

## Mark Scheme

Q1.

Question	Answer	Additional guidance	Mark
(a)	B1 e.g. Students who score highly in GCSE Mathematics also score highly in A-level Mathematics	B1 for a suitable hypothesis regarding relative performance in the two exams	(1)
(b)	B1 GCSE is sat first It is plotted on the x-axis	B1 for an acceptable reason. Allow equivalent wording. Condone 'horizontal' axis.	(1)
(c)	B2 ft The scatter graph shows <u>positive correlation</u> ... which supports the hypothesis	B2 for a correct conclusion for their hypothesis (ft) and mention of positive correlation. (Otherwise B1 for identifying positive correlation)	(2)
(d)	B1 straight line with correct gradient B1 straight line through (578, 78)	1 <sup>st</sup> B1 accept $0.6 < \text{gradient} < 0.8$ (not inclusive) If line does not extend at least from $x = 540$ to $x = 600$ then score max B1B0	(2)
(e)	B1 ft e.g. for every extra mark at GCSE an extra 0.7 is scored for A-level	B1 ft for correct equivalent interpretation. Allow ft from their line. (e.g. 7 extra A-level marks for every extra 10 GCSE marks)	(1)
(f)	B1 (Will not be reliable because) 540 is outside the range of data / it is extrapolation	B1 for assessing the appropriateness of the method	(1)

Q2.

Question	Answer	Additional guidance	Mark
(a)	B1 eg Data is bivariate/paired OR so she can see if there is correlation (between her variables)	B1 for justifying appropriateness of a scatter diagram.	(1)
(b)	B2 Line of best fit drawn through (25.2, 72), Otherwise B1 For plotting (25.2, 72) OR a sensible line of best fit which does not go through double mean point	B2 for an appropriate line which recognises that it should be drawn through the double mean point. B1 only for a partially correct answer	(2)
(c)	B1 Correct comment related to the graph (points 1 to 4) B1 Correct comment related to sample (points 5 to 8) B1 B1 B1 for any three further correct comments, eg 1. Points are close to a line (so strong correlation) OR some points not close to a line (so correlation not strong) 2. Line has positive gradient / points increasing left to right (so positive correlation) 3. Life expectancy is higher when age of mother is higher 4. Life expectancy increases by just over 2 years (eg 2.3) as age of mother increases by 1 year 5. Small sample of data / only valid for these 12 countries 6. Sample is not random / may not be representative 7. Source given is likely to be reliable OR secondary data may not be reliable	To gain all 5 marks at least one correct comment must be made interpreting the graph (equivalent to points 1 to 4) AND assessing the validity of conclusions based on the sample (equivalent to points 5 to 8) Accept equivalent comments for each example (1 to 11) 1. Assessing <i>strength</i> of correlation 2. Justifies <i>positive</i> correlation 3. <i>Interpreting</i> positive correlation 4. Interpreting gradient (accept 1sf from their line) 5. Assessing validity based on small sample 6. Recognising the sample may not be valid as not random 7. Considering the reliability of the source	(5)

8. Data may be out of date	8. Recognising that secondary data may not be up to date
9. Statement A is (OR is not) appropriate*	9. Assessing the appropriateness of statement A (*dependent upon a correct supporting reason)
10. Statement B is not appropriate *	10. Concluding that statement B is not valid (*dependent upon a correct supporting reason)
11. Correlation does not imply causation	11. Recognising that causation is not implied by correlation
	Ignore excess comments if not contradictory

Q3.

Question	Scheme	Marks
(a)	Quantitative Bivariate	B1 B1 (2)
*(b)	(Except for one point) there is <b>positive correlation</b> (between male and female earnings) Two further comments from: <ul style="list-style-type: none"> <li>○ Countries with higher male earnings also have higher female earnings</li> <li>○ Male earnings are higher (than female earnings for all these countries)</li> <li>○ One country / Ireland does not fit the pattern.</li> </ul>	B1  B2 (3) [5]
<b>Notes</b>		
(a)	B1B1 for two correct words clearly identified in some way. (Any <u>extra</u> word identified cancels out a correct answer)	
(b)	QWC 1 <sup>st</sup> B1: require 'positive correlation'  B2: Two clearly expressed comments based on the three options. (Allow only one from each.) Allow equivalent / converse comments. Ignore any figures.  Otherwise award B1 for one acceptable comment.	

Q4.

