Mark Scheme (Results)

June 2011

GCSE Statistics (5ST1H_01)
Higher Paper 01

Edexcel is one of the leading examining and awarding bodies in the UK and throughout the world. We provide a wide range of qualifications including academic, vocational, occupational and specific programmes for employers.

Through a network of UK and overseas offices, Edexcel's centres receive the support they need to help them deliver their education and training programmes to learners.

For further information, please call our GCE line on 0844 576 0025, our GCSE team on 0844 576 0027, or visit our website at www.edexcel.com.

If you have any subject specific questions about the content of this Mark Scheme that require the help of a subject specialist, you may find our **Ask The Expert** email service helpful.

Ask The Expert can be accessed online at the following link: http://www.edexcel.com/Aboutus/contact-us/

NOTES ON MARKING PRINCIPLES

1   Mark Schemes
These should be applied positively. Candidates should all receive the same treatment. They should be rewarded for what they have shown they can do rather than penalised for omissions.

2   Types of mark
M marks: method marks
A marks: accuracy marks   Note: you cannot give an A mark if you have given M0
B marks: unconditional accuracy marks (independent of M marks)

3   Abbreviations

cao – correct answer only	ft – follow through
isw – ignore subsequent working	SC: special case
oe – or equivalent (and appropriate)	dep – dependent
indep – independent	awrt – anything which rounds to
QWC – quality of written communication	( ) – brackets round words mean these are not essential

4   No working
If no working is shown then correct answers normally score full marks
If no working is shown then incorrect (even though nearly correct) answers score no marks.

5   With working
If there is a wrong 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.
If working is crossed out and still legible, then it should be given any appropriate marks, as long as it has not been replaced by alternative work.
If it is clear from the working that the "correct" answer has been obtained from incorrect working, award 0 marks. Send the response to review, and discuss each of these situations with your Team Leader.
If there is no answer on the answer line then check the working for an obvious answer.
Any case of suspected misread loses A (and B) marks on that part, but can gain the M marks. Discuss each of these situations with your Team Leader.
If there is a choice of methods shown, then no marks should be awarded, unless the answer on the answer line makes clear the method that has been used.

6   **Follow through marks**
        Follow through marks which involve a single stage calculation can be awarded without working since you can check the answer yourself, 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: e.g. incorrect cancelling of a fraction that would otherwise be correct
        It is not appropriate to ignore subsequent work when the additional work essentially makes the answer incorrect e.g. algebra.
        Transcription errors occur when candidates present a correct answer in working, and write it incorrectly on the answer line; mark the correct answer.

8   **Probability**
        Probability answers must be given a fractions, percentages or decimals. If a candidate gives a decimal equivalent to a probability, this should be written to at least 2 decimal places (unless tenths), unless it states otherwise on the mark scheme.
        Incorrect notation should lose the accuracy marks, but be awarded any implied method marks.
        If a probability answer is given on the answer line using both incorrect and correct notation, award the marks.
        If a probability fraction is given then cancelled incorrectly, ignore the incorrectly cancelled answer.

9   **Linear equations**
        Full marks can be gained if the solution alone is given on the answer line, or otherwise unambiguously indicated in working (without contradiction elsewhere). Where the correct solution only is shown substituted, but not identified as the solution, the accuracy mark is lost but any method marks can be awarded.

10   **Parts of questions**
        Unless allowed by the mark scheme, the marks allocated to one part of the question CANNOT be awarded in another.

11   **Range of answers**
        Unless otherwise stated, when an answer is given in a range (e.g. 3.5 - 4.2) then this is inclusive of the end points, and includes all the numbers in between.

12   **Quality of Written Communication**
        This is denoted by an asterisk near the question number/part (*). Mark schemes will indicate within the table how marks are to be allocated.
        In this subject we need to see that correct statistical terms are used.

