



Rewarding Learning

General Certificate of Secondary Education

2022

Statistics

Unit 2

Higher Tier

[GST22]

FRIDAY 24 JUNE, MORNING

**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:

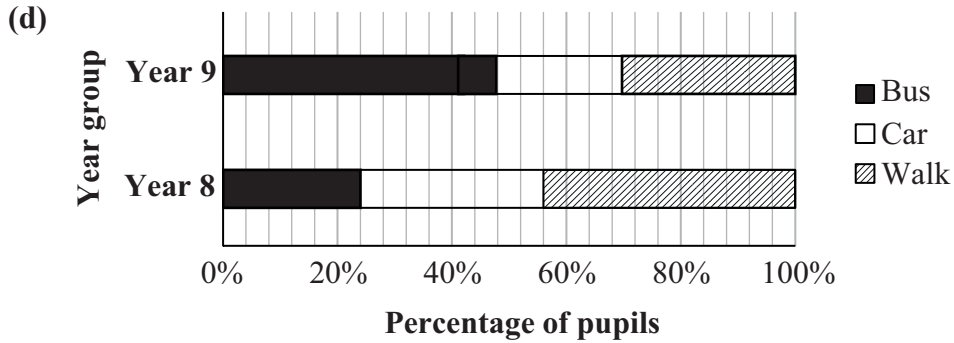
- (a) a numerical error in one entry in a table of values might lead to several answers being incorrect, but these might not be essentially separate errors;
- (b) readings taken from candidates' inaccurate graphs may not agree with the answers expected but might be consistent with the graphs drawn.

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)

COVID-19 Context

Given the unprecedented circumstances presented by the COVID-19 public health crisis, senior examiners, under the instruction of CCEA awarding organisation, are required to train assistant examiners to apply the mark scheme in case of disrupted learning and lost teaching time. The interpretation and intended application of the mark scheme for this examination series will be communicated through the standardising meeting by the Chief or Principal Examiner and will be monitored through the supervision period. This paragraph will apply to examination series in 2021–2022 only.

- 1 (a) Primary data A1
- (b) $30 - 22 = 8$ MA1
- (c) The numbers of pupils in the two year groups are different which makes comparisons using a bar chart difficult. A2



MA4

- (e) A greater proportion of Year 8 pupils travel by car than Year 9 pupils.
A greater proportion of Year 8 pupils walk to school than Year 9 pupils.
A greater proportion of Year 9 pupils travel by bus than Year 8 pupils. A2
- (f) Car A1

- (g) (i) 6 7 9 10 12 12 12 13 13 14 15 15 17 23 25 53 MA1
Median = 13 minutes A1

(ii)

Mean = $\frac{6 + 7 + 9 + 10 + 12 + 12 + 12 + 13 + 13 + 14 + 15 + 15 + 17 + 23 + 25 + 53}{16}$ MA1

= 16 minutes A1

- (iii) The median is more appropriate in this case since the mean is affected by one outlying value (53) whereas the median is not. A2

AVAILABLE MARKS

17

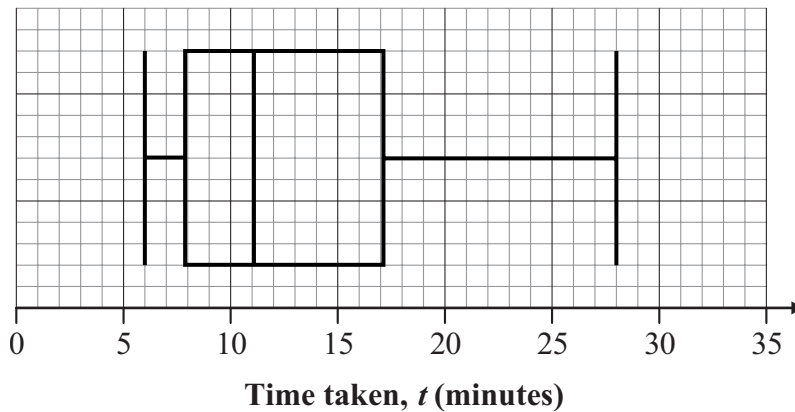
2 (a) 80

A1

(b) $5 \leq t < 10$

A1

(c)



M1
MA3

(d) Positively skewed.

A1

(e) No.
Nomally distributed data is not skewed.

A1

A1

(f) 49

A1

10

3 (a) April 2020

A1

(b) (i) $1513 - 1402 = 111$

MA1

(ii) Some part-time jobs may have been advertised and filled or may not have been advertised during the month so the vacancy did not get counted.

A2

(c) The months with more full-time vacancies have more part-time vacancies. A1

(d) Product moment correlation coefficient.

A1

(e) 0.925

A2

(f) There is a strong positive correlation between the number of full-time vacancies and part-time vacancies each month so months with more full-time vacancies have more part-time vacancies.

