



General Certificate of Secondary Education
2022

Centre Number

--	--	--	--	--

Candidate Number

--	--	--	--	--

Statistics

Unit 2
Higher Tier



MV18

[GST22]

FRIDAY 24 JUNE, MORNING

Time

2 hours, plus your additional time allowance.

Instructions to Candidates

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

Write your answers in the spaces provided in this question paper.

Answer **all ten** questions.

Any working should be clearly shown in the spaces provided since marks may be awarded for partially correct solutions.

You **may** use a calculator for this paper.

Information for Candidates

The total mark for this paper is 100.

Figures in brackets printed at the end of each question 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.

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
(Questions start overleaf)

Answer **all** questions

1 Megan is carrying out an investigation about how pupils travel to school.

She wants to compare the main methods of travel used by pupils in Year 8 and Year 9 at her school.

Megan collects the data from the pupils herself.

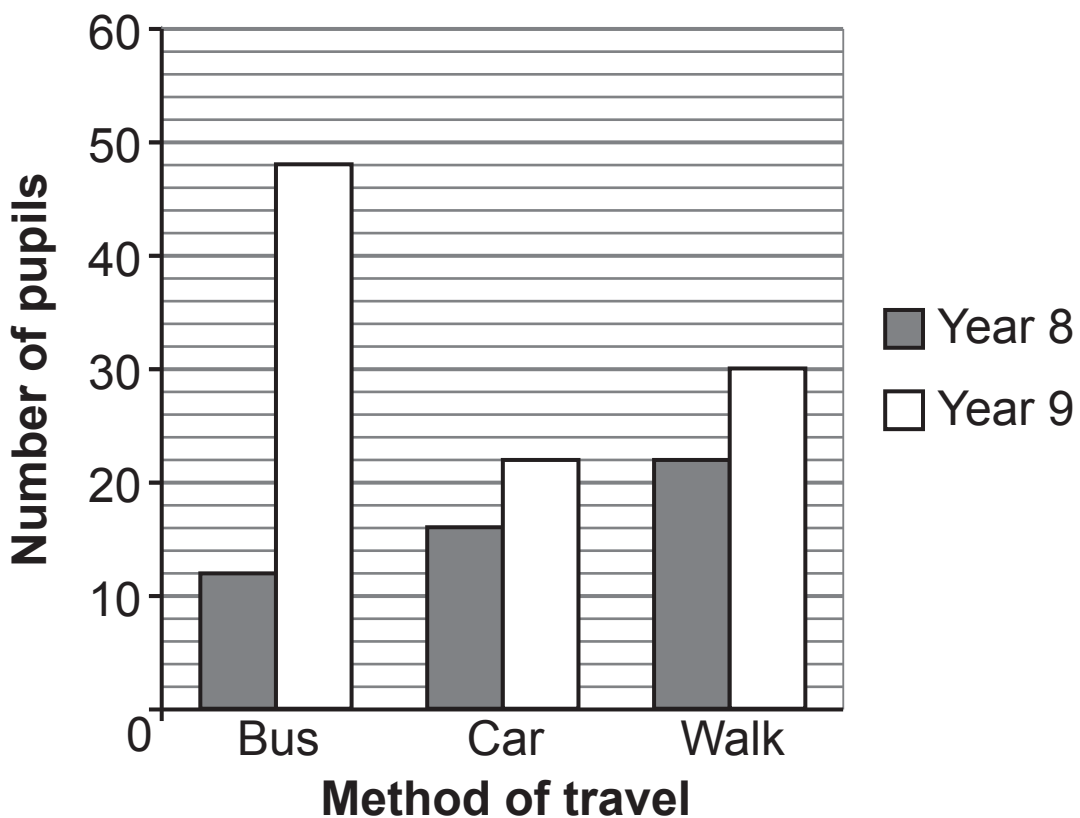
(a) What type of data is Megan using? [1 mark]

Circle the correct answer.

Primary data

Secondary data

The dual bar chart below shows the data Megan collected.

