TAKING A SAMPLE

AtoZrevision

TAKING A SAMPLE: ADVANTAGES AND DISADVANTAGES.

MARK SCHEME

Q1. AtoZrevision.com

(a)	B1 Advantage: • e.g. Convenient • e.g. Easy B1 Disadvantage • e.g. Not representative • e.g. Biased • e.g. Students arriving early may all travel to school using the same transport	1st B1 any one correct advantage 2nd B1 any one correct disadvantage	(2)
(b)	B1B1B1 for each of three aspects from: • Number the students in the database / use the position in the database • Generate numbers using a calculator / computer / random number table • Students with the matching number are selected • Ignore repeated numbers or numbers outside the range of those given to the students	B1B1B1 for demonstrating understanding of how to select a random sample	(3)

Q2. AtoZrevision.com

B1 ft for e.g. • Only 1 packet from Katy's sample has fewer than 12 lollies • The modal number of packets is 12 • $\frac{17}{18}$ is close to 1 (the probability of lollies >11)	B1 for a correct statement explaining the appropriateness of the packaging label	(1)
B1 for e.g. • It might be quicker/easier to take the sample from the same machine • Katy may have only wanted to check the quality of one of the machines B1 for e.g. • One of the machines might be faulty and you might not find out if you only sampled from one machine	B1 for a correct statement explaining the appropriateness of the data collection method (of sampling from one machine) B1 for a correct statement explaining the inappropriateness of the data collection method (of sampling from one machine)	(2)

Q3. AtoZrevision.com

B1 Quota B1 Quota sampling is useful when • time is limited • a sampling frame is not available • the research budget is tight • when detailed accuracy or randomness is not important	B1 for quota B1 for a correct explanation when it is appropriate to use quota sampling eg one of the bullet points	**(2)**

Q4. AtoZrevision.com

Any sensible problem with small sample size. e.g. may not be representative may be biased results may not be reliable results may not be accurate	B1	(1)
Cluster	B1	
Stratified	B1	(2)
Method 2 is…. • More representative • Random/allows all members of population equal chance of selection • Not/less biased • More reliable	B2	(2)

NOTES:

Any one problem identified (ignore extraneous non-contradictory responses)

B2 for any 2 reasons
(B1 for any one reason)

Just repeating wording in method 2 on its own is B0
e.g. selecting residents in proportion B0

Must have at least 3 response boxes
B2 for all 3 criteria
(B1 for at least 1 criteria)
Condone a useable table for response boxes

Note: For exhaustive allow
e.g. 0 – 1, 2 – 3, 4+
Allow correct use of inequalities

Q5.

(a)	B1 Quota	B1 Condone misspelling Condone opportunity/convenience	(1)

B1 B1 B1 Any three from… • Not random/ not everyone has an equal chance of been chosen/ Navine is picking the people • Easy to do • Allows for comparison in age groups, a range of people • no sample frame required • Only asking people's opinions on one day/ask opinions on another day • Small sample compared with number visitors each day/should ask more people • Face to face gets better response rate/explain question • Face to face expensive/ time consuming/interview bias/ pressured to give an	B1 for equivalent of each of three appropriate bullets from the list.	(3)
answer/may not be honest/should give a questionnaire instead • Only asks about favourite ride and no other opinions/ask other questions about the theme park • Asking them at the end of the day is sensible as they have had an opportunity to visit the theme park • May not want to answer the questionnaire as they leave (as they are tired) / ask them inside the theme park		

Q6.

