



Pearson
Edexcel

Mark Scheme

Summer 2023 (Results)

Pearson Edexcel GCSE (9 – 1)
In Statistics (1ST0)
Higher Paper 1H

Edexcel and BTEC Qualifications

Edexcel and BTEC qualifications are awarded by Pearson, the UK's largest awarding body. We provide a wide range of qualifications including academic, vocational, occupational and specific programmes for employers. For further information visit our qualifications websites at www.edexcel.com or www.btec.co.uk. Alternatively, you can get in touch with us using the details on our contact us page at www.edexcel.com/contactus.

Pearson: helping people progress, everywhere

Pearson aspires to be the world's leading learning company. Our aim is to help everyone progress in their lives through education. We believe in every kind of learning, for all kinds of people, wherever they are in the world. We've been involved in education for over 150 years, and by working across 70 countries, in 100 languages, we have built an international reputation for our commitment to high standards and raising achievement through innovation in education. Find out more about how we can help you and your students at: www.pearson.com/uk

Summer 2023

Publications Code 1ST0_1H_2306_MS

All the material in this publication is copyright

© Pearson Education Ltd 2023

General marking guidance

These notes offer general guidance, but the specific notes for examiners appertaining to individual questions take precedence.

- 1** All candidates must receive the same treatment. Examiners must mark the last candidate in exactly the same way as they mark the first.

Where some judgement is required, mark schemes will provide the principles by which marks will be awarded; exemplification/indicative content will not be exhaustive. When examiners are in doubt regarding the application of the mark scheme to a candidate's response, the response should be sent to review.

- 2** All the marks on the mark scheme are designed to be awarded; mark schemes should be applied positively. Examiners should also be prepared to award zero marks if the candidate's response is not worthy of credit according to the mark scheme. If there is a wrong answer (or no answer) indicated on the answer line always check the working in the body of the script (and on any diagrams), and award any marks appropriate from the mark scheme.

Questions where working is not required: In general, the correct answer should be given full marks.

Questions that specifically require working: In general, candidates who do not show working on this type of question will get no marks – full details will be given in the mark scheme for each individual question.

- 3** **Crossed out work**

This should be marked **unless** the candidate has replaced it with an alternative response.

- 4** **Choice of method**

If there is a choice of methods shown, mark the method that leads to the answer given on the answer line.

If no answer appears on the answer line then mark both methods **as far as they are identical** and award these marks.

- 5** **Incorrect method**

If it is clear from the working that the "correct" answer has been obtained from incorrect working, award 0 marks.

- 6** **Follow through marks**

Follow through marks which involve a single stage calculation can be awarded without working as you can check the answer, but if ambiguous do not award.

Follow through marks which involve more than one stage of calculation can only be awarded on sight of the relevant working, even if it appears obvious that there is only one way you could get the answer given.

7 Ignoring subsequent work

It is appropriate to ignore subsequent work when the additional work does not change the answer in a way that is inappropriate for the question or its context. (eg an incorrectly cancelled fraction when the unsimplified fraction would gain full marks).

It is not appropriate to ignore subsequent work when the additional work essentially makes the answer incorrect (eg incorrect algebraic simplification).

8 Probability

Probability answers must be given as a fraction, percentage or decimal. If a candidate gives a decimal equivalent to a probability, this should be written to at least 2 decimal places (unless tenths).

Incorrect notation should lose the accuracy marks, but be awarded any implied method marks.

If a probability fraction is given then cancelled incorrectly, ignore the incorrectly cancelled answer.

9 Range of answers

Unless otherwise stated, when an answer is given as a range (eg 3.5 – 4.2) then this is inclusive of the end points (eg 3.5, 4.2) and all numbers within the range.

Guidance on the use of abbreviations within this mark scheme

- M** method mark awarded for a correct method or partial method
- A** accuracy mark (awarded after a correct method; if no method is seen then full marks for the question are implied but see individual mark schemes for more details)
- B** unconditional accuracy mark (no method needed)
- oe** or equivalent
- cao** correct answer only
- ft** follow through (when appropriate as per mark scheme)
- sc** special case
- dep** dependent (on a previous mark)
- indep** independent
- awrt** answer which rounds to
- isw** ignore subsequent working

