



General Certificate of Secondary Education  
2024

Centre Number

--	--	--	--	--

Candidate Number

--	--	--	--	--

# Statistics

## Unit 2 (With Calculator)

Higher Tier

[GST22]



\*GST22\*

**MONDAY 17 JUNE, AFTERNOON**

### TIME

2 hours.

### INSTRUCTIONS TO CANDIDATES

Write your Centre Number and Candidate Number in the spaces provided at the top of this page.

**You must answer the questions in the spaces provided.**

**Do not write outside the boxed area on each page or on blank pages.**

Complete in black ink only. **Do not write with a gel pen.**

Questions which require drawing can be completed using an HB pencil.

Any working **must** be clearly shown in the spaces provided. Marks may be awarded for partially correct solutions.

Answer **all ten** questions.

### INFORMATION FOR CANDIDATES

The total mark for this paper is 100.

Figures in brackets printed down the right-hand side of pages indicate the marks awarded to each question or part question.

You should have a calculator, ruler, compasses and protractor.

The formula sheet is on page 2.

14285



\*28GST2201\*

## HIGHER TIER FORMULA SHEET

$$\text{Standard deviation} = \sqrt{\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2}$$

Spearman's Rank Correlation Coefficient

$$r_s = 1 - \left(\frac{6 \sum d^2}{n(n^2 - 1)}\right)$$





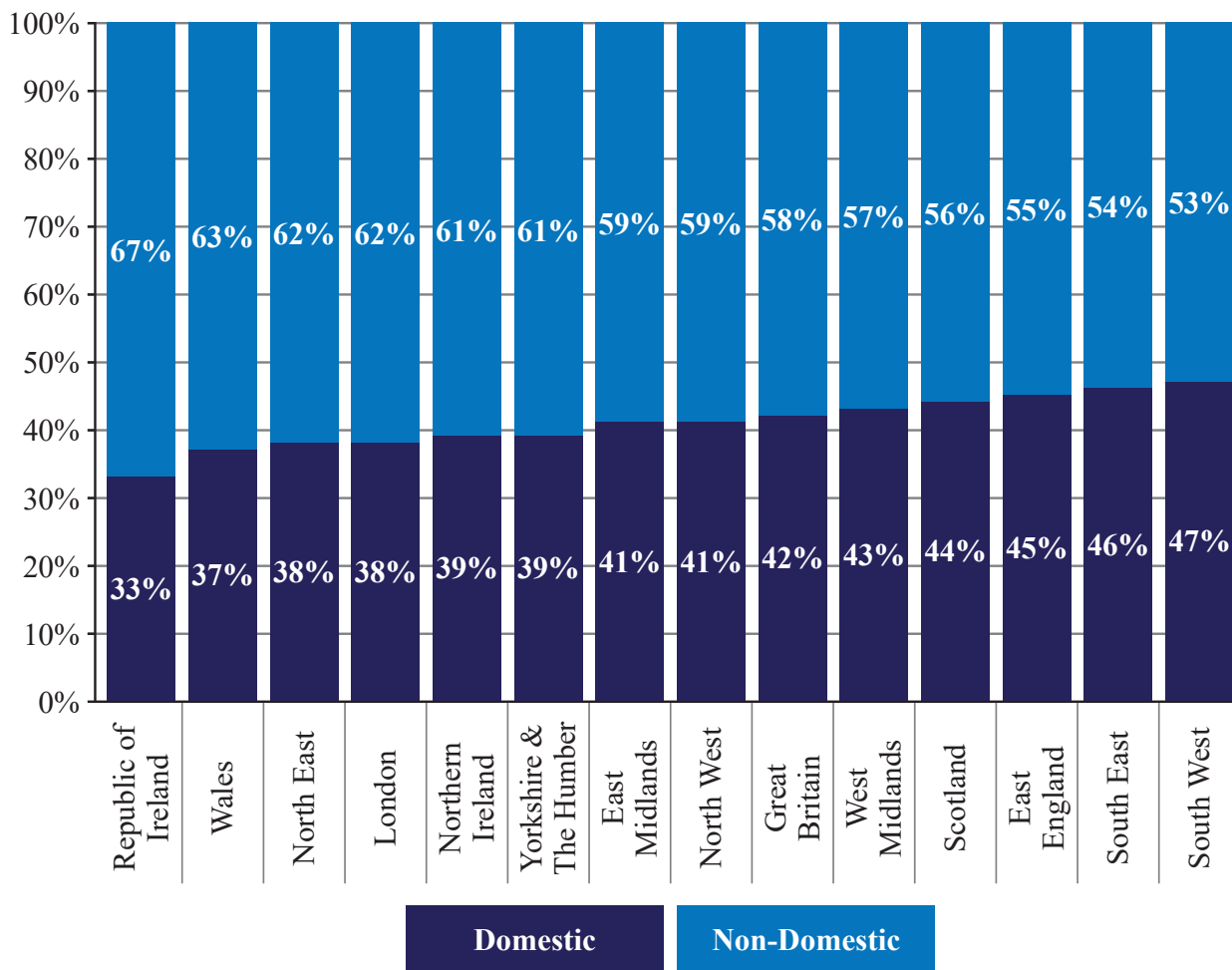
**BLANK PAGE**  
**DO NOT WRITE ON THIS PAGE**  
**(Questions start overleaf)**