(a)	B1 for random start point (from 1 to 20) B1 for selecting every 20th item		**(2)**
(b)	B1 for e.g. • not random • not representative • interval may coincide with some pattern in the population	B1 for identifying a disadvantage of systematic sampling	**(1)**

B2 Quota sampling may not be appropriate since • hammers likely to all have the same characteristics • there is no information given to put the hammers into different groups • it is not random/it is biased • you have a sampling frame (list of id numbers) (B1 for not appropriate with an attempt at a reason OR a correct reason with no conclusion or an incorrect conclusion).	B2 not appropriate with correct supporting reason (B1 not appropriate with any reason) Accept: e.g. 'No it isn't' in place of inappropriate.	(2)
B2 Systematic sampling is (more) appropriate since • it is easy to take every n^{th} item from a production line • it is less/not biased (compared to quota sampling) OR B2 Systematic sampling is not appropriate as there may be a fault in the machine that occurs at regular intervals. (B1 for appropriate with an attempt at a reason OR a correct reason with no conclusion or an incorrect conclusion).	B2 appropriate with correct supporting reason (B1 appropriate with any reason) Accept: e.g. 'yes it is' in place of appropriate. Do not accept: 'Appropriate, it is easy…' for B2 This will score B1 only.	(2)

Q8. AtoZrevision.com

M1 $\frac{140}{1200} \times 60$ o.e. A1A1 Bedrooms: 1 | 2 | 3 | 4 | 5+ Houses in sample: 7 | 15 | 21 | 12 | 5	Accept a correct equivalent calculation shown for any one class M1 implied by one correct answer OR an indication they need 1 in 20 1st A1 for any one value correct 2nd A1 for all correct	(3)
B1 Use a sampling frame for each strata B1 Select houses randomly or generate random numbers B1 For an aspect of detail	Each category/strata to be considered separately Samples have to be random e.g. How the random numbers are obtained and used	(3)

Q9. AtoZrevision.com

	5ST1H_01 Scheme	**Marks**
(b)(i) **(ii)**	tudents numbered - either of: • **number** the students OR use a **list/database/register** of students Selection: • Use of a **random number generator** Matching: • Use student corresponding to number selected For any relevant statistical problem about the sample selection Stratified (sampling) $^{240}/_{1200}$ × 40 = **8**	B1 B1 B1 B1 (4) B1 M1A1cao (3)
	Notes	
(a) (b) (c)(ii)	*For QWC: 1st B1 and 2nd B1 use of words in bold (oe) required *1st B1 for **number/list/data base/register/sampling frame/spreadsheet** *2nd B1 for use of a suitable **random number generator** e.g. **random number table**, **calculator**, **computer** oe (B0 for put the names/numbers in a hat) 3rd B1 for matching number to student. 4th B1 e.g. 'do not use repeats' 'discard numbers out of range' 'may be difficult to obtain a register of all 1200 students' 'register of students may not be up-to-date' 'students may be absent/may refuse to participate,' etc. M1 for any equivalent method A1 cao	

Q10.

Sample is (any two from): • Quicker • Cheaper / uses less resources • Easier (to do / to calculate ... etc) • Less data to handle	B1B1 (2)
All people/items have same/equal chance of selection	B1 (1)
• Leading/biased Open OR no answer boxes/options given	B1 B1 (2)

NOTES:

May have two reasons in one statement. **Must be from these four options, but each bullet point once only.**
Condone 'more convenient' as a separate point for B1 Only allow converse statements if they use the word 'census'. Note: possible non-response from census, ... **is B0**
completely fair / not biased / no control over choice OR just a description of how to take a random sample ... **alone are B0**

Q11.

B1 Advantage: • e.g. Convenient • e.g. Equal number of males and females represented B1 Disadvantage: • e.g. Not representative • e.g. Biased • e.g. Employees coming early may have similar pay	1st B1 any one correct advantage 2nd B1 any one correct disadvantage	**(2)**

