



General Certificate of Secondary Education
2023

Centre Number

--	--	--	--	--

Candidate Number

--	--	--	--	--

Statistics

Unit 1 (With Calculator)

Higher Tier



[GST12]

GST12

MONDAY 12 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.**

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.



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)

13827.05 R

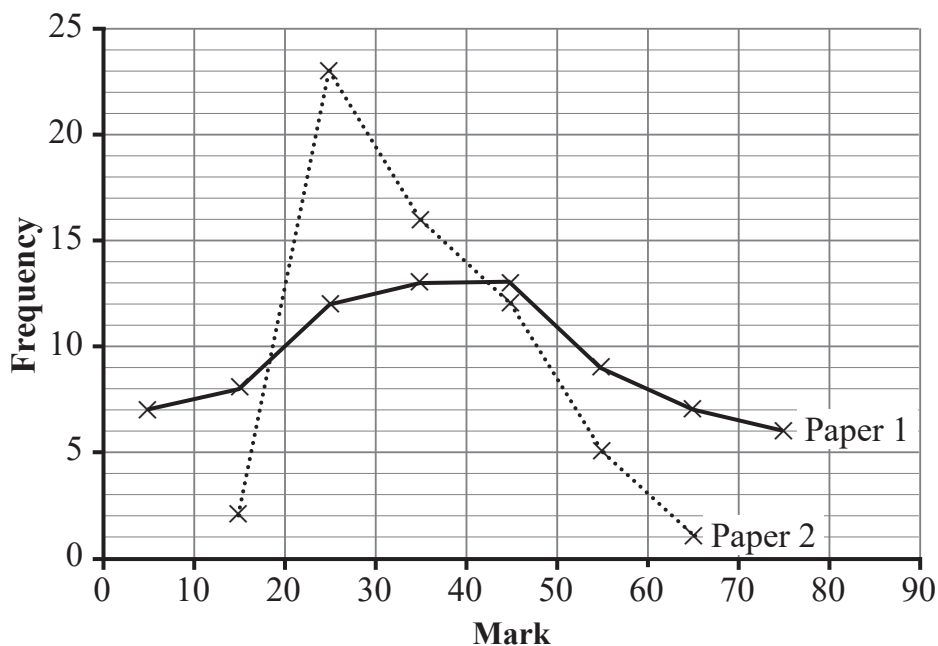
[Turn over



24GST1203

Answer **all** questions

- 1 The frequency polygons below show the distributions of marks in two examination papers, Paper 1 and Paper 2



- (a) In which paper did the marks show less variation?

Tick the correct box.

Paper 1

Paper 2

[1]

- (b) Write down the modal class interval for Paper 2

Answer _____ [1]

- (c) (i) Find the number of pupils who sat Paper 1

Answer _____ [2]



(ii) Hence write down the class interval which contains the median mark for Paper 1

Answer _____ [2]



- 2 Noeleen is investigating how long primary school children in Northern Ireland spend playing games online each week.

She intends to conduct a survey by asking some children to complete a questionnaire.

- (a) What type of data is the time spent playing games online?

Circle two words from the list below.

Quantitative
Control

Discrete
Bivariate

Qualitative
Continuous

[2]

- (b) Is Noeleen using primary data or secondary data?

_____ [1]

- (c) Describe how Noeleen could use cluster sampling in her investigation, and outline one problem she might encounter when using this method.

_____ [4]

Noeleen collects her data. She summarises it and finds that, on average, primary school pupils spend 3.4 hours per week playing games online.

- (d) Suggest one way in which Noeleen could improve her investigation using the same sample of pupils.

_____ [1]



3 ChipCorp is a company which produces computers.

During 2021, ChipCorp produced 8000 computers.

Of these, 60 computers were found to be faulty and had to be replaced.

(a) Work out the risk of ChipCorp producing a faulty computer.

Express your answer as a percentage.

Answer _____ % [2]

The cost to the company of each faulty computer was £140

(b) Calculate the cost of this risk to the company during 2021

Answer £ _____ [2]

ChipCorp plans to increase production during 2022 to 12 800 computers.

(c) Work out the expected number of faulty computers which will be produced in 2022

Answer _____ [2]

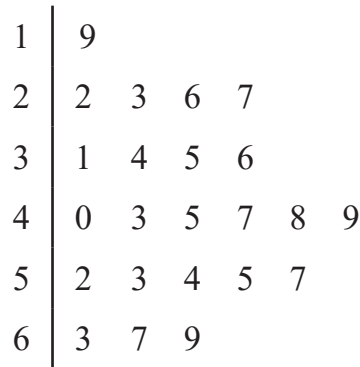
(d) Explain why the actual number of faulty computers produced in 2022 may differ from your answer to part (c).

