GCSE
STATISTICS
8382/1H

Higher Tier Paper 1

Mark scheme

June 2021

Version: 1.0 Final

216G8382/1H/MS

Mark schemes are prepared by the Lead Assessment Writer and considered, together with the relevant questions, by a panel of subject teachers. This mark scheme includes any amendments made at the standardisation events which all associates participate in and is the scheme which was used by them in this examination. The standardisation process ensures that the mark scheme covers the students' responses to questions and that every associate understands and applies it in the same correct way. As preparation for standardisation each associate analyses a number of students' scripts. Alternative answers not already covered by the mark scheme are discussed and legislated for. If, after the standardisation process, associates encounter unusual answers which have not been raised they are required to refer these to the Lead Examiner.

It must be stressed that a mark scheme is a working document, in many cases further developed and expanded on the basis of students' reactions to a particular paper. Assumptions about future mark schemes on the basis of one year's document should be avoided; whilst the guiding principles of assessment remain constant, details will change, depending on the content of a particular examination paper.

Further copies of this mark scheme are available from aqa.org.uk

Copyright information

AQA retains the copyright on all its publications. However, registered schools/colleges for AQA are permitted to copy material from this booklet for their own internal use, with the following important exception: AQA cannot give permission to schools/colleges to photocopy any material that is acknowledged to a third party even for internal use within the centre.

Copyright © 2021 AQA and its licensors. All rights reserved.

Level of response marking instructions

Level of response mark schemes are broken down into levels, each of which has a descriptor. The descriptor for the level shows the average performance for the level. There are marks in each level.

Before you apply the mark scheme to a student's answer read through the answer and annotate it (as instructed) to show the qualities that are being looked for. You can then apply the mark scheme.

Step 1 Determine a level

Start at the lowest level of the mark scheme and use it as a ladder to see whether the answer meets the descriptor for that level. The descriptor for the level indicates the different qualities that might be seen in the student's answer for that level. If it meets the lowest level then go to the next one and decide if it meets this level, and so on, until you have a match between the level descriptor and the answer. With practice and familiarity you will find that for better answers you will be able to quickly skip through the lower levels of the mark scheme.

When assigning a level you should look at the overall quality of the answer and not look to pick holes in small and specific parts of the answer where the student has not performed quite as well as the rest. If the answer covers different aspects of different levels of the mark scheme you should use a best fit approach for defining the level and then use the variability of the response to help decide the mark within the level, ie if the response is predominantly level 3 with a small amount of level 4 material it would be placed in level 3 but be awarded a mark near the top of the level because of the level 4 content.

Step 2 Determine a mark

Once you have assigned a level you need to decide on the mark. The descriptors on how to allocate marks can help with this. The exemplar materials used during standardisation will help. There will be an answer in the standardising materials which will correspond with each level of the mark scheme. This answer will have been awarded a mark by the Lead Examiner. You can compare the student's answer with the example to determine if it is the same standard, better or worse than the example. You can then use this to allocate a mark for the answer based on the Lead Examiner's mark on the example.

You may well need to read back through the answer as you apply the mark scheme to clarify points and assure yourself that the level and the mark are appropriate.

Indicative content in the mark scheme is provided as a guide for examiners. It is not intended to be exhaustive and you must credit other valid points. Students do not have to cover all of the points mentioned in the Indicative content to reach the highest level of the mark scheme.

An answer which contains nothing of relevance to the question must be awarded no marks.

Glossary for Mark Schemes

GCSE examinations are marked in such a way as to award positive achievement wherever possible. Thus, for GCSE Statistics papers, marks are awarded under various categories.

If a student uses a method which is not explicitly covered by the mark scheme the same principles of marking should be applied. Credit should be given to any valid methods. Examiners should seek advice from their senior examiner if in any doubt.

M	Method marks are awarded for a correct method which could lead to a correct answer.
A	Accuracy marks are awarded when following on from a correct method. It is not necessary to always see the method. This can be implied.
B	Marks awarded independent of method.
ft	Follow through marks. Marks awarded for correct working following a mistake in an earlier step.
SC	Special case. Marks awarded for a common misinterpretation which has some mathematical worth.
M dep	A method mark dependent on a previous method mark being awarded.
B dep	A mark that can only be awarded if a previous independent mark has been awarded.
oe	Or equivalent. Accept answers that are equivalent. eg accept 0.5 as well as $\frac{1}{2}$
[a, b]	Accept values between a and b inclusive.
[a, b)	Accept values $a \leq$ value $< b$
3.14…	Accept answers which begin 3.14 eg 3.14, 3.142, 3.1416
Use of brackets	It is not necessary to see the bracketed work to award the marks.

Examiners should consistently apply the following principles

Diagrams

Diagrams that have working on them should be treated like normal responses. If a diagram has been written on but the correct response is within the answer space, the work within the answer space should be marked. Working on diagrams that contradicts work within the answer space is not to be considered as choice but as working, and is not, therefore, penalised.