(a)	B1B1 Any two from: • Repeated random numbers • Random numbers out of range/may not correspond to students' numbers • Selected students may not (want to) participate • Some students may have left the university	B1 for each bullet point up to a maximum of 2. Accept each bullet point only once. Students may have joined the university is B0. There may not be 100 students at the university is B0. Random numbers may not be whole numbers is B0. Database may not be up to date **on its own** is B0. Do not accept (random) numbers may be more than 100 for the second bullet point. Ignore extraneous non-contradictory comments.	(2)
(b)	B1B1 Any two advantages from: • Easy/convenient/quick/efficient/cheap • Represents population (proportions) • Allows for comparison (between undergraduates and postgraduates) • No sample frame required	B1 for each bullet point up to a maximum of 2 Accept each bullet point only once For 2nd bullet point allow e.g. 'fair number of each (group)' 'Unbiased' on its own is B0. Ignore extraneous non-contradictory comments.	(2)
(c)	B1 Any one from: • Not every student has an **equal**(o.e.) chance of being selected • Only those in the main building can be selected/not every student has a chance of being selected • Robert is **choosing** the students	B1 for a reason which states or implies 'equal likelihood' of being selected **or** that Robert is doing the choosing Do not allow 'even' chance or 'its biased' for the first bullet point, but condone 'fair chance'.	(1)

Q13. AtoZrevision.com

(a)	B1 All people/items have the same/equal chance of being chosen.	B1 for demonstrating understanding of a random sample	(1)
(b)	B1B1B1 for each of three aspects from: • Get a list / register (as the sampling frame) of all the students in the school oe • Number the students in the (sampling frame) list / register • Generate random numbers using a calculator / computer / random number table • Students with the matching number are selected	B1×3 for demonstrating understanding of how to select a random sample	(3)
(c)	B2 for a complete answer e.g. the plan is appropriate **AND** e.g. the number of hours spent on homework might differ between school years OR if B2 not earned... B1 for an incomplete answer e.g. the plan is appropriate, with an attempt at a reason OR for correct reason without conclusion	B2 for complete answer assessing the appropriateness of the suggested plan OR if B2 not earned... B1 for an attempt at assessing the appropriateness of the suggested plan	(2)

Q14. AtoZrevision.com

	5ST1F_01 Mark Scheme	**Marks**
	One mark for each of three aspects: **1.** Number all the boys/pupils (from 0 to 159) or put all names on piece of paper, oe **2.** EITHER: Generate **eight random numbers** OR: draw out **eight** numbers/names from a hat, oe **3.** Boys/pupils (with the) drawn (numbers) are selected for sample.	B1 B1 B1 **[3]**
	Notes	
	If describing a method other than simple random, then score 0/3 **e.g. mention of 'picking every 20th boy/pupil' is B0B0B0** **e.g. 'choosing the same number from each class/group' is B0B0B0**'words' in brackets not needed. For aspect **1**. allow use of a **numbered** list/**numbered** register For aspect **3**. Accept e.g. Boys drawn represent the school / boys with those numbers are used, etc. **NB**: 'eight boys are drawn from a hat' on its own scores B1 for aspect 2 (**but** without 'names put on paper' etc is B0 for aspect 1, and, without 'are used' etc is B0 for aspect 3).	

Question	Scheme	Marks
(a)	Quota (sampling)	B1 (1)
(b)	Advantage: • Convenient • Easy • Same number of boys and girls selected/boys and girls are equally represented • Questions can be explained Disadvantage: • Takes a long time (to reach quota)/May not reach quota • Biased/not random • May not be representative	B1 B1 (2)
***(c)**	Use a **numbered list/spreadsheet/database/register** of all students Select a **random sample/random numbers** (using computer) The sample should be representative of the population (include the same proportion/percentage as the population of each gender)	B1 B1 B1 (3) **[6]**
	Notes	
(b)	B1 for any suitable advantage of quota sampling B1 for any suitable disadvantage of quota sampling	
(c)	Must include bold words for QWC oe Random numbers from a hat/box scores B0	

