



GCSE STATISTICS 8382/1H

Higher Tier Paper 1

Mark scheme

June 2025

Version: 1.0 Final



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.

No student should be disadvantaged on the basis of their gender identity and/or how they refer to the gender identity of others in their exam responses.

A consistent use of 'they/them' as a singular and pronouns beyond 'she/her' or 'he/him' will be credited in exam responses in line with existing mark scheme criteria.

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 © 2025 AQA and its licensors. All rights reserved.

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 $a \leq \text{value} < b$
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	Mark	Comments
1	Valid reason	B1	eg older people will be underrepresented
	Additional Guidance		
	Ignore any non-contradictory or irrelevant statements		
	Answer must reference ages		
	He will mostly/only get younger people		B1
	All/Most of the people he can ask are school age		B1
	Comments identifying a given age range or group eg retired people not being represented		B1
	Not representative because there are only children at the school		B1
	Not representative (does not mention ages)		B0
	There may be another school (comment misses that age is the main problem with his plan)		B0

Q	Answer	Mark	Comments
2	Valid reason	B1	eg there is no reason for there to be lots of ants there or only one nest and food source
	Additional Guidance		
	Ignore any non-contradictory or irrelevant statements		
	D1 is not on the route to and from food source (implies the nest)		B1
	It does not fit the pattern as there is only one food source and nest		B1
	It does not fit the pattern		B0
	Only one nest (needs to reference both nest and food source)		B0
	There is another ant's nest / food source		B0
	The surrounding cells have fewer ants		B0
	His data may be incorrect		B0

Q	Answer	Mark	Comments
3	$\frac{24}{354}$	B1	oe fraction, decimal or percentage
	Additional Guidance		
	Ignore incorrect simplification/conversion after correct answer seen		

Q	Answer	Mark	Comments
4	Valid improvement	B1	eg take his samples from different places in the lake or take some samples from the south end of the lake (as well)
	Additional Guidance		
	Ignore any non-contradictory or irrelevant statements		
	Change location (in the lake)		B1
	Take more samples		B0
	Not representative		B0

Q	Answer	Mark	Comments
5(a)	Two valid comments about the patterns in the data	B2	eg (on average the) income is rising (over time) or quarter 4 always has the highest (income for that year) or quarter 2 is always higher than quarter 1 or quarter 3 is higher each year B1 one valid comment with no other contradicting comment
	Additional Guidance		
	For marks to be awarded must reference income or quarters and not reference an individual point		
	Ignore any non-contradictory or irrelevant statements		
	Do not award B2 with an incorrect comment		
	First quarter is always lowest, last quarter is always highest		B2
	Income goes up (over time)		B1
	Income keeps changing / Income goes up and down (over time)		B0
	It always drops at the end of the year (end of the year could mean Q4)		B0
	Peaks in 4th quarters		B1
	Peaks in 4th quarter		B0
	The steepest line is always between Q3 and Q4		B1
	The biggest increase is always from Q3 to Q4		B1
	Condone it always increases between Q3 and Q4		B1
	Overall trend is increasing		B0
	Positive correlation between quarters and income		B0
	Gets higher and higher every year		B0
	Increasing		B0

Q	Answer	Mark	Comments
5(b)(i)	116 + 121 + 127 + 149 or 513	M1	oe
	their 513 ÷ 4 or 128.25	M1dep	oe
	128 (billion)	A1	
	Additional Guidance		
	Answer may be seen in table but answer line takes precedence		

Q	Answer	Mark	Comments
5(b)(ii)	Alternative method 1: calculating 2019 Quarter 4		
	$\frac{x + 75 + 89 + 96}{4} = 87$ or 88	M1	oe
	88 and No	A1	
	Alternative method 2: using 97 and comparing the moving average to 87		
	$(97 + 75 + 89 + 96) \div 4$ or 89.2(5) or 89.3	M1	oe
	89.2(5) or 89.3 and No	A1	accept 89 with M1 awarded
	Alternative method 3: calculating totals using 87 and 97		
	$87 \times 4 = 348$ and $97 + 75 + 89 + 96 = 357$	M1	oe
	348 and 357 and No	A1	
	Additional Guidance		
	Check table, working space takes precedence		
	Accept trial and improvement with a trial involving an income below 97 and a trial involving an income above 97 that both give moving averages above 87 eg $96 + 75 + 89 + 96$ with 89 and $98 + 75 + 89 + 96$ with 89.5 and No		M1A1
	$88 < 97$		M1A0