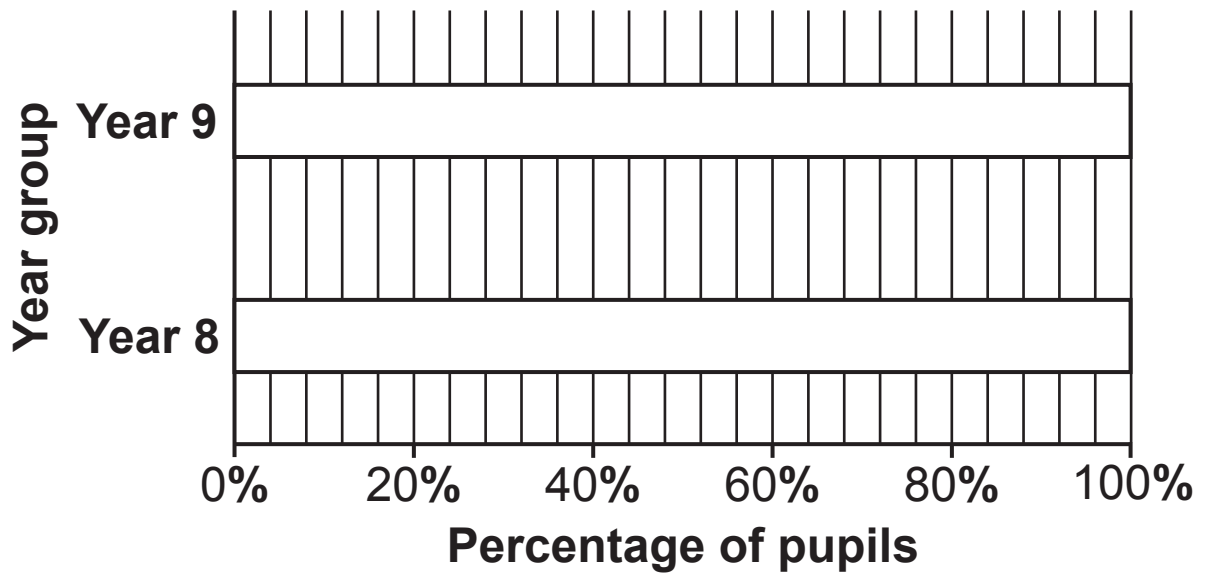


(b) How many more Year 9 pupils walk to school than Year 8 pupils? [1 mark]

Answer _____

(c) Give a reason why Megan's choice of chart may not be appropriate. [2 marks]

(d) On the grid below, draw a compound percentage bar chart for Megan's data. [4 marks]



(e) Use your chart to compare the methods of travel to school by pupils in Year 8 and Year 9 [2 marks]

Blank Page
(Questions continue overleaf)

Megan extended her investigation by asking the same pupils about their travel times to school.

Her results, in minutes, are given below for some of the Year 8 pupils.

These pupils all travelled to school by the same method.

7	9	15	23
25	12	53	14
12	10	15	12
13	17	13	6

(f) How did these pupils travel to school? [1 mark]

Answer _____

(g) For Megan's data above:

(i) find the median travel time; [2 marks]