Answer **all** questions

- 1 The chart below shows the distribution of domestic and non-domestic electricity consumption by country and region for 2020

**Distribution of domestic/non-domestic electricity consumption by country and region, 2020**



Source: © Crown copyright /BEIS; NIE Networks; Commission for Regulation of Utilities

- (a) What type of bar chart is this?

Circle the correct answer.

**Compound percentage bar chart**

**Multiple bar chart**

[1]



(b) For which country or region shown in the chart is the non-domestic electricity consumption approximately twice the domestic electricity consumption?

[1]

\_\_\_\_\_

Neil says that the North East uses the same amount of domestic electricity as London.

(c) Is Neil correct?

Tick the correct box.

Yes

No

Cannot tell

Explain your answer.

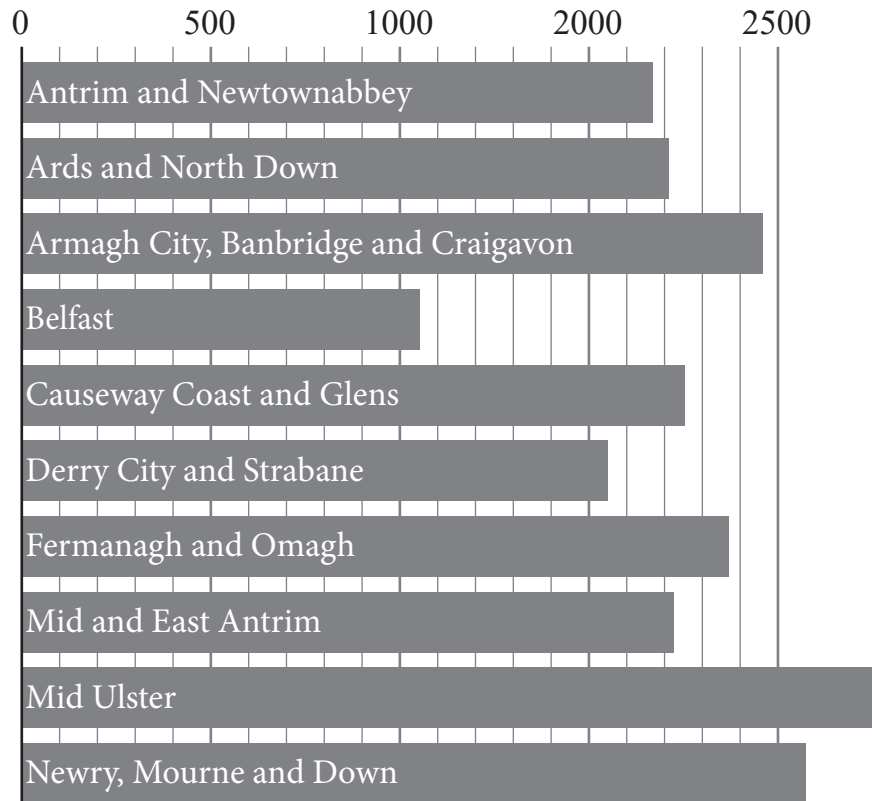
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_ [3]

[Turn over



- 2 Macy has drawn a bar chart to compare the average domestic electricity consumption per meter (kWh) for each of the 11 council areas in Northern Ireland over a period of time.

**Average domestic electricity consumption per meter (kWh)**



Identify three issues with this bar chart.

1. \_\_\_\_\_ [1]
2. \_\_\_\_\_ [1]
3. \_\_\_\_\_ [1]





**BLANK PAGE**

**DO NOT WRITE ON THIS PAGE**

**(Questions continue overleaf)**

**[Turn over**



3 Jack is investigating how often university students visit the library.

He asked 15 students from his course how many times they had visited the library in the past term.