Q	Answer	Mark	Comments
5(c)	All 5 moving averages plotted correctly	B2ft	ft their moving average from part (b)(i) B1 3 or 4 moving averages plotted correctly or correct vertical heights for all points with incorrect but consistent horizontal plots tolerance half a small square
	Appropriate single straight trend line with at least 9 moving average points	B1ft	ft their plotted points
	Additional Guidance		
	For a trend line to be appropriate it must extend across all plotted moving average points and be increasing with at least one point above and at least one point below		

Q	Answer	Mark	Comments
6(a)	(Lower quartile =) 5 and (Median =) 6 and (Upper quartile =) 13	B3	B2 two correct values B1 one correct value
	Box plot with values of their lower quartile, median and upper quartile correctly plotted and box drawn and minimum and maximum values correctly plotted and whiskers drawn to meet box	B1ft	ft their median and quartiles with up to B2 awarded where $3 < \text{their LQ} < \text{their median}$ and $\text{their median} < \text{their UQ} < 16$ tolerance half a small square
	Additional Guidance		
	5 and 6 and 13 could be seen in the original list, in the table, in the working or on the diagram even if contradicted		
	Condone a horizontal line drawn through the box		

Q	Answer	Mark	Comments
6(b)	Comparison of the medians in context with no incorrect values for the two medians seen	B1	eg on average the waiting times on Friday were longer
	7 and 5 or Monday has a wider box or Friday has a smaller interquartile range	B1	oe
	Correct contextualised comparison comment for the variation	B1	eg the waiting times on Monday were more varied or the ranges are the same so both days had the same variation (in waiting times)
	Additional Guidance		
	Ignore any non-contradictory or irrelevant statements		
	Check diagram		
	In general Monday took less time	B1	
	Friday has a larger average waiting time	B1	
	8 and 11 and Friday took longer on average	B1	
	8 and 12 and Friday took longer on average	B0	
	Friday has a larger waiting time	B0	
	Friday had a greater median (no interpretation of the median)	B0	
	The width of the box on Friday is narrower	B1	
	4 to 11 is bigger than 9 to 14	B1	
	Waiting times on Monday had a greater variation than on Friday	B1	
	Friday was more consistent	B1	
	Friday varied less	B1	