Question	Scheme	Marks
(a)	Quota (sampling)	B1 (1)
(b)	Advantage: ensures both genders / a variety of ages are represented	B1
	Disadvantage: Not random / is open to interviewer bias in selection oe	B1 (2) [3]
	Notes	
(b)	Allow equivalent wording for advantage and disadvantage if meaning is clear. **Advantage:** B1 for an answer indicating inclusion of more than one gender or age group. They must refer to gender and/or age B0 Do not accept (on their own), e.g. fair, gets enough data, equal numbers in each section, includes a variety of people, reliable/accurate representation, easy/quick/cheap. **Disadvantage:** B1 for stating method is not random / could be biased, or may not reflect the (proportions in the) population. e.g. 'may not be a fair representation' Accept: expensive / time consuming, or age range is incomplete (e.g. no U18) B0 Do not accept (on their own), e.g. less accurate, those selected may be biased, won't have correct proportions (too vague), not enough data	

(a)	(Simple) random (sample)	B1 (1)
(b)*	QWC Use random numbers OR random number table/generator Plus any **two** points from: • Number the customers OR get a list o.e. (ignore incorrect numbers here) • Use the customers with corresponding numbers • Ignore repeats OR ignore numbers out of range	B1 B1B1 (3) **[4]**
	Notes	
(a)	B1 for 'random' but 'stratified random' is B0	
(b)	Allow equivalent wording. 1st B1 is for appropriate use of random numbers. 2nd/3rd B1B1 for any two from the three options (each point once only). Condone each point within description of stratified sampling (up to 3 marks) BUT if describing systematic sampling award maximum 2 marks out of 3 Names in hat method is 0/3	

(a)	Advantage: - quicker / easier / cheaper (to get data) Disadvantage: - not random / may be biased (Any one of) - may not be representative (of all opinions)	B1 B1 (2)
(b)	Cluster	B1 (1)
(c)	Correct comments from two of these five aspects: **UNDERSTANDING:** Questions/responses can be explained in an interview OR may not be understood in a questionnaire **CANDOUR:** Employee may be less open/honest in an interview OR questionnaire can be done without pressure **RESOURCES:** Interviewing can be time consuming/expensive OR questionnaire can be done by all at the same time (or in their own time, or more cheaply) **INCLUSIVITY:** Interviews more likely to include all employees OR questionnaires might not be returned OR office manager views are not included by interview. **INTERVIEWER BIAS:** Possible bias from manager (in interview) / no interviewer bias with questionnaire.	B2,1,0 (2) [5]
	Notes	
(a)	**Advantage**: B1 for answer implying that the method will be quicker/easier/cheaper B0 lots of data / random / offices have a fair chance of inclusion **Disadvantage**: B1 for any equivalent comment implying one of the two options. Condone 'not accurate' for not representative. Condone 'small sample' / '10 offices too few' e.g. employees in one office may have similar opinions / employees in some areas may not be included B0 for time consuming / expensive / may be non-responses / not everyone included	
(c)	B2 for two different aspects from these five, or B1 for one aspect. Accept equivalent comments but **each aspect counts once only**. Ignore excess comments if not contradictory.	

	B2 All three sampling methods correctly identified: **A** – Quota **B** – Systematic **C** – Stratified B1 for an appropriate comment on sampling method **A** e.g. • every shop is represented • method is not random • biased / directors choose individuals to include • numbers sampled not in proportion to size of stores • may not give you 450 people B1 for an appropriate comment on sampling method **B** e.g. • easy to carry out • not every shop is represented / unrepresentative • method is not random • intervals may coincide with a pattern B1 for an appropriate comment on sampling method **C** e.g. • method is random (condone for this mark if this is identified as the sampling type) • every shop is represented / representative • each employee has an equal chance of being selected • not biased depB1 for identifying which method is most appropriate therefore chose **C**	B2 for identifying all three types of sampling correctly (B1 for identifying one or two types of sampling correctly) For **A** condone judgement sampling Ignore additional comments unless contradictory. Do <u>not</u> accept for A: • it is in proportion to store size Condone for A: • not everyone has an equal chance of being selected Do <u>not</u> accept for B: • it is quick Condone for B: • not everyone has an equal chance of being selected dep on at least B1B1 scored out of the 3rd to 5th marks Do not award for multiple methods identified as appropriate.	**(6)**