Section A: Statistics

Qu	Scheme	Marks	AO
1 (a)	Positive (correlation)	B1 (1)	1.2
(b)	Every extra point gives £4.5(0) more on pay (o.e.)	B1 (1)	3.4
(c)	e.g. For points < 11 it would give pay < 0 which is ridiculous	B1 (1)	2.4
		(3 marks)	
	Notes		

(a)	B1 for "positive".
	    Allow an interpretation e.g. "as points increase pay increases" is B1
	    Read whole answer: contradictory comments such as "positive correlation, as points increase pay decreases" scores B0		

(b)	B1 for any correct comment conveying idea of <u>£s per point</u> and including a correct value; must have idea of <u>rate</u>. Can condone missing £ sign. Accept 4.5 e.g. "every 10 points earns an <u>extra</u> (or increase) of £45" is B1 BUT "every point earns £4.5(0)" is B0 *doesn't have idea of rate*		

(c)	B1 for a suitable comment mentioning "points" or "pay" (o.e. e.g. "amount")
	    <u>or</u> commenting on "small sample" or "range of points" used to find line
	 The following examples would score B1
	    Can say that *n* <u>points</u> (for *n* < 10.$\dot{4}$) would give <u>negative pay</u> so not suitable
	    Any comment suggesting that some jobs would end up with <u>negative pay</u>
	    Don't know the <u>range of points</u> used to find the <u>regression line</u>
	    A <u>small sample</u> of <u>size</u> 8 may not be <u>representative</u> to cover all jobs
	
	 B0 for a focus on "qualifications" or "hours" worked only
	 The following examples would score B0
	    Some jobs require no (or low) skills or qualifications (*need negative pay*)		

Question	Scheme	Marks	AOs
1(a)	Label **each** year group	B1	1.1b
	Use <u>random</u> numbers to select a ...	B1	1.1b
	Simple random sample of <u>24 Year 12s</u> and <u>16 Year 13s.</u>	B1	1.1b
		(3)	
(b)	<u>Increase</u> by <u>2.8</u> marks	B1	3.4
		(1)	
(c)	e.g. 'the best performance is predicted for the students who never wake up'	B1	3.5b
		(1)	
			(5 marks)
	Notes		
(a)	**B1:** for a suitable numbered/labelled/ordered(o.e.) list/database/register(o.e.) for **each** year group. Condone poor numbering but if just one list, then the Year 12s must be distinguishable from the Year 13s		
	B1: for use of random numbers/sample/selection to choose students		
	B1: for <u>24 Year 12s</u>, and <u>16 Year 13s</u>		
Note:	A description of a systematic sample: only allow access to the first mark and therefore may score maximum B1B0B0		
(b)	**B1:** Using the gradient of the regression equation must include <u>increase</u>(o.e.) and <u>2.8</u> 'Increase by approximately 3 marks' is B0 but isw if 2.8 is seen $5.6 \div 2$ is not sufficient		

(c)	**B1:** for any suitable limitation of the model
	 e.g. the idea that the longer you sleep the better performance in the test
	     or only valid between 0 and 24 hours (within range of the data)
	     or only applicable to the amount of sleep the night before the test
	     or only takes sleep into consideration/does not include other variables (factors)
	     or cannot score below 26.1 marks on the test
	     or the model might not be linear over the entire range
	     or the model might predict more than the maximum mark
	
	 B0: e.g. might not be correlation between s and p
	     or individual student performance may vary		

Question	Scheme	Marks	AOs
2(a)	0 to 500 m	B1	1.2
		(1)	
(b)	$1100 + 1600 + 1.5 \times 1600$ [$= 5100$]	M1	2.1
	$5300 > 5100$ therefore outlier	A1	1.1b
		(2)	
(c)	As the humidity increases the mean visibility decreases	B1	2.4
		(1)	
(d)	(Hours of) sunshine	B1	2.2b
		(1)	
		(5 marks)	

		Notes
(a)	B1:	For realising it is the maximum distance and distance given with correct units. Allow 0 to 50dm or < 500m or < 50dm
(b)	M1:	Attempt to find Q_3 and the upper limit
	A1:	5100, if a value for the point is stated it must be above 5100 otherwise it is A0. For a statement comparing and conclusion it is an outlier or it is above $Q_3 + 1.5$IQR. Allow accept the point circled is greater than 5100 oe
(c)	B1:	For a suitable interpretation of a negative correlation mentioning humidity and visibility
(d)	B1:	A correct deduction that the unlabelled variable is the hours of sunshine. Condone missing hours. Do not allow if more than one variable given. Must be quantitative variable Not cloud cover since values bigger than 8 Not wind speed since values not integers Not daily mean temperature since mean temperature near to zero are unlikely in June

Question	Scheme	Marks	AOs
4(a)	IQR $= 2.3$ and $20.6 \gg 2.4 + 1.5 \times 2.3$ $(= 5.85)$ (Compare correct values)	B1	1.1b
		(1)	
(b)(i)	e.g. It is a piece of data and we should consider all the data o.e.	B1	2.4
(ii)	e.g. It is an extreme value and could unduly influence the analysis **or** It could be a mistake	B1	2.4
		(2)	
(c)	e.g. "as humidity increases rainfall increases"	B1	2.2b
		(1)	
(d)	e.g. a 10% increase in humidity gives rise to a 1.5 mm increase in rainfall **or** represents 0.15mm of rainfall per percentage of humidity	B1	3.4
		(1)	
(e)(i)	Not a good method since only uses 11 days from one location in one month	B1	2.4
(ii)	e.g. She should use data from more of the UK locations and more of the months **or** using a spreadsheet or computer package she could use all of the available UK data	B1	2.4
		(2)	
		(7 marks)	