

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 st 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
	B1B1B1 for planning for each of 3 uses of technology Rebecca can use technology to Plan for Processing and Presenting - e.g. sort/order data columns by age/number of appointments - e.g. generate random numbers - e.g. remove extraneous data (phone number/address) columns - e.g. identify missing data - e.g. remove extraneous symbols - e.g. automate the calculation of summary statistics - e.g. automate the production of visual representations - e.g. remove outliers B1B1B1 for justifying why it is appropriate/advantageous Advantages of using technology - e.g. saves time - e.g. can reduce errors (in lengthy calculations) - e.g. facilitates use of all data (unbiased) - e.g. easier to correct a mistake - e.g. more visually appealing	B1B1B1 for each use of technology B1B1B1 for each advantage over processing data by hand Ignore extraneous non-contradictory comments	(6)

Q3.

Question number	Answer	Additional guidance	Mark
(a)	B1 e.g. Cost of computer storage reduces over time.	B1 for a suitable hypothesis including cost and time (B0 for a question)	(1)
(b)	M1 for either $\frac{24.87}{29.00} \times 100 (= 85.75 \dots)$ OR $\frac{24.33}{24.87} \times 100 (= 98.2 \dots)$ A1 for awrt 85.8 and awrt 98.2	M1 for correct calculation of chain-based index number. May be implied by one correct answer. A1 both correct	(2)
(c)(i)	M1 for $\sqrt[4]{93.8 \times 92.8 \times '85.8' \times '98.2'}$ A1ft awrt 92.5	M1 for correct calculation of the geometric mean of the four chain-based index numbers A1ft correct answer ft their (a)	(2)
(c)(ii)	B1ft for (average) 'decrease' B1ft ... (at a rate) of '7.5'% <u>per year</u>	B1ft for correct contextual interpretation which must follow through their answer to c(i) B1ft for complete correct contextual interpretation of their value for geometric mean.	(2)

Q4.

Question	Answer	Additional guidance	Mark
(a)	B1 e.g. countries with higher percentage of urban population have higher life expectancy	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)	B1 Urban population (%) is the explanatory variable... depB1 ...because Irina believes this affects life expectancy, or ...because life expectancy is determined by this, or ...because life expectancy is the response variable	B1 for identifying the explanatory variable. B1 (dependent on previous B1) for correct reasoning. Accept equivalent comments if meaning is clear	(2)
(d)	B2ft e.g. <u>positive correlation</u> , so hypothesis is supported	B2ft for correct conclusion with reference to positive correlation and consistent with their hypothesis in (a). If B0 scored in part (a), then max score is B1 (B1ft for an incomplete answer, either missing vocab. OR correct reasoning with incorrect/no conclusion)	(2)

(e)	B2 Line of best fit drawn through (63.9, 77.8)	B2 for appropriate line of best fit with positive gradient through given mean point. (Mean point need not be plotted). (B1 for plotting the double mean point or for an appropriate line with positive gradient not through double mean point)	(2)
(f)	B1 e.g. for every extra 1% in urban population , life expectancy increases by 0.19 years (accept appropriate use of their gradient if found)	B1 for a complete equivalent numerical interpretation of gradient including population (oe) and years (oe)	(1)

(g)	B1 For any one from • lower life expectancy value than expected • (65, 63) plotted correctly • life expectancy for 65% should be around 77/78 (years) • life expectancy read off their line of best fit at 65% B1 Anomaly / does not fit with other data	B1 for correct reasoning B1 for correct conclusion Condone 'outlier'	(2)
-----	---	--	-----

Q5.

Question number	Answer	Additional guidance	Mark
(a)	B1 Not appropriate as it is a question / is not a statement	B1 for recognising that a question is not a hypothesis	(1)
(b)	B1 e.g. Likely to be representative (with a small sample), or Age groups and genders will be fairly represented	B1 for recognising that stratifying will help give a representative sample	(1)
(c)	B1 e.g. makes it easy to analyse responses / put into graphs B1 B1 for any two problems from: - List is not exhaustive / too few options (e.g. no 'radio' / no 'other') - Students may use more than one option - Students may not follow the news - Accept: students' ages not recorded - Not investigating effect of gender so no need to split into male and female	1st B1 for recognising the advantage that responses will be easy to work with 2nd/3rd B1 for each of two appropriate problems identified	(3)

(d)	B1 The number of students at the university will be represented by the size of the pie charts	B1 for recognising that the 'size' of pie chart should represent the sample/ population size.	(1)
(e)	M1 $10 \times \sqrt{\frac{18}{12}}$ o.e. A1 12.2 (cm)	M1 for an appropriate method to find new diameter or radius (if labelled radius) or area (if labelled area) A1 for answer in the range 12.2 – 12.3 (Accept 12 following correct working)	(2)
(f)	B1 B1 for any two correct comments from e.g.: - Ensure sample size is not too small - Use similar criteria in selecting students (e.g. by gender/age) / same sampling method / similar proportions of ages and genders - Ask the question in the same way / ask the same question / give the same options / same method of data collection - Do at similar times - Ask in similar venues	B1 for each of two appropriate points (maximum 2) which would help improve reliability or validity	(2)

Q6.