5ST1H_01				
Question	**Working**	**Answer**	**Mark**	**Notes**
1		3D Broken bar No vertical scale	3	B1 Any equivalent expression accepted e.g. Diagram is at an angle. B1 B1 Note 1.'It does not have any values' is equivalent to ' No scale' 2. Bars have gaps between them B0 3. No labels B0.
2 (a)		Point plotted at (26.0, 12.5)	1	B1 Allow half a small square tolerance.
(b)		Negative (Correlation) The greater the weight the lower the life expectancy	2	B1 Do not accept negative skew. B1 Converse accepted. Any equivalent statement in context accepted
(c)(i)		Point plotted at (34.1, 11.2)	2	B1 Allow half a little square tolerance.
(ii)		Ruled straight Line of best fit through their mean point.		B1 Needs to lie between : (25,12) and (25,12.8) and also (55, 8) and (55, 8.8) and needs to cover x from 12 to 54
(d)		Any one of: It involves extrapolation. It is outside the data set. It uses data from only 8 dogs. Border terriers are not one of the types of dog used so may not fit in to the data set.	1	B1 Any equivalent answer relating to it being beyond the given data is acceptable. B0 It is far away from the other points.

5ST1H_01				
Question	Working	Answer	Mark	Notes
3 (a)		7.1	1	B1 No other answer accepted
(b)		West Midlands	1	B1 No other answer accepted
(c)	40.2 + 19.5 + 1.2	60.9	2	M1 For effort to add correct numbers or to do 100 – (19.6 + 9.1) could be implied by 71.3 A1 Correct answer only Note: 1. Do NOT ignore subsequent working 2. 60.9 no working gets both marks. 3. 67.968 no working gets M0A0 4. 60.9 in working but 60/61 on answer line is M1A1.
(d)	42.8 + 20.1 + 2.1+ 19.8 + 6.8	91.6	2	M1 For effort to add five correct numbers A1 Correct answer only. Do NOT ignore subsequent working Note: 91.6 in working but 91/92 on answer line is M1A1.
(e)		Any one of: Some graduates might not have been contacted or replied or have given the required information	1	B1 Any equivalent answers accepted. Not all graduated or some dropped out or did not study in UK all B0 Reference to rounding errors B0.

5ST1H_01				
Question	**Working**	**Answer**	**Mark**	**Notes**
4 (a)		Drinking cocoa (before bedtime) reduces/may help blood pressure.	1	B1 Converse accepted Any equivalent expression accepted (ignore reference to numbers). B0 if written as a question
(b)		Any two of: It is quick to do. It is easier to do It is cheap to do. It is convenient There is less data.	2	B1 B1 Accept equivalent statements. Accept the opposites if the word census is included. e.g. A census is slower. Note: You may get two reasons in one comment.
(c)		**A list/register/database of the students** (at the university) OR A register of the university	1	B1 Any equivalent answer accepted. Note: Do not give this mark if it refers to the sample. e.g. A list of the students in the sample B0
(d)		Stratified	1	B1
(e)		It helps to check whether the effects felt is due to **cocoa** or is **psychological** oe. OR Comparing a group having cocoa with one not having **cocoa** (the control group) helps to assess the effect of having **cocoa**.	1	B1 Equivalent answers that suggest comparing a group with and without cocoa makes assessment easier is acceptable Allows you to compare genders B0

5ST1H_01				
Question	**Working**	**Answer**	**Mark**	**Notes**
5 (a)	<table><tr><td></td><td>V</td><td>R</td><td>Total</td></tr><tr><td>G</td><td></td><td></td><td>50</td></tr><tr><td>P</td><td></td><td></td><td>44</td></tr><tr><td>Total</td><td>56</td><td>38</td><td>94</td></tr></table>	50, 44, 56, 38 94	2	B1 for two figures correct B1 B1 for all correct. All figures must be in the correct place
(b)		Paris and Venice	2	B1
		32 choose this holiday (all the others are less than this) OR This is the highest figure excluding totals OR Because the highest number is 32 OR The highest number of holidays booked.		B1 Equivalents accepted. Do not accept comments about the attractions of these places.
(c)		$\dfrac{26}{'94'}$ or $\dfrac{13}{47}$ or awrt 0.28 or 28%	1	B1ft This is a follow through for their 94 Do not accept ratios. Must be a fraction, decimal or percentage.
(d)	$\dfrac{32}{their 56}$	$\dfrac{32}{'56'}$ or $\dfrac{16}{28}$ or $\dfrac{8}{14}$ or $\dfrac{4}{7}$ or awrt 0.57 or 57%	2	M1 ft for picking 56 for denominator or their 56 (in table) or seeing 32:56 A1 cao Do NOT accept ratios. Must be a fraction, decimal or percentage.