Question number	Answer	Additional guidance	Mark																														
1(a)	<p>B2 for correctly completing the 2017 season on the stem and leaf diagram B1 for a suitable key</p> <table border="1" data-bbox="421 379 1216 852"> <thead> <tr> <th>2017 season</th> <th></th> <th>2018 season</th> </tr> </thead> <tbody> <tr><td></td><td>1</td><td>9</td></tr> <tr><td></td><td>2</td><td>5 6 9</td></tr> <tr><td></td><td>3</td><td>5 7 9</td></tr> <tr><td>9 6</td><td>4</td><td>3 3 5 5 5 7</td></tr> <tr><td>6 5 5 5 4 2 2</td><td>5</td><td>0 2 2 4 6 8 8</td></tr> <tr><td>6 4 2 1 1</td><td>6</td><td>4 5 6</td></tr> <tr><td>5 4 4 2</td><td>7</td><td>0 2</td></tr> <tr><td>5</td><td>8</td><td></td></tr> <tr><td>4</td><td>9</td><td>0</td></tr> </tbody> </table> <p>Key: 4 9 0 represents a score of 94 in the 2017 season and 90 in the 2018 season</p>	2017 season		2018 season		1	9		2	5 6 9		3	5 7 9	9 6	4	3 3 5 5 5 7	6 5 5 5 4 2 2	5	0 2 2 4 6 8 8	6 4 2 1 1	6	4 5 6	5 4 4 2	7	0 2	5	8		4	9	0	<p>B2 for a fully correct back-to-back stem and leaf diagram</p> <p>(B1 for ordered diagram with at most 2 errors or for unordered diagram)</p> <p>AND</p> <p>B1 for a suitable key for the stem and leaf diagram.</p> <p>Accept a key given as two parts.</p> <p>If key given in two parts then this must be complete and there must be reference to 2017 and 2018 or it must be clear how this is interpreted for the two sides.</p> <p>e.g. 4 9 represents a score of 94 in the 2017 (season) and 9 0 represents a score of 90 in the 2018 (season)</p> <p>or 1 9 represents 19, 4 9 represents 94.</p>	(3)
2017 season		2018 season																															
	1	9																															
	2	5 6 9																															
	3	5 7 9																															
9 6	4	3 3 5 5 5 7																															
6 5 5 5 4 2 2	5	0 2 2 4 6 8 8																															
6 4 2 1 1	6	4 5 6																															
5 4 4 2	7	0 2																															
5	8																																
4	9	0																															
(b)	<p>M1 for $58 - 39$</p> <p>A1 19</p>	<p>M1 for $58 - k$ or $k - 39$ or for both 39 and 58 identified</p> <p>or for $59.5 - k$ or $k - 38.5$ or for both 38.5 and 59.5 identified (leads to IQR=21)</p> <p>or for $58 - k$ or $k - 38$ or for both 38 and 58 identified (leads to IQR=20)</p> <p>A1 accept 20 or 21</p> <p>If working is seen then it must be correct and consistent with their answer for award of M1A1.</p> <p>Accept if answer given in table.</p>	(2)																														

(c)(i)	<p>B1 for the median in the 2017 season is higher than the median in the 2018 season</p> <p>B1ft for the interquartile range for the 2017 season is lower than the interquartile range in the 2018 season</p> <p>B1ft (dep) so the team had higher scores and were more consistent in the 2017 season (so no Naomi is not correct)</p>	<p>B1 for correct comparison of medians. Accept use of average rather than median. Does not need to be in context. Allow e.g. median for 2017 is higher.</p> <p>B1ft for correct comparison of interquartile range. Must be correct for their IQR. Does not need to be in context. Allow e.g. IQR for 2017 is lower.</p> <p>B1ft for assessing the appropriateness of the conclusion. Dep on B1B1 scored (question indicates reasons required) Final B mark needs to have contextual interpretation seen in the answer e.g. ...so higher scores, ... so more consistent. Do not award the 3rd B mark if the reasoning is linked with the wrong statistic e.g. linking median with consistency.</p>	(3)
(ii)	<p>B1 for giving a limitation of using the data provided to comment on team improvement</p> <p>e.g.</p> <ul style="list-style-type: none"> • we don't know if the matches were played against the same teams in both seasons • we don't know what the scores were for the opposition teams • opposition team performance may also have changed 	<p>B1 for a limitation of using the data provided for the conclusion</p> <p>Do not accept:</p> <ul style="list-style-type: none"> • team members might have changed <p>Accept:</p> <ul style="list-style-type: none"> • we only know the scores, we don't know the results of the matches. • we don't know how many times they have won/lost each season. 	(1)

