



Mark Scheme (Results)

Summer 2025

Pearson Edexcel GCSE (9 – 1)
In Statistics (1ST0)
Foundation Paper 2F

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 2025

Question Paper Log Number P76386A

Publications Code 1ST0_2F_2506_MS

All the material in this publication is copyright

© Pearson Education Ltd 2025

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 number	Answer	Additional guidance	Mark																
1 (a)	B1 <table border="1"> <thead> <tr> <th></th><th>Black</th><th>White</th><th>Total</th></tr> </thead> <tbody> <tr> <td>Star</td><td>3</td><td>1</td><td>4</td></tr> <tr> <td>Circle</td><td>3</td><td>3</td><td>6</td></tr> <tr> <td>Total</td><td>6</td><td>4</td><td>10 ...</td></tr> </tbody> </table>		Black	White	Total	Star	3	1	4	Circle	3	3	6	Total	6	4	10 ...	B1 all 6 values correct.	(1)
	Black	White	Total																
Star	3	1	4																
Circle	3	3	6																
Total	6	4	10 ...																
(b)	B1 $\frac{4}{10}$	B1 accept any equivalent $\frac{2}{5}$, 0.4 or 40% Do not allow 4 out of 10 or 4:10	(1)																

Question number	Answer	Additional guidance	Mark
2	B2 Not appropriate with a correct reason • All are to the same scale/height/no vertical scale • No labels /no key/no years • Different sized devices (Ipads, computers and controllers) • No title (B1 for not appropriate with an attempt at a reason or for correct reason with an incorrect conclusion/no conclusion)	Acceptable answers include • Can't measure the percentage by the size of the icons B0 if stating it is hard to read unless it has one of the supporting reasons.	(2)

Question number	Answer	Additional guidance	Mark
3 (a)	M1 2, 12, 9, 4 or $\frac{2+12+9+4}{12}$ A1 2.25 oe	M1 frequency $\times$ number of films written in table or seen in working. Condone one error. (or) Sum of frequency $\times$ number of films divided by the total frequency of 12. Condone one error in their frequency $\times$ number of films. A1 for 2.25 Accept a rounded answer of 2 if 2.25 is seen.	(2)
(b)	B1 e.g. • data is only for number of films watched and not general television • pensioners may have watched more television than the students that did not include films • the films may not have been watched on the television / films may have been watched on a different device • the data is only for the number of films and some films may be longer than others/we don't know the length of time of the films • Small sample size	B1 for a comment suggesting we do not have enough information about the total number of hours of television watched in the week Do not accept reference to more/less students/pensioners in the sample or that one week is not long enough.	(1)

Question number	Answer	Additional guidance	Mark
4 (a)	B1 for Median	B1 cao	(1)
(b)	B1 for Mode	B1 cao	(1)
(c)	B1 for Median	B1 cao	(1)

Question number	Answer	Additional guidance	Mark
5 (a)	B1 $\frac{3}{10}$ oe	B1 allow for 0.3 or 30% Do not allow 3 out of 10 or 3:10	(1)
(b)	B1 $\frac{4}{50}$ oe	B1 allow for $\frac{2}{25}$, 0.08, or 8% Do not allow 4 out of 50 or 4:50	(1)
(c)	B2 Dhiyan's estimate is more reliable since - it is based on more trials - he has more data - he spun the spinner more (B1 for Dhiyan and an attempt at a reason)	B2 for Dhiyan and correct supporting reason (B1 for Dhiyan and an attempt at a reason)	(2)
(d)	B2 No/ the spinner is not fair with a supporting reason - the frequencies are not all (approximately) the same size - expected number of all colours would be 10 if fair - there is (significantly) more green's (and blue's) than the other colours - fewer pink/red than the other colours - big difference between the highest and lowest frequencies. - landed on some colours a lot more/less than other colours. (B1 A correct reason with no or incorrect conclusion)	Allow for a comment on the variation of size of the frequencies. B0 for just stating the number of times the spinner landed on a colour. B0 for just stating that green is the highest and/or pink is the lowest.	(2)

Question number	Answer	Additional guidance	Mark
6 (a)	M1 for 41 44 68 68 72 74 80 97 or 68, 72 A1 70	M1 is for ordering the numbers OR for identifying the two middle numbers. Condone at most one error or omission. A1 cao	(2)
(b)	B1ft ('70' + 2 =) 72	B1ft 72	(1)
(c)	B1 e.g. • there are more pigs than cows (on the eight farms) • the (eight) farms have more pigs (on average) • The mean is greater for pigs • The mean is smaller for cows • On average there are 22 more pigs (on each farm than cows) • the total number of pigs is greater than cows	B1 comment comparing the two mean values in context B0 the difference between pigs and cows is 22. Condone there are more pigs on each farm.	(1)
(d)	B1 median	B1 Condone misspelling Do not accept medium	(1)