Q	Answer	Mark	Comments																																																																																																																																																																																									
7(a)	Correct diagram with ordered leaves and numbers vertically aligned.	B3	B2 for at least three correct, ordered rows or all values in correct rows but not ordered B1 for at least two rows correct regardless of order																																																																																																																																																																																									
	Additional Guidance																																																																																																																																																																																											
	Correctly completed diagram (December side is already completed):																																																																																																																																																																																											
	<table><tr><th colspan="10">May</th><th></th><th colspan="10">December</th></tr><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td>1</td><td>0</td><td>2</td><td>0</td><td>6</td><td>7</td><td>8</td><td>9</td><td></td><td></td><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td><td></td><td></td><td></td><td>6</td><td>5</td><td>2</td><td>1</td><td>3</td><td>1</td><td>2</td><td>2</td><td>4</td><td>5</td><td>6</td><td>7</td><td>7</td><td>9</td><td></td></tr><tr><td></td><td></td><td></td><td></td><td>9</td><td>8</td><td>6</td><td>5</td><td>4</td><td>1</td><td>4</td><td>0</td><td>2</td><td>3</td><td>6</td><td>7</td><td></td><td></td><td></td><td></td><td></td></tr><tr><td></td><td></td><td>9</td><td>8</td><td>7</td><td>7</td><td>6</td><td>4</td><td>2</td><td>0</td><td>5</td><td>1</td><td>2</td><td>5</td><td>9</td><td></td><td></td><td></td><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td><td></td><td>9</td><td>8</td><td>7</td><td>7</td><td>3</td><td>1</td><td>6</td><td>0</td><td>1</td><td>2</td><td>4</td><td></td><td></td><td></td><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td>7</td><td>4</td><td>2</td><td>7</td><td>0</td><td>4</td><td>5</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td>6</td><td>0</td><td>8</td><td>6</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr></table>																				May											December																		1	0	2	0	6	7	8	9												6	5	2	1	3	1	2	2	4	5	6	7	7	9						9	8	6	5	4	1	4	0	2	3	6	7								9	8	7	7	6	4	2	0	5	1	2	5	9											9	8	7	7	3	1	6	0	1	2	4														7	4	2	7	0	4	5																6	0	8	6									
	May											December																																																																																																																																																																																
									1	0	2	0	6	7	8	9																																																																																																																																																																												
							6	5	2	1	3	1	2	2	4	5	6	7	7	9																																																																																																																																																																								
					9	8	6	5	4	1	4	0	2	3	6	7																																																																																																																																																																												
			9	8	7	7	6	4	2	0	5	1	2	5	9																																																																																																																																																																													
					9	8	7	7	3	1	6	0	1	2	4																																																																																																																																																																													
							7	4	2	7	0	4	5																																																																																																																																																																															
								6	0	8	6																																																																																																																																																																																	
Key 3 4 7 represents 43 cars on a day in May and 47 cars on a day in December																																																																																																																																																																																												
Values for May are in correct rows but are consistently ordered from ascending from left to right rather than right to left.																			B2																																																																																																																																																																									
Condone lack of vertical alignment for B2 and B1																																																																																																																																																																																												
Condone lack of vertical alignment if cell/value crossed out and values would have correct vertical alignment otherwise eg <table><tr><td>6</td><td>5</td><td>2</td><td>1</td><td>2</td><td>3</td></tr></table>																			6	5	2	1	2	3																																																																																																																																																																				
6	5	2	1	2	3																																																																																																																																																																																							
Marks can be scored for work in working space if grid blank or crossed out																																																																																																																																																																																												

Q	Answer	Mark	Comments
7(b)	Valid comparison	B1ft	ft their stem and leaf diagram eg the average (number of cars passing each day) is greater in May (than in December) or the modal group for May is higher
	Additional Guidance		
	If values are seen in all rows ft their stem and leaf diagram in (a)		
	Condone use of the median instead of the mode		
	Any values stated must be correct		
	Ignore use of the mean		
	Ignore any reference to skew		
	The modal group for May was (around) 50 and for December it was only (around) 30		B1
	The modal group for May was (around) 50 and for December it was (around) 30		B0
	The mode for May was (around) 50 and for December it was only (around) 30		B0
	More cars in May		B1
	More days have more cars in May		B1
	Less cars in December		B1
	Most days have more cars in May		B1
	It is bigger for May		B1
	More cars in the middle in May but December it is at the beginning		B0
	All days have more cars in May		B0
	Incorrect months referenced		B0

Q	Answer	Mark	Comments
8	One correct reason for each statement	B3	B2 for two correct reasons B1 one correct reason eg it will depend on how often you have used the sunbed (reference to risk being frequency linked – Andi) eg frequently is undefined (reference to categorisation - Kav) eg you may not have been born in 1961 (reference to base year of data - Frankie)
	Additional Guidance		
	If person is named then reason must match their statement		
	Ignore any non-contradictory or irrelevant statements		
	Referencing sunbed use (Andi)		
	No reference to time spent /frequency		B1
	May not use a sunbed		B1
	We do not know the risk (for people who haven't used sunbeds)		B1
	We don't know when they used the sunbed		B0
	Referencing lotion (Kav)		
	May not use (perfumed body) lotion		B1
	Does not say how much greater the risk is		B1
	Frequency is not defined		B1
	We do not know the risk (for people who haven't used lotion)		B1
	Referencing year of birth (Frankie)		
	Only mentions people born in 1961/ one year		B1
	We do not know the risk for people not born in 1961		B1
	The information is from 1961 and is too old		B0
	None/Some of the statements (might) apply to them		B1