5ST1H_01				
Question	**Working**	**Answer**	**Mark**	**Notes**
6 (a)		Method 2/Questionaire . Any one of: Method 1 only allows those people already using facilities to give their views. Method 1 is biased. Method 2 gives all residents the chance to say what they think Method 2 is more reliable. Method 2 could mean more people (participate). Method 2 (probably) uses a greater variety of people.	2	B1 could be implied by dismissing method 1 B1 **dependent on first B1** Equivalent expressions accepted. Note: Method 2 is more accurate B0.
(b)		Any one from: This is an open question. This will allow for too many answers. This has no answer boxes. Assumes use of facilities Can't answer if do not use facilities. Difficult to show/process the results.	1	B1 Equivalent expressions accepted. Note: B0 for 'It is a leading question' on it's own. B0 for 'Only uses for recycling'.

5ST1H_01				
Question	**Working**	**Answer**	**Mark**	**Notes**
(c)		e.g. How often do you use the recycling facilities (each month)? ☐ Never ☐ Once a week/month ☐ More than once a week/month. ☐ Other.	2	B1 An unbiased question related to recycling. B1 Must have at least three non-overlapping all inclusive answer boxes **and** reference to a time frame either in the question or in the response boxes. Actual boxes do not need to be shown Note: This is discrete data so boxes such As 0 – 2, 3 – 5, 6+ are acceptable if there is a time frame in the question. They cover all options and include 6 and above.

5ST1H_01					
Question		**Working**	**Answer**	**Mark**	**Notes**
7	(a)(i)	60,78,78,79,82,84,86,88,89, 91,94,96,97,99,100	88	2	M1 attempt at ordering – allow omissions A1 88 only Note: 1. Look at crossed out work as students often cross out to find the median. 2. An attempt at ordering can be ranking the numbers in the list.
	(ii)		Lower 79 Upper 96	2	B1 for 79 B1 for 96 Must be in correct place on answer line
	(b)			3	M1 box plot – a box with 2 whiskers.(ruled lines not required) ignore outliers. A1 ft for their three quartiles correctly placed A1 All correct. (no ft)

5ST1H_01				
Question	Working	Answer	Mark	Notes
*(c)		o The **median** for sale time is greater (than the median for non sale time)	4	B1
		o The **IQR/Range** for the sale time is less (than the **IQR/Range** for non sale time).		B1
		o Both have negative **skew.** OR No sale has **negative skew** whereas New year sale is **symmetrical**.		B1 Converses and equivalents accepted. Correct statistical words are needed. e.g. do not accept medium/spread/average. Median, IQR/Range/Skew need to be obviously correct. e.g. Do not accept medium but do accept mediun, skuw, etc.
		The new year sale increased sales (or the number of shoes sold)		B1

5ST1H_01				
Question	**Working**	**Answer**	**Mark**	**Notes**
8 (a)		Any one of o It is biased o Only people with land lines will be included. o Not everyone has equal chance of being asked o The sample is not big enough o Only 10 towns are used PLUS It is not satisfactory/No	2	B1 Note: Converses are acceptable. Simple reference to car performance B0 People not answering phone B0 B1 dependent on the first B.
(b)		Number all the names (from 0 to 9999 accept 1 to 10000) Use random number generator/tables to generate 100 numbers The sample will be the (100) people that corresponds to the (100) numbers	3	B1 for numbering – figures not required. B1 for generating 100 random numbers Note: Do not accept picking numbers out of a hat. B1 for explaining correspondence