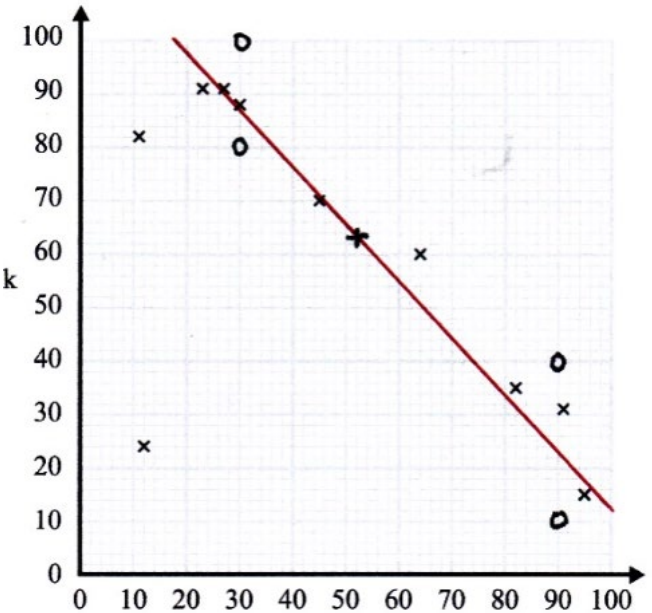
Question number	Answer	Additional guidance	Mark
7 (a)	B1 e.g. - the midpoints have to be used for the calculation - we do not have the raw data/the true values are not known - the data has been grouped - we don't know exact values - we don't know exactly how long the butterflies have lived for	B1 for a comment identifying a loss of accuracy through use of midpoints or of grouped data Accept it is in intervals B0 for it is not accurate or that it has been rounded.	(1)
(b)	B1 e.g. - there is an extreme value which represents a butterfly that has lived a lot longer than the others. - The mean would be affected by outliers	B1 for stating there is one butterfly that lives longer than the other 14 Allow for extreme value reference to outlier/anomaly in the data.	(1)

Question number	Answer	Additional guidance	Mark
8	B1 e.g. - conducting a census would use up all (Fran's) stock/nothing left to sell - lightbulbs expire so cannot be used again after being sampled - the company would lose money for taking a census - quicker/take less time - would be cheaper - less lightbulbs to test/less data to handle - easier to do	Do not accept 'more accurate' or 'unbiased.'	(1)

Question number	Answer	Additional guidance	Mark
9 (a)	B1 for categorical circled or otherwise unambiguously identified	B0 if more than one word is circled	(1)
(b)	B1  B1 a bar for goats of height 13	B1 correct tallying B1 correct bar Ignore the width of the bar, condone lines not ruled and bar not shaded.	(2)
(c)	B1B1 eg • Mandeep has tallied a sample / 55 is not all of the animals / there are over 600 animals on the farm and he only saw 55 • sample was not random/is biased • he may have counted some of the animals more than once • might not have visited the whole farm/stayed in the same area of the farm/some areas of the farm may have had more/less animals • some animals may have been out of sight/some might not have been counted	B1 for each correct statement	(2)

Question number	Answer	Additional guidance	Mark
10 (a)	B1 e.g. • The short survey will be quicker/easier to do (than conducting an interview) • students less likely to lie / will be more honest/feel less pressured • get results back quicker	B1 for a correct statement. Allow a disadvantage of conducting an interview if it is made clear they are talking about interviewing e.g. an interview would take longer B1. Do not accept 'it will give the teacher more data or 'more responses'	(1)
(b)	B1 Open (question)	B1 cao	(1)
(c)	B1 e.g. • to put data into the same units/inconsistent units • missing units • to make all data quantitative/ times are given in different formats • to remove anomalous answers • Inexact data e.g. about an hour	B1 for a correct reason Do not allow timings are incorrect. B0 for it is easier to read. B0 for not in order , not organised or difficult to understand.	(1)
(d)	B1 for scatter diagram circled or otherwise unambiguously identified	B0 if more than one word is circled	(1)

