



*Rewarding Learning*

**General Certificate of Secondary Education  
2023**

---

**Statistics**

Unit 1

Higher Tier

[GST12]

**MONDAY 12 JUNE, AFTERNOON**

---

**MARK  
SCHEME**

## General Marking Instructions

### Introduction

The mark scheme normally provides the most popular solution to each question. Other solutions given by candidates are evaluated and credit given as appropriate; these alternative methods are not usually illustrated in the published mark scheme.

The marks awarded for each question are shown in the right hand column and they are prefixed by the letters **M**, **A** and **MA** as appropriate. The key to the mark scheme is given below:

**M** indicates marks for correct method.

**A** indicates marks for accurate working, whether in calculation, readings from tables, graphs or answers.

**MA** indicates marks for combined method and accurate working.

The solution to a question gains marks for correct method and marks for an accurate working based on this method. Where the method is not correct no marks can be given.

A later part of a question may require a candidate to use an answer obtained from an earlier part of the same question. A candidate who gets the wrong answer to the earlier part and goes on to the later part is naturally unaware that the wrong data is being used and is actually undertaking the solution of a parallel problem from the point at which the error occurred. If such a candidate continues to apply correct method, then the candidate's individual working must be **followed through** from the error. If no further errors are made, then the candidate is penalised only for the initial error. Solutions containing two or more working or transcription errors are treated in the same way. This process is usually referred to as "follow-through marking" and allows a candidate to gain credit for that part of a solution which follows a working or transcription error.

It should be noted that where an error trivialises a question, or changes the nature of the skills being tested, then as a general rule, it would be the case that not more than half the marks for that question or part of that question would be awarded; in some cases the error may be such that no marks would be awarded.

### Positive marking

It is our intention to reward candidates for any demonstration of relevant knowledge, skills or understanding. For this reason we adopt a policy of **following through** their answers, that is, having penalised a candidate for an error, we mark the succeeding parts of the question using the candidate's value or answers and award marks accordingly.

Some common examples of this occur in the following cases:

- (a) a numerical error in one entry in a table of values might lead to several answers being incorrect, but these might not be essentially separate errors;
- (b) readings taken from candidates' inaccurate graphs may not agree with the answers expected but might be consistent with the graphs drawn.

When the candidate misreads a question in such a way as to make the question easier only a proportion of the marks will be available (based on the professional judgement of the examiner)

			AVAILABLE MARKS
<b>1</b>	<b>(a)</b> Paper 2	A1	6
	<b>(b)</b> 20 – 30	MA1	
	<b>(c) (i)</b> $7 + 8 + 12 + 13 + 13 + 9 + 7 + 6$ = 75	MA1 A1	
	<b>(ii)</b> $\frac{75 + 1}{2} = 38$ 30 – 40	MA1 MA1	
<b>2</b>	<b>(a)</b> Quantitative, continuous	A2	
	<b>(b)</b> Primary data.	A1	
	<b>(c)</b> Obtain a list of all primary schools in Northern Ireland. Select a primary school at random from the list. Survey all pupils in the chosen school. If the school is very small it may not be representative of all primary schools.	A1 A1 A1 A1	
	<b>(d)</b> Noeleen could separate the pupils into different groups, e.g. age or gender.	A1	8
<b>3</b>	<b>(a)</b> $\frac{60}{8000} \times 100$ = 0.75%	MA1 A1	7
	<b>(b)</b> $60 \times 140$ = £8400	MA1 A1	
	<b>(c)</b> $\frac{0.75}{100} \times 12\,800$ = 96	MA1 A1	
	<b>(d)</b> Any valid reason, e.g.		
	• changing to more reliable machinery.		
	• staff may become more/less reliable from one year to the next.	A1	

4 (a)

Minimum	Lower quartile	Median	Upper quartile	Maximum
19	31	45	54	69

AVAILABLE MARKS

MA3

(b)  $54 - 31 = 23$

MA1  
A1

(c) The middle 50% of the data has a range of 23 years.