(d)	<p>B1 for e.g.</p> <ul style="list-style-type: none"> • appropriate as if there are outliers in the data then it might impact further diagrams or calculations • appropriate as there appear to be outliers in the data / 90 is much higher than the other values for 2018 season • not appropriate as there is nothing to suggest that the data is not genuine • not appropriate as using all data would be fully representative 	<p>B1 for assessing the appropriateness of checking for outliers by calculation Accept 'yes' for appropriate and 'no' for not appropriate. Allow reference to specific diagrams or calculations, but not median or interquartile range as these have already been calculated. Do not accept as reasons:</p> <ul style="list-style-type: none"> • ...making it more accurate/precise • ...make results fair • ...an outlier can affect the <u>data</u> with no other comment / clarification • ...using median/IQR which has already removed outliers (this has already been done) • ...they should use all the data (needs further explanation) 	(1)
-----	---	---	-----

Question number	Answer	Additional guidance	Mark
2(a)	B1 for continuous		(1)
(b)	M1 for use of correct scale (may be implied by either correct answer) A1 for 87 and 45		(2)
(c)	M1 for correctly plotting one bar using their scale A1 for both bars (35 and 14) correct on histogram with a correct scale	Do not accept frequency density values for scale unless label changed from frequency to frequency density.	(2)
(d)	B1 for correct interpretation of negative skew e.g. <ul style="list-style-type: none"> the weights of basketball players (from 2000 to 2009) below the median have a greater spread the mean of the basketball players weights is lower than the median of the basketball players weights more than half of the basketball players weigh more than the mean 	B1 for correct contextual interpretation of skew	(1)
(e)(i)	M1M1 $\frac{12 \times 175 + 146 \times 185 + 175 \times 195 + 323 \times 205 + 146 \times 215 + 8 \times 225}{810}$ (= 200.79 ...) A1 for 200.79...	M1 for consistent use of fx with x within interval (including end points) and attempt to sum. Use of mid-interval gives $2100 + 27010 + 34125 + 66215 + 31390 + 1800 = 162640$ If multiplications not shown then allow one incorrect product for both M marks. M1 for correct use of fx with x the mid-interval value, attempt to sum and division by 810 A1 for awrt 200.8 OR answer of 201 from correct working	(3)

(e)(ii)	B1ft for e.g. mean height of basketball players has increased (by 9.9 cm)	B1ft for correct comparison of means Accept e.g. increased (FT their value from (i)) Allow e.g. players have gotten taller (FT their value from (i)) Ignore figures when marking.	(1)
---------	---	--	-----

Question number	Answer	Additional guidance	Mark
3(a)	B1 for e.g. the number of motorcycles (first registered in the UK) is decreasing over time	B1 for a suitable hypothesis. B0 for a question. Require reference to: motorbike/motorcycle/vehicle; time e.g. over time / comparing years / comparing time of year; registrations or sales changing/increasing/decreasing/staying the same	(1)
(b)	<p>B1 for a correct statement identifying any seasonality e.g.</p> <ul style="list-style-type: none"> • the greatest values are in quarter 2 • the least values are in quarter 4 <p>depB1 for a correct interpretation in context for the identified seasonality e.g.</p> <ul style="list-style-type: none"> • ...which means more motorcycle registrations (Q2) • ...which means less motorcycle registrations (Q4) 	<p>B1 for a correct statement identifying seasonality Allow quarters to be identified by reference to correct season or months. Condone if quarter referenced and an incorrect attempt to interpret as a season / months. If more than one quarter commented on then ignore extra non-contradictory comments and interpretations. depB1 for a correct interpretation of seasonality</p> <p>B1B1 may be scored in a single comment e.g. more motorcycle registrations in spring Condone sales for registrations. Allow vehicles for motorcycles.</p>	(2)
(c)	<p>B1 for e.g.</p> <ul style="list-style-type: none"> • moving averages allow us to see the trend • the pattern in the data repeats after four quarters • this removes the seasonal variation 	<p>B1 for a correct statement assessing the appropriateness of using (4-point) moving averages Accept e.g. there are 4 quarters / 4 seasons.</p>	(1)
(d)	<p>B1 for a correct comparison of seasonality e.g.</p> <ul style="list-style-type: none"> • greatest values for motorcycles are in quarter 2, but the greatest values for cars are in quarter 1 • lowest values for <u>both</u> motorcycles and cars are in quarter 4 	<p>B1 for a correct statement comparing seasonality Must reference / indicate both motorcycles and cars in discussion of quarters. Do not accept reference to one year and quarter e.g. in 2017 quarter 1...</p>	(1)

