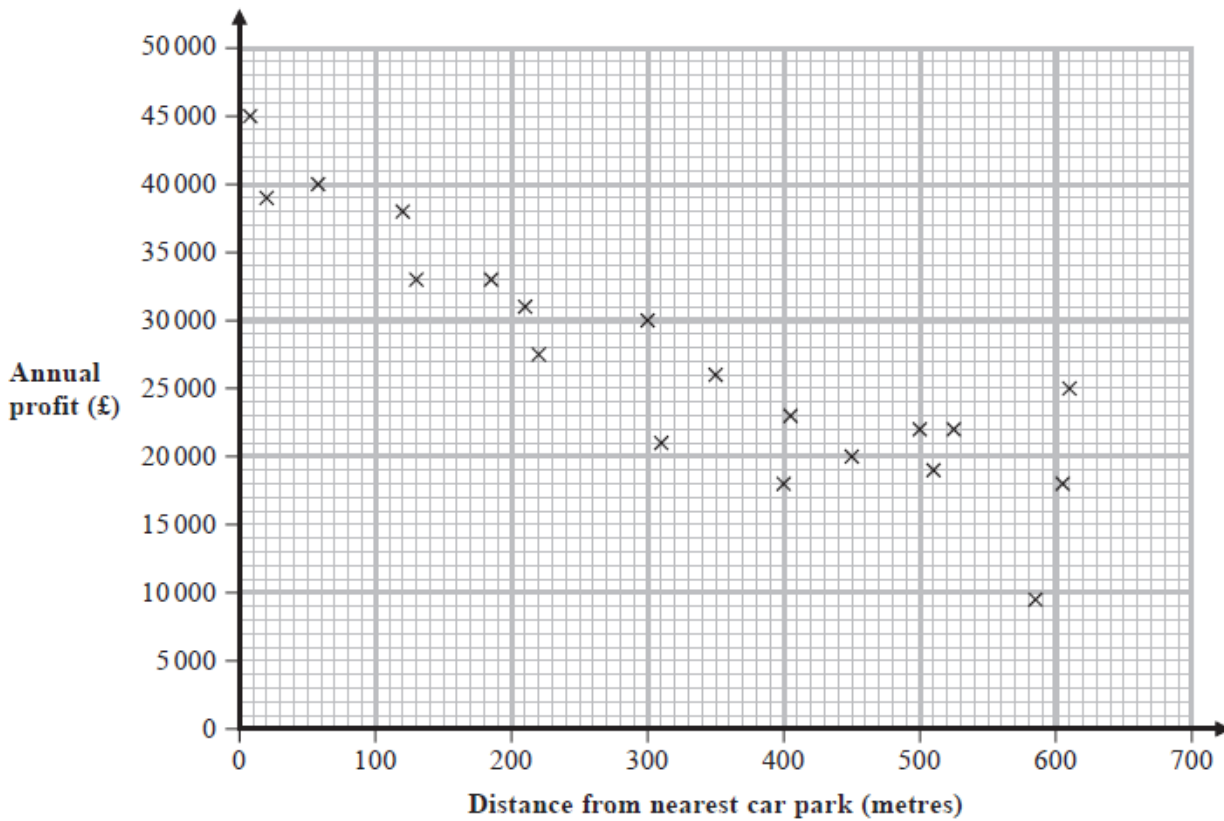
**(Total for question = 7 marks)**

Q8.

Mike is investigating whether or not there is a relationship between the distance from the nearest car park to a restaurant and the annual profit made by the restaurant.

He finds out the distance, in metres, from the nearest car park to a restaurant and the annual profit, in £, made by the restaurant. He does this for twenty restaurants in a city.

Mike then draws a scatter diagram for this information.



(a) Explain why annual profit is the response variable for this scatter diagram.

.....

.....

(1)

Mike's hypothesis is that, for these restaurants, the further the restaurant is from its nearest car park the less is its annual profit.

(b) Explain, giving a statistical reason, whether or not the scatter diagram supports Mike's hypothesis.

.....

.....

.....

