Mark Scheme (Results)

November 2021

Pearson Edexcel GCSE
In Statistics (1ST0) Higher Tier
Paper 1H

General Marking Guidance

- All candidates must receive the same treatment. Examiners must mark the first candidate in exactly the same way as they mark the last.
- Mark schemes should be applied positively. Candidates must be rewarded for what they have shown they can do rather than penalised for omissions.
- Examiners should mark according to the mark scheme not according to their perception of where the grade boundaries may lie.
- There is no ceiling on achievement. All marks on the mark scheme should be used appropriately.
- All the marks on the mark scheme are designed to be awarded. Examiners should always award full marks if deserved, i.e. if the answer matches the mark scheme. Examiners should also be prepared to award zero marks if the candidate's response is not worthy of credit according to the mark scheme.
- Where some judgement is required, mark schemes will provide the principles by which marks will be awarded and exemplification may be limited.
- When examiners are in doubt regarding the application of the mark scheme to a candidate's response, the team leader must be consulted.
- Crossed out work should be marked UNLESS the candidate has replaced it with an alternative response.

Question number	Answer	Additional guidance	Mark
1(a)	B1 for one of • Checks response rate • See if questions are understood • Makes sure questionnaire gets relevant answers • Identifies likely responses • Checks how long it will take		(1)
(b)	B1 e.g. 'not appropriate since a pilot test is small scale study'	B1 for not appropriate and correct supporting reason Condone reasons relating to time and cost.	(1)
(c)	B1 e.g. 'not a suitable suggestion since histograms require quantitative data'	B1 for not suitable and correct supporting reason (allow grouped data) Accept 'qualitative data is not suitable for a histogram'	(1)

Question number	Answer	Additional guidance	Mark
2(a)(i)	M1 Reading off the graph at 0.75×48 (36) A1 answer in the range 3200 to 3600	M1 for reading off graph at 75% A1 for answer in range Condone use of $n + 1$	(2)
(ii)	B1 e.g. '75% of counties have an area of '3400' sq km or less'	B1 for correct interpretation in context	(1)
(b)	M1 Reading a cumulative frequency off graph at 2000 M1 '19' + 24 (= 43) A1 answer in the range $4400 < k < 4800$	M1 may be implied by 19 identified. M1 for adding 24 to their value A1 for answer in range Note: working may be seen on or next to the graph	(3)

Question number	Answer	Additional guidance	Mark
3(a)(i)	B1 <u>all</u> of the students in John's school		**(1)**
(a)(ii)	B1 <u>all</u> of the (types of) films (in UK cinemas last year)	'All' is required, but condone omission in (a)(ii) if omitted in (a)(i).	**(1)**
(b)	B1 e.g. 'use a trusted website', 'use up-to-date / recent data'	B1 for a suitable suggestion Accept 'use reliable website / reliable source'	**(1)**
(c)	B1 Method A: e.g. 'each student may not have the same chance of being selected' B1 Method B: e.g. 'selecting at a particular time/place so not all students have an equal chance of being selected'	B1 for any suitable reason as to why this quota sampling method is not random B1 for any suitable reason as to why this opportunity sampling method is not random	**(2)**
(d)	B2 Method A/quota sampling should be less biased since it is more likely to be representative (OR if B2 not scored B1 Method A/quota sampling should be less biased with an attempt at a reason)	B2 for Method A and identifying that quota sample aims to represent the characteristics of the population (OR if B2 not scored B1 for Method A and an attempt at a supporting reason)	**(2)**

Question number	Answer	Additional guidance	Mark
4(a)	B1 2009 Quarter 3		(1)
(b)	M1 $\dfrac{456663}{444292} \times 100$ A1 103	M1 correct equivalent calculation A1 awrt 103	(2)
(c)	M1 444292×0.973 A1 (£) 432 296 (million)	M1 correct equivalent calculation A1 awrt (£)432 000 (allow 432 296.116, awrt 432 300)	(2)
(d)	M1 The index numbers are increasing / the GDP is increasing (in 2010) A1 therefore Marc is incorrect.	M1 for understanding that the index numbers are increasing / GDP is increasing A1 correct assessment of Marc's conclusion	(2)

Question number	Answer	Additional guidance	Mark
5(a)	B1 4		(1)
(b)	B1B1 • Lowest in the Quarter 1 (each year) • Highest in Quarter 2 (each year) • Downward from Q2 to Q4 (each year) • Quarters 2 & 3 are above the overall trend line (each year) / quarters 1 & 4 are below the overall trend line (each year)	B1 for each correct description of the seasonal trend For third bullet point accept 'downward from Q2 to Q1' Do not accept descriptions of overall trend e.g. 'there is an upwards trend'	(2)
(c)	B1 4 B1 since the data is quarterly/repeats every 4 quarters		(2)
(d)	B1 Not reliable due to extrapolation	B1 for not reliable and correct supporting reason	(1)

Question number	Answer	Additional guidance	Mark
6(a)	B1 events *A* and *C* B1 they do not intersect	B1 for identifying the two events B1 for correct supporting reason	(2)
(b)	B1 0.58	Allow equivalent fraction or percentage	(1)
(c)	M1 P(*A* or *C*) = P(*A*) + P(*C*) or (0.38 + 0.08) + (0.05 + 0.23) A1 0.74	M1 for correct expression A1 allow equivalent fraction or percentage	(2)
(d)	B1 0.08 and 0.38 placed correctly B1 0.2 and 0.34 placed correctly		(2)