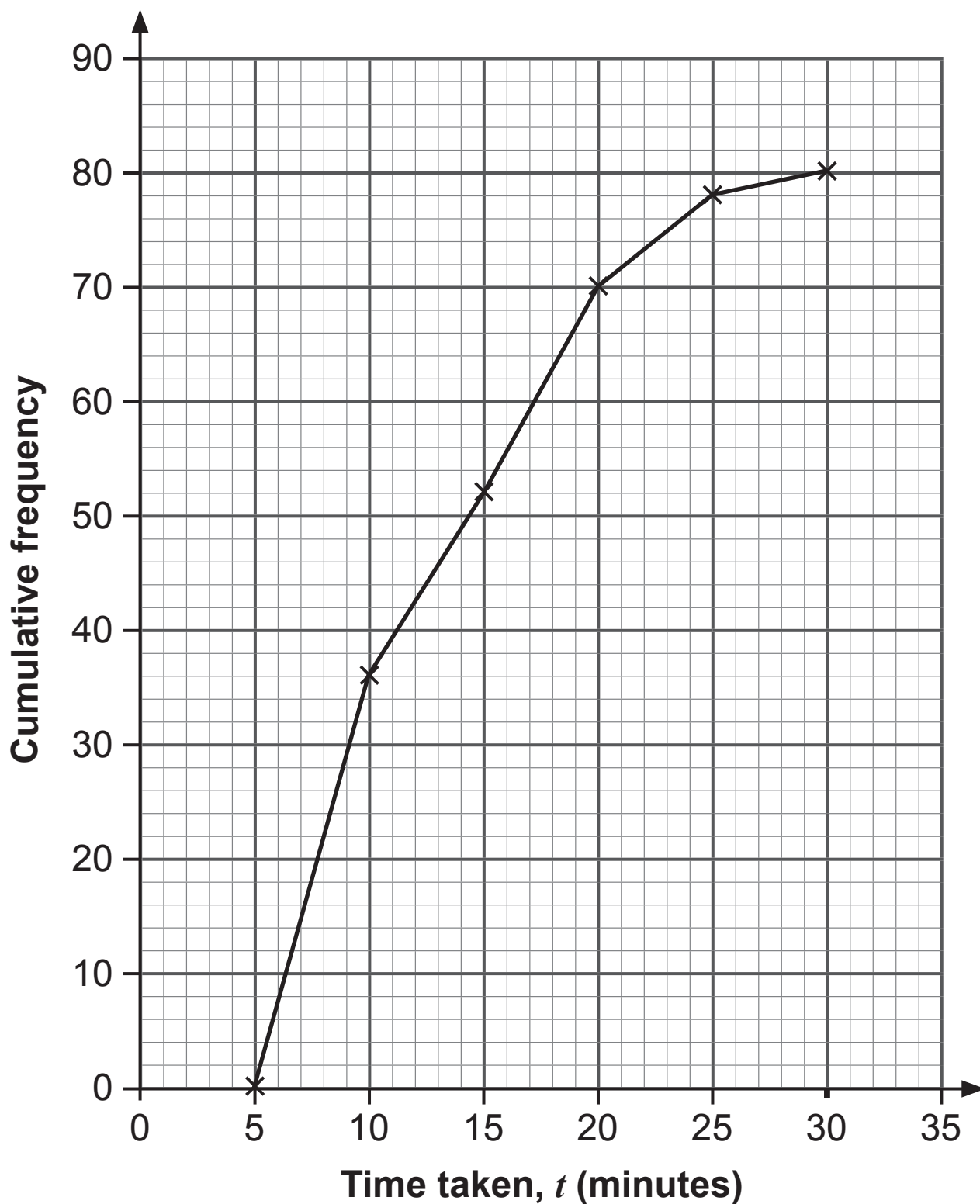
Answer _____ minutes

(ii) calculate the mean travel time. [2 marks]

Answer _____ minutes

(iii) Is the median or mean a more appropriate measure of average for Megan's data? Give a reason for your answer. [2 marks]

- 2 The times taken by a sample of people to complete a puzzle in a competition are summarised in the cumulative frequency diagram below.



(a) What is the size of the sample? [1 mark]

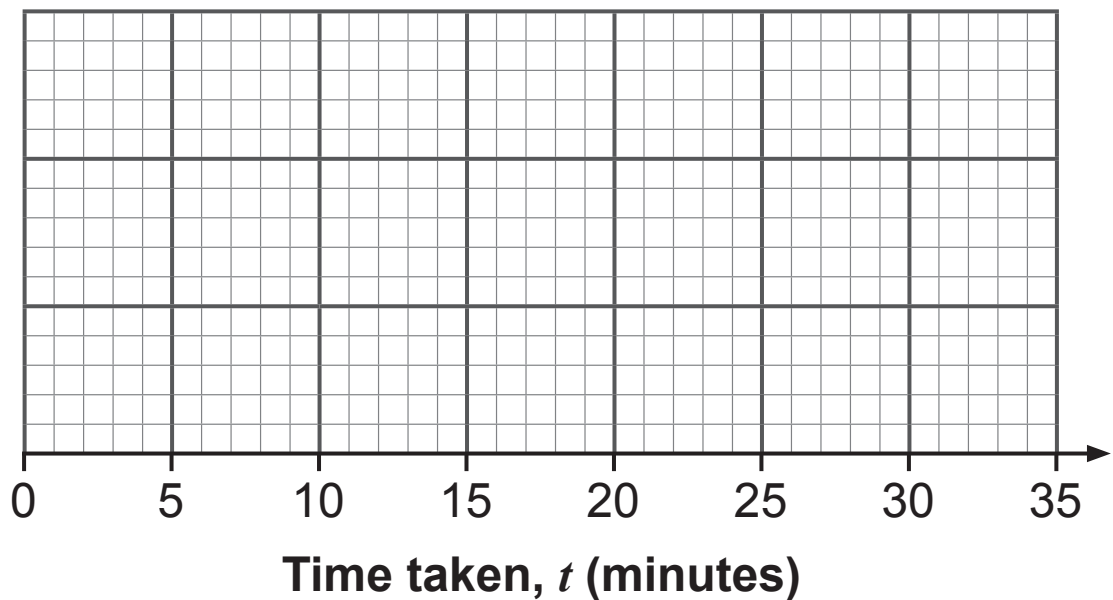
Answer _____

(b) Write down the modal class. [1 mark]