Question	Scheme	Marks
(a)	Points plotted at (2,3), (2.5,2.3)	B1B1 (2)
(b)	Negative (correlation) Interpretation: As the distance (from the source of the river) increases, the width (of the stone) decreases	B1 B1 (2)
(c)	Single straight line drawn from at least $x=0.5$ to at least $x=1.7$ which if extended would pass between: (0.5,5) and (0.5, 5.6) and (2.0, 2.7) and (2.0, 3.3)	B1 (1)
(d)	Answer in range [3.8,4.2]	B1ft (1)
(e)	Extrapolation/3.2km is beyond range of data set/Trend may not continue	B1 (1)
<b>Notes</b>		<b>[7]</b>
(a)	$\frac{1}{2}$ small square tolerance	
(b)	B1 negative (negative skew is B0) B1 must be in context and mention 'distance' oe and 'width' oe	
(d)	B1 anything in the range [3.8,4.2] or if answer is not in range, then fit from value read off their 'line' of best fit with negative gradient and $\frac{1}{2}$ small square tolerance	
(e)	B1 for an answer which states that the distance is beyond/outside the data set. 'Far away from other points' is B0 'The line doesn't extend that far' is B0	

Q5.

<b>5ST1F_01 Mark Scheme</b>		<b>Marks</b>
<b>(a)</b>	Point (30,95) uniquely identified.	B1 (1)
<b>(b)</b>	Sensible attempt at <u>ruled</u> line of best fit, neither above nor below all points.	) B1 (1)
<b>(c)</b>	Positive The greater the (body) length the greater the (brain) weight, oe	) B1 (2)
<b>(d)(i)</b>	Line on graph from 120 to their lobf Answer in Range 52 to 64 (answer in range scores both marks)	) B1 (2)
<b>(d)(ii)</b>	Reliable/yes AND interpolation/in range of data (oe), or <u>strong</u> correlation	M1 A1 (2)
<b>(e)(i)</b>	Line on graph from 180 to their lobf Answer in Range 80 to 92 (answer in range scores both marks)	) B1 (1)
<b>(e)(ii)</b>	Unreliable/no Extrapolation or outside range of <u>data</u> or mammal may not be the same as the type used for the graph oe.	) M1 A1 (1)

		) B1 B1 )	(2)  (2)
<b>[11]</b>			
<b>Notes</b>			
<b>(b)</b>	Line should be between tramlines on overlay & extend at least as far as indicated.		
<b>(c)</b>	Converse comments are fine.		
<b>(d)(i)</b>	If answer not in range and no line seen from 120, then award M1A0 if their answer follows correctly from their lobf (½ square tolerance)		
<b>(ii)</b>	Must have both conclusion and reason. Reason should relate to interpolation or to correlation being strong. (Do not accept eg 'inside range of <u>graph</u> ')		
<b>(e)(i)</b>	If answer not in range and no line seen from 180, then award M1A0 if their answer follows correctly from their lobf (½ square tolerance)		
<b>(ii)</b>	Reason should relate to extrapolation, or mammal being of different type. Do NOT accept eg 'outside range of <u>graph</u> ', 'no points at/near 180'.		

Q6.

Question	Answer	Additional guidance	Mark
(a)	B1 e.g. countries with higher percentage of <b>urban population</b> have higher <b>life expectancy</b>	B1 for an appropriate statement linking urban population (oe) and life expectancy (oe). Condone e.g. 'people living in urban areas live longer' for this mark. A question scores B0.	(1)
(b)	B1 Either: Data is <u>paired</u> / <u>bivariate</u> Or: Scatter diagram will show any <u>correlation</u>	B1 for an appropriate justification for a scatter diagram. Underlined words are needed.	(1)

(c)	<p>B1 <b>Urban population(%)</b> is the explanatory variable...</p> <p>depB1 ...because Irina believes this affects life expectancy, or ...because life expectancy is determined by this, or ...because life expectancy is the response variable</p>	<p>B1 for identifying the explanatory variable. B1 for correct reasoning Accept equivalent comments if meaning is clear</p>	(2)
(d)	<p>B2 e.g. <u>positive correlation</u>, so hypothesis is supported</p>	<p>B2ft for correct conclusion with reference to positive correlation and consistent with their hypothesis in (a).</p> <p>If B0 scored in part (a), then max score is B1</p> <p>(B1ft for an incomplete answer, e.g. missing vocab.</p> <p>OR correct reasoning with incorrect/no conclusion)</p>	(2)

(e)	<p>B2 Line of best fit drawn through (63.9, 77.8)</p>	<p>B2 for appropriate line of best fit with positive gradient through given mean point. (Mean point need not be plotted).</p> <p>(B1 for plotting the double mean point or for an appropriate line with positive gradient not through double mean point)</p>	(2)
(f)	<p>B1 e.g. for every extra 1% in urban <b>population</b>, life expectancy increases by 0.19 years (accept appropriate use of their gradient if found)</p>	<p>B1 for a complete equivalent numerical interpretation of gradient including population (oe) and years (oe)</p>	(1)
(g)	<p>B1 For any one from</p> <ul style="list-style-type: none"> <li>• lower <b>life expectancy</b> value than expected</li> <li>• (65, 63) plotted correctly</li> <li>• life expectancy for 65% should be around 77/78(years)</li> <li>• life expectancy read off their line of best fit</li> </ul> <p>B1 Anomaly / does not fit with other data</p>	<p>B1 for correct reasoning</p> <p>B1 for correct conclusion</p> <p>Condone 'outlier'</p>	(2)