A1

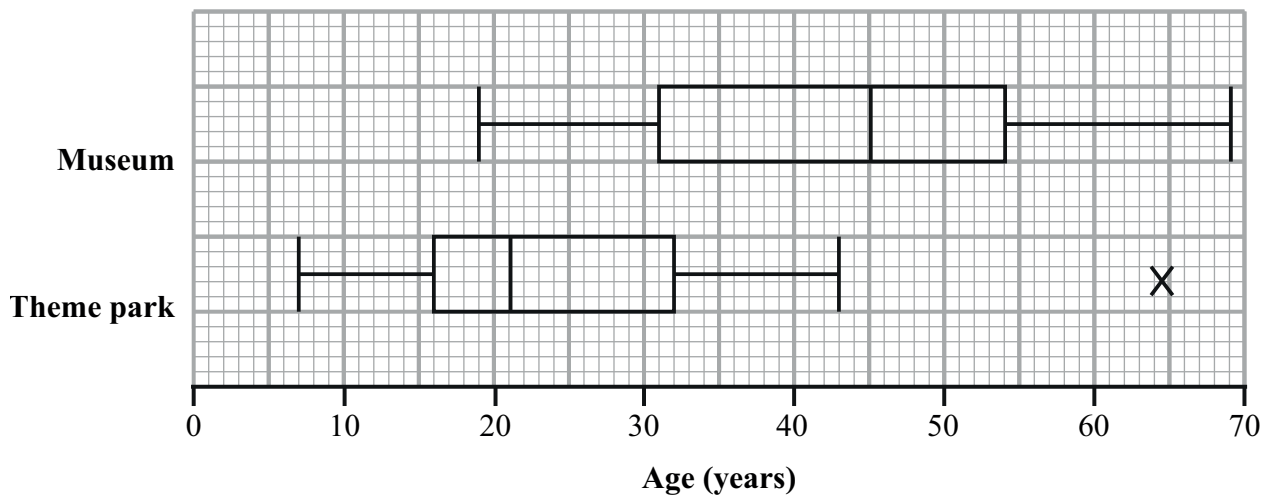
(d) There is no mode for this data as each value occurs only once.

A2

(e) It represents an outlier which is an extreme value in the dataset.

A2

(f)



MA2

(g) The museum has a higher median age so visitors to the museum are older on average.

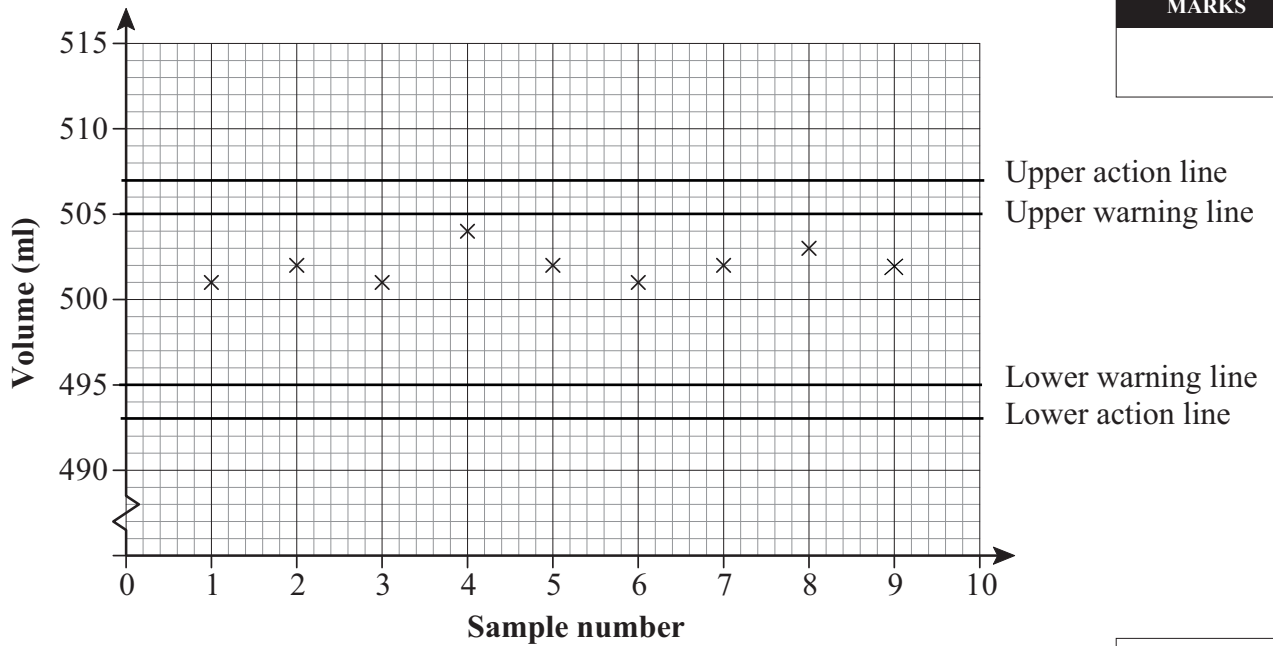
A2

The museum has a larger range/interquartile range so the ages of visitors to the museum are more varied.