Question number	Answer	Additional guidance	Mark
11 (a)	B1 e.g. • It is difficult to identify the department who makes the most profit. • Sectors are similar sizes so difficult to decide which one is the biggest • No angles/percentages on each sectors • Doesn't give exact amounts	B1 for a comment describing the difficulty to interpret the information presented in the pie chart. 'Difficult to read' on its own is B0, we need some reference to it being difficult to identify the size of the sectors. B0 for not accurate	(1)
(b)	B1 Bar chart or line chart	B1 for suggesting a diagram where the heights of the values can be compared. Condone pictogram.	(1)

Question number	Answer	Additional guidance	Mark
12 (a)	B1 (12,24) indicated on scatter diagram	B1 for correct data point circled.	(1)
(b)	B1 Negative (correlation) Or e.g. as Maths marks increase, Music marks decrease.	B1 Ignore references to strength.	(1)
(c)	B2 Line of best fit drawn through (52,63)  (B1 for plotting the double mean point or for an appropriate line of best fit with negative gradient not through double mean point)	B2 for appropriate line of best fit with negative gradient through double mean point. (Mean point need not be plotted) Must be a ruled line and should extend from between (30,80) and (30,100) and (90,10) and (90,40) to include most of the data and approximately half the points on either side	(2)
(d)	B1 e.g. - plot the results for more than 10 students - Include all 30 children in the class on the scatter diagram	B1 for a correct statement	(1)

Question number	Answer	Additional guidance	Mark
13 (a)	B1 Population pyramid	B1 Condone misspelling but must include both words.	(1)
(b)	B2 45 -54 as it has the largest percentage of males and females combined with 26.5% (B1 for 45-54 as it has the largest percentage of males and females combined without the correct supporting figure)	Accept 26.2% to 26.8% B1 answer of 45-54 as the males have the biggest bar	(2)

Question number	Answer	Additional guidance	Mark
14 (a)	B1B1 e.g. - Carried out under controlled conditions (distances/time) can be measured more accurately on a treadmill/will have accurate measuring equipment - Easy to repeat the experiment under the same conditions/environment - Can control other factors that might affect the outcome/same running conditions e.g. weather, traffic, ground conditions - Less time consuming/less expensive	B1B1 for each of two correct advantages to conducting a laboratory experiment over one in nature/field B0 for more accurate or more reliable on its own.	(2)
(b)	B1 e.g. - the distance (the participants) run in the 6 minutes - the running performance (of the participants) - Speed of the participant	B1 for a statement about the distance achieved. B0 for an answer that includes both distance/speed and music e.g. the distance they ran with music on.	(1)
(c)	B1 e.g. - The music each participant listens to should be the same - same genre of music for each participant - all listen to the same playlist	B1 for a comment about the music being constant for each experiment	(1)
(d)	B1 e.g. - to reduce their effect on the outcome on the experiment - as different music types may make a runner run faster/slower - the experiment can then be repeated easily - reduce the chance of any bias towards a favoured outcome / make the test fair	B1 for a correct comment B0 for 'to see if it affects the results'	(1)

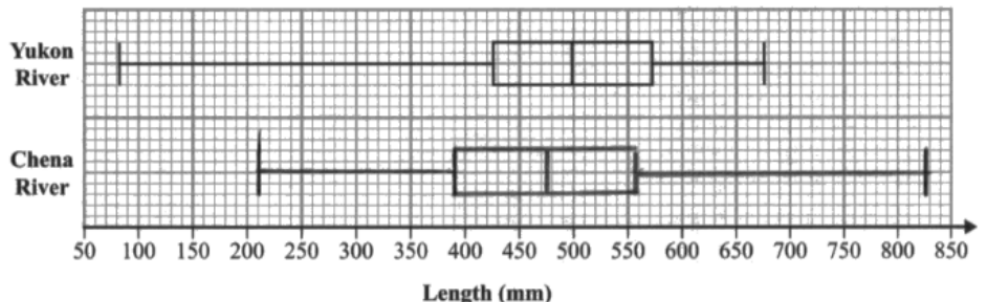
Question number	Answer	Additional guidance	Mark
15 (a)	M1 $100 - 71$ (= 29%) A1 29% decrease or -29(%)	M1 $100 - 71$ or 29(%) A1 29% decrease, or other indication that it has reduced by 29% must be given. M1A0 for 29 seen but spoilt by further working.	(2)
(b)	M1 95×0.77 A1 73 (.15) p	M1 for correct method seen or implied A1 awrt 73	(2)
(c)	B2 e.g. • Inappropriate because the prices may have changed significantly since 2020/the prices could have gone back up since 2020. • Inappropriate because the price change of the ingredients is not known/can't predict what is going to happen to future prices • Inappropriate as there may be a cost of living crisis/may be inflation. (B1 for a correct reason with no or incorrect decision)	B2 for a correct assessment of the appropriateness together with a reason	(2)