Responses which appear to come from incorrect methods

Whenever there is doubt as to whether a student has used an incorrect method to obtain an answer, as a general principle, the benefit of doubt must be given to the student. In cases where there is no doubt that the answer has come from incorrect working then the student should be penalised.

Questions which ask students to show working

Instructions on marking will be given but usually marks are not awarded to students who show no working.

Questions which do not ask students to show working

As a general principle, a correct response is awarded full marks.

Misread or miscopy

Students often copy values from a question incorrectly. If the examiner thinks that the student has made a genuine misread, then only the accuracy marks (A or B marks), up to a maximum of 2 marks are penalised. The method marks can still be awarded.

Further work

Once the correct answer has been seen, further working may be ignored unless it goes on to contradict the correct answer.

Choice

When a choice of answers and/or methods is given, mark each attempt. If both methods are valid then M marks can be awarded but any incorrect answer or method would result in marks being lost.

Work not replaced

Erased or crossed out work that is still legible should be marked.

Work replaced

Erased or crossed out work that has been replaced is not awarded marks.

Premature approximation

Rounding off too early can lead to inaccuracy in the final answer. This should be penalised by 1 mark unless instructed otherwise.

Continental notation

Accept a comma used instead of a decimal point (for example, in measurements or currency), provided that it is clear to the examiner that the student intended it to be a decimal point.

Q	Answer	Marks	Comments
1	$\frac{1}{64}$	B1	

Q	Answer	Marks	Comments
2	Sample frame	B1	

Q	Answer	Marks	Comments
3	C	B1	

Q	Answer	Marks	Comments
4	C	B1	

Q	Answer	Marks	Comments
5(a)	Curry sauce	B1	

Q	Answer	Marks	Comments
5(b)	Any two valid reasons, eg (It's not representative as) they've only asked adults The sample size is (too) small (It's biased as) they've only asked people at fish and chip shops (fish and chips are available elsewhere) Each area will have a different population There isn't a response from all areas Not all side orders were an option / No other box	B2	oe B1 any one valid reason
	Additional Guidance		
	Only asked 670 adults		B1
	Only asked adults		B1
	Only asked 670		B1
	Asked adults		B0
	Asked 670		B0

Q	Answer	Marks	Comments
6(a)	Any reasonable hypothesis relating Year 7, Year 11 and homework eg Year 11 receive more homework than Year 7	B1	oe
	Additional Guidance		
	Allow older (students) to imply Year 11, younger (students) to imply Year 7		
	Y11 students get more homework (than Y7 students)		B1
	Older students get more homework (than younger students)		B1
	16-year-olds have more homework (than 11-year-olds)		B1
	Year 11 spend more time on their homework than Year 7		B1
	Year 11 homework takes longer (on average) than Year 7 homework		B0

Q	Answer	Marks	Comments
6(b)	All Year 11 and All Year 7 students (in Tom's school)	B1	oe
	Additional Guidance		
	All Year 11 and Year 7 (students)		B1
	The Year 11 and Year 7 students		B1
	The Year 11s and Year 7s		B1
	Year 11 and Year 7 students		B0
	(All) students (at Tom's school)		B0

Q	Answer	Marks	Comments
6(c)	Method A named correctly as random (sampling) and any advantage or disadvantage given about Method A eg In A every student has an equal chance of being selected (which is not true of B and/or C)	B2	B1 Method A named correctly as random (sampling) or any advantage or disadvantage given about Method A
	Method B named correctly as convenience/opportunity (sampling) and any advantage or disadvantage given about Method B eg In B this excludes any students who do not go to the dinner hall	B2	B1 Method B named correctly as convenience/opportunity (sampling) or any advantage or disadvantage given about Method B
	Method C named correctly as quota (sampling) and any advantage or disadvantage given about Method C eg In C we do not know the selection method to be used	B2	B1 Method C named correctly as quota (sampling) or any advantage or disadvantage given about Method B
	Method chosen with a correct advantage given and a correct advantage or a correct disadvantage given for the other two methods	B1	
	Additional Guidance		
	Do not award the final B1 with an incorrect (or contradictory) advantage or disadvantage seen for any of the three methods		
	‘At random’ does not imply the name of Method A		
	‘Avoids bias’ is an advantage for Method A		
	Time can be an advantage for Method B, a disadvantage for Method A, a disadvantage for Method C		

Q	Answer	Marks	Comments
6(d)	Any two valid problems, eg There is no time frame given 'How much' is unclear – hours/pieces/nights There are no options given (so answers might be hard to collate)	B2	oe B1 any one valid problem
	Additional Guidance		
	It is an open question meaning data is harder to process		B0
	Reference to some people not getting any homework		B0
	There isn't a place to answer the question (implies no response section)		B1