[1]

[Turn over



- 4 The ages, in years, of visitors entering a museum during a one-hour period are shown in the stem and leaf diagram below.



Key:

3 | 4 means 34

- (a) Use the information in the stem and leaf diagram to complete the table below.

Minimum	Lower quartile	Median	Upper quartile	Maximum
	31			69

[3]

- (b) Work out the interquartile range for the ages.

Answer _____ [2]

- (c) Explain the meaning of your answer to part (b).

_____ [1]

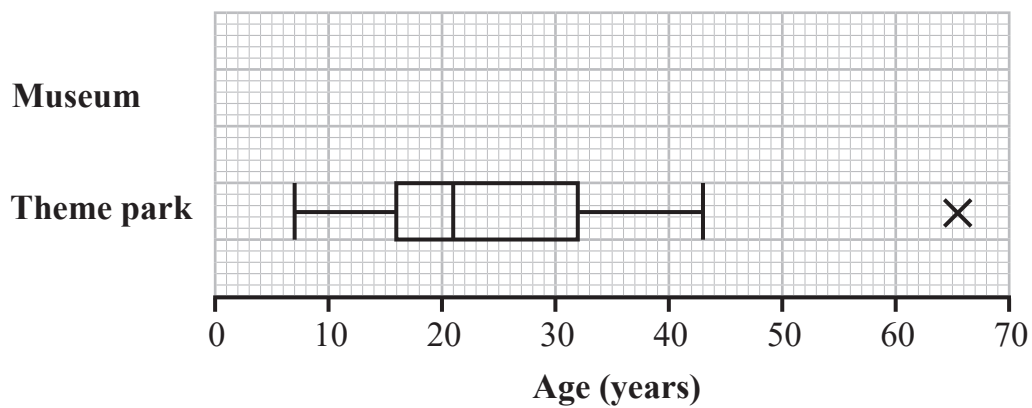


- (d) Explain why the mode would not be a suitable average to represent the data shown in the stem and leaf diagram.

[2]

The ages, in years, of visitors entering a theme park during the same one-hour period were also recorded.

The results are shown in the box plot below.



- (e) Explain the meaning of the value marked \times in the diagram above.

[2]

- (f) On the same diagram, draw a box plot for the ages of the visitors to the museum.

[2]

[Turn over



(g) Use the information in the box plots to make two comparisons between the ages of the visitors to the museum and the ages of the visitors to the theme park.

1. _____

_____ [2]

2. _____

_____ [2]





BLANK PAGE

DO NOT WRITE ON THIS PAGE

(Questions continue overleaf)

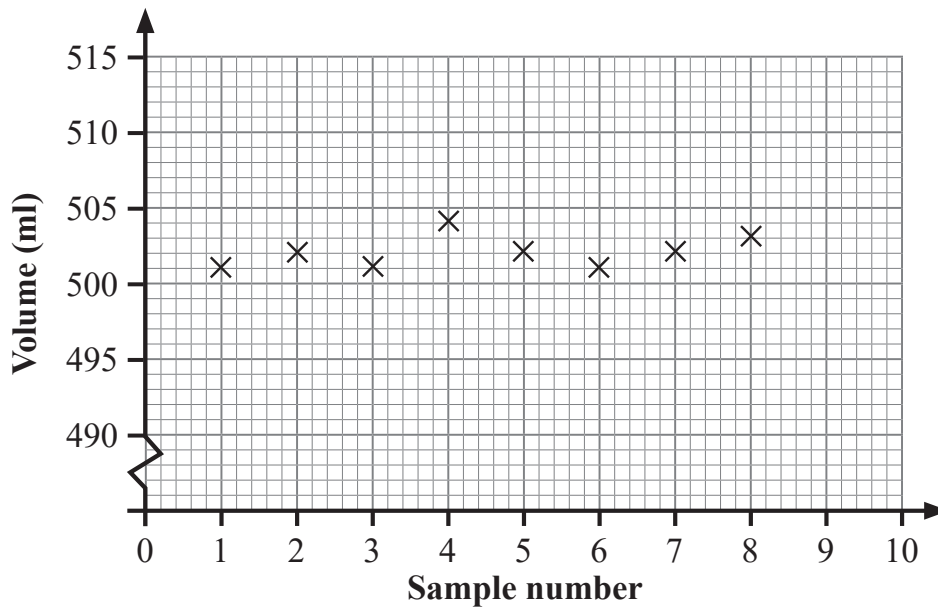


- 5 In a production process, a machine is programmed to dispense 500 ml of mouthwash into each bottle.