		Allow quarters to be identified by reference to correct season or months. Condone if quarter referenced and an incorrect attempt to interpret as a season / months. Ignore additional non-contradictory comments.	
(e)	<p>M1 correct plotting of at least one moving average OR all moving averages with correct horizontal plotting OR all moving averages with correct vertical plotting A1 fully correct</p> <p>B1 straight trend line through moving averages within tolerance</p>	<p>M1 for correctly plotting one moving average OR for recognising correct horizontal or vertical positions Horizontal plots should be on the appropriate grid line. Vertical plots between 570 & 580 inclusive. A1 for accurately plotting the three moving averages B1 their line should extend in the horizontal direction at least from 2017 Q3 to 2019 Q2 and be vertically within two squares of 620 at Q3 2017 (600 to 640 inclusive) and two squares of 570 at Q2 2019 (550 to 590 inclusive). Must be a ruled line. B1 may be awarded without plotting of the moving averages or with incorrect plots.</p>	(3)
(f)	<p>B2 for e.g. the number of cars (first registered in the UK) is decreasing (over time)</p> <p>(B1 for downwards/negative/correct description of the trend with missing or incorrect contextual interpretation)</p>	<p>B2 for a correct description of the trend with contextual interpretation</p> <p>(B1 for downwards/correct description of the trend with missing or incorrect contextual interpretation) Allow falling. Do not allow negative correlation alone, but condone if accompanied by e.g. decreasing / downwards trend. Ignore reference to figures.</p>	(2)

Quarters	Months	Condone for season
Quarter 1	January – March	Winter (Dec – Feb)
Quarter 2	April – June	Spring (March – May)
Quarter 3	July – September	Summer (June – August)
Quarter 4	October - December	Autumn (Sept – Nov)

Question number	Answer	Additional guidance	Mark
4	<p>B1 for each of five correct comments Maximum 3 marks for collecting data. Maximum 3 marks for processing and presenting data. eg Collecting data</p> <ul style="list-style-type: none"> • suitable sample size for processing/presenting planned • does not plan to make sure that there are enough male students and enough female students sampled • some students may not agree to take part / may have absent students • asking students to measure their own data / distances is likely to lead to inaccuracy / inconsistency / students may lie about their distances / data would need to be cleaned after collection • should repeat the jump several times for each student and use their mean distance / to help identify anomalous results in the jumps for the individual • good to split data into male students and female students as gender may impact on distance jumped • recording age is unnecessary as this is not being used in the processing and presenting / age will impact how far they can jump so he should pick just one age group • good idea to consider outliers/may not be appropriate to remove outliers as they could be genuine data • should plan for what to do with outliers if/when identified. • does not include athletes outside the school • not appropriate to use (to predict results) for athletes at the Olympic games as the sample is not representative of them • should plan to control other factors that may impact performance / extraneous variables e.g. time of day <p>Processing and presenting data</p> <ul style="list-style-type: none"> • scatter diagrams are appropriate for investigating how height affects distance jumped • scatter diagrams are appropriate as the data is bivariate / they show correlation 	<p>B1 for each correct comment on the appropriateness of the plans for collecting, processing and presenting the data Do not accept: number of male and female students may differ</p> <p>Each bullet point can be awarded once only.</p> <p>Accept for 9th bullet point: e.g. removing outliers would improve reliability</p> <p>Comments about appropriate /</p>	(5)

	<ul style="list-style-type: none">• (the height is the explanatory/independent variable and the distance jumped is the response/dependent variable) so height should be on horizontal axis and distance jumped on the vertical axis• age may also affect how far the students jump so should consider splitting by age• a line of best fit is appropriate if correlation is seen – should check for correlation first• double mean point will help to ensure that the line of best fit is suitable/appropriate• may not be appropriate to use line of best fit to predict results for athletes at the Olympic games as their heights may be outside the range of data collected / use of line of best fit may require extrapolation• use of software / spreadsheet is appropriate due to sample size	<p>not appropriate alone are not sufficient. There needs to be a correct reason.</p>	
--	---	--	--