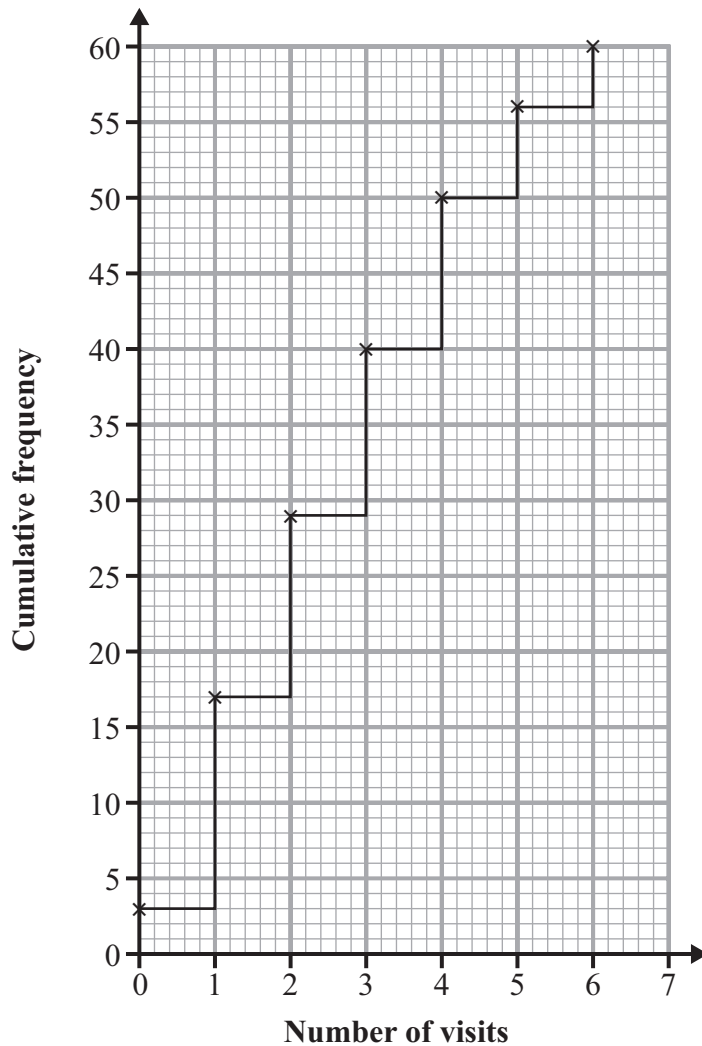
A2

10

4 (a)

Number of visits	Number of shoppers	Cumulative frequency
0	3	3
1	14	17
2	12	29
3	11	40
4	10	50
5	6	56
6	4	60

MA1



MA1
A1

(b) 3 visits

A1

(c) $4 - 1 = 3$ visits

MA1 A1

(d) The median number of visits for both towns was the same so, on average, people in both towns visit the greengrocer's shop the same number of times.

A2

The interquartile range for the number of visits in Strabane is smaller than that for Omagh so the number of visits to the greengrocer's by shoppers in Strabane is more consistent.

A2

AVAILABLE
MARKS

10

5 (a) Angle = 144°
 Number of cars = $\frac{360}{144} \times 62$
 = 155 cars

A1
 MA1
 A1

(b) For comparative pie charts:

$$\frac{R^2}{r^2} = \frac{N}{n}$$

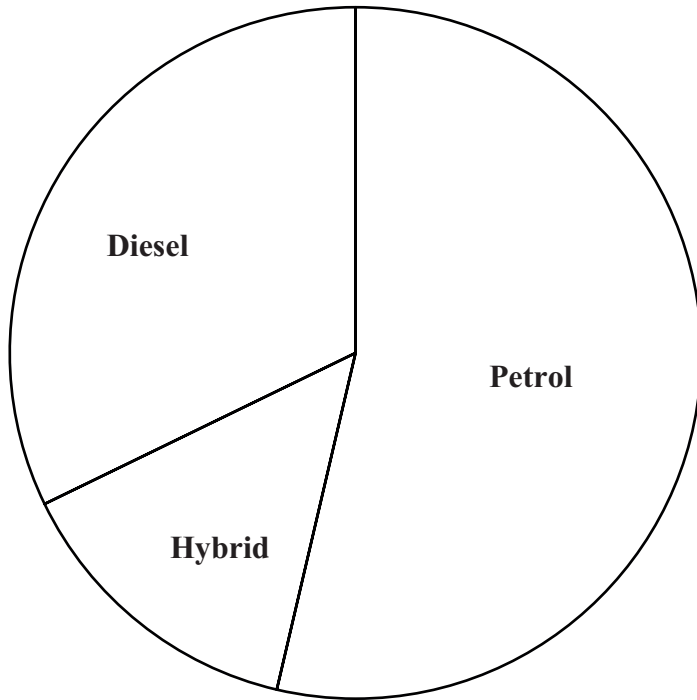
$$R^2 = 4^2 \times \frac{205}{155}$$

$$R = 4.6 \text{ cm}$$

MA1
 A1

Angles: $193^\circ, 51^\circ, 116^\circ$

MA1



A2

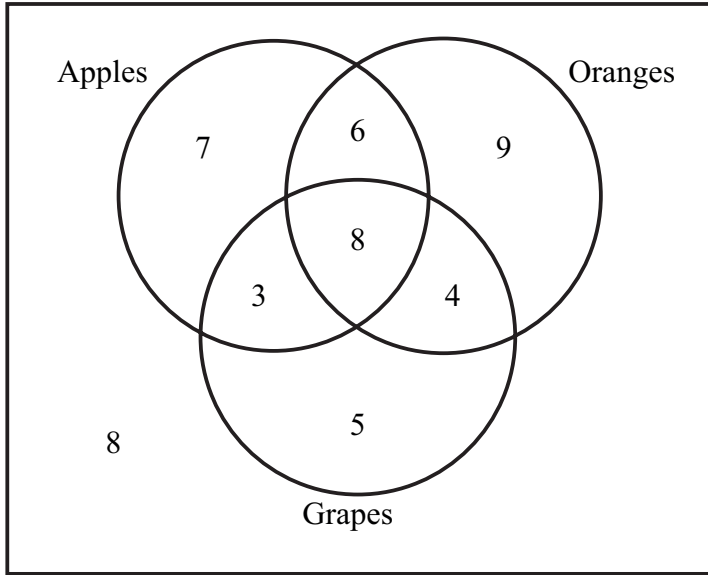
(c) Categorical or qualitative

A1

AVAILABLE
 MARKS

9

6 (a)



MA3

(b) (i) $P(\text{exactly two types of fruit}) = \frac{6+3+4}{50} = \frac{13}{50}$

MA2

(ii) $P(A|G) = \frac{8+3}{8+3+4+5}$
 $= \frac{11}{20}$

MA2

(c) (i) $75 \times \frac{27}{50} = 40.5$
 41 (or 40)

MA1

A1

(ii) The value may not be reliable as the purchases made by the customers on the first Friday may not be representative of the purchases made every Friday.

A2

11

7 The eighth sample mean lies between the lower warning line and the lower action line which indicates that the process may be out of control. The machine operator should take another sample straight away and calculate the mean of this sample. If this sample mean is within the warning lines, the process can continue. If it is outside the warning lines again, the process needs to be stopped as it is out of control.

A2

A3

5

AVAILABLE MARKS

	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
d	-3	-1	-2	7	-2	-1	4	-2
d^2	9	1	4	49	4	1	16	4

$$\Sigma d^2 = 88$$

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

$$= 1 - \left(\frac{6 \times 88}{8(8^2 - 1)} \right)$$

$$= -0.048$$

(b) Since r_s is close to zero there is unlikely to be any correlation between the boys' rankings of the challenges.

(c) The data is ordinal.

$$\begin{aligned} \text{(d)} \quad y &= 63 - 0.192x \\ &= 63 - 0.192(152) \\ &= 33.8 \text{ seconds} \end{aligned}$$

(e) The equation may not be reliable because the equation is based on data from just 10 participants, so the sample is quite small; there is no information on whether or not this sample is representative; the scatter diagram is not shown so it is unclear whether there is sufficiently strong correlation to justify a line of best fit at all.

9 (a) The binomial distribution is an appropriate model because there are only two possible outcomes (members pay cash or they do not), the probability that a member pays by cash is fixed at 35% for all members and the method of payment is independent between members.

(b) Let M be the number of members who pay by cash.

$$\begin{aligned} P(M = 3) &= 20p^3q^3 \\ &= 20(0.35)^3(0.65)^3 \\ &= 0.235 \end{aligned}$$

$$\begin{aligned} \text{(c)} \quad P(M \geq 4) &= p^6 + 6p^5q + 15p^4q^2 \\ &= (0.35)^6 + 6(0.35)^5(0.65) + 15(0.35)^4(0.65)^2 \\ &= 0.117 \end{aligned}$$

AVAILABLE
MARKS

MA1

MA1

MA1

A1

A2

A1

MA1

A1

A3

12

A3

M1 MA2

A1

M1 MA1

A1

10

10

Number of errors	Number of pages	Cumulative frequency
0–2	2	2
3–5	15	17
6–8	23	40
9–11	5	45
12 or more	0	45

$$\begin{aligned} \text{Median} &= 5.5 + \left(\frac{23-17}{23}\right)3 \\ &= 6.28 \end{aligned}$$

The median has increased from 5.61 errors to 6.28 errors so the editor’s concern is justified.

MA1

MA2

A1

A2

Total

AVAILABLE
MARKS

6

100