For quality assurance, warning and action limits are set at 5 ml and 7 ml.

Samples of bottles of mouthwash are taken at regular intervals and the mean volume of each sample is calculated.

The means of a series of eight samples are shown on the control chart below.



- (a) On the control chart, draw and label the warning and action lines. [4]

For the ninth sample, five bottles are taken and the volume of mouthwash in each bottle is as follows:

501 ml 504 ml 503 ml 501 ml 501 ml

- (b) Calculate the value of the ninth sample mean and plot it on the control chart. [3]



(c) Comment on what the control chart shows.

[4]

[Turn over



- 6 For GCSE Music, candidates are assessed in three components: Performing, Composing and Listening.

The contribution to the full course by each component is as follows:

Performing	Composing	Listening
35%	30%	35%

In her mock assessments, Geri scores 71% in Performing, 88% in Composing and 65% in Listening.

To get a grade A, candidates need to score 73% overall.

Michael says that Geri has got a grade A because $\frac{71 + 88 + 65}{3}$ is more than the 73% required.

- (a) Explain why Michael's calculation is not correct.

[1]

- (b) Use an appropriate calculation to decide if Geri has got a grade A.

[5]



Joe has completed the Performing and Composing assessments of his mock and has scored 65% and 80%, respectively.

- (c) Calculate the minimum score required by Joe in the Listening assessment in order for him to get a grade A.

Answer _____% [4]



- 7 The grouped frequency table below shows the duration, to the nearest second, for a group of people to find a hidden word in a word search.

Duration (nearest second)	Frequency		
$0 < t \leq 10$	3		
$10 < t \leq 20$	14		
$20 < t \leq 30$	22		
$30 < t \leq 40$	9		
$40 < t \leq 50$	2		

Jane said that someone took 50 seconds to find the hidden word.

- (a) Explain why Jane may not be correct.

[2]

- (b) Calculate an estimate of the mean time taken to find the hidden word.

You may use the blank columns to help with your working.

Answer _____ seconds [4]



(c) Using the statistical functions on your calculator, or otherwise, calculate an estimate of the standard deviation of the times taken.

Answer _____ seconds [2]

Some time later, all members of the same group did another word search.

On this occasion, the mean time taken to find a hidden word was 17.9 seconds and the standard deviation was 7.28 seconds.

(d) Compare the results for the two word searches.

[4]

[Turn over



8 Andrew wishes to estimate the number of seals in a body of water.

He plans to use the capture/recapture method to do this.

(a) What three assumptions must Andrew make to use the capture/recapture method?

[3]



One day, Andrew catches 90 seals, marks their skin with a harmless dye and releases them back into the water.

The following day, he catches 40 seals and finds that 7 of them have been marked with dye.

(b) Use this information to estimate the number of seals in the body of water.

Answer _____ [4]



9 At a village fair, ten jams, A–J, are given a score between 0 and 50 by two judges.

The judges' scores are given in the table below.

	A	B	C	D	E	F	G	H	I	J
Judge 1	27	45	29	34	31	29	35	38	44	30
Judge 2	24	23	29	36	40	17	41	38	15	18

(a) Calculate Spearman's rank correlation coefficient for the scores in the table.

Answer _____ [5]

(b) Give an interpretation, in context, of your answer to part (a).

[2]



Jason suggests calculating the equation of a regression line to model the relationship between the judges' scores.

(c) Explain why Jason's suggestion is not appropriate.

[2]



10 Roisin is a keen gardener and has plants A and B in her garden.

The lengths of the leaves on plant A are normally distributed with mean 9.2 cm and standard deviation 1.4 cm.

- (a) Within what range of lengths would you expect about two-thirds of the leaves on plant A to be?

Answer _____ cm and _____ cm [3]

- (b) Find the probability that a leaf, selected at random from plant A, is between 9.2 cm and 12 cm long.

Answer _____ [3]

The lengths of the leaves on plant B are normally distributed with mean 10.6 cm and standard deviation 2.1 cm.

The standardised score for a leaf, selected at random from plant B, is -0.7

- (c) Calculate the actual length of the leaf.

Answer _____ cm [3]



Roisin takes a leaf at random off the ground, but is unsure which plant it has come from.

The leaf measures 10 cm.

- (d) Using standardised scores, determine which plant this leaf is more likely to have come from, giving a reason for your choice.

Plant _____

Reason _____

_____ [5]

THIS IS THE END OF THE QUESTION PAPER



DO NOT WRITE ON THIS PAGE

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

Total Marks	
--------------------	--

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.

13827.05 R



24GST1224