Mark Scheme (Results)

Summer 2022

Pearson Edexcel GCSE
In Statistics (1ST0) Foundation Tier
Paper 1F

Edexcel and BTEC Qualifications

Edexcel and BTEC qualifications are awarded by Pearson, the UK's largest awarding body. We provide a wide range of qualifications including academic, vocational, occupational and specific programmes for employers. For further information visit our qualifications websites at www.edexcel.com or www.btec.co.uk. Alternatively, you can get in touch with us using the details on our contact us page at www.edexcel.com/contactus.

Pearson: helping people progress, everywhere

Pearson aspires to be the world's leading learning company. Our aim is to help everyone progress in their lives through education. We believe in every kind of learning, for all kinds of people, wherever they are in the world. We've been involved in education for over 150 years, and by working across 70 countries, in 100 languages, we have built an international reputation for our commitment to high standards and raising achievement through innovation in education. Find out more about how we can help you and your students at: www.pearson.com/uk

Summer 2022
Question Paper Log Number 68619
Publications Code 1ST0_1F_2206_MS

General marking guidance

These notes offer general guidance, but the specific notes for examiners appertaining to individual questions take precedence.

1 All candidates must receive the same treatment. Examiners must mark the last candidate in exactly the same way as they mark the first.

Where some judgement is required, mark schemes will provide the principles by which marks will be awarded; exemplification/indicative content will not be exhaustive. When examiners are in doubt regarding the application of the mark scheme to a candidate's response, the response should be sent to review.

2 All the marks on the mark scheme are designed to be awarded; mark schemes should be applied positively. Examiners should also be prepared to award zero marks if the candidate's response is not worthy of credit according to the mark scheme. If there is a wrong answer (or no answer) indicated on the answer line always check the working in the body of the script (and on any diagrams), and award any marks appropriate from the mark scheme.

Questions where working is not required: In general, the correct answer should be given full marks.
Questions that specifically require working: In general, candidates who do not show working on this type of question will get no marks – full details will be given in the mark scheme for each individual question.

3 Crossed out work

This should be marked **unless** the candidate has replaced it with an alternative response.

4 Choice of method

If there is a choice of methods shown, mark the method that leads to the answer given on the answer line.

If no answer appears on the answer line then mark both methods **as far as they are identical** and award these marks.

5 Incorrect method

If it is clear from the working that the "correct" answer has been obtained from incorrect working, award 0 marks.

6 Follow through marks

Follow through marks which involve a single stage calculation can be awarded without working as you can check the answer, but if ambiguous do not award.

Follow through marks which involve more than one stage of calculation can only be awarded on sight of the relevant working, even if it appears obvious that there is only one way you could get the answer given.

7 Ignoring subsequent work

It is appropriate to ignore subsequent work when the additional work does not change the answer in a way that is inappropriate for the question or its context. (eg an incorrectly cancelled fraction when the unsimplified fraction would gain full marks).

It is not appropriate to ignore subsequent work when the additional work essentially makes the answer incorrect (eg incorrect algebraic simplification).

8 Probability

Probability answers must be given as a fraction, percentage or decimal. If a candidate gives a decimal equivalent to a probability, this should be written to at least 2 decimal places (unless tenths).

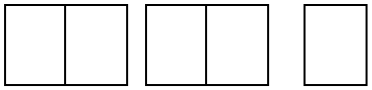
Incorrect notation should lose the accuracy marks, but be awarded any implied method marks.

If a probability fraction is given then cancelled incorrectly, ignore the incorrectly cancelled answer.

9 Range of answers

Unless otherwise stated, when an answer is given as a range (eg 3.5 – 4.2) then this is inclusive of the end points (eg 3.5, 4.2) and all numbers within the range.