The results are as follows:

58	36	55	85	55	69	48	65
105	42	54	56	34	55	37	

(a) In the space below, draw a stem and leaf diagram to illustrate Jack's data.

[3]



The lower quartile for Jack's data is 42

(b) Explain the meaning of this value in Jack's investigation.

---

---

[2]

(c) Calculate the interquartile range for the number of visits to the library.

Answer \_\_\_\_\_ [2]

(d) Using your answer to part (c), determine whether or not Jack's data contains any outliers.

[4]

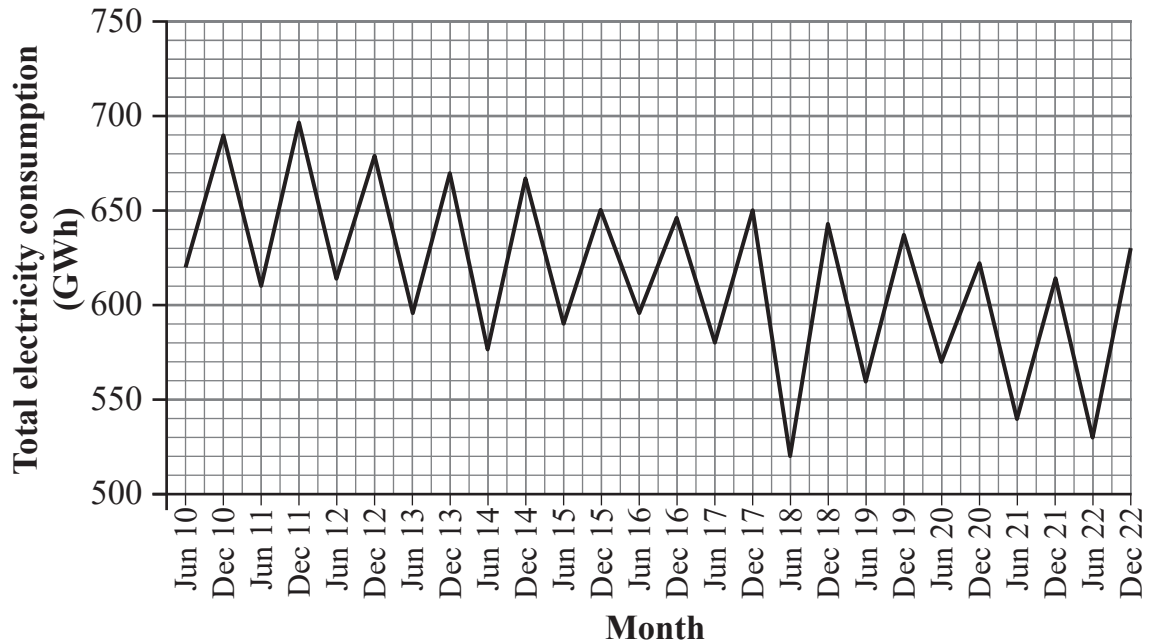
[Turn over



4 Graham works for a large international company.

He records the company's total electricity consumption every six months.

His results for 2010 to 2022 are shown in the time series graph below.



(a) What is a time series?

---

---

---

[2]

(b) In what month and year was the total electricity consumption the lowest?

---

[1]

(c) Draw a trend line on the graph above.

[1]



(d) What does your trend line show about the company's total electricity consumption over this period?

---

---

[1]

Graham wants to use a trend line to predict the total electricity consumption for 2023

(e) What assumption must Graham make in order to do this?

---

---

[1]

[Turn over



- 5 The table below shows data relating to domestic and non-domestic electricity consumption at district council level in Northern Ireland for 2020 to 2021

**Electricity consumption statistics at district council level, 2020 to 2021**

Council name	Domestic			Non-domestic		
	Total consumption (GWh)	Total number of meters	Average consumption per meter (kWh)	Total consumption (GWh)	Total number of meters	Average consumption per meter (kWh)
Antrim & Newtownabbey	234	63 647	3 676	332	4 130	80 392
Ards & North Down	282	75 867	3 715	183	4 847	37 774
Armagh City, Banbridge & Craigavon	357	90 030	3 965	529	7 347	71 942
Belfast	505	165 019	3 061	827	13 681	60 482
Causeway Coast & Glens	255	67 723	3 758	226	5 630	40 215
Derry City & Strabane	231	65 108	3 550	406	5 373	75 496
Fermanagh & Omagh	194	49 995	3 877	321	5 077	63 258
Lisburn & Castlereagh	242	64 233	3 767	241	4 363	55 337
Mid & East Antrim	234	62 887	3 729	248	4 622	53 698
Mid Ulster	245	57 243	4 283	476	6 348	74 933
Newry, Mourne & Down	304	74 520	4 081	283	6 689	42 317
Unallocated	6	2 628	2 273	16	190	82 403
<b>Northern Ireland</b>	<b>3 089</b>	<b>838 900</b>	<b>3 682</b>	<b>4 088</b>	<b>68 297</b>	<b>59 861</b>