Question	Scheme	Marks
(a)*	Positive correlation. The higher the price the more pages printed.	B1 B1 (2)
(b)	Line of best fit through mean point.	B1 (1)
(c)(i)	Points plotted (18, 200) and (25, 680)	B1, B1
(ii)*	Cartridge B (is better value) as it is above the line. OR Cartridge A is worse value as it is below the line.	B2ft  (4) [7]
Notes		
(a)	QWC 1 <sup>st</sup> B1 Need both <b>positive</b> and <b>correlation</b> 2 <sup>nd</sup> B1 for context interpretation equivalent to more expensive cartridge giving more pages. (Accept converse.)	
(b)	Appropriate <u>straight</u> line through mean point (it must at least cut the circle) extending horizontally at least from £15 to £35	
(c)(i)	B1 for each point (condone mislabelling or no labels)	
(ii)	QWC B2ft for a correct conclusion using correct interpretation of scatter diagram. (Must state B is <b>above</b> line OR A is <b>below</b> line – allow ft from their points.) Condone 'B is above the line' on its own for B2  If B2 not scored then: B1ft for a partially correct answer. e.g. B with <u>any</u> reason or A is expensive but gives few pages. or A is 9p/page (or 11 pages/£) <b>and</b> B is 4p/page (or 27 pages/£) o.e. (accept 1sf for these)  Note: Conclusion with no reason scores B0.	

Q8.

Question number	Answer	Additional guidance	Mark
(a)	B1 Annual profit is the response variable because... <ul style="list-style-type: none"> <li>• It depends on the distance to the car park</li> <li>• It is plotted on the y-axis</li> </ul>	B1 for a correct reason. Allow equivalent wording. Condone 'vertical' axis.	(1)
(b)	B2 Mike is correct as the scatter graph shows negative correlation	B2 for a correct conclusion and mention of negative correlation. Allow a description of negative correlation provided it is does not simply restate the question (B1 for correct conclusion with attempt at reason)	(2)
(c)(i)	B1 Straight line with intercept 40 000 B1 Straight line through (325, 27 000)		(2)
(c)(ii)	B1 (£40000 is) the profit when the distance (from the car park) is 0 (metres)	B1 for correct interpretation of intercept 40 000 in context	(1)
(d)	B2 Restaurant A/250m estimate is more reliable because... <ul style="list-style-type: none"> <li>• involves interpolation</li> <li>• 250 is inside the range of data</li> </ul>	B2 for conclusion that A is more reliable (or B less reliable) with correct reasoning. (B1 for correct conclusion with attempt at reason)	(2)
(e)	B1 Conclusion is not valid because correlation does not imply causation	B1 for correct conclusion of not valid with a correct supporting reason	(1)
(f)	B1 64(%)	Allow awrt 64% Do not all – 64%	(1)

Q9.



Question	Scheme	Marks
(a)	Point plotted at (62, 118)	B1 (1)
(b)	The price of the house is dependent upon the size of a house.	B1 (1)
(c)	Positive As the size of a house increases, the price increases.	B1 B1 (2)
(d)(i)	[142–144]	B1
(d)(ii)	Interpolation/80 is within the data set/80 lies between 62 to 79 and 81 to 103/It has been estimated from a size within the data set.	B1 (2)
(e)	Any one from, e.g. : <ul style="list-style-type: none"> <li>Extrapolation/Estimating using a size outside the range of data (62 to 103)</li> <li>Estimating a house price from a different city</li> </ul>	B1 (1)
<b>Notes</b>		<b>[7]</b>
(a)	Allow $\frac{1}{2}$ small square tolerance	
(b)	Allow 'size is the independent variable' or 'size of the house is fixed' Condone e.g. 'the larger the house, the more expensive it is' must be in that order	
(c)	For 2 <sup>nd</sup> B1, must be in context. May be in either order. Allow e.g. 'as house size decreases, price decreases' 'Larger houses are more expensive' is B0.	
(d)(i)	Also allow answers in the range 142000 to 144000	
(d)(ii)	Allow 'because the correlation is strong' or 'all of the points lie close to the line best fit' for B1 'As it is on the line of best fit' is B0	
(e)	Must be an example. Answers which only refer to the houses in the data set score B0.	

Q10.

Question	Scheme	Marks
(a)	Scatter (diagram)	B1 (1)
(b)	2.8 4.3	B1 B1 (2)
(c)	Circle drawn around (1.6, 1.6)	B1 (1)
*(d)	There is no/weak correlation (or the points do not lie close to a line/linear pattern), so it is not a good decision to draw a line of best fit.	B2 (2)
<b>Notes</b>		<b>[6]</b>
(d)	B2 NOT a good a decision and correct supporting reason which describes lack of (linear) correlation in data. (B1 for a correct description of the points on the scatter diagram (e.g. points are scattered) with no/incomplete conclusion)	