Q	Answer	Mark	Comments
	Alternative method 1: comparing expected probabilities or outcomes		
9	$\frac{1}{6} \times \frac{1}{6} (\times 150)$	M1	oe accept a fraction, decimal or percentage
	$\frac{1}{36}$ or [4.16, 4.17]	A1	oe fraction, decimal or percentage
	No and their $\frac{1}{36}$ and $\frac{11}{150}$ as comparable values or (11 and) [4.16, 4.17]	A1ft	oe fraction, decimal or percentage ft their $\frac{1}{36}$ or their [4.16, 4.17] providing both values are written in a comparable form allow 4 if [4.16, 4.17] already seen
	Alternative method 2: calculating expected probability of blue		
	$\frac{11}{150} \div \frac{1}{6}$ or $\frac{1}{6} \times x = \frac{11}{150}$ or $11 \times 6 \div 150$	M1	oe accept a fraction, decimal or percentage
	$\frac{11}{25}$	A1	oe fraction, decimal or percentage
	their $\frac{11}{25}$ and $\frac{1}{6}$ as comparable values and No	A1ft	oe ft their $\frac{11}{25}$ providing both values are written in a comparable form

Question 9 Additional Guidance continues on the next page

9 cont.	Additional Guidance	
	Equations and expressions may involve words or other informal symbols	
	Use the scheme that awards the best mark	
	Accept the blue dice is unfair for No	
	For the values to be in comparable form they could both be given as decimals or percentages. If the two values are given as fractions they must have either a common denominator or a common numerator.	
	$\frac{1}{6} \times \frac{11}{150}$	M0

Q	Answer	Mark	Comments
10(a)	Points plotted at (0,6), (1,6), (1,16), (2,16), (2,31) and (3,31) and straight lines used to complete step polygon	B3	B2 4 or more points plotted correctly but not necessarily joined or joined incorrectly or horizontal lines at 6, 16 and 31 on vertical axis but consistently in the wrong place on the x -axis B1 2 or 3 points plotted correctly
	Additional Guidance		
	Their points can be implied by their lines		
	Ignore line joining (0, 0) to (0, 6) if seen		

Q	Answer	Mark	Comments
10(b)	Alternative method 1: working with 5 or more goals		
	$64 - 49$ or $6 + 3 + 2 + 2 + 1 + 1$ or 15	M1	oe
	$\frac{\text{their } 15}{64} (\times 100)$	M1dep	oe
	Yes and (20(%)) and [23, 24](%)	A1	oe comparable forms for 20% and $\frac{15}{64}$ eg $[0.23, 0.24]$ and 0.2 $\frac{75}{320}$ and $\frac{64}{320}$ $\frac{1}{5}$ and $\frac{1}{4.266...}$
	Alternative method 2: working with under 5 goals		
	$\frac{49}{64}$ or [76, 77](%)	M1	oe
	$\left(1 - \text{their } \frac{49}{64}\right) (\times 100)$ or $100(\%) - \text{their } [76, 77](\%)$	M1dep	oe
	Yes and (20(%)) and [23, 24](%)	A1	oe comparable forms for 20% and $\frac{15}{64}$ eg $[0.23, 0.24]$ and 0.2 $\frac{15}{64}$ and $\frac{12.8}{64}$ $\frac{1}{5}$ and $\frac{1}{4.266...}$ 5 and 4.266...