Q	Answer	Marks	Comments
6(e)	On average, Year 11 have (3 hours) more homework (than Year 7) or On average, Year 7 have (3 hours) less homework (than Year 11)	B1	oe but must state 'on average' or similar, eg generally
	Additional Guidance		
	Condone spend/spent for received/receive		
	Year 11 have more homework than Year 7		B0

Q	Answer	Marks	Comments
6(f)	Students at Tom's school have more homework (on average) than students at other UK schools	B1	oe
	Students at Tom's school have less homework (on average) than students in Shanghai (– China) (schools)	B1	oe Any correct comparison of Tom's data with one of the other countries
	Additional Guidance		
	Ignore any specific times (per week) that are given unless clearly incorrect		

Q	Answer	Marks	Comments
6(g)	The secondary data is for 15-year-olds whereas Tom's data is for Year 11 (who are 15 and 16-year-olds) or The chart could be from several years ago	B1	oe
	Additional Guidance		
	The data from the Internet had no source		B0

Q	Answer	Marks	Comments
6(h)	Tom's as the internet chart had no source (so we have no idea where the data has come from) or The internet data as we don't know how Tom eventually collected his data or The internet data as it is likely to have been collected from more than one school in those countries (whereas Tom's is just from one school)	B1	oe
	Additional Guidance		
	The internet data as it has been collected from more than one school in those countries (this is not known for sure)		B0

Q	Answer	Marks	Comments
7(a)	1000 × 0.003	M1	oe
	3	A1	

Q	Answer	Marks	Comments
7(b)	24.6 and 0.2 seen as part of a calculation	M1	
	123	A1	
	Additional Guidance		
	1.23 (implies 24.6 ÷ 0.2)		M1A0

Q	Answer	Marks	Comments
7(c)	Two correct reasons eg the data is out of date/15 years old (reference to age of data) eg It might be hard to know whether you are a 'heavy smoker' or not (reference to categorisation) eg The chance of developing lung cancer will also be conditional on age…. The chance that a 90 year old heavy smoker will get lung cancer will be smaller than for 40 year old heavy smoker (reference to the generalisation of the Information) eg The estimates do not take into account how long it is since they quit smoking (or how long they smoked before they quit) eg The estimates do not take into account what type of cigarettes people smoke.	B2	oe B1 one correct reason

Q	Answer	Marks	Comments
8(a)	I : (The number of) minutes after 9am	B1	oe
	D : waiting time (to be served)	B1	oe
	Additional Guidance		
	The time after 9am		B0
	The number of staff		B0

Q	Answer	Marks	Comments
8(b)(i)	$y = -0.06 + 0.023x$	B1	

Q	Answer	Marks	Comments
8(b)(ii)	5 plots in the rectangle bounded by the coordinates (210, 5), (270, 5), (270, 4) and (210, 4)	B1	Must be exactly 5 plots added
	6 plots between 270 and 360 minutes (inclusive) in a strong negative correlation	B1	oe Must be exactly 6 plots added
	Additional Guidance		
	If more or less than 11 points added in total, then maximum of B1 awarded		

Q	Answer	Marks	Comments
8(b)(iii)	Any period of exactly two hours between 11.30am and 2pm	B1ft	Do not accept incorrect time notation

Q	Answer	Marks	Comments
8(c)	The data is for one day only	B1	oe
	Take further samples on other days	B1dep	oe eg Take more samples over the next few weeks
	Additional Guidance		
	Ask all the customers		B0B0
	Ask at a different time		B0B0
	Any reference to before 9am		B0B0

Q	Answer	Marks	Comments
9(a)	Median = 1	B1	
	Upper Quartile = 2 and Lower Quartile = 1	M1	oe
	IQR = 2 – 1 = 1	A1	Work for M1 must be seen

Q	Answer	Marks	Comments
9(b)	Cumulative frequencies 190, 335, 425, 485, 500	B1	Allow one error, which may be compounded May be implied by correct answer
	Cumulative step polygon drawn	B1	oe
	All points correct and joined (1, 190), (2, 335), (3, 425), (4, 485), (5, 500)	B1	Ignore graph before (1, 190) and after (5, 500)
	Axes labelled and scaled correctly	B1	

Q	Answer	Marks	Comments
9(c)	Second sample : min = 1, max = 5, median = 2	B1	Probably seen in box plot
	Second sample : LQ = 1, UQ = 3	B1	Possibly seen in box plot (LQ implied by lack of left whisker)
	Two correct box plots drawn	B1ft	ft their answer to (a)
	Both box plots labelled and scale correctly marked and labelled as ‘Number of rewards’	B1	Accept Sample A and Sample B for Sample 1 and Sample 2
	Additional Guidance		
	If only one box plot drawn, treat as second sample		
	Answers in ‘frequencies’ instead of ‘rewards’		B0B0B0B0