Answer _____

The fastest time to complete the puzzle was 6 minutes and the slowest time was 28 minutes.

- (c) On the grid below, draw a box plot to illustrate the information in the cumulative frequency diagram.
[4 marks]



- (d) Describe the skewness shown in the box plot.
[1 mark]

(e) Could the normal distribution be used as a model for this data? [2 marks]

Tick the correct box:

Yes

No

Explain your answer.

Competitors who finish the puzzle in 14 minutes or less go through to the next round of the competition.

(f) Estimate the number of competitors from the sample who will go through to the next round. [1 mark]

Answer _____

- 3 The table below shows the number of monthly notified vacancies between July 2019 and June 2020 in Northern Ireland.

Vacancies notified: July 2019 to June 2020

Month	Notified vacancies			
	Full-time	Part-time	Casual	Total
Jul-19	4415	1404	648	6467
Aug-19	4555	1817	501	6873
Sep-19	4240	1872	609	6721
Oct-19	5756	1922	591	8269
Nov-19	5295	1510	602	7407
Dec-19	2541	632	372	3545
Jan-20	4583	1513	633	6729
Feb-20	3375	1402	546	5323
Mar-20	2927	893	533	4353
Apr-20	1792	329	238	2359
May-20	2108	325	342	2775
Jun-20	2030	443	304	2777
Total	43 617	14 062	5 919	63 598

- (a) Which month had the fewest notified vacancies in total?
[1 mark]

Answer _____

Tom says that 111 part-time vacancies were filled between January 2020 and February 2020

(b) (i) How did Tom calculate this number? [1 mark]

(ii) Give a reason why Tom may not be correct.
[2 marks]

Tom decides to investigate if there is a relationship between the number of full-time vacancies each month and the number of part-time vacancies in the same month.

(c) Suggest a suitable hypothesis for Tom's investigation.
[1 mark]

(d) Write down the name of one calculation which Tom could use to decide whether or not his hypothesis is valid. [1 mark]

- (e) Using the table on page 14, calculate the value of the statistical calculation you identified in part (d), giving your answer correct to three decimal places.
[2 marks]

Answer _____

- (f) Give an interpretation of the value you calculated in part (e) in relation to the hypothesis you stated in part (c).
[2 marks]