(2)

Mike wants to draw a line of best fit on the scatter diagram. Using statistical software he obtains the following information about these restaurants.

Mean distance from the nearest car park	325 m
Mean annual profit	£27000
Intercept of the line of best fit on the Annual profit axis	40000

(c) (i) Using this information, draw a line of best fit on the scatter diagram.

(2)

(ii) Interpret the value of the intercept of the line of best fit on the Annual profit axis.

.....  
.....

(1)

Restaurant A and restaurant B are two other restaurants in the city.

Restaurant A is 250 m from its nearest car park.

Restaurant B is 700 m from its nearest car park.

Mike uses the scatter diagram to find an estimate for the annual profit of each of these restaurants.

(d) Explain which of these two estimates will be the more reliable estimate.

.....  
.....  
.....

(2)

Mike finds a positive correlation between the number of tables at a restaurant and its annual profit.

He concludes that as the number of tables increases this causes the annual profit to increase.

(e) Explain whether or not this conclusion is valid.

.....  
.....

(1)

Mike reads an article in a newspaper that says that restaurant profits for the top 100 restaurants had fallen from £345 million to £125 million in the past year.

(Source: *telegraph.co.uk*)

(f) Using the data in the newspaper article, calculate the percentage decrease in restaurant profits for the top 100 restaurants in the past year.

..... %

(1)

**(Total for question = 10 marks)**

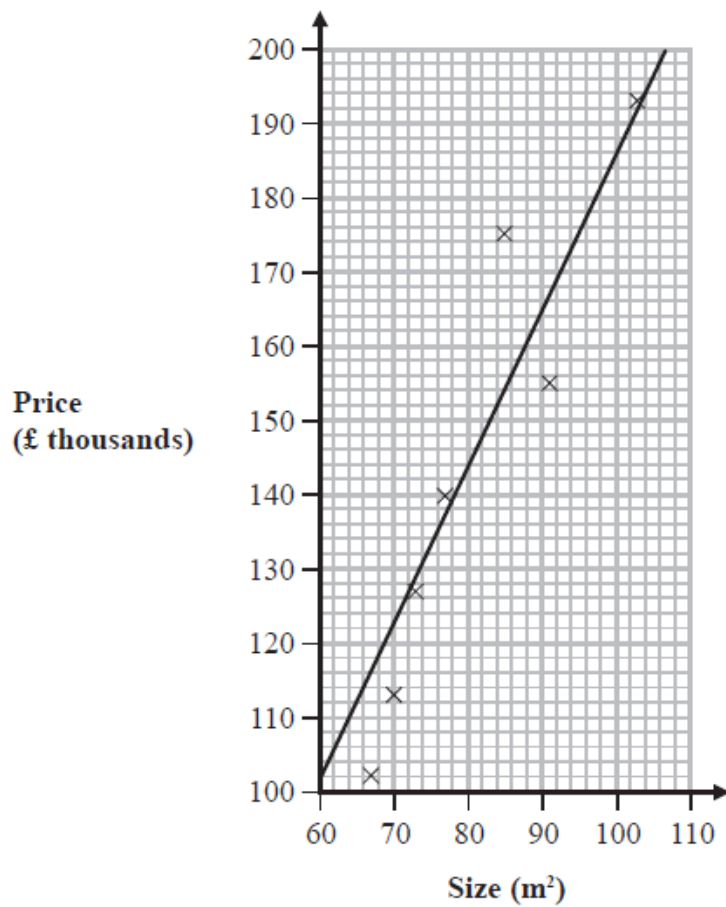


Q9.

The size, in  $m^2$ , and the price, in £ thousands, of 8 houses sold on a particular street in Leeds are shown in the table.

House	A	B	C	D	E	F	G	H
Size ( $m^2$ )	67	70	85	91	73	77	103	62
Price (£ thousands)	102	113	175	155	127	140	193	118

The scatter diagram shows this information for houses **A** to **G**.



Source: *rightmove.co.uk*

(a) Plot the information for house **H** on the scatter diagram.

(1)

(b) Explain why size is the explanatory variable.

.....  
 .....

(1)

(c) Describe and interpret the correlation.