Question 10 (b) continues on the next page

Q	Answer	Mark	Comments
10(b) cont.	Alternative method 3: comparison of 20% of the total to number scoring 5 or more goals		
	64 – (6 + 10 + 15 + 15 + 3) or 15	M1	oe
	0.2×64 or $\frac{20}{100} \times 64$ or 12.8	M1	oe
	Yes and 12.8 and 15	A1	allow 12 or 13 if 12.8 already seen
	Additional Guidance		
	If answer is given as fraction or decimal the final mark can only be awarded if 20% is converted into a comparable form		
	70 – 55 = 15 15 followed by $\frac{15}{70}$		M0 M0

Q	Answer	Mark	Comments
10(c)	$\frac{5}{64}$ or [0.07, 0.0782] or $\frac{4}{64-1}$ or [0.06, 0.0635]	M1	oe fraction, decimal or percentage ft their 64 from 10(b)
	$\frac{5}{1008}$ or [0.004, 0.005]	A1ft	oe fraction, decimal or percentage ft their 64 from 10(b) SC1 $\frac{10}{64}$ oe
	Additional Guidance		
	SC1 using $\frac{5}{64}$ for two games		

Q	Answer	Mark	Comments
11(a)	December 2022 and June 2023 and valid comparison	B1	either order eg because the difference (between these index numbers) is the largest
	Additional Guidance		
	Candidates must state which years in their answer		
	Any values stated must be correct		
	December 2022 has the highest number and June 2023 is the lowest	B1	
	December 2022 and June 2023 and $27.3 > 21.2$	B1	
	December 2022 and June 2023 and a change of 27.3	B0	
	December 2022 and June 2023 as the values are increasing but this one then drops to below the baseline	B1	
	Dec 22 and Jun 23 as differences are 21.2, 4.2, 27.3 and 27.3 is biggest	B1	
	December 2022 and June 2023 as differences are 21.2, 4.2, 27.3	B0	

Q	Answer	Mark	Comments
11(b)	Alternative method 1: comparing index numbers		
	101.9	B1	
	$\frac{153.7}{150.4} (\times 100)$ or [102, 102.2]	M1	oe
	[102, 102.2] and 101.9 and statement that claim is incorrect	A1ft	oe ft their 101.9 providing it is a value from the table and a correct statement for their values
	Alternative method 2: percentage change		
	101.9	B1	implied by 1.9% oe
	$\frac{153.7 - 150.4}{150.4} (\times 100)$ or [2.1, 2.2](%)	M1	oe
	[2.1, 2.2](%) and 1.9(%) and statement that claim is incorrect	A1ft	oe ft their 101.9 providing it is a value from the table and a correct statement for their values
	Alternative method 3: comparison of prices (working forwards)		
	101.9	B1	implied by 1.019
	150.4×1.019 or [153.2, 153.3]	M1	oe
	[153.2, 153.3] and 153.7 and statement that claim is incorrect	A1ft	ft their 101.9 providing it is a value from the table and a correct statement for their values

Question 11 (b) continues on the next page

Q	Answer	Mark	Comments
11(b) cont.	Alternative method 4: comparison of prices (working backwards)		
	101.9	B1	
	$\frac{100}{\text{their } 101.9}$ or [0.981,0.982] or $\frac{150.4}{153.7}$ or [0.978,0.979]	M1	oe
	[0.981,0.982] and [0.978,0.979] and statement that claim is incorrect	A1ft	oe values ft their 101.9 providing it is a value from the table and a correct statement for their values
	Additional Guidance		
	For a correct comparison, values must be in a comparable format 1.9% and 2.2 and the claim is incorrect 101.9 and 2.2% and the claim is incorrect Condone 1.9 and 2.2 and the claim is incorrect		B1M1A0 B1M1A0 B1M1A1
	Wales = 100.6, $\frac{153.7}{150.4} \times 100 = 102.2$ The statement is wrong because it increased by more in Somerset.		B0M1A1ft