A2

16

5 (a)



AVAILABLE MARKS

A4

(b) Mean =  $\frac{501 + 504 + 503 + 501 + 501}{5} = 502$  ml  
 Sample Mean plotted at (9, 502)

M1 A1

A1

(c) All nine sample means are within the target volume and the upper warning line, so the process is under control.

A2

However, the mean volume is consistently above the target volume which may suggest consistent slight over-filling.

A2

11

6 (a) Michael has assumed equal contributions by each component.

A1

(b) Weighted mean =  $\frac{35 \times 71 + 30 \times 88 + 35 \times 65}{100} = 74\%$

M1 MA1

A1

Since  $74 > 73$  Geri has got a grade A in the mock.

A2

(c)  $\frac{35 \times 65 + 30 \times 80 + 35 \times m}{100} = 73$

M1 MA1

$$4675 + 35m = 7300$$

$$35m = 2625$$

$$m = 75$$

MA1

Therefore, Joe needs 75% in the Listening assessment to get a grade A.

A1

10

7 (a) The data is in groups so there is no way of telling exactly what the values are in the stated ranges. A2

(b)

Duration	Frequency	$t$	$ft$
$0 < t \leq 10$	3	5	15
$10 < t \leq 20$	14	15	210
$20 < t \leq 30$	22	25	550
$30 < t \leq 40$	9	35	315
$40 < t \leq 50$	2	45	90
	<b>50</b>		<b>1180</b>

MA2

$$\begin{aligned} \text{Mean} &= \frac{1180}{50} \\ &= 23.6 \text{ seconds} \end{aligned}$$

M1

A1

(c) Calculator: 9.167 seconds MA2

(d) The mean decreased from 23.6 seconds to 17.9 seconds, so the participants found the hidden word more quickly in the second word search. A2

The standard deviation decreased from 9.167 seconds to 7.28 seconds so the times taken by the participants were more consistent in the second word search than in the first.

A2

12

8 (a) The population of seals must be the same throughout the process (no births or deaths). A1

The marked seals must distribute themselves evenly throughout the population. A1

The ink used to mark the seals must remain visible for the duration of both sampling procedures. A1

(b)  $\frac{90}{N} = \frac{7}{40}$  M1

$7N = 3600$  MA1

$N = 514.28\dots$  A1

The estimated number of seals is 514 A1

7

	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
Rank 1	10	1	8.5	5	6	8.5	4	3	2	7
Rank 2	6	7	5	4	2	9	1	3	10	8
$d$	4	-6	3.5	1	4	-0.5	3	0	-8	-1
$d^2$	16	36	12.25	1	16	0.25	9	0	64	1

A2

$$\Sigma d^2 = 155.5$$

MA1

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

$$= 1 - \left( \frac{6 \times 155.5}{10(10^2 - 1)} \right)$$

MA1

$$= 0.0576$$

A1

- (b) The value of  $r_s$  is close to zero so there is no agreement between the judges' scores.

A2

- (c) Since  $r_s$  is close to zero, the points do not lie close to a straight line so a line of best fit would not be appropriate for this data.

A2

9

- 10 (a)  $9.2 \pm 1.4$   
7.8 cm to 10.6 cm

MA2

A1

- (b)  $9.2 + 2 \times 1.4 = 12$  cm  
 $95\% \div 2 = 0.475$

MA2

MA1

- (c)  $-0.7 = \frac{x - 10.6}{2.1}$   
 $x = 9.13$  cm

M1 MA1

A1

- (d)  $z_A = \frac{10 - 9.2}{1.4} = \frac{4}{7} = 0.571$

M1 MA1

$$z_B = \frac{10 - 10.6}{2.1} = -\frac{2}{7} = -0.286$$

MA1

The leaf is more likely to have come from plant B, since  $z_B$  is closer to 0 than  $z_A$ .

A2

14

**Total****100**