Question number	Answer	Additional guidance	Mark
16	B1 for a correct comment on the hypothesis eg - the hypothesis suggested is a question not a statement (so is inappropriate). B1 B1 B1 B1 for comments on data collection Sampling method - no indication of how (Elise) will sample the skulls from the museums/we do not know if this will be primary data (that she collects/measures herself) or secondary (given the data by museum staff/records) - collecting skull measurements from museums is appropriate because it is less time consuming/easier/reliable/accurate/quicker OR collecting the data from all the museums may take a long time OR museums may have biased/non-representative samples - the sample size per mammal/species is small (only two males/females per species) so not representative / she should collect more female and males samples per species OR. males and females for each species often differ in size (so it is sensible to include both)/collecting data on male and female per species is representative - only collecting data on adults of the species and she needs to include young mammals - The species of mammals is not specified/ambiguous (Marsupials/Monotremes/Placentals etc so data may be skewed/not representative) - Large sample (100 mammals) is appropriate - Elise has made no plan to collect data on body mass (as per her hypothesis) - Elise also needs to record the type of seeds and fruits (as per her hypothesis) not just the size. Issues with the collected data - there may be errors in recording of measurements/the data may not be reliable/accurate (in the databases) OR collecting data from the existing databases/museums is likely to be a reliable/accurate source - The skulls in the museum may be damaged/degraded (so cannot get data from them) - the data may need 'cleaning' before it can be processed/data may be in the wrong format - there may be missing data in the databases (too)	B1 for a comment on the hypothesis B1 for each of up to four comments on data collection Award a maximum score of B1 per bullet point Contradictions within the same bullet point written scores B0 Do not accept 'Data from museum/internet may be out of date'	(5)

	Use of secondary data • do not know how the internet data was collected (may not be reliable/accurate/may be incorrect/biased) • the sources of the secondary data should be included / we do not know which museums/databases (she is using) OR secondary data is appropriate if a reliable source is used • skulls / seeds and fruit may not be measured by the same people or in the same way or the measurements need to be consistent and measure what was intended • Using secondary data is appropriate due to the difficulty of collecting this (“missing” primary) data (herself)/impossible to collect all information as this is primary data OR finding/collecting all the data required will be difficult/time consuming/ Elise may find it difficult to access all the data required (from the museums/databases) / databases might cost money to access OR using secondary data is appropriate because it will be cheaper/quicker/easier		
--	---	--	--

Question number	Answer	Additional guidance	Mark
17 (a)	B1 Secondary (data)	B1 cao	(1)
(b)	B1 e.g. • Clippers won more games/matches than Lakers (each year/season) • Lakers won fewer games/matches than the Clippers (each year/season) • Clippers performed/ were/did better than the Lakers (each year/season) • Clippers have won the most games (each year/season) • Bars for clippers are higher/taller so did better (than Lakers each year/season)	B1 for a comparison made between the two teams for each year. Condone any use of the word ‘average/mean’ in candidates response B0 if a comparison is made of just one season e.g. 2015-2016 or comparison of each team’s performance from one season to another. Note: In this question we can ignore any figures.	(1)
(c)	B1 e.g. • (The number of games won by the Lakers year on year is) increasing/improving/rising/upward (trend)/ positive trend • The Lakers are winning more games each season than in the previous. • The Lakers are getting/doing better each season.	B1 for a comment made on the overall trend seen. B1 for Lakers won more and (then) more games B0 for Lakers won more games Do not allow positive correlation alone, but condone if accompanied by e.g. increasing /upwards trend Ignore reference to figures.	(1)
(d)	B2 for 57 and 53 (B1 for 57 or 53)	B2 for awrt 57 and awrt 53 (B1 for awrt 57 or awrt 53) Note: The reference to round to nearest whole number in the question was merely to assist candidates but is not a requirement for the B1 marks. For 57 allow 56.5(217...) For 53 allow 52.5(641...)	(2)

(e)	B1 e.g. • Clippers (because) they won more (54(%)) (of their games) whereas the Lakers only won 52/53(%). • (54(%)) > 52/53(%) Clippers won more (games). • Clippers won more games by 1%/2% • Clippers won more (44(.28)) games whereas Lakers only won 42(.64)/43 games • Clippers won 44 games which is 1 / 2 more (than Lakers) • Lakers won 42/ 43 games which is 1 / 2 less (than Clippers) • Clippers (won) 54%, Lakers (won) 52%/53% so Clippers (won more).	B1 for correctly comparing the percentage of games won by both teams with correct conclusion Note: for the Lakers, accept a value stated between 52 and 53% inclusive. A comparison of the (54(%)) and 52(%)/(53(%)) must be seen or implied as part of their answer. Condone missing 54% in their comparison if their 52/53(%) is seen and correct answer is therefore implied. 52/53(%) may be indicated on the graph Just stating figures is B0.	(1)
-----	--	--	-----