Question number	Answer	Additional guidance	Mark
5(a)	<p>B1 for giving a reason why Spearman rank correlation coefficient may not be appropriate e.g.</p> <ul style="list-style-type: none"> • data may not be bivariate • may not be the same entrants to each competition 	<p>B1 for correctly identifying reason why Spearman rank correlation coefficient may not be appropriate</p> <p>Accept:</p> <ul style="list-style-type: none"> • may not be the same number of entrants to each competition <p>Do not accept:</p> <ul style="list-style-type: none"> • correlation does not imply causation • different competitions / different skills 	(1)
(b)	<p>M1 for difference in ranks: 2, 0, -3, 1, -1, -3, 1, 3</p> <p>M1 $(r_s =) 1 - \frac{6 \times 34}{8 \times (8^2 - 1)}$</p> <p>A1 0.595(...)</p> <p>B1ft Positive correlation</p> <p>B1ft Judges had (some) agreement with each other</p>	<p>M1 for difference in ranks (condone one slip and allow \pm). Can be implied by $\sum d^2 = 34$</p> <p>M1 for demonstrating correct use of formula for Spearman's rank correlation coefficient.</p> <p>Must have attempted sum of squares (condone one error in squares, intention to add) and substitution into correct formula. Must use 8.</p> <p>A1 for awrt 0.6</p> <p>B1ft for statistical interpretation of '0.6'.</p> <p>Allow reference to $0.6 > 0$</p> <p>Condone indication of positive without reference to correlation</p> <p>For ft $-1 \leq \textit{their} \text{ SRCC} \leq 1$</p> <p>B1ft for correct contextual interpretation of correlation coefficient.</p> <p>Ignore references to strength in interpretations of correlation and agreement.</p> <p>For ft $-1 \leq \textit{their} \text{ SRCC} \leq 1$</p> <p>Note: for the B marks there must be a correlation coefficient seen as they are interpreting this.</p>	(5)

Question number	Answer	Additional guidance	Mark
6(a)	M1 for $\frac{1071.9}{1000.0} \times 100 (= 107.2)$ A1 for 107.2	M1 for correct calculation of chain base index number A1 for correct answer awrt 107.2	(2)
(b)(i)	M1 for $\sqrt[5]{101.3 \times 97.4 \times 99.9 \times 101.3 \times '107.2'}$ A1ft for 101.4	M1 for correct calculation of the geometric mean of their five chain base index numbers A1ft 101.36... or better follow through their (a) Note: arithmetic mean is 101.42 so working must be shown (as required in question).	(2)
(b)(ii)	B2ft for (average) rate of <u>'increase'</u> (of gold prices) is <u>'1.4'% per month</u> (B1ft for (average) rate of 'increase' (of gold prices))	B2ft for complete correct contextual interpretation as rate of 'increase' including interpretation of the value (B1ft for incomplete contextual interpretation as rate of 'increase')	(2)

Question number	Answer	Additional guidance	Mark
7(a)	M1 for $\frac{683}{n} = \frac{148}{587}$ oe A1 for 2709	Allow alternative equivalent methods Statement of the two fractions alone is not sufficient, these must be appropriately equated oe. M1 implied by 2708.(9...) Answer must be integer	(2)
(b)	B1B1B1 for each of three comments on the appropriateness of the method for data collection e.g. <ul style="list-style-type: none"> • sample sizes are small / she should capture more • sufficient time is allowed for the frogs to mix back into the population • population may change due to births / deaths / migration / predation during the time frame • tags may fall off as (only stuck on) • the recapture is done at the same time of day (controlling extraneous variables) • recapture should be done at the same location 	B1B1B1 for each of three comments on the appropriateness of the method for data collection Ignore additional non-contradictory comments.	(3)

Question number	Answer	Additional guidance	Mark
8(a)	B1 for 0.82, 0.26, 0.74 oe in correct places on tree diagram	B1 for correct probabilities on the tree diagram	(1)
(b)	M1 for $\frac{0.18 \times 0.84}{0.18 \times 0.84 + "0.82" \times "0.74"}$ A1 for 0.199(4...)	M1 for correct method to calculate conditional probability. Allow use of their figures from (a) as long as these are probabilities (between 0 and 1). A1 for awrt 0.20 Accept 0.2 with working.	(2)
(c)(i)	M1 for $\frac{"0.26"}{0.16}$ A1 for 1.625	M1 for correct method to calculate relative risk A1 cao (allow 1.63)	(2)
(c)(ii)	B1ft for the risk of a driver who has not had advanced training breaking the speed limit is "1.625" times/"162.5%" of the risk of a driver who has had advanced training breaking the speed limit oe	B1ft for correct interpretation of their relative risk in context Allow use of their figure rounded in comment e.g. 1.63, 1.6 Allow e.g. the risk of a driver who has not had advanced training breaking the speed limit is " 62.5% " more than the risk of a driver who has had advanced training breaking the speed limit Do not allow for incorrect interpretation e.g. the risk of a driver who has not had advanced training breaking the speed limit is " 162.5% " more than risk of a driver who has had advanced training breaking the speed limit – this is not equivalent to 162.5% of the risk of... Condone if a correct statement seen followed by an incorrect attempt to write in a different form.	(1)