Guidance on the use of abbreviations within this mark scheme	
M	method mark awarded for a correct method or partial method
A	accuracy mark (awarded after a correct method; if no method is seen then full marks for the question are implied but see individual mark schemes for more details)
B	unconditional accuracy mark (no method needed)
oe	or equivalent
cao	correct answer only
ft	follow through (when appropriate as per mark scheme)
sc	special case
dep	dependent (on a previous mark)
indep	independent
awrt	answer which rounds to
isw	ignore subsequent working

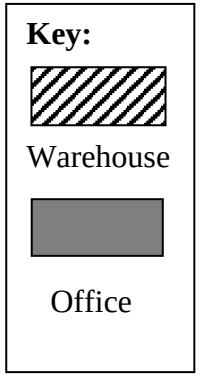
Question	Answer	Additional Guidance	Marks
1 (a)	B1 – for two and a half images drawn 		**(1)**
(b)	B2 – There were more cars sold in August because there were 35 cars sold in August and 20 cars sold in September.	Accept for example; There were 15 more cars sold in August than September. (B1 – There were more cars sold in August (than September) OR correct figures for August and September with no comparison OR one correct figure for either August or September with a correct comparison).	**(2)**
(c)	B1 – Numbers of cars can only be shown (accurately) in multiples of 5 cars. The response must include a reference to 5 cars per box or 10 cars per box.	Using the given key on 5 cars per box it will be difficult/can't/not possible to show 17 (or 7 or 2) cars. OR It is difficult to show 17 because each box has 5 cars.	**(1)**

Question	Answer	Additional Guidance	Marks
2 (a)	B1 – It is equally likely that the spinner will land on each face. OR – The probability that the spinner will land on one of the sides is the same for all sides.	Accept on **any** face. Accept use of the phrase 'equal chance' Do not accept 'even chance'	**(1)**
(b)	B1 - Unlikely – underlined or any other unambiguous indication.		**(1)**
(c)	B1 - Impossible – underlined or any other unambiguous indication.		**(1)**
(d)	B1 for ONE of the following • There is a disproportionate large number of 2's • There are fewer 5's than expected dB1 – for stating that the dice is not fair. OR B1 – for stating that this is a random process so a fair dice could produce the results given in the table. dB1 – for stating the dice might be/is fair.	Or any other reasonable explanations	**(2)**

Question	Answer	Additional Guidance	Marks
3 (a)	(i) B1 – Accept in the range 24% – 26% Inclusive (ii) B1 – Accept in the range 89% – 91% Inclusive		**(2)**
(b)	B1 for point plotted correctly at (2019, 93)	Allow tolerance of 92 – 94	**(1)**
(c)	B1 2005		**(1)**
(d)	B2 for the two **correct** reasons with a conclusion • The percentage of households with internet access from 2001 to 2005 increased from 35/36/37 % to 54/55/56 % OR there is an increase in the percentage of 18/19/20% • The percentage of households with internet access from 2015 to 2019 increased from 85/86/87 % to 92/93/94 % OR there is an increase in the percentage of 6/7/8% So the increase was greater from 2001 to 2005.	B1 for one reason only **with** a conclusion, OR B1 for only one correct reason **with** a conclusion OR B1 two reasons without a conclusion.	**(2)**

Question	Answer	Additional Guidance	Marks
4 (a)	B1 for - It is secondary data because it has been collected by another researcher/agency Or States or implies that Alexa did not collect the data herself.	Or similar wording. Accept any wording implying that Alexa was given the data/did not collect it herself. Ignore spurious references to reliability of the data.	**(1)**
(b)	B1 for (i) Advantage - Any one from • It is economical (less expensive to collect) • A large amount of data can be accessed easily • Historical data can been looked at easily • It is convenient and easy/quick to collect/access. B1 for (ii) Disadvantage – Any one from • The data **may** not be in the form needed. • Lack of control over the quality of the data • Data **could** be out of date.	Or any other reasonable advantage or disadvantage.	**(2)**

Question	Answer	Additional Guidance	Marks
5 (a)	B1 for – No mode because there is not more than one of any piece of data. OR No mode because all the numbers are different.		**(1)**
(b)	M1 for data ordered correctly 15 17 19 24 25 27 29 32 35 A1 for 25	25 without any working scores M1A1	**(2)**
(c)	B1 for 35 – 15		**(1)**
(d)	B1 Yes, it is more consistent, because the range is lower to change the tyres/higher to change the oil	Accept: Yes, because the range is lower.	**(1)**

