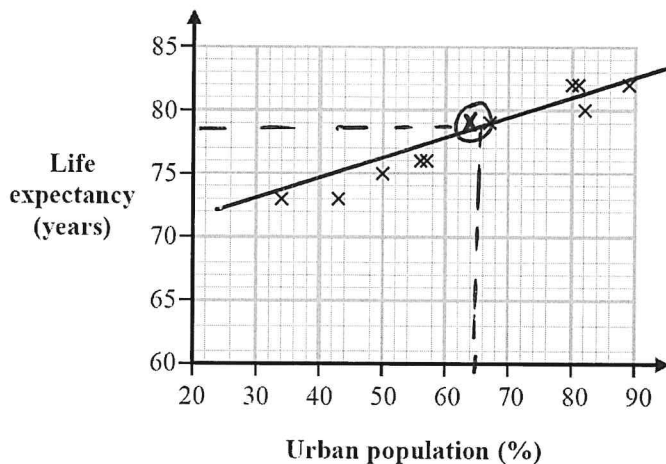


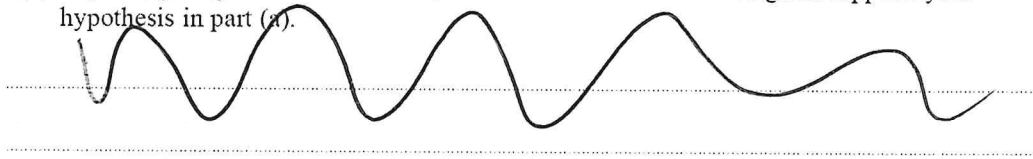
Q1

Scattergraphs and the double mean point

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).



For these 10 countries, the double mean point of the data is  $(\bar{x}, \bar{y})$  (2)

(e) Using this information, draw a line of best fit on the scatter diagram.  
 Plot the point then draw through this point  
 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.  
 $0.19$  is gradient =  $\frac{\text{life exp.}}{\text{urban \%}}$   
 For every 1% more urban pop the life expectancy goes up by 0.19 years (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.

Test by putting urban pop 65% into the graph.  
 You would be expecting 78 years or 79  
 The result is actually 63 years  
 This is a result that doesn't follow the line  
 This result is an **OUTLIER**

# Scattergraphs and the double mean point

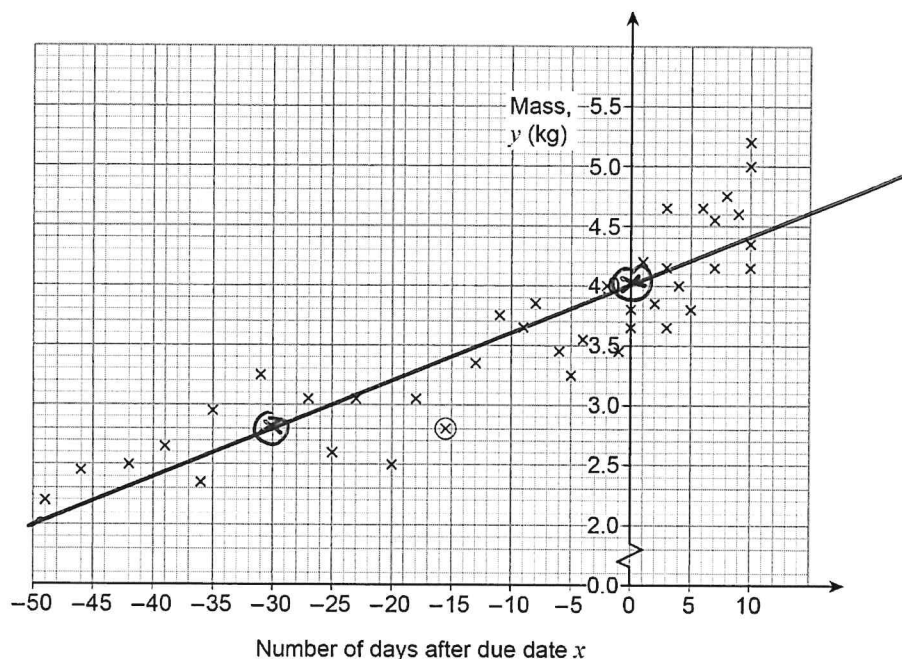
## Q2

The due date of a baby is the date on which it is expected to be born.

The scatter diagram shows the mass of 40 new-born babies born on March 20th 2021 plotted against the number of days the babies were born after their due date in whole days.

For example,

Value of number of days after due date	Interpretation
-10	Baby born 10 days before due date
5	Baby born 5 days after due date



- (a) How many of these babies had a due date of March 10th 2021?  
Circle your answer.

1      2      3      4

The equation of the line of best fit for the data is  $y = 4.01 + 0.04x$

- (c) (i) Interpret the value 4.01 in the context of this scatter graph.

4.01 is the mass of a baby born on exactly the due date.

- (c) (ii) Interpret the value 0.04 in the context of this scatter graph.

