



Cambridge O Level

STATISTICS

4040/23

Paper 2

October/November 2022

MARK SCHEME

Maximum Mark: 100

Published

This mark scheme is published as an aid to teachers and candidates, to indicate the requirements of the examination. It shows the basis on which Examiners were instructed to award marks. It does not indicate the details of the discussions that took place at an Examiners' meeting before marking began, which would have considered the acceptability of alternative answers.

Mark schemes should be read in conjunction with the question paper and the Principal Examiner Report for Teachers.

Cambridge International will not enter into discussions about these mark schemes.

Cambridge International is publishing the mark schemes for the October/November 2022 series for most Cambridge IGCSE™, Cambridge International A and AS Level components and some Cambridge O Level components.

This document consists of **10** printed pages.

Generic Marking Principles

These general marking principles must be applied by all examiners when marking candidate answers. They should be applied alongside the specific content of the mark scheme or generic level descriptors for a question. Each question paper and mark scheme will also comply with these marking principles.

GENERIC MARKING PRINCIPLE 1:

Marks must be awarded in line with:

- the specific content of the mark scheme or the generic level descriptors for the question
- the specific skills defined in the mark scheme or in the generic level descriptors for the question
- the standard of response required by a candidate as exemplified by the standardisation scripts.

GENERIC MARKING PRINCIPLE 2:

Marks awarded are always **whole marks** (not half marks, or other fractions).

GENERIC MARKING PRINCIPLE 3:

Marks must be awarded **positively**:

- marks are awarded for correct/valid answers, as defined in the mark scheme. However, credit is given for valid answers which go beyond the scope of the syllabus and mark scheme, referring to your Team Leader as appropriate
- marks are awarded when candidates clearly demonstrate what they know and can do
- marks are not deducted for errors
- marks are not deducted for omissions
- answers should only be judged on the quality of spelling, punctuation and grammar when these features are specifically assessed by the question as indicated by the mark scheme. The meaning, however, should be unambiguous.

GENERIC MARKING PRINCIPLE 4:

Rules must be applied consistently, e.g. in situations where candidates have not followed instructions or in the application of generic level descriptors.

GENERIC MARKING PRINCIPLE 5:

Marks should be awarded using the full range of marks defined in the mark scheme for the question (however; the use of the full mark range may be limited according to the quality of the candidate responses seen).

GENERIC MARKING PRINCIPLE 6:

Marks awarded are based solely on the requirements as defined in the mark scheme. Marks should not be awarded with grade thresholds or grade descriptors in mind.

MARK SCHEME NOTES

The following notes are intended to aid interpretation of mark schemes in general, but individual mark schemes may include marks awarded for specific reasons outside the scope of these notes.

Types of mark

- M** Method marks, awarded for a valid method applied to the problem.
- A** Accuracy mark, awarded for a correct answer or intermediate step correctly obtained. For accuracy marks to be given, the associated Method mark must be earned or implied.
- B** Mark for a correct result or statement independent of Method marks.

When a part of a question has two or more 'method' steps, the M marks are in principle independent unless the scheme specifically says otherwise; and similarly where there are several B marks allocated. The notation 'dep' is used to indicate that a particular M or B mark is dependent on an earlier, asterisked, mark in the scheme.

The symbol \surd implies that the A or B mark indicated is allowed for work correctly following on from previously incorrect results. Otherwise, A and B marks are given for correct work only.

Abbreviations

AG	answer given on question paper
awrt	answer which rounds to
cao	correct answer only
dep	dependent
ft	follow through after error
oe	or equivalent
SC	special case
soi	seen or implied
www	without wrong working

Question	Answer	Marks	Partial Marks
1(a)	Qualitative	3	B1
	Quantitative continuous		B1B1
1(b)	7 and 15 seen or 15 – '7' where $5 < '7' < 10$	2	M1
	8		A1
1(c)	$25\% \times 144$ oe	2	M1
	36		A1

Question	Answer	Marks	Partial Marks
2(a)	Sectional or composite bar chart	2	B1
	It will show the totals [for each sport]/which is the most popular [sport] oe		B1
2(b)	$6/25$ oe Correct numerator in a probability	2	B1
	Correct denominator in a probability		B1
2(c)	$3/11$ oe	1	B1
2(d)	$1/2$ oe	1	B1

Question	Answer	Marks	Partial Marks
3(a)	Sample has 2 of each battery size oe	3	M1
	[To be representative sample should have] $100/600 \times 6$ [of D], $300/600 \times 6$ [of AA], $200/600 \times 6$ [of AAA] oe		M1
	Should have 1 [of D], 3 [of AA] and 2 [of AAA] so not representative		A1
3(b)	$1/6 \times 5$, $3/6 \times 5$ and $2/6 \times 5$ oe	2	M1
	Leading to 1, 2, 2		A1