Q	Answer	Marks	Comments
9(d)	Their LQ – 1.5 × their IQR or their UQ + 1.5 × their IQR	M1	1 – 1.5 × 2 or 3 + 1.5 × 2
	Their LQ – 1.5 × their IQR and their UQ + 1.5 × their IQR	M1	1 – 1.5 × 2 or 3 + 1.5 × 2
	Correct evaluation of their correct calculations and the correct decision	A1ft	Below – 2 or above 6 and No
	Additional Guidance		
	Values may be seen on box plot or in part (b) Condone use of values from their box plot in part (c) if frequency used If only one unlabelled box plot in part (c) treat as second sample		

Q	Answer	Marks	Comments
9(e)	Correct decision and reason for median target	B1ft	ft their medians
	Correct decision and reason for interquartile range target	B1ft	ft their IQRs

Q	Answer	Marks	Comments
10(a)	The higher the finishing position, the lower the resting heart rate or The higher the finishing position, the higher the resting heart rate	B1	oe Allow hypothesis to be either way around

Q	Answer	Marks	Comments
10(b)	Sight of n = 15	B1	May be seen in calculation
	$(1-)\frac{6\times 60}{15(15^2-1)}$	M1	
	0.892… or 0.893 (= 0.89)	A1	At least 3 decimal places

Q	Answer	Marks	Comments
10(c)	Appropriate conclusion based on answer to parts (a) and (b) or 0.89 shows strong positive correlation	B1ft	

Q	Answer	Marks	Comments
11(a)(i)	To enable the tagged squirrels to mix fully with the others	B1	oe
	Additional Guidance		
	To allow the squirrels to return home	B1	
	To ensure a fair/unbiased result	B0	

Q	Answer	Marks	Comments
11(a)(ii)	As the population size may change due to births and/or deaths or migration or The tags will probably come off over time	B1	oe
	Additional Guidance		
	To make sure there are as few changes in the population as possible	B0	
	Squirrels may leave the forest	B0	

Q	Answer	Marks	Comments
11(b)	Correct equation with sample and population proportions equated eg $\frac{11}{40}=\frac{50}{n}$ eg $x=\frac{50\times 40}{11}$	M1	oe
	181.8 or better	A1	
	181 or 182	B1ft	Their decimal rounded or truncated to a whole number of squirrels with M1 awarded

Q	Answer	Marks	Comments
12(a)	Take additional samples	B1	oe Idea of 'warning' not 'action'

Q	Answer	Marks	Comments
12(b)	Stop production or Have the machine checked/repaired or Reset the machine	B1	oe Idea of 'action' not 'warning'
	Additional Guidance		
	Restart the machine		B1

Q	Answer	Marks	Comments
12(c)	1 plot for each sample (10 in total), all between the warning lines	B1	

Q	Answer	Marks	Comments
13(a)(i)	Geometric mean	B1	oe

Q	Answer	Marks	Comments
13(a)(ii)	$\sqrt[5]{1.067\times1.013\times1.008\times1.005\times1.126}$ or 1.0427…	M1	
	104.27(%)… or 1.0427 or 4.27… and 4.3	A1	Must show to at least 2dp

Q	Answer	Marks	Comments
13(a)(iii)	$200\,000\times1.043^4$	M1	oe For 1.043 allow 1.0427 or better
	[236 410, 236 684]	A1	

Q	Answer	Marks	Comments
13(b)	[25, 29] (÷ 360)	M1	
	$\frac{5}{72}$ or $\frac{13}{180}$ or $\frac{3}{40}$ or $\frac{7}{90}$ or $\frac{[25, 29]}{360}$ or [0.0694, 0.0806] or [6.94%, 8.06%]	A1	oe
	Additional Guidance		
	If correct answer seen, do not ignore further working other than an incorrect simplification		

Q	Answer	Marks	Comments
13(c)	One correct calculation for an angle or one correct angle value eg $\frac{60}{160}\times 360$ or 135 eg $\frac{38}{160}\times 360$ or 85.5 eg $\frac{22}{160}\times 360$ or 49.5 eg $\frac{40}{160}\times 360$ or 90	M1	oe may be implied by correct answer
	All angles correct	A1	oe May be implied by correct answer
	Correct full method for calculating new radius eg $\sqrt{2}\times 4$	M1	oe May be implied by correct answer
	New radius = 5.7 or better	A1	May be implied by correct answer
	New pie chart drawn with • correct radius • correct angles • correct labels	B2	B1 New pie chart drawn with two of the three bullet points correct

Q	Answer	Marks	Comments
13(d)	One correct comparison between the proportion of types of houses	B1	
	Additional Guidance		
	There are more bungalows on Maisy's estate than the new one (not a proportion)	B0	
	There is a higher proportion of bungalows	B0	