Question number	Answer	Additional guidance	Mark
18 (a)	M1 80 – ‘24’ (= 56) or sight of a value between 23- 24 A1 56 or 57	M1 for partial method. Where their ‘24’ is a value between 20 and 40. A1 56 or 57 M1A0 for a non-integer value between 56 - 57	(2)
(b)	B1 e.g. • Easy (to do)/convenient • Takes less time/more efficient • Can change location/flexible (along the river)	B1 for a correct comment about the advantages of using opportunity sampling B0 for ‘cheaper’ or more accurate.	(1)
(c)	M1 for a box with two whiskers AND at least two values plotted correctly.  A1 for all correct	M1 for box with two whiskers and an attempt at a single median line AND at least two values plotted correctly A1 Fully correct plot (210, 390, 476, 556, 825)	(2)

(d)	B1ft B1ft B1ft B1ft		Allow ft for their box plot. Also allow recovery if candidates refer back to values in the table. For contextual comments we are looking for reference to lengths/size (of pike) Do not accept eg Yukon river is longer B1ft For a correct statistical statement comparing medians. Condone longer for comparison of medians. Must refer to <u>median</u> , not average for this B mark. Condone misspelling but medium is B0. B1ft For a correct comparison of the IQRs/ranges. Condone wider for comparison of IQR/range. B1ft For correct comparison of the skews. For Chena/Yukon also accept no skew but not neutral/normal/symmetrical skew. Condone use of 'normally distributed' instead of symmetric Accept 'symmetrical distribution' B1ft One correct interpretation Do not awarded mark for interpretation if clearly linked to incorrect statistical measure Allow equivalent/converse statements but underlined words must be seen. Note: in this question ignore any numerical values in comparisons	(4)
	Comparison	Interpretation		
	Yukon (pike) greatest <u>median</u> or Chena (pike) <u>median</u> < Yukon (pike) <u>median</u>	eg Yukon (pike) are larger/wider/longer/bigger (in length).		
	Chena (pike) has the greatest <u>IQR</u> or Yukon (pike) (≈ 145) <u>IQR</u> < Chena (pike) (≈ 165) <u>IQR</u> or Chena has the greatest <u>range</u> or Yukon (≈ 595) <u>range</u> < Chena <u>range</u> (≈ 615)	eg Yukon (pike) lengths are the most consistent		
	Yukon (pike) have <u>negative skew</u> /are symmetrical and Chena (pike) are <u>symmetrical</u> / have positive/negative skew	eg The spread of <u>lengths</u> below the average (median) for Yukon is smaller than the spread of <u>lengths</u> above the average or The spread of <u>lengths</u> above the average (median) is the same as the spread of <u>lengths</u> below the average for (both) Yukon (and) Chena. or eg the spread of lengths below the average for Chena is larger than the spread of lengths above the average.		

Question number	Answer	Additional guidance	Mark
19 (a)	B1 $\frac{12}{120}$	B1 accept any equivalent $\frac{1}{10}$, 0.1 or 10% Do not allow 12 out of 120 or 12:120	(1)
(b)	M1 $40 = 4 + 3 + 12 + x$ or $40 - (4 + 3 + 12) = (21)$ or $120 = 30 + 30 + 5 + 12 + 4 + 3 + x + y$ or $120 - (30 + 30 + 5 + 12 + 4 + 3 + x) = (15)$ A1 $x = 21$ or $y = 15$ A1 $x = 21$ and $y = 15$	M1 for correct sum seen for drums or all sample space. A1 for $x = 21$ or $y = 15$ A1 for $x = 21$ and $y = 15$ If no answer on answer line check Venn diagram	(3)
(c)	M1 $\frac{k}{51}$ or $\frac{17}{m}$ or $\frac{\frac{17}{120}}{\frac{51}{120}}$ A1 $\frac{17}{51}$ oe	M1 for realising 51 is the denominator where, $\frac{k}{51}$, $0 < k < 51$ or for realising 17 is the numerator where, $\frac{17}{m}$, $17 < m \leq 120$ A1 allow equivalent fraction e.g. $\frac{1}{3}$, awrt 0.33 or awrt 33%	(2)

		For “17 out of 51” or 17:51 or “1 out of 3” or 1:3 award M1A0	
--	--	--	--