Q	Answer	Mark	Comments
12(a)	Valid statement including age and gender	B1	eg (there is a) greater percentage of males in most of the age groups (compared to females)
	Additional Guidance		
	Ignore any non-contradictory or irrelevant statements		
	Statement must refer to both age and gender		
	Condone statement referring to number rather than percentage/proportion		
	Greater percentage of males in groups between 35 and 79 years	B1	
	More males in groups between 35 and 74 years (of age) (than females)	B1	
	Lists of individual age groups and genders showing they do not match	B1	
	Greater percentage of males across most ages	B1	
	More males across most ages	B1	
	Greater percentage of males (not referencing age)	B0	
	There is a (large) difference between the percentages of males and females in most/some age groups	B1	
	There is a (large) difference between the percentages of males and females in the/all age groups	B0	
	No two bars for the same age are the same	B0	
	The percentages vary a lot in terms of age and gender	B0	
	More male than females / Less females than males	B0	
	The diagram shows both age and gender	B0	

Q	Answer	Mark	Comments
12(b)	$8 + 8.3 + 7.1 + 4.6$ or 28(%) or 0.28	M1	oe
	$4200 \times \frac{100}{\text{their } 28}$ or $4200 \div \frac{\text{their } 28}{100}$	M1dep	oe
	15000	A1	SC2 30 000 SC1 [1 206 829, 1 206 830]
	Additional Guidance		
	SC2 for calculating number of females SC1 for calculating 28% of the total population		
	$(0.08 + 0.083 + 0.071 + 0.046) \times 4200$ (finding individual percentages)		M1

Q	Answer	Mark	Comments
12(c)	Yes and the 10th percentile is 19 or less and the 90th percentile is above 59 so the difference must be more than 40	B1	oe
	Additional Guidance		
	Any values stated must be correct		
	Yes and $60 - 19 = 41$		B1
	Yes and $79 - 19 = 60$		B0
	Yes and difference between the 26 th and 89.8 th percentiles is (at least) 40 so the difference between the 10 th and 90 th (percentiles) must be more than 40		B1
	Yes and $59 - 19 = 40$ but 90 th percentile must be greater than 59 so difference must be greater than 40 Yes and $59 - 19 = 40$		B1 B0

Q	Answer	Mark	Comments
13	$\frac{54}{80} = \frac{189}{N}$ or $\frac{80 \times 189}{54}$	M2	oe accept use of ratio M1 $\frac{54}{80}$ or $\frac{80}{54}$ oe
	280	A1	
	Additional Guidance		
	<i>N</i> can be any letter for the population		
	280 followed by $280 \times 80 = 22\,400$		M2A0
	$\frac{189}{N} = \frac{80}{54}$		M1A0

Q	Answer	Mark	Comments
14(a)	[4.5,5]	B1	

Q	Answer	Mark	Comments
14(b)(i)	The line is steeper (between 10% and 20%) for (Country) A (than it is for Country B) or the line goes up by more (between 10% and 20%) for (Country) A (than it is for Country B) or there are [14,16](%) of towns in this interval for (Country) A and [9,11](%) for (Country) B	B1	oe
	Additional Guidance		
	Ignore any non-contradictory or irrelevant statements		
	Allow comments about gradient between 10% and 20% providing interpretation is correct		
	Condone comments about there being more towns between 10% and 20% for Country A		
	Higher number of towns in Country A as there are 15% compared to 10%	B1	
	Higher number of towns in Country A as there are 15 compared to 10	B1	
	Higher number of towns in Country A between 10% and 20%	B0	
	The (cumulative) frequency increases more for (Country) A	B1	
	The graph for B is above A	B0	

Q	Answer	Mark	Comments
14(b)(ii)	[6.5, 7.5] and [25.5, 26.5]	B2	either order B1 one correct value

Q	Answer	Mark	Comments
15(a)	Any two reasons in context eg the probability of getting a rare card will not change or she opens a fixed number of packs or the contents of any pack are independent (from the contents of the other packs) or each pack either has a rare card or does not	B2	B1 one reason in context or 2 reasons with no context
	Additional Guidance		
	Ignore any non-contradictory or irrelevant statements		
	The odds of getting one (rare card) are fixed		B1
	Probability of a rare card is known (must state that it does not change)		B0
	Constant probability (not linked to context)		B0
	Two outcomes (not linked to context)		B0
	Fixed number of attempts (not linked to context)		B0
	Fixed/Discrete number of cards in a pack		B0
	Each rare card is independent from each other		B0