5ST1H_01

Question	Working	Answer	Mark	Notes
*(c)		Any 3 different well written comments that involve the consideration of time, cost, truthfulness and response rate. e.g. Not everyone will fill in the questionnaire. o Questionnaire could be less embarrassing o You are likely to get a response on face to face interview o Face to face offers the opportunity to explain questions/see peoples reactions. o Face to face is time consuming. o Using a questionnaire will be faster than face to face interviews. o People might not speak the truth o Face to face could be more expensive than a questionnaire. o The costs will be different.	3	B1 B1 B1 Other satisfactory comments are possible Make sure that the same statement written twice in different format does not get credit twice. This is a QWC question so statistical expressions should be used if appropriate..

5ST1H_01				
Question	Working	Answer	Mark	Notes
9 (a)	$64/2 = 32$ th value	$0.2 < h \le 0.5$	2	M1 for seeing 64/2 or 65/2 or 32 or 32.5 A1 cao
(b)	$251.9/64 = 3.9$		4	M1 For attempt at finding $\sum fx$ where all x are in the given interval
		251.9		A1 for 251.9 M1 for attempt at a suitable fraction with denominator of 64 **dependent on getting first M mark** A1 cao (question asks for 1 dp)
		3.9		Note: Using frequency density gets M0A0M0A0 NB: if you see an awrt 31.49 this implies the first M1 A1 NB: if you see an awrt 31.5 this implies the first M1.
10 (a)		There are less getting married in 2005 OR It has decreased.	2	B1 Note: Must talk about totals and not refer to individual groups.
		The **area/radius/diameter** of the circle is less for 2005 than 1991 OR The pie chart is smaller		B1 **dependent on getting first B mark** Note: 1. It is smaller in 2005 B0 2. If converses used then dates needed.
(b)		20 – 24	1	B1
(c)		The **area** of the sector in 2005 is greater than the area for the sector in 1991	2	B1 Note: 'because the angle is bigger' is <u>incorrect</u> and get B0
		The **number** has increased.		B1 **dependent on getting first B mark** Equivalent expressions accepted but do not allow % or fraction or proportion

5ST1H_01					
Question	**Working**	**Answer**	**Mark**	**Notes**	
11 (a)	$(31.8+44.2+41.8+28.6)\big/4$ $(44.2+41.8+28.6+31.4)\big/4$	36.6 (thousands) 36.5 (thousands)	3	M1 for showing four numbers, at least one set correct, added and divided by 4. This may be implied by 1 correct answer. A1 36.6 and 36.5 cao (but can be in either space) B1 For correctly plotting points ft (tolerance half a small square)	
(b)		Ruled Line passing between 32.5 and 34.5 at q2 2005 and 35 and 37 at q4 2007 Must cover the all moving averages plotted.	1	B1	
(c)		**Positive** or **rising** or **upwards** or **increasing** As time/years pass more motor cycles are registered or ft	2	B1 ft SC Accept 'level' if line is horizontal B1 Both marks could be gained in the same sentence.	
(d)		2	1	B1 cao Do NOT allow reference to a specific year.	