Source: © Crown copyright /BEIS (Sub-national electricity consumption statistics in Northern Ireland, gov.uk website)



(a) (i) Show that the mean total consumption of non-domestic electricity for the 11 council areas (i.e. excluding Unallocated) is 370 GWh to the nearest whole number.

[2]

(ii) Write down the name of the council area in which the total consumption of non-domestic electricity is closest to the mean for all 11 council areas.

[1]

\_\_\_\_\_

Ryan is investigating whether there is a link between the total consumption of domestic electricity and the total consumption of non-domestic electricity for each of the 11 council areas.

He calculates the product moment correlation coefficient to be 0.793

(b) How might Ryan interpret this value?

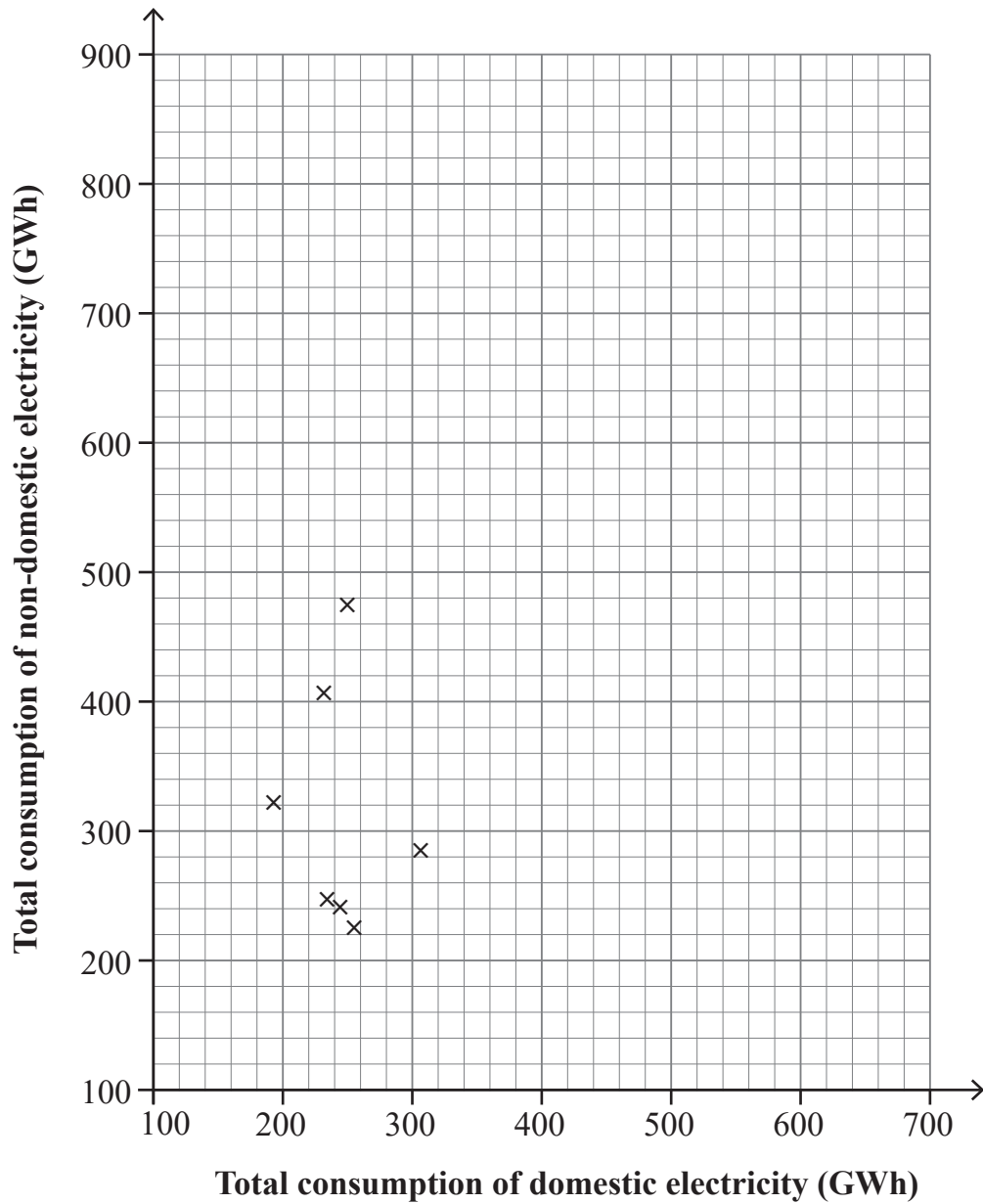
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