0.04 is gradient  
For every day overdue mass increases by 0.04kg  
For every day premature mass decreases by 0.04kg

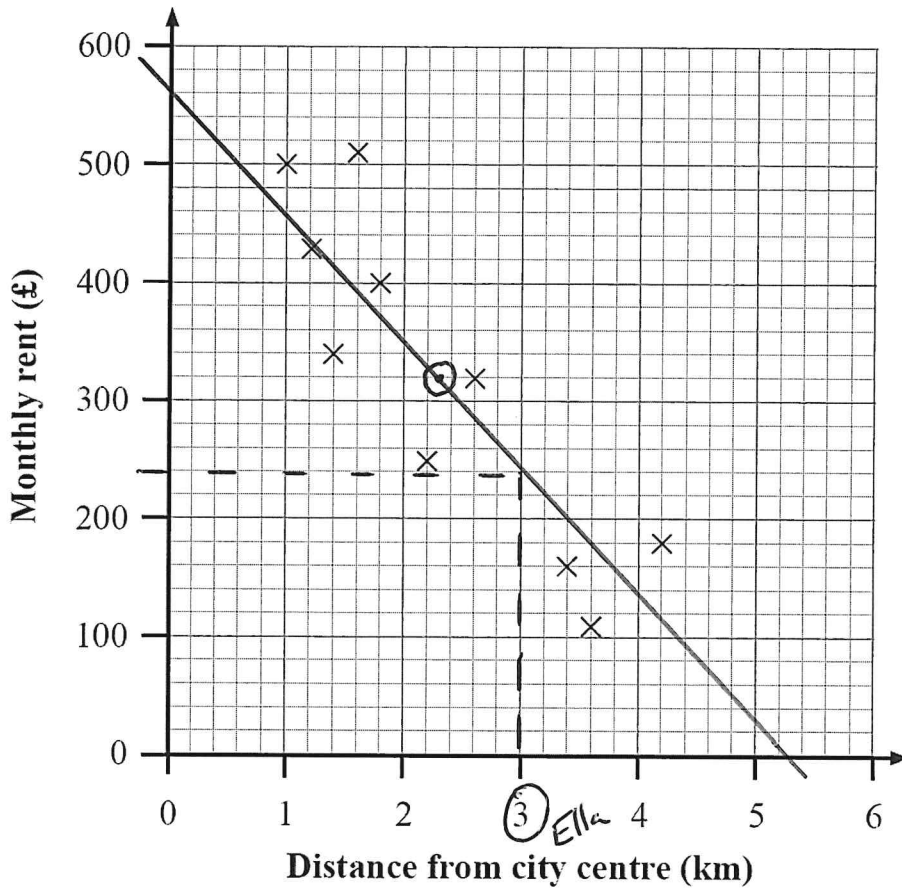
- (c) (iii) Draw the line of best fit on the graph.

⊗ is intercept. Put in  $x = -30$  get point  $(-30, 2.81)$

2 points on the line ⊗ and ⊗. Join up.

The scatter diagram below shows some information about the monthly rent and distance from the city centre for 10 apartments in a city.

**Q3**



(a) Show that the mean distance from the city centre is 2.3 km.

$$\frac{\text{Add } 1 + 1.2 + 1.4 + 1.6 + 1.8 + 2.2 + 2.6 + 3.4 + 3.6 + 4.2}{10} = \frac{23}{10} = 2.3$$

The mean monthly rent for the 10 apartments is £320

(b) Plot the double mean point on the scatter diagram and draw a line of best fit. [2]

⊙ plotted

1) Ella lives 3.0 km from the city centre.

(i) Could Ella use the line of best fit on the scatter diagram to estimate a price for her monthly rent?

Yes

No

[1]

(ii) Give a reason for your answer.

This result is within data values.

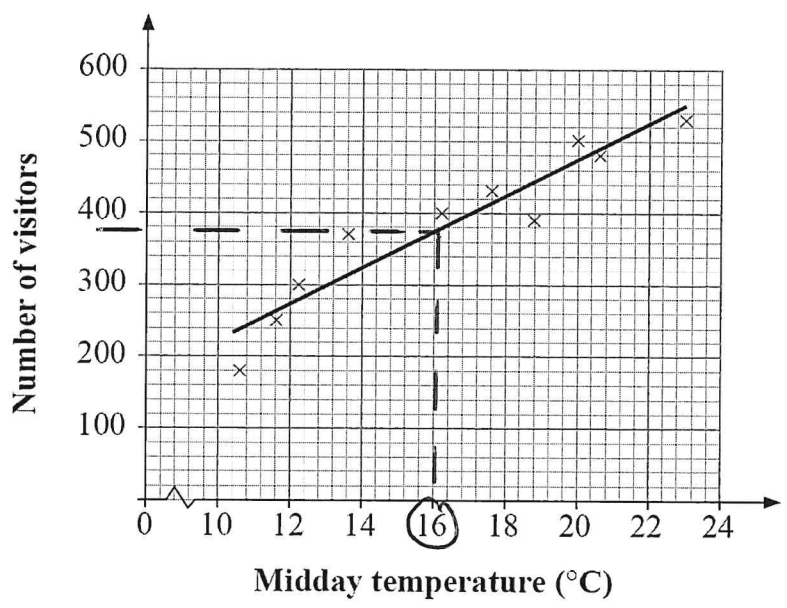
Estimate  
£240

Scattergraphs and the double mean point

Q4

A travel agent recorded the midday temperature,  $x$ , and the number of visitors,  $y$ , to a park over a period of ten Saturdays between April and June.