Question number	Answer	Additional guidance	Mark
7(a)	B1 can be used to compare relative frequencies (areas) rather than just proportions	B1 for a suitable advantage of using comparative pie charts	(1)
(b)(i) (ii)	B1 2005 sector will have a bigger angle B1 2018 sector will have a bigger area	Allow converse statements	(2)
(c)	M1 $\dfrac{\sqrt{12.5}}{\sqrt{8.3}} \times 5$ A1 6.1 (cm)	M1 for a correct calculation A1 awrt 6.1	(2)

Question number	Answer	Additional guidance	Mark
8	M1 IQR = 8.75 – 7.25 (= 1.5) M1 8.75 + 1.5 × '1.5' (= 11) or 7.25 – 1.5 × '1.5' (= 5) A1 11.5 is the only outlier B1 A box with at least one whisker drawn B1ft 5.15, 7.25, 8.05, 8.75 and upper tail at 10 or 11 or 11.5 all correct B1 upper tail at 10 or 11 and single outlier plotted at 11.5	M1 correct expression for IQR M1 use of Q3 + 1.5 × IQR or Q1 – 1.5 × IQR A1 for identifying 11.5 as the only outlier May be seen on box plot Ft their outlier limits	(6)

Question number	Answer	Additional guidance	Mark
9	M1 $2\times72+3\times84+5\times88$ $(=836)$ M1 $90\times(2+3+5+8)\,(=1620)$ M1 $\dfrac{1620-836}{8}$ A1 Yes, it is possible (if he achieves 98 (or greater)).	M1 for using weighting with first 3 assignments M1 for attempt at total score required M1 for calculating score needed on final exam A1 for correct conclusion from correct supporting working Alternative: M1 $2\times72+3\times84+5\times88$ $(=836)$ M1 '836' + 100×8 $(=1636)$ M1 $\dfrac{836+100\times8}{(2+3+5+8)}$ $(=90.88\ldots)$ Alternative: M1 $72\times\frac{2}{18}+84\times\frac{3}{18}+88\times\frac{5}{18}\,(=46.44)$ M1 $x\times\frac{8}{18}=90-'46.44'$ M1 for correct method to solve for x $x=\frac{18}{8}\times(90-'46.44')$	(4)

Question number	Answer	Additional guidance	Mark
10(a)	<table><tr><td>**Time ranks**</td><td>***d*** **(difference in ranks)**</td></tr><tr><td>5</td><td>–4</td></tr><tr><td>2</td><td>0</td></tr><tr><td>7</td><td>–4</td></tr><tr><td>(1)</td><td>3</td></tr><tr><td>8</td><td>–3</td></tr><tr><td>6</td><td>0</td></tr><tr><td>4</td><td>3</td></tr><tr><td>3</td><td>5</td></tr></table> M1 M1 $\sum d^2 = 84$ A1 $r_s = 1 - \dfrac{6 \times 84}{8(8^2 - 1)} = 0$ B1ft No correlation, so… depB1ft Amelia's hypothesis is not supported	M1 at least 5 correct time ranks (may be implied by 2nd M1) Allow if one rank misplaced but then subsequent ranks in correct order. M1 attempt at calculating sum of d^2 for *their* ranks A1 $r_s = 0$ B1ft for no correlation ft their r_s provided $-1 \leq r_s \leq 1$ depB1ft not supported (dep on at least 1 previous M mark being scored and an attempt at identification of correlation) allow follow through their value of r_s	(5)
(b)	B1 e.g. 'collect more data', 'repeat 100 metre race more than once'	B1 for a suitable reason to improve the reliability of her results Condone 'collect primary data'	(1)

Question number	Answer	Additional guidance	Mark
11(a)	B1 recognising that the standard deviation is 2 M1 One warning line or action line A1ft both warning lines correct $\pm 2\,\sigma$ A1ft both action lines correct $\pm 3\,\sigma$ A1 correct scale	Allow ft on their identified standard deviation	(5)
(b)	B1 not appropriate since the machine should be stopped/reset the first time the action limit is exceeded	B1 for not appropriate and correct supporting reason Condone for 'first time' reference to immediately / straight away	(1)
(c)	B1 Means will be the same/similar B1 Standard deviation of individual values will be greater dB1 So conclusion not supported	B1 for assessment of means B1 for assessment of standard deviations dB1 for conclusion (dep on 2nd B1)	(3)

Question number	Answer	Additional guidance	Mark
12(a)	B1 Point circled at (15, 11000)	No other points circled	(1)
(b)	B1ft The value of this car is significantly higher than other cars (around the same age)	B1 correct interpretation in context	(1)
(c)(i) (ii)	B1 e.g. 'may be an error in the data', 'doesn't fit the trend' B1 e.g. 'includes all data', 'genuine value'	B1 for a suitable appropriate reason for not including the outlier B1 for a suitable appropriate reason for including the outlier	(2)
(d)	B2 closer to –1 (smaller/lower) (B1 will still be negative) B1 (since as age increases, value of car decreases but) not at a constant rate / linear pattern		(3)

Question number	Answer	Additional guidance	Mark
13(a)	B1 3.0	B1 allow 3	(1)
(b)	B1 55%	B1 allow 0.55	(1)
(c)	B1 distribution is symmetric		(1)
(d)	B1 distribution symmetric so $\dfrac{5.7+5.3}{2}=5.5$ B1 (95% of data should fall between) $5.5 \pm 2{\times}0.75$ (from 4 to 7) B1 97.5% $-$ 2.5% = middle 95% dB1 so claim is supported	B1 for using symmetry to identify the mean Allow this B mark for demonstration that two appropriately chosen percentiles are equidistant from 5.5 e.g. $6.1 - 5.5 = 0.6$ and $5.5 - 4.9 = 0.6$ B1 for use of mean +/- 2 s.d. Must show calculation using 2×0.75 not just 4 and 7 B1 for comparing boundaries with data dB1 for identifying claim is supported (dep on 1st & 2nd B1)	(4)