[2]

[Turn over



Ryan notices that there may be an outlier in the data, so he decides to plot a scatter diagram.



(c) On the scatter diagram above, plot the points for the first four council areas from the table overleaf.

The points for the other council areas have already been plotted.

[2]



(d) (i) On the scatter diagram opposite, draw a circle around the point which represents the outlier and write down the name of the council area.

\_\_\_\_\_ [2]

(ii) Suggest a possible explanation for this.

\_\_\_\_\_  
\_\_\_\_\_ [2]

(e) With the outlier removed, calculate the product moment correlation coefficient for the remaining 10 council areas, giving your answer correct to three decimal places.

Answer \_\_\_\_\_ [2]

(f) What conclusion could Ryan draw from the value calculated in part (e)?

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_ [2]

[Turn over



6 The table below shows the total annual energy consumption in Northern Ireland (GWh) from 2009 to 2014

Some simple and chain base index numbers have been calculated.

The base year is 2009

Year	2009	2010	2011	2012	2013	2014
<b>Total annual energy consumption in Northern Ireland (GWh)</b>	8049		8235	8095	8181	8022
<b>Chain base index number</b>	100	104.76	97.66	98.30		98.06
<b>Simple index number</b>	100	104.76	102.31		101.64	99.66

(a) Based on the information in the table, explain what a simple index number shows.

---

---

[2]

The chain base index number for 2014 is 98.06

(b) Give an interpretation of this value.

---

---

---

[3]



(c) Showing your working clearly, calculate:

(i) the simple index number for 2012;

Answer \_\_\_\_\_ [2]

(ii) the chain base index number for 2013;

Answer \_\_\_\_\_ [2]

(iii) the total annual energy consumption in 2010

Answer \_\_\_\_\_ GWh [2]

(d) Describe the overall pattern in total annual energy consumption between 2009 and 2014, and suggest one factor which may have contributed to this pattern.

---

---

---

---

[2]

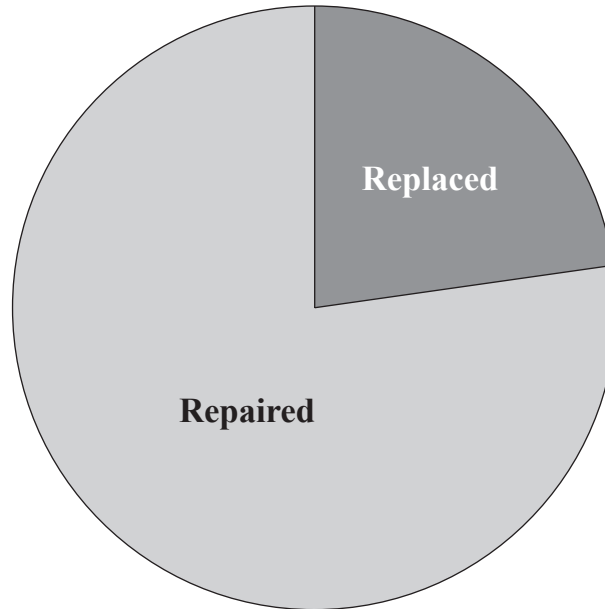
[Turn over



7 Easiview is a company specialising in windscreen replacement and repair.

In 2022, Easiview replaced or repaired a total of 3178 windscreens.

The pie chart below shows the proportions of windscreens which were replaced or repaired.



(a) How many windscreens were repaired in 2022?

Answer \_\_\_\_\_ [2]

In 2023, Easiview replaced 912 windscreens. The angle for the sector in the 2023 pie chart representing this value is  $77^\circ$

(b) Calculate the number of windscreens which were repaired by Easiview in 2023

Answer \_\_\_\_\_ [2]



(c) Calculate the radius of the pie chart for 2023

Answer \_\_\_\_\_ cm [3]

When Easiview replace a windscreen, they fit one of two types:  
Crystalclear and Omniglaze.