His results are shown on the scatter diagram below.



(a) What type of data is shown in the scatter diagram?

Circle the two correct words below.

- Qualitative
- Univariate
- Bivariate
- Ordinal
- Quantitative
- Categorical

*"Bi" means 2 values*

*number values*

The equation of the line of best fit is  $y = -25.5 + 24.9x$

(d) Give an interpretation, in context, of the number 24.9 in the equation.

*24.9 is gradient*      *For every 1°C increase*  
*the number of*  
*visitors increase [2] by 24.9*  
*(roughly 25 extra)*

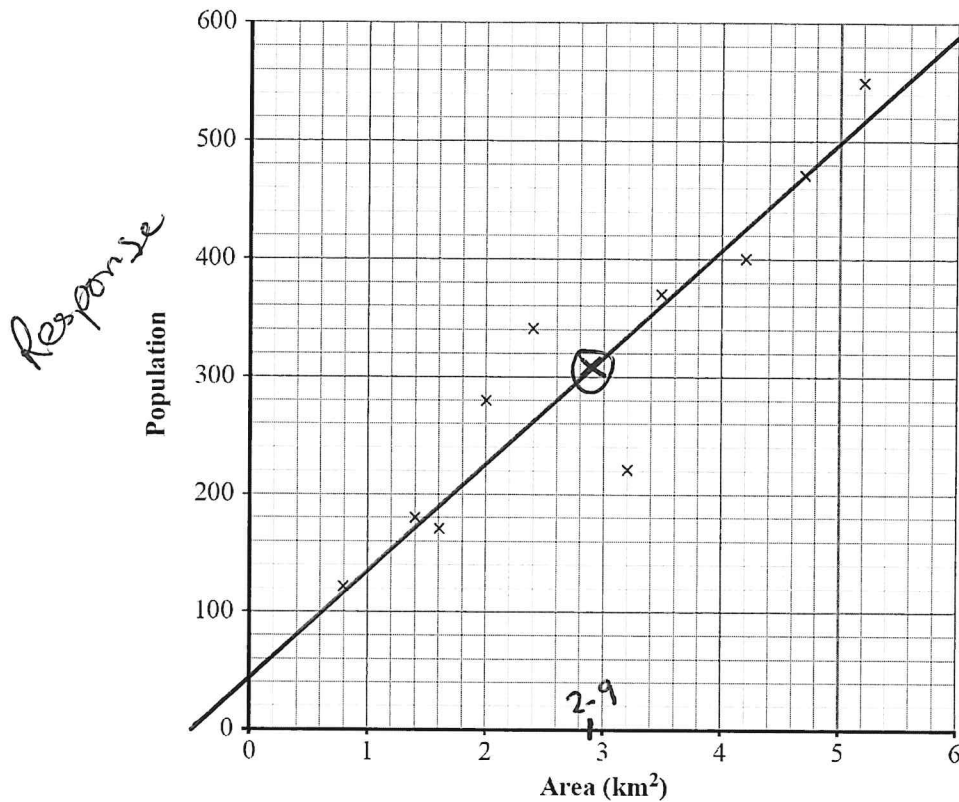
(e) Use the equation of the line of best fit to estimate the number of visitors to the park when the midday temperature is 16°C

*Use*  
*Visitors = -25.5 + 24.9 (temperature)*  
*= -25.5 + 24.9 (16)*      *Answer* \_\_\_\_\_ [2]  
*= 372.9*      *approx 373.*

# Scattergraphs and the double mean point

**Q5** 7 Aoife uses the internet to investigate the land area and population of 10 small villages in Northern Ireland.

Her results are shown in the scatter diagram below.



(a) (i) Identify the response variable in the scatter diagram.

Population

[1]

The mean area of the 10 small villages was 2.9 km<sup>2</sup> and the mean population was 310 people.

Plot (2.9, 310)  
⊗ point

(b) Plot the double mean point on the scatter diagram and draw a line of best fit.

[2]

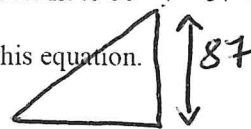
*Just eyeball it*

(c) Describe and interpret the correlation shown in the scatter diagram.

Aoife calculates the equation of her line of best fit to be  $y = 57 + 87x$

(d) Interpret, in context, the number 87 in this equation.

*87 is gradient*



*For every 1 km<sup>2</sup> extra the pop. rises by 87*

(e) Use the equation of her line of best fit to calculate an estimate of the population of a small village with a land area of 4.4 km<sup>2</sup>

*Use eq (not graph)*

$$\text{Pop} = 57 + 87(4.4) = 439.8$$

*Ans 440*