Paper 2 Higher tier mark scheme

Question number	Answer	Additional guidance	Mark
1(a)	B1 for Transport	B1 cao	**(1)**
1(b)	B1 for upwards	B1 for upwards oe, eg increasing	**(1)**
1(c)	B1 for a correct answer, e.g. rounding error	B1 for a correct answer	**(1)**
2	B1 Top right of grid B1 Squares are shaded darkest in this region.	B1 for a correctly identifying the region (e.g. NE corner) (But simply listing individual squares is B0) B1 for a statistical reason relating to use of the key.	**(2)**
3(a)	B1 quantitative B1 continuous		**(1)** **(1)**
3(b)(i) **3(b)(ii)**	M1 16 44 A1 30 M1M1A1 $\frac{2\times10+4\times16+6\times44+8\times30}{100}=5.88$	M1 for correctly applying a scale to find a second value (implied either by labelling a scale or correctly finding one other frequency). A1 for all values correct. M1 for consistent use of fx with x within interval M1 for correct use of Σfx with x the mid-interval value A1 for 5.88	**(5)**
4(a)(i) **4(a)(ii)** **4(a)(iii)**	B1 (0.3 + 0.4=) 0.7 B1 0.3 M1 $\frac{0.3}{0.5}$ A1 = 0.6	For probability answers accept equivalent fractions, decimals or percentages	**(4)**
4(b)	M1 0.8×0.5 A1 = 0.4		**(2)**

Question number	Answer	Additional guidance	Mark
5	**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
5 ***continued***	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	

Question number	Answer	Additional guidance	Mark
6(a)	B1 Male median > female median OR Male IQR > female IQR OR Male range > female range B1 Male cats weigh more than female cats OR there is a greater variation in the weights of male cats compared with female cats	B1 for comparison of median or range/IQR values B1 for contextual interpretation of comparison of median or range/IQR values	**(2)**
6(b)	M1 for $\frac{3(47+1)}{4}$ $(= 36)$ accept 47×0.25 $(= 11.75)$ A1 =12	M1 for $\frac{3(47+1)}{4}$ $(= 36)$ A1 for 12	**(2)**
7(a)	B2 for Yes / good choice of diagram AND reference to the data being bivariate OR if B2 not earned B1 for Yes / good choice of diagram and reason that does not refer to the type of data OR B1 for referring to the data as bivariate but without commenting on whether the diagram is appropriate or not	B2 for complete assessment of the appropriateness of the diagram with a reason OR if B2 not earned B1 for an incomplete assessment of the appropriateness of the diagram	**(4)**
7(b)	B1 for Lata is not right – (the test results are correlated, but) the English test mark does not cause the Maths test mark oe	B1 for assessment of the given conclusion, including reference to correlation not implying causation	**(1)**
7(c)	B1 for correct interpretation in context e.g. • Correlation for Maths test scores and Science test scores is stronger than the correlation between Maths test score and English test score (or second correlation is stronger) oe • There is a greater association between Maths test score and Science test score than between Maths test score and English test score	B1 for statistical interpretation in context	**(1)**

Question number	Answer	Additional guidance	Mark
8(a)	M1 $(P_{90} - P_{10} =) 72 - 8$ A1 64	M1 for subtracting two percentiles identified from graph, at least one correct. May be seen on graph. A1 Accept [63~65] if working shown and P_{10} or P_{90} correct.	**(2)**
8(b)	B1 UK median is 38 (from graph) B1ft Manchester has a smaller 10 to 90 percentile range B1ft for either • Manchester population (9 years) younger on average • Manchester population has less variation of ages	B1 for stating UK median is 38 (or 9 years higher than Manchester). B1ft for correct statistical reasoning. (Follow through their answer to (a) for 2nd and 3rd B1 mark) B1ft for one contextual interpretation of average or spread comparisons. Accept equivalent/converse statements about UK population.	**(3)**
8(c)(i) **8(c)(ii)**	M1 (Skew =) $\frac{3(32-29)}{19.3}$ A1 = 0.4663... B1ft (Positive skew means that) the majority in the population are younger with fewer older people, OR there is a greater spread of ages at the upper end	M1 for demonstrating correct use of formula A1 for 0.47 or better B1ft allow any equivalent wording for a correct contextual interpretation of their skew value. (This mark may be gained independently following comparison of mean/median)	**(3)**