<p>(d)</p>	<p>B2 for e.g. no / not possible to tell from the data, it could be that the proportion of newly qualified drivers speeding has reduced or it could be that the proportion of not newly qualified drivers speeding has increased</p> <p>(B1 for e.g. it could be that the proportion of newly qualified drivers speeding has reduced or it could be that the proportion of not newly qualified drivers speeding has increased)</p>	<p>B2 for a complete reason assessing the appropriateness of the conclusion with statistical reason Accept for B2 no as the relative risk may have decreased due to drivers who are not newly qualified exceeding the speed limit more than before.</p> <p>(B1 for correct interpretation of the change in relative risk with incorrect or no conclusion) Accept for B1 the relative risk may have decreased due to drivers who are not newly qualified exceeding the speed limit more than before with no or incorrect conclusion.</p>	<p>(2)</p>
------------	--	---	------------

Question number	Answer	Additional guidance	Mark
9(a)	M1 for $\frac{3 \times (19.6 - 20.0)}{1.33}$ A1 for -0.902	M1 for a method to calculate skewness A1 for awrt -0.90 Accept -0.9 or $-\frac{120}{133}$	(2)
(b)	B1B1B1B1B1 for each of five comparisons or contextual interpretations Comparison of average e.g. <ul style="list-style-type: none"> • mean for male handspans is greater • median for male handspans is greater Interpretation e.g. <ul style="list-style-type: none"> • ...on average males have greater / wider handspans Comparison of measure of spread e.g. <ul style="list-style-type: none"> • standard deviation for male handspans is greater • range for male handspans is greater Interpretation e.g. <ul style="list-style-type: none"> • ...male handspans are more varied / less consistent / more spread out Comparison of skew e.g. <ul style="list-style-type: none"> • male handspan data is positively skewed, female handspan data is negatively skewed Interpretation e.g. <ul style="list-style-type: none"> • ...neither male nor female handspans are normally distributed • ...for male handspans the mean handspan is greater than the median handspan for female handspans the mean handspan is less than the median handspan • ...for male handspans the spread of handspans greater than the median handspan is greater than the spread of handspans less than the median, but for female handspans the spread of handspans greater 	B1B1B1B1B1 for each of five comparisons or contextual interpretations B1 for comparison of measure of central tendency. Award mark for comparison of means or for comparison of medians. B1(dep) for correct interpretation of comparison of central tendency Accept e.g. males have bigger hands. B1 for comparison of a measure of spread Condone wider range. B1(dep) for correct interpretation of comparison of a measure of spread Do not allow more distributed or wider for contextual interpretation of spread. B1ft for comparison of skew Follow through their (a). Allow correct or ft (may judge from (a) or histogram). B1(dep) for correct interpretation of comparison of skew Follow through their (a) For comparisons correct statistical terminology is required i.e. mean, median, standard deviation, range, skew.	(5)

	<p>than the median handspan is less than the spread of handspans less than the median</p> <ul style="list-style-type: none"> • ... more than half of the males have handspans lower than the mean for male handspans, however more than half of the females have handspans greater than the mean for female handspans. 	Condone use of 'average' when comparing mean or median if values are given.	
(c)	<p>B1 for e.g.</p> <ul style="list-style-type: none"> • reference to unequal numbers of male and female handspan measurements • sample size for females is greater • sample sizes are different 	B1 for correct justification of the appropriateness of using weighted mean	(1)

Question number	Answer	Additional guidance	Mark
10(a)	<p>B1 for correct interpretation in context of the figure 19.61 e.g. for each 1cm increase in snout-vent length the weight of the snake would increase by 19.61g / by this amount</p>	<p>B1 for a correct contextual interpretation of the gradient of the regression line Accept if converse is given e.g. for each 19.61g increase in weight the snout-vent length would increase by 1cm.</p>	(1)
(b)	<p>B1B1B1 for each of three correct comparisons of regression equations</p> <p>B1 for a correct comparison of regression equations of all snakes e.g.</p> <ul style="list-style-type: none"> • all snakes increase in weight as the snout-vent length increases <p>B1 for a correct comparison of regression equations between male of species and female of species e.g.</p> <ul style="list-style-type: none"> • male Timber Rattlesnake weight increases more per unit increase in snout-vent length than female Timber Rattlesnake weight • male Eastern Racer weight increases more per unit increase in snout-vent length than female Eastern Racer weight • male snakes weight increases faster with increasing snout-vent length <p>B1 for a correct comparison of regression equations between the two species e.g.</p> <ul style="list-style-type: none"> • Timber Rattlesnake weight increases more per unit increase in snout-vent length than Eastern Racer weight • male Timber Rattlesnake weight increases more per unit increase in snout-vent length than male Eastern Racer weight • female Timber Rattlesnake weight increases more per unit increase in snout-vent length than female Eastern Racer weight 	<p>B1 for overall comparison of regression equations.</p> <p>B1 for comparison of regression equations for male of species and for female of species</p> <p>B1 for comparison of regression equations between species</p>	(3)