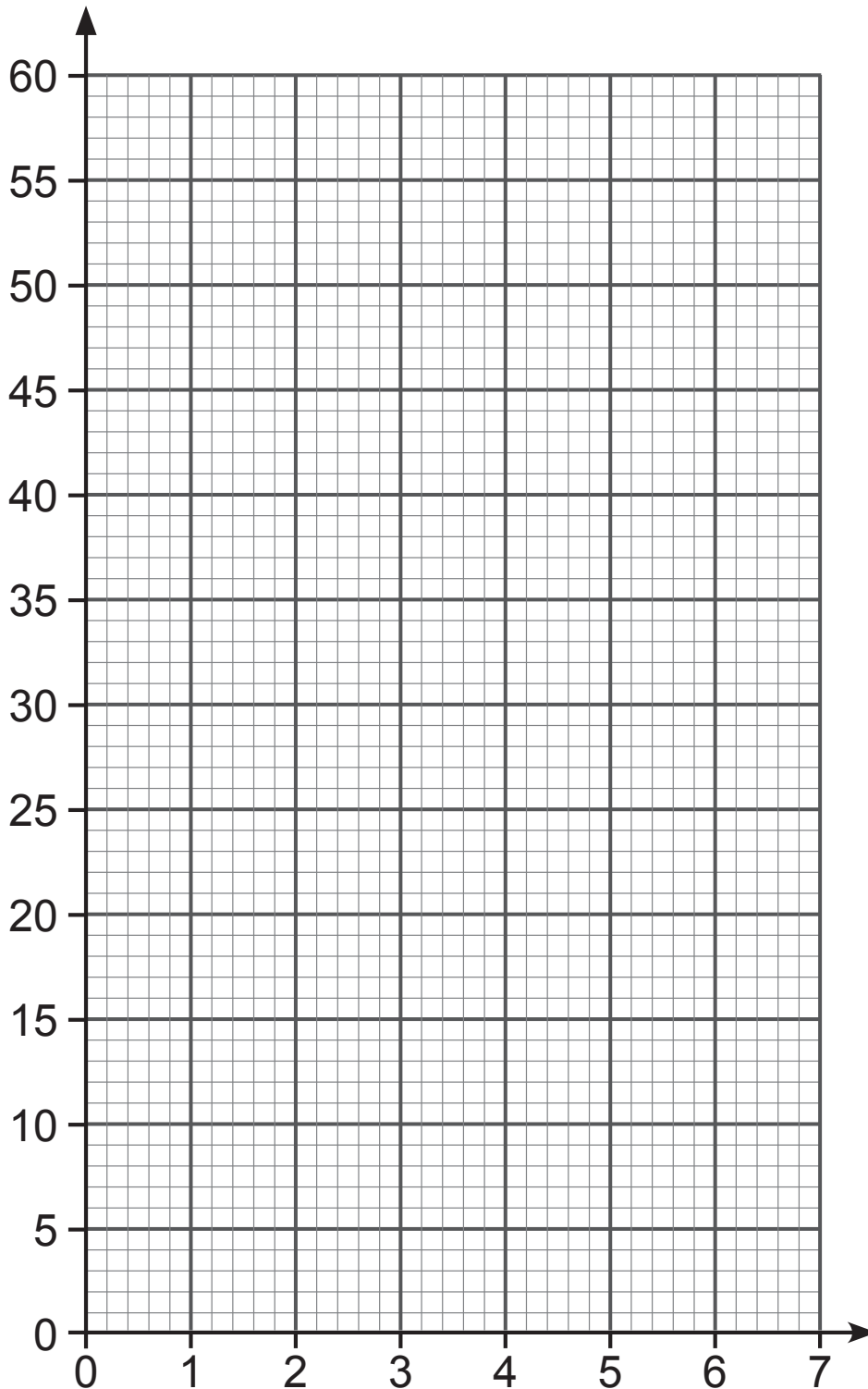
Blank Page
(Questions continue overleaf)

- 4 The table below shows the results of a survey of 60 shoppers in Omagh one afternoon.

Each person was asked how often they had visited the greengrocer's shop in the last week.

Number of visits	Number of shoppers
0	3
1	14
2	12
3	11
4	10
5	6
6	4

(a) Draw a stepped cumulative frequency diagram for the data on page 18. [3 marks]



(b) Find the median number of visits. [1 mark]

Answer _____

(c) Calculate the interquartile range for the number of visits.
[2 marks]

Answer _____

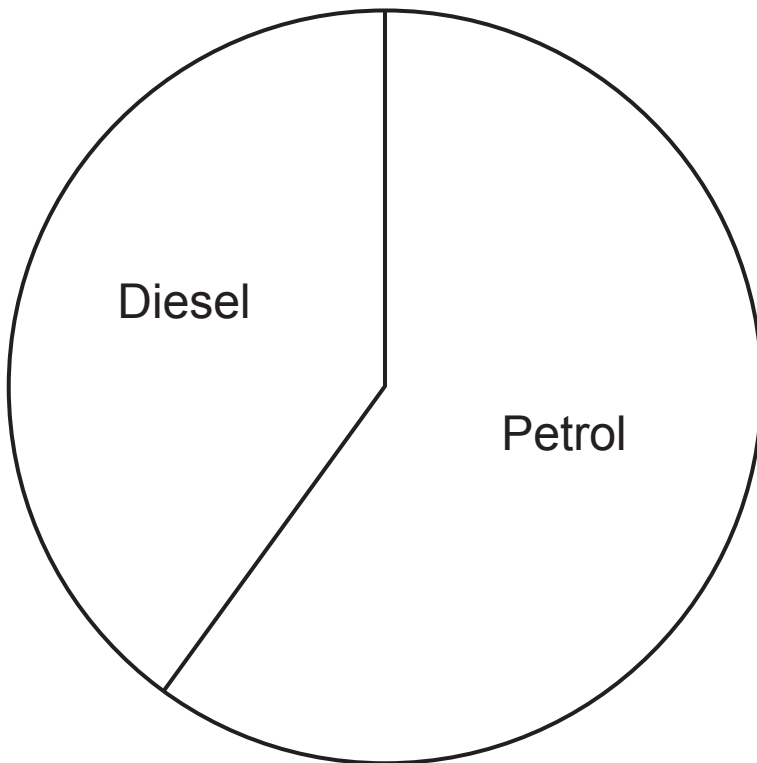
A similar survey was carried out in Strabane.