Question number	Answer	Additional guidance	Mark
9(a)	M1A1A1 for calculating standardised scores M1 Long jump: $\frac{6.43-6.10}{0.26}$, High jump: $\frac{1.86-1.79}{0.066}$ A1 Long jump: 1.3 A1 High jump: 1.1 B2 for e.g. Better performance in long jump (relative to their competitors) as there is a higher standardised score in long jump OR (if B2 not scored) B1 for e.g. Better performance in long jump (relative to their competitors) with an attempt at a reason	M1 for either correct calculation A1 for awrt 1.3 A1 for awrt 1.1 B2 for a correct contextual interpretation of results with a correct reason using standardised scores OR (if B2 not scored) B1 for a correct contextual interpretation of results with an attempt at a reason	**(5)**
9(b)	M1 $-0.32=\frac{42.51-x}{5.85}$ A1 (x =) 44.382	M1 for demonstrating correct use of formula. A1 for awrt 44.38	**(2)**
10(a)	B1 More people live in the Lake District B1 Larger pie chart. o.e.	B1 Accept converse contextual statements about Snowdonia. B1 for any statistical reasoning implying larger chart (may refer to radius/diameter/area/size, etc)	**(2)**
10(b)	B1 More (people aged 40-59) live in the Lake District + reason B1 Larger area.	B1 Accept converse contextual statements about Snowdonia. Answer without reason scores B0 B1 for statistical reasoning with reference to area clearly implied. (Accept complete clear argument based on proportions of different sized populations.)	**(2)**
10(c)	B1 for a correct answer, e.g. populations (of regions), angles (of sectors), numbers of people (in each sector)	B1 for a correct answer	**(1)**
10(d)	B1 for $3 \div \sqrt{3.9}$ (=1.519…)	B1 for a complete method to find radius of circle for Exmoor	**(1)**

Question number	Answer	Additional guidance	Mark
11(a)	B1 for LQ = 14, UQ = 34 M1 for 34 + 1.5 × ('34' – '14') (= 64) 70 is an outlier as 70 > 64	B1 for 14 or 34 M1 for 34 + 1.5 × ('34' – '14') with their quartiles A1 for demonstrating understanding of calculations for outliers	**(3)**
11(b)	B1 B1 B1 B1 B1 for each of five correct aspects e.g. • Use of statistical words, e.g. average/spread is too vague in this context (conclusion not appropriate) • Comparing medians (conclusion appropriate) • Comparing ranges or IQRs, e.g. IQRs are the same (conclusion appropriate) or Pine Wood range greater than Acorn Wood range (conclusion not appropriate) • Identifying Acorn Wood as having a negative skew (conclusion not appropriate) • Identifying Pine Wood as having no skew or is symmetrical (conclusion not appropriate)	B1 B1 B1 B1 B1 for each of five correct comments assessing the appropriateness of the conclusions	**(5)**

Question number	Answer	Additional guidance	Mark
12(a)	B1 lower warning line added at 4.0 and labelled. B1 both action lines correct and labelled (Upper line at 4.25, lower line at 3.95)	B1 for knowing warning lines should be symmetrical. B1 for knowing that action lines should be 3 × s.d. from the mean. Both lines correct with correct labels.	**(2)**
12(b)	B1ft Another sample should be taken immediately B1 If further sample is also beyond warning line then production should be stopped. (If within warning lines then production can continue.)	B1ft for attempting correct practical decision based on position of their value relative to warning/action lines. B1 (cao) for complete description of actions needed including criteria for stopping after a second sample.	**(2)**
12(c)	B1 David is incorrect because sample means will be more closely grouped than individual values OR standard deviation will be >0.05	B1 for demonstrating understanding that sample means will have a closer grouping.	**(1)**

Question number	Answer	Additional guidance	Mark
13(a)	M1 for correct answer, e.g. may be out of date, sources may be unreliable, etc.	B1 for a correct disadvantage of collecting data using technology	**(1)**
13(b)	B1 for $\frac{131334.5}{43}$ or $\frac{2361}{43}$ B1 for $\sqrt{\frac{131334.5}{43} - \left(\frac{2361}{43}\right)^2}$ or $\frac{131334.5}{43} - \left(\frac{2361}{43}\right)^2$ leading to correct answer	B1 for a correct use of 43 in a calculation B1 for showing a correct complete calculation	**(2)**
13(c)	e.g. M1 LB = mean – 6.29 (= 48.62) M1 UB = mean + 6.29 (= 61.2) M1 (mean – 52) × 2.6 + (52 – LB) × 2.2 (= 15.002) (57 – mean) × 2.6 + (61.2 – 57) × 2 (= 13.834) '15.002' + '13.834' (= 28.836) M1 28.836 ÷ 43 A1 = 0.671	M1 for correct method to find number of presidents 1 sd below mean M1 for correct method to find number of presidents 1 sd above mean M1 for correct method to find total number of presidents within 1 sd of mean M1 for correct method to find proportion of presidents within 1 sd of mean A1 for answers rounding to 0.67 or 67%	**(5)**
13(d)	B1 for e.g. yes, close to 68%	B1 for referring to 68%	**(1)**

Question number	Answer	Additional guidance	Mark
14	B2 for yes with a correct complete reason, e.g. $P(X \mid Y) = 0.4$ implies $P(Y) = 0.5$, and as $P(Y \mid X) = 0.5$, so X and Y are independent events OR if B2 not earned B1 for using $P(X \text{ and } Y) = P(X) \times P(Y)$ or $P(Y \mid X) = P(Y)$ or $P(X \mid Y) = P(X)$	B2 for a correct comment assessing the appropriateness of the conclusion OR if B2 not earned B1 for an incomplete attempt to assess the appropriateness of the conclusion	**(2)**
15	B1 B1 B1 for each of three correct aspects • e.g. diagrams are not appropriate for discrete data • e.g. amount of data is not appropriate as too small for the number of intervals used • e.g. it is not appropriate to compare the two diagrams as different class intervals / different frequency density scales are used	B1 B1 B1 for each of three correct comments assessing the appropriateness of the diagrams used	**(3)**