(c)(i)	B1 for e.g. a histogram would allow you to check for a bell shape to the distribution	B1 for correct comment on the appropriateness of using a histogram Accept reference to identifying if the data is skewed or not. Accept for checking if distribution is symmetrical.	(1)
(c)(ii)	<p>B1B1 for each of two from:</p> <p>B1 for e.g. checking that mean, median and mode are equal</p> <p>B1 for e.g.</p> <ul style="list-style-type: none"> • checking that 68% of data are within 1 standard deviation of the mean • checking that 95% of data are within 2 standard deviations of the mean • checking that almost all of data are within 3 standard deviations of the mean <p>B1 for e.g.</p> <ul style="list-style-type: none"> • use the mean, median and standard deviation to calculate the skewness (a normal distribution would have a skewness of 0) 	<p>B1 for correct comment on the appropriateness of using mean, median and mode Allow mean, median and mode should be similar / close.</p> <p>B1 for correct comment on the appropriateness of using mean with standard deviation. If more than one comment on amounts of data within 1/2/3 standard deviations of the mean then award B1 if any one correct. For 3 s.d. comment allow reference to 99.7% or 99.8% for ‘almost all’. Must refer to within x standard deviations of the <u>mean</u>, do not condone use of ‘average’.</p> <p>B1 for correct comment on the appropriateness of using mean, median and standard deviation to calculate skewness.</p>	(2)

Question number	Answer	Additional guidance	Mark
11(a)	B1 for e.g. <ul style="list-style-type: none"> • this is a sensitive question / people may want to lie • people are more likely to be truthful (using random response) • people may not want to answer it otherwise 	B1 for a correct comment on the appropriateness of the random response technique in this context. Condone 'it is a personal question'	(1)
(b)	B2 for not appropriate together with reference to two of: <ul style="list-style-type: none"> • the numbers not being randomly generated • lack of time frame in the question • personal errands are not defined / people may interpret differently (B1 for not appropriate together one of the reasons OR for two or three of the reasons given with no or incorrect conclusion)	B2 for a complete assessment of the appropriateness of the question with statistical reasons (B1 for a partial assessment of the appropriateness of the question)	(2)
(c)	M1 for $\frac{93}{126+874+610+93+615+208} \times n = 14.5$ A1 for 394	M1 for setting up an appropriate equation or inequality to determine minimum total sample size. M1 implied by 393.(8...) or 407.(4...) Allow for working with 15 A1 394 cao	(2)

Modifications to the mark scheme for Modified Large Print (MLP) papers: 1ST0 1H

Only mark scheme amendments are shown where the enlargement or modification of the paper requires a change in the mark scheme.

The following tolerances should be accepted on marking MLP papers, unless otherwise stated below:

Angles: $\pm 5^\circ$

Measurements of length: ± 5 mm

PAPER: 1ST0_1H

Question		Modification	Mark scheme notes
1		<p>Wording added 'Look at the diagram for Question 1 in the Data Booklet. It shows a back-to-back stem and leaf diagram.'</p> <p>Numbers left aligned.</p> <p>Wording added 'in the Data Booklet'.</p> <p>Diagram enlarged.</p> <p>Key box moved above diagram and enlarged.</p> <p>Table enlarged and left aligned.</p>	Standard mark scheme.
2		<p>Q2(b) and 2(c)</p> <p>Wording added 'Look at the diagram for Question 2(b) and 2(c) in the Data Booklet. It shows an incomplete histogram.'</p> <p>Wording added 'in the Data Booklet'.</p> <p>Wording added 'on the following page'.</p> <p>Diagram enlarged.</p> <p>Small squares removed.</p> <p>Grey shading removed and replaced with dotted shading.</p> <p>Axes labels moved to top of vertical axis and to left of horizontal axis.</p> <p>Open headed arrows</p> <p>Black grid lines.</p> <p>Table enlarged and left aligned.</p> <p>Wording added 'in the Data Booklet'.</p> <p>Wording added 'on the previous page.'</p> <p>Wording added 'There are two spaces to fill'.</p> <p>Wording added 'on the previous page'.</p> <p>Wording added 'in the Data Booklet.'</p> <p>Q2(e)</p> <p>Wording added 'Look at the table for Question 2(e) in the Data Booklet'.</p> <p>Wording 'below' removed and replaced with 'in the Data Booklet'.</p> <p>Table enlarged.</p>	Standard mark scheme.