In this survey, the median number of visits was 3 and the interquartile range was 1

(d) Compare the results for Omagh and Strabane.
[4 marks]

5 A mechanic notes the fuel type of each car he works on.

The pie chart below shows his results for 2011



62 of the cars the mechanic worked on in 2011 used diesel.

(a) Calculate the total number of cars the mechanic worked on in 2011 [3 marks]

Answer _____

The table below shows the mechanic's results for 2021

Petrol	Hybrid	Diesel
110	29	66

(b) In the space below, draw a comparative pie chart to show the data in the table. [5 marks]

(c) Complete the statement below. [1 mark]

Fuel type is a _____ variable.

6 The manager of a fruit and vegetable shop kept a record of the type of fruit 50 customers bought one Friday morning.

27 customers bought oranges

20 customers bought grapes

24 customers bought apples

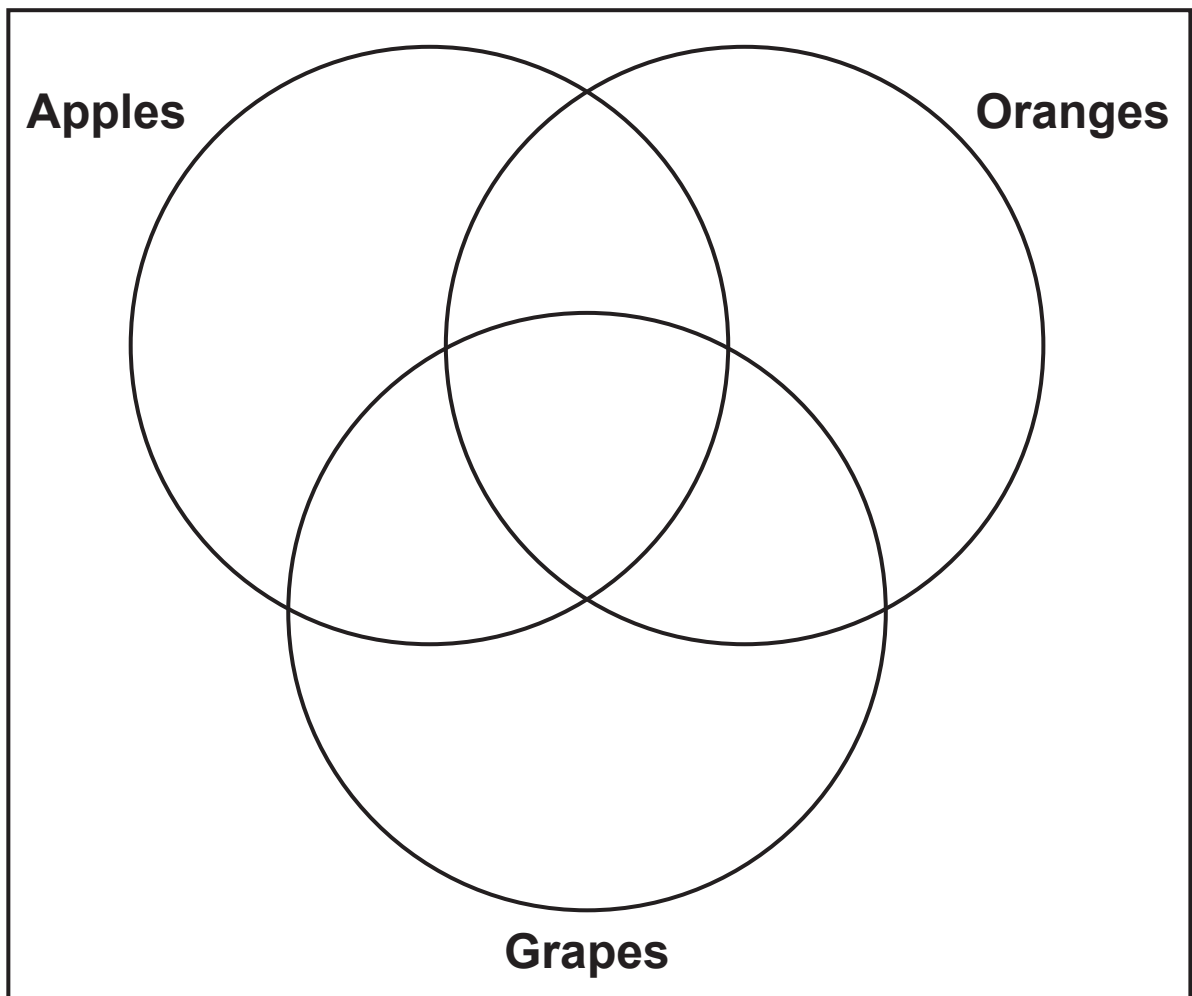
14 customers bought apples and oranges

12 customers bought oranges and grapes

11 customers bought apples and grapes

8 customers bought apples, oranges and grapes

(a) Complete the Venn diagram below to show this information. [3 marks]



(b) Use the information in the Venn diagram to find the probability that a customer, selected at random, bought:

(i) exactly two types of fruit; [2 marks]

Answer _____

(ii) apples if they bought grapes. [2 marks]

Answer _____

The following Friday morning, 75 customers come to the shop.

(c) (i) How many of these customers would you expect to buy oranges? [2 marks]

Answer _____

(ii) Comment on the reliability of your answer to part **(c)(i)**. [2 marks]

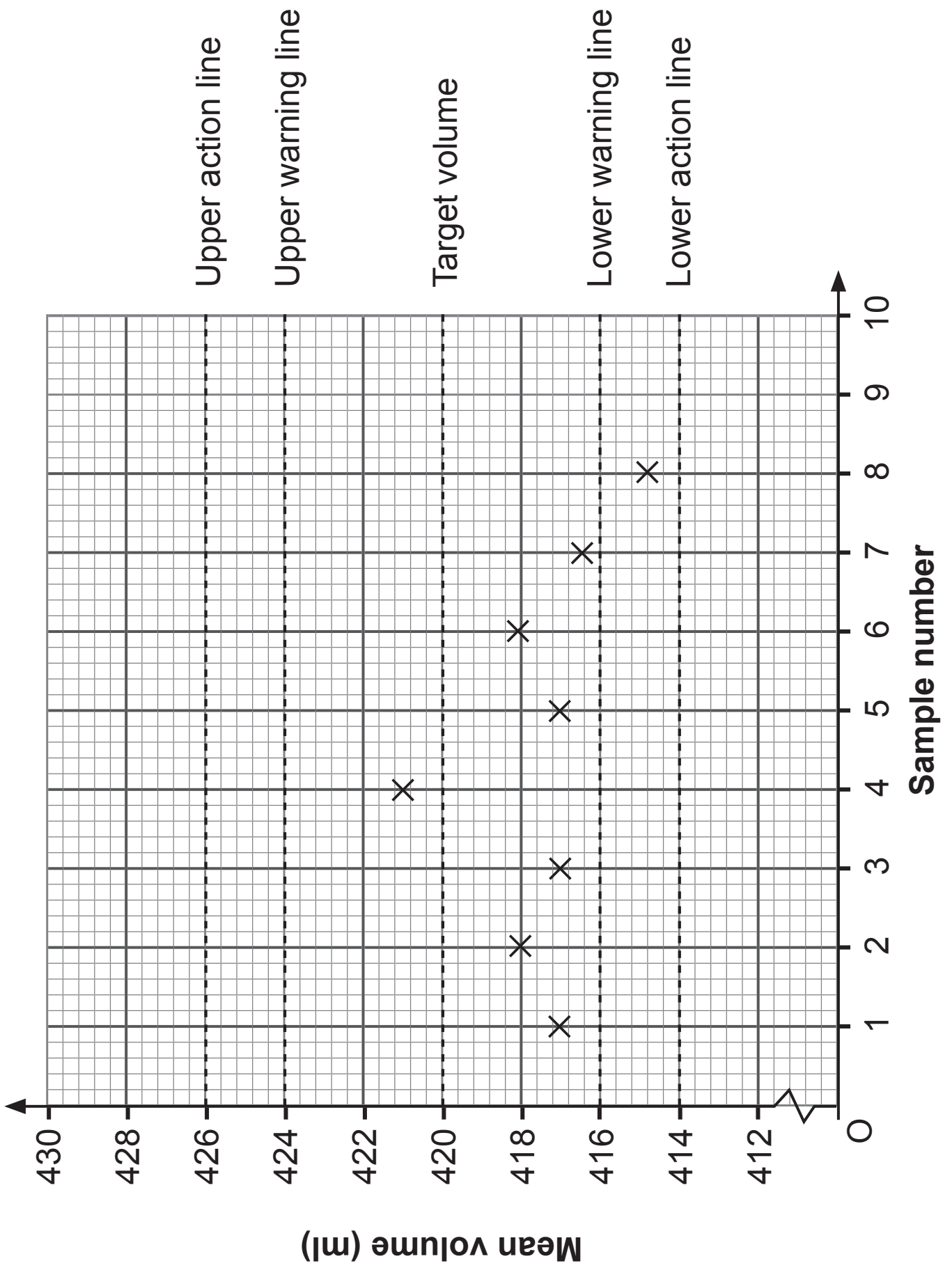
7 A machine is designed to fill containers with soup.