Question	Answer	Marks	Partial Marks
4(a)	Use of $P(A \cap B) = P(A) \times P(B)$	3	M1
	Use of $P(A \cup B) = P(A) + P(B) - P(A \cap B)$		M1
	0.58 oe		A1
4(b)	Use of $P(A \cup B) - P(A \cap B)$ oe	2	M1
	0.46 oe		A1
4(c)	$(1 - 0.3) \times (1 - 0.4)$ or $1 - \textit{their (a)}$ or $1 - 0.3 - 0.4 + 0.3 \times 0.4$	2	M1
	0.42		A1

Question	Answer	Marks	Partial Marks
5(a)	250×12 or 3000 or 0.8×1200 or 960	2	M1
	$3000 : 960 : 360$ or e.g. $250 \times 12 : 0.8 \times 1200 : 360$ leading to $25 : 8 : 3$ AG		A1
5(b)	109 seen [for rent]	5	B1
	100 seen [for electricity]		B1
	98 seen [for other costs]		B1
	$(25 \times '109' + 8 \times '100' + 3 \times '98') / (25 + 8 + 3) [= 3819/36]$ oe		M1
	106.1		A1
5(c)	For example, 'the amount of electricity/ number of units used may have changed/increased/decreased.'	1	B1

Question	Answer	Marks	Partial Marks
6(a)	$600 + (50 - 28)/(90 - 28) \times (1000 - 600)$ 600 + a fraction of the class width (1000 – 600) or 400	4	M1
	Division of the class width by (90 – 28) or 62		M1
	Multiplication of the class width by (50 – 28) or 22		M1
	Or $1000 - (90 - 50)/(90 - 28) \times (1000 - 600)$ 1000 – a fraction of the class width (1000 – 600) or 400		(M1)
	Division of the class width by (90 – 28) or 62		(M1)
	Multiplication of the class width by (90 – 50) or 40		(M1)
	Or $(600 \times (90 - 50) + 1000 \times (50 - 28))/(90 - 28)$ Correct numerator oe		(M2)
	Correct denominator oe		(M1)
	742 cao		A1
6(b)	That the masses of the parcels are increasing linearly (between 600 g and 1000 g) oe	1	B1

Question	Answer	Marks	Partial Marks
7(a)	15 correct entries added to left of diagram, correctly ordered (from right to left) and equally spaced. $\begin{array}{cccccccc c} & & & & & & & & 0 & 3 \\ & & & & & & & & 9 & 6 & 4 \\ 7 & 6 & 5 & 4 & 2 & 0 & 0 & & & & 5 \\ & & 5 & 4 & 2 & 1 & 1 & & & & 6 \end{array}$ (B2 for 1 error, B1 for 2 errors)	4	B3
	Key: 6 4 1 represents a 41 year old female and a 46 year old male		

Question	Answer	Marks	Partial Marks																					
7(b)	<table border="1"> <thead> <tr> <th></th> <th>Males</th> <th>Females</th> </tr> </thead> <tbody> <tr> <td>LQ</td> <td>50</td> <td>39</td> </tr> <tr> <td>M</td> <td>55</td> <td>48</td> </tr> <tr> <td>UQ</td> <td>61</td> <td>63</td> </tr> </tbody> </table> <p>(B2 for 4 or 5 correct entries, B1 for 2 or 3 correct entries)</p>		Males	Females	LQ	50	39	M	55	48	UQ	61	63	3	B3									
	Males	Females																						
LQ	50	39																						
M	55	48																						
UQ	61	63																						
7(c)	<p>Azeeb: Not correct and the median for the males is higher (than that of the females) or $55 > 48$ oe</p> <p>Tebogo: Correct and the interquartile range for the males is smaller (than that of the females) or $11 < 24$ oe</p>	2	B1ft																					
7(d)	<table border="1"> <thead> <tr> <th>Age</th> <th>M</th> <th>F</th> </tr> </thead> <tbody> <tr> <td>$20 \leq x < 30$</td> <td>0</td> <td>0</td> </tr> <tr> <td>$30 \leq x < 40$</td> <td>1</td> <td>4</td> </tr> <tr> <td>$40 \leq x < 50$</td> <td>2</td> <td>4</td> </tr> <tr> <td>$50 \leq x < 60$</td> <td>7</td> <td>2</td> </tr> <tr> <td>$60 \leq x < 70$</td> <td>5</td> <td>5</td> </tr> <tr> <td>$70 \leq x < 80$</td> <td>0</td> <td>0</td> </tr> </tbody> </table>	Age	M	F	$20 \leq x < 30$	0	0	$30 \leq x < 40$	1	4	$40 \leq x < 50$	2	4	$50 \leq x < 60$	7	2	$60 \leq x < 70$	5	5	$70 \leq x < 80$	0	0	1	B1ft
Age	M	F																						
$20 \leq x < 30$	0	0																						
$30 \leq x < 40$	1	4																						
$40 \leq x < 50$	2	4																						
$50 \leq x < 60$	7	2																						
$60 \leq x < 70$	5	5																						
$70 \leq x < 80$	0	0																						
7(e)	<p>Correctly plotted points joined with straight-line segments</p> <p>Suitable key</p> <p>Labels 'age (years)' horizontally and 'number of people/workers' vertically</p>	3	B1ft																					
7(f)	<p>Original ages/data can be seen Exact medians etc. can be calculated</p>	1	B1																					
7(g)(i)	3	1	B1																					
7(g)(ii)	They must all be 57 or over	1	B1																					