Q	Answer	Mark	Comments
15(b)	Alternative method 1: calculating probability of no rare cards		
	${}^5C_0 \times 0.85^5 (\times 0.15^0)$ or [0.44, 0.444]	M1	oe fraction or percentage
	$1 - \text{their [0.44, 0.444]}$	M1dep	oe fraction or percentage
	[0.55, 0.56]	A1	oe fraction or percentage
	Alternative method 2: calculating the probability of 1, 2, 3, 4 and 5 rare cards		
	${}^5C_1 \times 0.85^4 \times 0.15^1$ or [0.39, 0.392] or ${}^5C_2 \times 0.85^3 \times 0.15^2$ or [0.138, 0.14] or ${}^5C_3 \times 0.85^2 \times 0.15^3$ or [0.024, 0.024 4] or ${}^5C_4 \times 0.85^1 \times 0.15^4$ or [0.002, 0.0022] or ${}^5C_5 (\times 0.85^0) \times 0.15^5$ or [0.000 075 9, 0.000 08]	M1	oe fraction or percentage
	their [0.39, 0.392] + their [0.138, 0.14] + their [0.024, 0.0244] + their [0.002, 0.0022] + their [0.000 075 9, 0.000 08]	M1dep	oe fraction or percentage
	[0.55, 0.56] with all calculated values in range	A1	oe fraction or percentage

Q	Answer	Mark	Comments
16(a)(i)	Point plotted correctly	B1	tolerance half a small square

Q	Answer	Mark	Comments
16(a)(ii)	Horizontal lines drawn at 25.5 and 29.5 at on vertical axis	B2	B1 one correct line tolerance half a small square
	Additional Guidance		
	Intention must be for straight horizontal lines		
	Line can be solid or dashed/dotted and must extend beyond plotted points		
	Ignore any labels unless used to indicate choice		

Q	Answer	Mark	Comments
16(b)	6(th sample) or sample 6	B1	

Q	Answer	Mark	Comments
16(c)	The machine should be stopped/serviced/reset/restarted and it is above the (upper) action limit	B2	oe B1 the machine should be stopped/serviced/reset/restarted
	Additional Guidance		
	Ignore any non-contradictory or irrelevant statements		
	Reset then check the machine		B1
	Check and reset the machine (choice)		B0
	Stop and adjust the machine		B1
	Adjust the machine		B0
	Reduce the mass of crisps per packet		B0
	It is an outlier		B0
	Take another sample/retest/recheck		B0
	Needs to be within (warning/action/both) limits		B0
	Outside acceptable limits/range		B0
	Action must be taken		B0
	It is too heavy		B0

Q	Answer	Mark	Comments
17	Geometric mean	B1	

Q	Answer	Mark	Comments
18(a)	Collecting the data herself is impractical or it is quicker/cheaper/convenient	B1	oe
	Additional Guidance		
	Ignore any non-contradictory or irrelevant statements		
	It would be too difficult to collect the data herself		B1
	It was taken over two years so would be impractical to repeat It was taken over two years		B1 B0
	The data has already been collected		B0
	It's a reliable source (because they are a doctor)		B0
	The data could be outdated		B0

Q	Answer	Mark	Comments
18(b)(i)	4 and the value for the mass is missing (which is what she is investigating)	B1	
	Additional Guidance		
	Ignore any non-contradictory or irrelevant statements		
	4 and the mass and the gender are missing		B1
	4 and the row is incomplete (does not reference the mass)		B0
	4 and the gender is missing		B0
	4 and the data is missing/blank		B0

Q	Answer	Mark	Comments
18(b)(ii)	48 and the gender is missing or one cell is empty / the data is incomplete	B1	oe
	Additional Guidance		
	Ignore any non-contradictory or irrelevant statements		
	48 and only the mass and year are recorded		B1
	48 and only the mass is recorded		B1
	48 and only the year is recorded		B0
	48 and it's an outlier and there is no gender		B0