Question	Answer	Additional Guidance	Marks	Type
6 (a)	B1 – for the correct labelling. From the origin **upwards** 10, 20, 30, 40, 50, (60)	**All** numbers from 10 through to 50 must be in place. Allow omission of 60.	**(1)**	1
(b)	B1 for the correct keys in the correct place. 		**(1)**	
(c)	B2 – for the correct two bars drawn and shaded. Height of black/solid bar is 25 (5 squares) Height of hatched bar is a further 10 [up to 35] (a further 2 squares, going up to 7 squares)	(B1 for the correct total height (7 squares) with or without shading). OR (B1 for either correct bar with shading provided an attempt has been made on the other).	**(2)**	
(d)	B1 for 80		**(1)**	3
(e)	M1 for $\frac{(35-20)}{125}$ A1 for $\frac{15}{125}$ or 0.12	Accept $\frac{k}{125}$ where $0 < k < 35$ Accept any equivalent fraction e.g. $\frac{3}{25}$	**(2)**	1

Question	Answer	Additional Guidance	Marks
7	**<u>Data Collection</u>** • B1 – for using a questionnaire • depB1 – because e.g., it is easy to send a questionnaire to a large number of people. **OR** • B1 – for using a tally chart/table. • depB1 – because e.g. a tally chart sorts the data into groups or (it is a simple way of recording and counting data) **OR** • B1 – for interviewing pupils to find out (how they use the internet for homework) • depB1 – because e.g. high response rate **<u>Choosing a sample</u>** • B1 - for using a (simple) random sample. • depB1 - because e.g., every student has an equal chance of being chosen **OR** • B1 – for using systematic sampling • depB1 – because e.g., it is easy to choose every *n*th person in a large population **OR** • B1 – For using Quota sampling • depB1 – because e.g., we can group the population by year group/gender and ask a number from each group **OR** • B1 – For using Cluster sampling • depB1 – because e.g .she can split the school into year groups (and choose children from each cluster) **OR** • B1 – For using Stratified sampling. • depB1 – because e.g., each student has an equal chance of being chosen within a strata.	Please be generous on spelling as long as the intention is clear with no ambiguity. B1 – A data collection method stated B1 – explain the key characteristic of why it is appropriate. B1 - A sampling method stated B1 – For why it is suitable with reference to the key characteristics of that method.	**(6)**

	<u>Diagrams</u> • B1 – For example, Bar Chart, Line graph, Pictogram • depB1 – because a bar chart/line diagram/pictogram shows the shape of the distribution or (it is easy to show numbers of students in each category) or (easy to make comparisons in categories) **OR** • B1 - Pie Chart • depB1 – because a pie chart shows proportions/percentages of how students use the internet	B1- A suitable diagram B1 – For stating a key characteristic of the diagram which makes it suitable. **Do not accept** – 'easy to read' as that is not a specific characteristic to one of these diagrams.	

Question	Answer	Additional Guidance	Marks
8	B2 Two reasons from • The sample size is too small • Not representative – only taking a sample from one area of the Councils operations • Not random • The method of transport might not be the same each day. • Asked only one day. • Only asked people who work mornings.	(B1 for one acceptable reason).	**(2)**

Question	Answer	Additional Guidance	Marks
9 (a)(i)	B2 Two reasons from: • Grouping data can help to spot patterns in the data. • Makes it easier to process large amounts of data. • Easy to compare different groups • The data is easier to read. • The data is easier to represent on graphs.	Or any other valid reason (B1 for only one reason)	**(2)**
(ii)	B1 – We lose the detail/accuracy when grouping data e.g. we do not know the maximum and minimum values or we can only calculate estimates of statistical values.	Or any other valid reason.	**(1)**
(b)	B2 **a conclusion** with 2 of the reasons given below for any of the following points up to a maximum of 2 marks **Table A** • There is no need for 0 – 20 or 80 – 100 because there are no people in those age groups. • All the data is concentrated into 3 groups **Table B** • Each group has a smaller width showing more detail. • The table starts at 20 because the youngest person is 21 and finishes at 70 because the oldest person is 68 Conclusion: Dylan's claim is justified	(B1 for Dylan's **claim is justified** with one reason only, or two with no conclusion).	**(2)**
(c)	(i) M1 – for $\sum fa = 25\times6+35\times67+45\times53+55\times13+65\times11=150+2345+2385+715+715$ depM1 – $\frac{\sum fa}{150}=\frac{'6310'}{150}=\left(42.0\dot{6}\right)$ A1 – 42.1	For $\sum f\times a$ with correct midpoints consistently within interval. Allow one slip. Award M1 for 6310 seen Award provided first M mark scored. awrt 42.1	**(3)**