Question	Answer	Marks	Partial Marks
8(a)	Any two from <ul style="list-style-type: none"> To eliminate/reduce the variation To see the trend/find the trend line To calculate the seasonal components To make predictions (B1 for one correct reason)	2	B2
8(b)	$x = 1331.25$	3	B1
	$y = 1306.25$		B1
	$z = 1305.5$		B1
8(c)	$332 - 1306.75 [= -974.75]$ or $320 - 1297.75 [= -977.75]$	3	M1*
	$(-974.75 + -977.75)/2$ dep		DM1
	Or $(332 + 320)/2$ and $[-](1306.75 + 1297.75)/2$		(M1)
	$'326' - '1302.25'$ dep		(DM1)
	-976.25 or -976.3 or -976		A1
8(d)	Correct plots vertically	3	B1ft
	Correct plots horizontally		B1
	Appropriate trend line		B1ft
8(e)	The ranger is not correct. The number of swans is falling over time/each year, but not each quarter.	1	B1
8(f)	Reading from their straight trend line at Q2 of 2022 + <i>their (c)</i> . E.g.: $1292 - 976$	2	M1
	316 (Follow through only their trend line)		A1ft

Question	Answer	Marks	Partial Marks
9(a)(i)	$2/9$ oe	1	B1
9(a)(ii)	$4/9$ oe	1	B1
9(a)(iii)	$1/9$ oe	1	B1
9(a)(iv)	$1/3 \times 2/3 + 2/3 \times 2/3 - 1/3 \times 1/3$ or $'2/9' + '4/9' - '1/9'$	2	M1
	$5/9$ oe		A1

Question	Answer	Marks	Partial Marks
9(b)	[P(identical) =] 2/9 or [P(not identical) =] 7/9	4	B1
	Use of P(not identical) = 1 – P(identical)		M1
	'P(identical)' $\times x =$ 'P(not identical)' [$\times 1$]		M1
	[\$]3.50		A1
9(c)(i)	$4/6 \times 2/5 \times 2$ $n \times (n - 1)$ in denominator	3	M1
	product of two probabilities $\times 2$		M1
	8/15 oe		A1
9(c)(ii)	2/15 oe	1	B1
9(c)(iii)	1/3 oe or $1 - '8/15' - '2/15'$	1	B1ft
9(d)	$3 \times '8/15' + 9 \times '2/15' + 6 \times '1/3'$ oe	2	M1
	[\$]4.80		A1

Question	Answer	Marks	Partial Marks
10(a)	$(52 - 62)/8$ or $(54 - 68)/10$	3	M1
	-1.25 and -1.4 (or correctly converting each to any common mean and standard deviation)		A1
	Hazeema (ft , provided one correct scaled score)		A1ft
10(b)	$\Sigma x = 62 \times 159 + 68 \times 141$ [= 19 446]	7	M1
	'19 446'/(159 + 141)		M1
	Mean = 64.82 or 64.8		A1
	$8 = \sqrt{(\Sigma x^2/159 - 62^2)}$ or $10 = \sqrt{(\Sigma x^2/141 - 68^2)}$ or better		M1*
	'621 372' + '666 084' [=1 287 456] dep		DM1
	'1287456'/(159 + 141) – '64.82' ² [=89.8876] dep		DM1
	Standard deviation = 9.48 awrt		A1
10(c)	$\Sigma x/(141 + 149) = 73.4$ or better	3	M1
	'21 286' – 68×141 [= 11 698]		M1
	$11\ 698/149 = 78.5$ awrt		A1

Question	Answer				Marks	Partial Marks					
10(d)	68.2 and 8.8				1	B1					
10(e)	<table border="1" style="width: 100%; text-align: center;"> <tr> <td></td> <td>increase</td> <td>decrease</td> <td>same</td> <td>Not enough info</td> </tr> </table>					increase	decrease	same	Not enough info	2	B1
		increase	decrease	same	Not enough info						
	Mean	✓				B1					
S. d.		✓			B1						