Q	Answer	Mark	Comments
18(c)	$\frac{2099472500}{151} - \left(\frac{558800}{151}\right)^2$ or $\frac{2099472500}{151} - (\text{their } 3701)^2$ or [206 390, 213 792]	M1	oe
	457(.0...)	A1	accept 457 or better
	(Mean =) $\frac{558800}{151}$ or [3700.6, 3701]	M1	oe implied by 5072 or 2330 accept 3700
	their 3701 + 3 × their 457 or 5072 or their 3701 – 3 × their 457 or 2330	M1dep	dep on M2 ft their 3701 and their 457
	There are no outliers (as the lightest and heaviest penguins are within these limits) with $3701 \pm 3 \times 457$ seen	A1ft	ft a correct decision if M3 awarded

Q	Answer	Mark	Comments
18(d)(i)	No and valid reason	B1	eg the peak for the Gentoo curve is further along the axis (than the Chinstrap curve)
	Additional Guidance		
	Ignore statements giving values for the mean unless they mention how it is linked to the graph		
	Ignore any non-contradictory or irrelevant statements		
	Answers must reference at least one of the curves		
	Accept reference to mean, mode, median or average		
	Gentoo curve's peak is a higher value/mass (than the Chinstrap curve)	B1	
	Gentoo curve's peak is higher (than the Chinstrap curve)	B0	
	The Gentoo peak is further to the right	B1	
	The mean for Chinstrap is [3600,3800] and the mean for Gentoo is [5000,5200]	B0	
	The median is lower for the Chinstrap	B0	
	Gentoo curve goes further up the graph	B0	
	Gentoo have a larger mass	B0	
	The Gentoo curve is wider	B0	
	Reference to skew/range	B0	

Q	Answer	Mark	Comments
18(d)(ii)	Yes and valid reason	B1	eg the (Gentoo) curve is wider
	Additional Guidance		
	Ignore statements giving values for the standard deviation/IQR unless they mention how it is linked to the graph. Ignore any non-contradictory or irrelevant statements.		
	Answers must reference at least one of the curves		
	The Chinstrap curve is narrower		B1
	Yes and the Gentoo curve's peak is above a higher value (than the Chinstrap curve)		B0
	Yes and Chinstrap graph is steeper		B0
	Yes and it is more spread out (implies the curve)		B1
	Yes and they are more spread out (does not reference the curve)		B0
	Yes and larger distribution (implied comparison)		B1
	Yes and large distribution (no comparison)		B0
	Yes and the range is higher because the curve is wider		B1
	Yes and the range is higher		B0

Q	Answer	Mark	Comments
18(e)	Lightest and valid reason	B1	eg the standardised score (for the lightest penguin) is further from zero/mean
	Additional Guidance		
	Ignore any non-contradictory or irrelevant statements		
	Lightest and 2.42 is closer to mean/zero		B1
	Lightest and the lightest penguin is 0.45 further from the mean		B1
	Lightest and 2.87 is bigger than 2.42		B1
	Lightest and it is negative		B0
	Lightest and highest standardised score		B0

Q	Answer	Mark	Comments
18(f)	It was reliable for the Adelie (penguins) and it is less/fairly reliable for the Chinstrap (penguins) and it is unreliable for the Gentoo (penguins)	B2	oe B1 one correct comment
	Additional Guidance		
	Ignore any non-contradictory or irrelevant statements		
	Allow larger value for skew		
	It is only reliable for Adelie penguins		B1
	Adelie is (very) reliable, the Chinstrap is less/somewhat reliable and the Gentoo is unreliable		B2
	Adelie is most reliable, then the Chinstrap and then the Gentoo		B1
	Adelie is (the most) reliable		B1
	Adelie is the most reliable and the Gentoo is the least reliable		B1
	All three are reliable (one correct and two incorrect statements)		B0