0.5% of Crystalclear windscreens are defective.  
1 in 800 Omniglaze windscreens are defective.

(d) Show that the relative risk that a Crystalclear windscreen is defective compared to an Omniglaze windscreen being defective is 4

[3]

(e) Give an interpretation of the value of the relative risk stated in part (d).

---

---

---

[2]

[Turn over



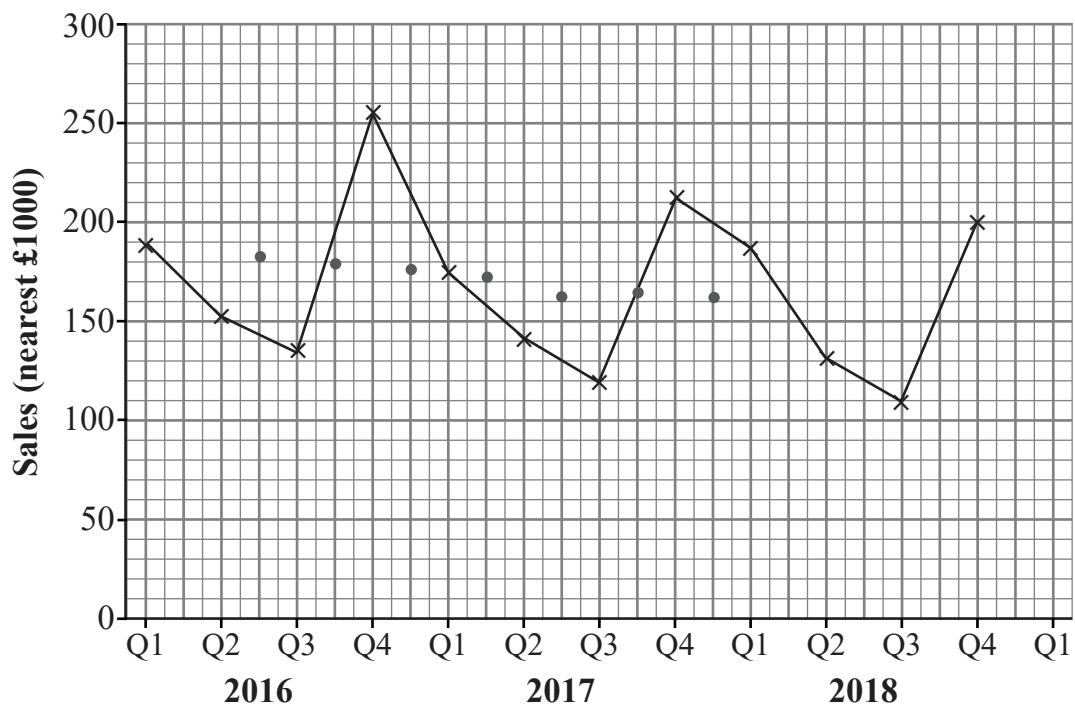
- 8 The table below shows the sales (nearest £1000) for a factory in Northern Ireland between 2016 and 2018

		Sales (nearest £1000)			
		Q1	Q2	Q3	Q4
Year	2016	188	152	134	255
	2017	174	140	118	212
	2018	186	130	108	200

- (a) For which quarter are sales highest each year?

Answer \_\_\_\_\_ [1]

The time series graph below illustrates the data from the table.



The first seven 4-point moving averages have been plotted on the graph.

(b) Calculate the last two 4-point moving averages.

Answer £ \_\_\_\_\_ thousand

£ \_\_\_\_\_ thousand [2]

(c) Plot the moving averages calculated in part (b) on the graph and draw a trend line. [2]

(d) Comment on the trend in the time series graph.

\_\_\_\_\_  
\_\_\_\_\_ [1]

(e) Showing clearly where any reading is taken, use your trend line to estimate the sales, to the nearest £1000, for Quarter 1 of 2019

Answer £ \_\_\_\_\_ thousand [3]

[Turn over



- 9 A Mathematics teacher is looking at the assessment results for a small group of students.

The table below shows the results for the Geography, Mathematics and Science assessments for this group of students.

<b>Geography</b>	57	78	88	80	85	81	91	92	77	90	100	95	89
<b>Mathematics</b>	55	62	85	66	66	75	75	80	72	92	95	92	89
<b>Science</b>	54	65	76	67	65	74	72	74	65	83	82	85	80

- (a) What type of data is the teacher using?