Question	Answer	Additional Guidance	Marks
10 (a)	B1 for 6		**(1)**
(b)	B1 for $\frac{42}{120}$ oe		**(1)**
(c)	M1 for $\frac{6+1+2+4+39+20}{120}$ A1 for $=\frac{72}{120}$ oe	Accept any equivalent. E.g., 3/5, 0.6, 60% Isw any incorrect subsequent simplification. Award M1A1 when $=\frac{72}{120}$ oe is seen	**(2)**
(d)	M1 for $\frac{6+2}{20+4+6+2}$ A1 for $=\frac{8}{32}$ oe	Isw any incorrect subsequent simplification. Award M1A1 for $=\frac{8}{32}$ oe seen. $\frac{1}{4}$ or 0.25	**(2)**
(e)(i)	B1 for Qualitative		**(1)**
(ii)	B2 for Quantitative **and** discrete OR if B2 not earned B1 for either Quantitative or discrete	Either order B1 for only one	**(2)**

Question	Answer	Additional Guidance	Marks	Type
11	B1 (i) Yes because there is some (negative) correlation The candidate **must** assess the appropriateness of the vet's claim. B1 (ii) 10 breeds of dogs is an insufficient amount of data so may be unreliable	Allow: Yes because the heavier the dog, the shorter is its lifespan, or the lighter the dog the longer its lifespan. Any other reasonable statistical explanation	**(2)**	1
		Total	**2 marks**	

Question	Answer	Additional Guidance	Marks
12	B2 for (median) height of seedlings grown in sunlight is 41 cm **and** the (median) height of seedlings grown in the shade is 27 cm (B1 for either median correct, must be identified as median). **OR** B2 for (mean) height of seedlings grown in sunlight is 40.2 cm **and** the (mean) height of seedlings grown in the shade is 26.3 cm (B1 for either mean correct, must be identified as mean). **OR** B2 for (modal) height of seedlings grown in sunlight is 35 cm **and** the (modal) height of seedlings grown in the shade is 31 cm (B1 for either mode correct, must be identified as mode). depB1ft for any of the above interpreted **in context** • Seedlings grown in sunlight are on average taller than seedlings grown in the shade.	For B2 identification of which average used is not required. For B1 median / mean / mode must be indicated. Accept 40 cm Accept 26 cm Accept 40.235… or 26.294… rounded or truncated to 1dp or greater If multiple averages given with some correct and some incorrect then award marks for those correct. E.g. means – both incorrect, medians – both correct then B2 Dep on two average values stated (may be incorrect). The statement must refer to, or imply heights (by using the word taller/shorter) Condone bigger / smaller trees.	**(3)**

Question	Answer	Additional Guidance	Marks
13 (a)	B1 for 3.7		**(1)**
(b)	M1 for 2.8 + 2.9 =…. A1 for 5.7	Not as part of a larger calculation.	**(2)**
(c)	M1 $\frac{13310}{65\,511\,097}$ [× 100] A1 0.0203… (%) Accept awrt 0.02(%) A1dep All of the data in the table has been rounded to 1 decimal place and 0.02(%) correct to 1 decimal place is 0.0(%), (so the table is correct) **ALT** M1 0.05% of 65 511 097 $= \left[\frac{0.05}{100} \times 65\,511\,097\right] = (32\ 755(.5485))$ A1 32 755 A1dep As 0.05% of 65 511 097 [32 755] is greater than 13 310 that means that 13 310 is 0.0(%) rounded to 1 decimal place. If 0 scored SCB1 for 65 511(.097)	Dep on previous A mark Must refer to rounding to 1 decimal place or state rounds to 0.0(%) For awrt 13 755 Dep on previous A mark	**(3)**
(d)	Without percentages quoted: B2 for • From the age of 40 upwards, each year group in the population has a greater percentage of women than men, so Jamie's claim is incorrect. (B1 for from the age of 40 upwards, each year group in the population has a greater percentage of women than men with no or incorrect conclusion) With percentages quoted: B2 for • The percentage of males (age 40 upwards) is 24% and the percentage of females (age 40 upwards) is 26.3% OR there are 2.3% more females (age 40 upwards), so Jamie's claim is incorrect	May also refer to number of males / females: Male 15722663(.28) Female 17229419(.51)	**(2)**