Question	Working	Answer	Mark	Notes
(e)	$\frac{10+7+8}{3}=8\frac{1}{3}$	$8\frac{1}{3}$ or 8333	2	M1 for 3 correct numbers from their graph ft, added and divided by 3. A1 Awrt 7, 8 or 9 OR awrt 7000,8000,9000 **Note** awrt to 7, 8 or 9 OR awrt 7000,8000,9000 gets M1 A1
(f)	37 + 'their e'	45 or 45000	2	M1 for 37 from graph (accept 36 to 38) A1 ft Note: 1. Sight of 31.4 gets M0A0 – it gets an answer in range from incorrect working. 2. This question needs to have working shown ie correct answer no working M0A0
12 (a)	1 2 3 4 5 6 7 8 9 10 11 3 7 2 6 10 9 11 1 5 4 8 OR reverse ranks 11,10,9,8,7,6,5,4,3,2,1 9,5,10,6,2,3,1,11,7,8,4 $\sum d^2 = 194$ 1- $\frac{6\times194}{11(121-1)}$ 1 - $\frac{1164}{1320}=1-0.882$	0.118	3	M1 for finding correct ranks. M1 for 1- $\frac{6\times194}{11(121-1)}$ ft for d^2 A1 for awrt 0.12
(b)		There is no /positive correlation. RPI and Mortgage rate are not associated (if 'no' used) As mortgage rate goes up so does RPI (if positive used)	2	B1ft (only if values in a between −1 and +1) B1ft (only if values in a between −1 and +1) Equivalent expressions accepted

5ST1H_01				
Question	**Working**	**Answer**	**Mark**	**Notes**
13 (a)		Freq density 0.4 and 0.2	3	M1 for attempt at a calculation of freq density = $\frac{freq}{width}$ (formula only not enough). May be implied by a bar of correct height. A1 0.4 and 0.2 (can be implied from graph) A1 for correct histogram – tolerance ½ square.
(b)	$5 \times 1.2 = 6$ (or ½ x 12 = 6) $5 \times$ '0.2' = 1 (or ¼ x4 = 1) 6+1+8 Alternative method 20 small squares = 1 King/Queen '300'/20	15	3	M1 for seeing 5 x 1.2 or 6. M1 for seeing 5 x their 0.2 or 1 Other correct methods accepted. A1 cao Alternative method mark scheme M1 M1 their 300 A1 cao

Question	Working	Answer	Mark	Notes
14 (a)		0.9, 0.2, 0.05, 0.95 in correct places 0.02, 0.045, and 0.855 in correct places	2	B1 B1ft All must be in the correct places and must be probabilities ie between 0 and 1 If 0.8 and 0.2 used in both places they will get 0.02, 0.72, 0.18 in final column – this would get B0 B1 on ft. In this case we require the exact answers.
(b)	$\frac{0.08}{0.08+'0.045'}$	0.64 or 64%	2	M1 ft Note: ft from their 0.9 x 0.05 (16/25) or their 0.045 A1 cao Note: 0.8 x 0.8 = 0.64 M0A0 but correct answer no working gets both marks.
(c)(i)		Binomial Distribution	1	B1 Reasonable incorrect spelling accepted.
(ii)		Any Two Of Independent trials Fixed number of people Probability remains constant Only two outcomes OR Has allergy or does not have allergy	2	B1 B1
(iii)	$P = 0.1$ $q = 0.9$ Term $10p^3q^2$ or $10p^2q^3$ $10 \times 0.1^3 \times 0.9^2$ or $^5C_3\, 0.1^3\, 0.9^2$	0.0081 or 8.1 x 10^{-3} or equivalent	3	M1 for **using** $10p^3q^2$ or $10p^2q^3$ where p+q =1 This may be part of a larger expression. M1 Need to see the correct term clearly shown **on its own** with p+q=1 A1 cao

5ST1H_01					
Question		**Working**	**Answer**	**Mark**	**Notes**
15	(a)		The Paper has been a success so far. The percentage smoking in 2007 is already below the target or The trend/line is down	2	B1 Any equivalent B1 The percentage/amount/number of smokers has decreased is also accepted
	(b)		There is a lower percentage of women smokers (than men) OR Each year the women are below the men.	2	B1 Note: Accept converses. Accept there are fewer woman smokers (than men).
			Both men and women have decreased OR Men decreased more than women OR Men have dropped by 8% and women by 5/6% OR Men have decreased from 30 to 22 and women have decreased from 25.5 to 19.5 OR The gap between the men and woman has decreased.		B1 (if figures used they should be correct) .