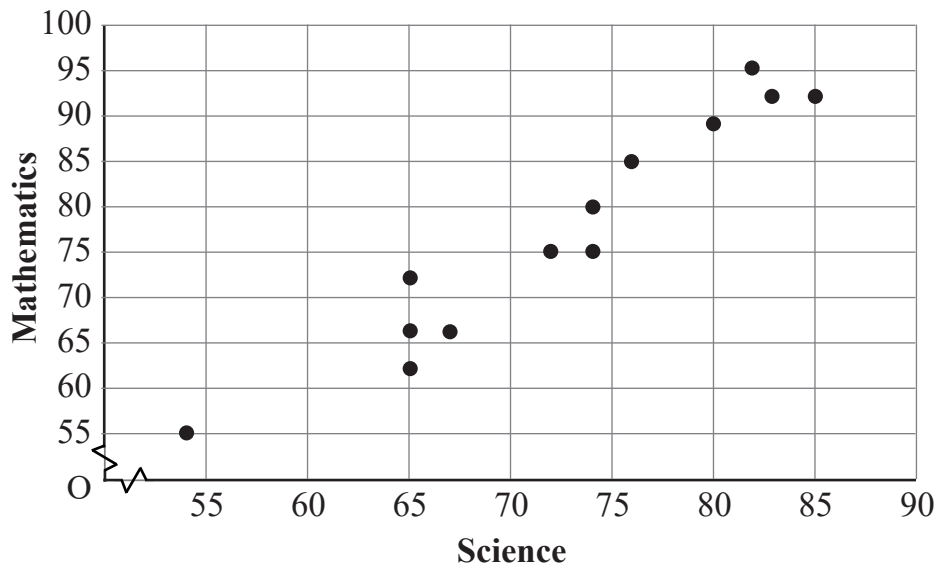
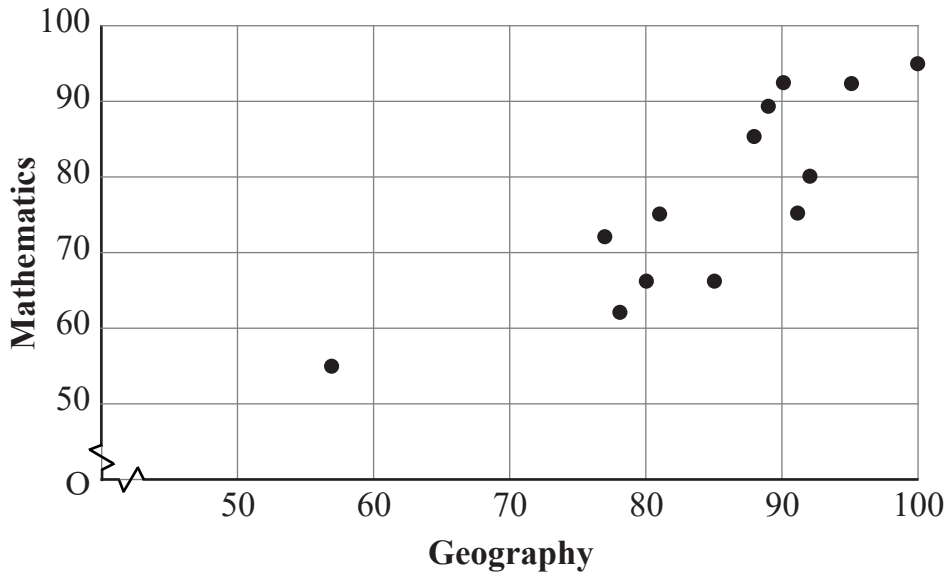
Circle two answers.

**Univariate      Bivariate      Multivariate      Discrete      Categorical**

[2]



The teacher draws two scatter diagrams: one for Geography and Mathematics and another for Science and Mathematics.



(b) Comment on how these scatter diagrams have been drawn.

---



---

[2]

[Turn over



The teacher calculates Spearman's rank correlation coefficient for each of the pairs of subjects shown in the scatter diagrams.

The results are as follows:

Geography and Mathematics:  $r_s = 0.850$

Science and Mathematics:  $r_s = 0.957$

(c) What do these values show?

---

---

---

[2]



The teacher then finds an equation for the line of best fit for each scatter diagram.

The results are as follows:

Geography and Mathematics:  $y = -22.491 + 1.376x$

Science and Mathematics:  $y = -7.783 + 1.002x$

A student misses the Mathematics assessment but obtains 78 in Geography and 78 in Science.

- (d) (i)** Which of the equations should the teacher use to estimate the student's Mathematics result?

Explain your answer.