Question number	Answer	Additional guidance	Mark
(a)	B1 Graph supports the hypothesis + reason B1 Scatter shows negative correlation	B1 for conclusion supported by sensible reason B1 for statistical reasoning using words in bold	(2)
(b)	M1 e.g. $83.5 - 0.7 \times 4 = 80.7$ or $83.5 - 0.7 \times 8 = 77.9$ or sensible straight line with correct gradient A1 correct straight line within tolerance	M1 for one pair of coordinates correctly identified, or for a sensible attempt at straight line with correct gradient. A1 Their line should extend horizontally at least from 4.2 to 7 and (if extended) be vertically within one square of 80.7 at $x = 4$ and one square of 77.9 at $x = 8$	(2)
(c)	B1 79.6 from graph	B1 for answer between 79.5 and 79.7 (may use equation OR vertical line drawn from $x = 5.6$)	(1)
(d)	B1 e.g. change in life expectancy as unemployment increases. B1 0.7 years fall in life expectancy (per 1% increase in unemployment)	B1 for recognising in context that gradient indicates a rate. Accept equivalent wording. No figures needed for 1 st B1 B1 for interpreting in context the <u>value</u> . Mark may be gained for correct equivalent figures used within their comment. (e.g. 1.4% more unemployment results in 1 year reduction in life expectancy)	(2)
(e)	B1 B1 for any two comments from • Involves extrapolation / 8% is outside of range • It is a for a different region so may be affected by other factors • Correlation does not look very strong	Accept equivalent statistical reasoning. B1 for each correct point but allow each bullet point once only.	(2)

Q7.

Question number	Answer	Additional guidance	Mark
	Collecting data B1 for identifying one appropriate thing that should be included in the plan for collecting data and B1 for explaining why this aspect is appropriate OR B1 for deciding what data to collect and/or how to collect and record it and B1 for an appropriate reason OR B1 for a strategy to process data and B1 for an appropriate reason OR B1 for designing a collection method for primary/secondary data and B1 for an appropriate reason OR B1 for appreciating the importance of acknowledging sources and B1 for an appropriate reason OR B1 for recognising where issues of sensitivity may influence data availability and B1 for an appropriate reason Processing and presenting B1 for planning to organise and/or process data and B1 for an appropriate reason OR	B1B1B1 for each of three planned elements and B1B1B1 for each of three appropriate reasons from their three things in the statistical enquiry cycle. Maximum 4 marks if only one aspect (from Collecting data, Processing and presenting, Interpreting and Evaluating) is referenced. B1 for e.g. use amount of time measured to the nearest minute and B1 for e.g. this is sufficient as there will be a large range of times B1 for e.g. collect data for 23 boys and 23 girls and B1 for e.g. this will make the calculation of quartiles easier B1 for e.g. use random sampling and B1 for e.g. this reduces bias as Gary's friends/class/peers may generally watch the same programs B1 for e.g. use primary data and B1 for e.g. this increases reliability as Gary will know how the data was collected B1 for e.g. A student (Gary) should collect the data and B1 for e.g. students are more likely to give an honest answer to a fellow student (less threatening) B1 for e.g. use a grouped frequencies table for the data and B1 for e.g. as this will enable a quick way of estimating the mean or this can be used to draw a histogram	(6)

Question number	Answer	Additional guidance	Mark
<i>continued</i>	B1 for planning to generate diagrams and/or visualisations to represent the data and B1 for an appropriate reason OR B1 for planning to generate statistical measures to compare data and B1 for an appropriate reason Interpreting B1 for planning to interpret diagrams and/or calculations/measures and B1 for an appropriate reason OR B1 for planning to make an inference and/or prediction and B1 for an appropriate reason Evaluating B1 for planning to identify weaknesses in approach or representation and B1 for an appropriate reason OR B1 for planning to refine the processes to elicit further clarification of the hypothesis and B1 for an appropriate reason	B1 for e.g. use box plots and B1 for e.g. these will enable the comparison of both the medians and the IQRs (i.e. the distributions) of the data B1 for e.g. interpret results for each individual school year and B1 for e.g. as different years could have different watching habits B1 for e.g. by comparing means and B1 for e.g. you can see whether the amount of time that boys spend watching TV is greater, in general, than the amount of time that girls spend watching TV B1 for e.g. use the results from the school to predict the results nationally and B1 for e.g. as students in different parts of the country are likely to have the same watching habits B1 for e.g. choose not to display the information in histograms and B1 for e.g. as "the target audience" may not know how to interpret them B1 for e.g. consider using more than one type of visual representation (for the same information) and B1 for e.g. as different representations focus on different aspects of the data	

Q8.

Question number	Answer	Additional guidance	Mark
	B1 for each of five correct comments eg - the hypothesis should be a statement not a question - pre-test a good idea as it can identify issues/problems with the reaction time test - people might do better on the caffeine test because they have done the test before – biased - need to give more detail on test protocol e.g. how much coffee to drink, how long to wait after drinking coffee before doing second test - need to plan to control extraneous variables e.g. what time of day - people may not record data in spreadsheet accurately / may lie about their time - median/interquartile range are suitable measures to compare (as data may be skewed/have outliers)/needed to draw a box plot - good idea to consider outliers/may not be appropriate to remove outliers as they could be genuine data - box plots are suitable to e.g. show the distribution of the times - need to draw comparative box plots and not just one for after the coffee has been drunk	B1 for each correct comment on the appropriateness of the hypothesis or plans for collecting, processing and presenting the data	(5)