.....  
.....  
.....

(2)

The line of best fit has been drawn on the scatter diagram.

(d) (i) Use the line of best fit to estimate the price of a house on the same street in Leeds with a size of 80 m<sup>2</sup>.

£ ..... thousands

(ii) Give a reason why the estimate should be reliable.

.....  
.....

(2)

(e) Give an example where it may **not** be reliable to use the line of best fit to estimate the price of a house.

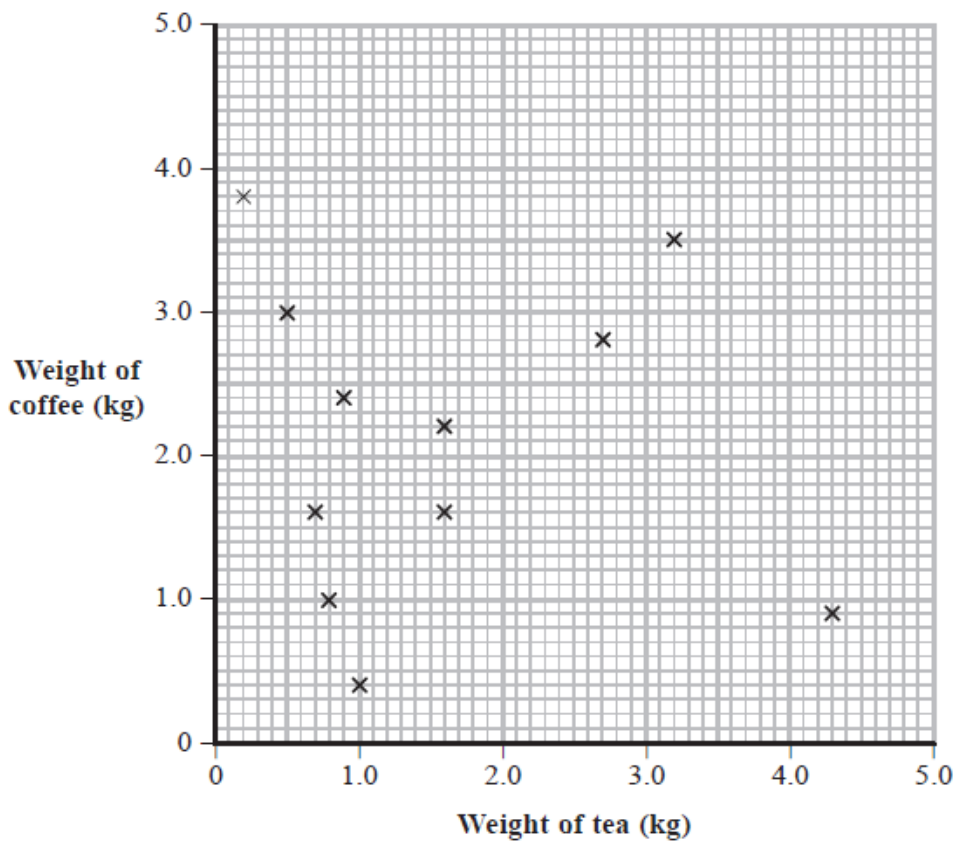
.....  
.....  
.....

(1)

**(Total for question = 7 marks)**

Q10.

The diagram shows information for eleven different countries about the mean annual weight of tea and of coffee used per person to make drinks.



(Data source: World Factbook)

(a) Write down the name given to this type of diagram.

.....  
(1)

Two of the eleven countries are the United Kingdom and the Republic of Ireland.

(b) Use the information in the diagram to complete the table.

Country	Weight of tea (kg)	Weight of coffee (kg)
United Kingdom	2.7	.....
Republic of Ireland	.....	0.9

(2)

For one of the eleven countries, the annual weight of tea is equal to the annual weight of coffee.

(c) On the diagram, circle the point that shows the data for this country.

(1)

Andy wants to make predictions by drawing a line of best fit on the diagram.

\*(d) Discuss whether or not this is a sensible thing to do.

.....

.....  
.....

(2)

**(Total for question = 6 marks)**