	(B1 for percentage of males (age 40 upwards) is 24% and the percentage of females (age 40 upwards) is 26.3% OR there are 2.3% more females (age 40 upwards) with no or incorrect conclusion) (B1ft for only one correct figure with a correct conclusion (ft their percentages))	B0 if neither percentage total (or neither number of males/females) for age 40 upwards is correct	

Question	Answer	Additional Guidance	Marks
14 (a)	B1 for one from • Data must be in the same form before it can be used / she has data in different formats / the percentage of motorcycles says two not 2 • The data set must be complete before it can be used / there is missing data • (Percentages) don't add up to 100	Ignore additional comments unless contradictory.	**(1)**
(b)	B1 for one of • the total of the numbers in the motorcycle column must be 100 so 124 has to be 24 • it is greater than 100 and is a percentage • because the total is 100	Do not accept reference to outlier / anomaly	**(1)**
(c)	Uses 50 as the median M1 For identifying $70 \leq s < 80$ M1 $\frac{17}{56} \times 10$ A1 for 73.04 (mph) Accept awrt 73 (mph) **ALT** Uses 50.5 as the median M1 For identifying $70 \leq s < 80$ M1 $\frac{17.5}{56} \times 10$ A1 for 73.125 (mph) Accept awrt 73 (mph)	May be implied by $70+...$ Allow first M mark for $\frac{70+80}{2}$ For reference: The estimated **mean** speed of the motorcycles is 71.7 mph. This scores M0M0A0	**(3)**
(d)	B2 fully correct frequency polygon – all 6 points correctly plotted and joined with straight lines (B1 all 6 points plotted correctly but not joined OR at least 3 points correct and joined with straight lines) OR for joining the points with line segments at the correct heights consistent within intervals (including end points).	For B2 condone line joining (85,11) to *x*-axis, but not joining start to end For B2 or B1 allow for 5 points used if (35,0) is <u>omitted</u> Points are: [(35, 0)] (45, 2) (55, 7) (65, 24) (75, 56) (85, 11)	**(2)**

(e)	B1 for • Distribution of the cars is symmetrical / (weak) positive skew whereas the distribution of the motorcycles is negatively skewed. B1 for • This means that for the cars the speeds are equally spread out on either side of the median and for the motorcycles the speeds are mainly at the upper end of the distribution with those speeds less than the median more spread out. • For the cars there are 50% of speeds above the mean and for the motorcycles there are more than 50% of the speeds above the mean.	Accept 'not skewed' for 'symmetrical' Do not accept 'normally distributed' or 'symmetrical skew' for symmetrical Comment must be on spread of data within distribution. B0 for motorcycles are faster than cars Allow the 2nd B1 for interpreting one of the skews in context: • for the cars the speeds are equally spread out on either side of the median • for the motorcycles the speeds are mainly at the upper end of the distribution with those speeds less than the median more spread out • for the cars there are 50% of speeds above the mean • for the motorcycles there are more than 50% of the speeds above the mean	**(2)**

Question	Answer	Additional Guidance	Marks
15 (a)	B1 for any one from 1, 2 or 4 with **NO** incorrect regions identified.		**(1)**
(b)	B2 Adam's comment is valid because regions 15 and 16 are the darkest regions on the map (or the regions indicating more than 1500 orangutans) showing there are the greatest numbers of orangutans	B2 for valid **and** correct reason (B1 for correct reason without stating that Adam's claim is valid).	**(2)**

Pearson Education Limited. Registered company number 872828
with its registered office at 80 Strand, London, WC2R 0RL, United Kingdom