PAPER: 1ST0_1H

Question	Modification	Mark scheme notes
3	<p>Wording added 'Look at Diagram 1 for Question 3 in the Data Booklet. It shows a'.</p> <p>Wording removed 'The'.</p> <p>Wording removed 'shows' and replaced by 'with'.</p> <p>Diagram enlarged.</p> <p>Small squares removed.</p> <p>Black grid lines.</p> <p>Open headed arrows.</p> <p>Axes labels moved to top of vertical axis and to left of horizontal axis.</p> <p>Right axis labelled.</p> <p>Dashed lines made longer and thicker.</p> <p>Wording added 'in Diagram 1.'</p> <p>Wording added 'Look at Diagram 2 for Question 3 in the Data Booklet.'</p> <p>Wording added 'in the Data Booklet'.</p> <p>Diagram enlarged.</p> <p>Small squares removed.</p> <p>Black grid lines.</p> <p>Open headed arrows.</p> <p>Axes labels moved to top of vertical axis and to left of horizontal axis.</p> <p>Right axis labelled.</p> <p>Dashed lines made longer and thicker.</p> <p>Crosses changed to solid dots.</p> <p>Numbers stacked vertically and left aligned.</p> <p>Wording added 'on Diagram 2'.</p> <p>Leeway needed.</p>	<p>3e</p> <p>Changes to guidance: For M and A marks Horizontal plots should be on the appropriate grid line. Vertical plots between 550 & 600 inclusive.</p> <p>B1 their line should extend in the horizontal direction at least from 2017 Q3 to 2019 Q2 and be vertically between 600 and 650 at Q3 2017 and between 550 and 600 at Q2 2019 Must be a ruled line.</p>
4	<p>Wording added 'Look at the information for Question 4 in the Data Booklet.'</p> <p>Wording 'Here is' removed and replaced with 'It shows in the Data Booklet'.</p> <p>Frame removed.</p>	<p>Standard mark scheme.</p>

PAPER: 1ST0_1H

Question		Modification	Mark scheme notes
5		Wording added 'Look at the table for Question 5(b) in the Data Booklet'. Wording removed 'following'. Wording added 'shown in the table in the Data Booklet'. Table enlarged. Wording added 'in the Data Booklet'.	Standard mark scheme.
6		Wording added 'Look at the table for Question 6 in the Data Booklet. It'. Wording removed 'The table'. Table enlarged and turned vertical. Solid answer line added. Shading changed. Wording added 'in the Data Booklet.'	Standard mark scheme.
7		Wording added 'Look at the table for Question 7 in the Data Booklet'. Wording removed 'following'. Wording added 'shown in the table in the Data Booklet'. Table enlarged. Frame removed.	Standard mark scheme.
8		Wording added 'Look at the diagram for Question 8 in the Data Booklet. It shows a probability tree diagram.' Wording added 'in the Data Booklet'. Diagram enlarged. Answer lines made solid.	Standard mark scheme.

PAPER: 1ST0_1H

Question	Modification	Mark scheme notes
9	<p>Wording added ‘Look at Diagram 1 and Diagram 2 for Question 9 in the Data Booklet.’</p> <p>Wording removed ‘here are’ and replaced by ‘and’.</p> <p>Wording added ‘are shown in the Data Booklet. Diagram 1 shows male handspans. Diagram 2 shows female handspans.’</p> <p>Diagrams enlarged.</p> <p>Small squares removed.</p> <p>Black grid lines.</p> <p>Open headed arrows.</p> <p>Axes labels moved to top of vertical axis and to left of horizontal axis.</p> <p>Right axes labelled.</p> <p>Shading changed.</p> <p>Frames removed and information moved above diagram.</p>	Standard mark scheme.
10	<p>Wording added ‘Look at the table for Question 10 in the Data Booklet.’</p> <p>Wording removed ‘below’ and replaced by ‘in the Data Booklet’.</p> <p>Table enlarged.</p>	Standard mark scheme.
11	<p>Q11(b)</p> <p>Wording added ‘Look at the information for Question 11(b) in the Data Booklet. It is a random response question designed by the manager.’</p> <p>Frame removed and wording left aligned.</p> <p>Answer boxes enlarged.</p> <p>Q11(c)</p> <p>Wording added ‘Look at the table for Question 11(c) in the Data Booklet’.</p> <p>Wording added ‘in the Data Booklet’.</p> <p>Table enlarged.</p>	Standard mark scheme.

Pearson Education Limited. Registered company number 872828
with its registered office at 80 Strand, London, WC2R 0RL, United Kingdom