The label on the container states that it holds 420 ml of soup.

For quality control, a sample of containers of soup is taken at regular intervals and the mean volume of soup is calculated.

The control chart opposite shows the mean volume of soup in the first eight samples during one day's production.

How might the value of the eighth sample mean be interpreted and what action, if any, needs to be taken?
[5 marks]



- 8 John and Jack each took part in some challenges at a Fun Day.

Each boy ranked the challenges in order of difficulty from 1 (easiest) to 8 (hardest).

Their results are given in the table below.

	A	B	C	D	E	F	G	H
John	2	1	4	8	5	3	7	6
Jack	5	2	6	1	7	4	3	8

- (a) Calculate Spearman's rank correlation coefficient for the data in the table.
Give your answer correct to three decimal places.
[4 marks]

Answer _____

(b) Give an interpretation of your answer to part **(a)**.
[2 marks]

(c) Explain why the use of Spearman's rank correlation coefficient is appropriate for the type of bivariate data in the table. [1 mark]

For one of the challenges at the Fun Day, participants had to climb to the top of a rope as quickly as possible.

Anita thinks that taller people perform better at this challenge.

She records the height, x (cm), and time, y (seconds), for the first 10 participants and draws a scatter diagram of her results.

The equation of the line of best fit on Anita's scatter diagram is

$$y = 63 - 0.192x$$

(d) Niall is 152 cm tall.

Use the equation of the line of best fit to estimate how long it might take for Niall to climb to the top of the rope.
[2 marks]

Answer _____ seconds

(e) Comment on the reliability of Anita's equation.
[3 marks]

- 9** Members of a large fitness club can use different methods to pay their membership fees.

Over time, it is estimated that 35% of members pay their fees using cash.

During one afternoon, six members pay their membership fees.

- (a)** Explain why the binomial distribution can be used to model the number of these six members who will pay their membership fees using cash. [3 marks]

- (b) Calculate the probability that exactly three of these six members will pay their membership fees using cash.
[4 marks]

You may use

$$(p + q)^6 = p^6 + 6p^5q + 15p^4q^2 + 20p^3q^3 + 15p^2q^4 + 6pq^5 + q^6$$

Answer _____

(c) Calculate the probability that most of these six members will pay their membership fees using cash. [3 marks]

You may use

$$(p + q)^6 = p^6 + 6p^5q + 15p^4q^2 + 20p^3q^3 + 15p^2q^4 + 6pq^5 + q^6$$

Answer _____

Blank Page
(Questions continue overleaf)

10 A newspaper editor is concerned that the median number of spelling errors has increased from 5.61 errors per page.

She selects a recent newspaper and counts the number of spelling errors on each page.

Her results are shown in the table below.

Number of errors	Number of pages
0–2	2
3–5	15
6–8	23
9–11	5
12 or more	0

Use the data in the table on page 36 to decide if the editor's concern is justified. [6 marks]

This is the end of the question paper

For Examiner's use only	
Question Number	Marks
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
Total Marks	

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.