---

---

[2]

- (ii)** Use the equation you chose in part **(d)(i)** to estimate this student's Mathematics result.

Answer \_\_\_\_\_ [2]

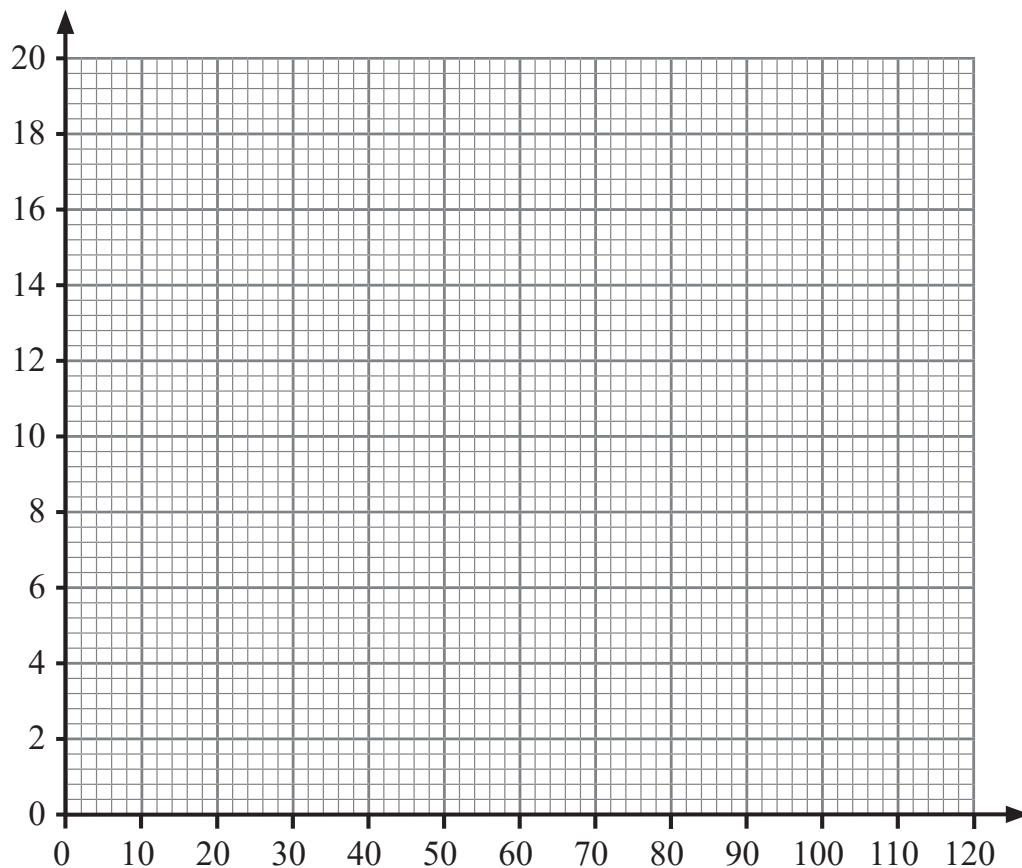
[Turn over



- 10 The table below shows the mileage, in thousands of miles, of the cars for sale at Rory's garage.

Mileage, $m$ (1000 miles)	Frequency		
$0 \leq m < 20$	8		
$20 \leq m < 40$	17		
$40 \leq m < 60$	19		
$60 \leq m < 80$	9		
$80 \leq m < 100$	4		
$100 \leq m < 120$	1		

- (a) On the grid below, draw a frequency polygon to illustrate the distribution of the mileages of the cars for sale in the garage.



[4]



- (b) Comment on the skewness shown in the frequency polygon and state whether or not the normal distribution would be an appropriate model for this data.

---

---

---

[2]

An advertisement for Rory's garage states:

*"The average mileage of the cars for sale in this garage is under 45 000 miles."*

- (c) Using the data in the table and appropriate calculations, decide whether Rory should use the mean or median mileage to support this statement.

You may use the blank columns in the table on the opposite page.

Answer \_\_\_\_\_

---

[10]



---

**THIS IS THE END OF THE QUESTION PAPER**

---

**DO NOT WRITE ON THIS PAGE**

<b>For Examiner's use only</b>	
<b>Question Number</b>	<b>Marks</b>
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	

<b>Total Marks</b>	
--------------------	--

Examiner Number

--

Permission to reproduce all copyright material has been applied for.  
In some cases, efforts to contact copyright holders may have been unsuccessful and CCEA will be happy to rectify any omissions of acknowledgement in future if notified.

14285/7



\*28GST2228\*