2.1 Statistics: Sampling

EDEXCEL 9MA0 · Calculator allowed · about 230 minutes · Write-in edition

Total Marks

Name: _____ Date: ____ / ____ / _____

Answer ALL questions. Show all your working.

1 A gym has 3 types of membership: Standard (240 members), Premium (160 members) and Family (100 members). A stratified sample of 25 members, proportional to membership type, is taken to survey satisfaction scores out of 10.

(a) Calculate the number of Premium members that should be included in the sample. **(2)**

(b) The satisfaction scores of the 8 Premium members sampled are: 7, 9, 6, 8, 10, 7, 9, 8. Calculate the mean satisfaction score for this stratum. **(2)**

(c) The mean satisfaction scores for the Standard and Family strata are 6.5 (from the 12 Standard members sampled) and 7.6 (from the 5 Family members sampled) respectively. Calculate an estimate of the overall mean satisfaction score across the whole sample of 25 members. **(3)**

(Total for Question 1 is 7 marks)

2 A school has 240 students. The head wants a sample of 48 students to complete a survey about the canteen. The population is split as follows: Year 7: 80, Year 8: 50, Year 9: 110.

(a) Describe how the sample could be selected using stratified sampling, and find the number to be taken from each group. **(3)**

(b) State one advantage of stratified sampling over simple random sampling. **(1)**

(c) Explain how systematic sampling could be used to select a sample of 48 from the whole population of 240. **(2)**

(Total for Question 2 is 6 marks)

3 A school has 840 students. The head wants a sample of 42 students to complete a survey about the canteen. The population is split as follows: Year 7: 360, Year 8: 120, Year 9: 360.

(a) Describe how the sample could be selected using stratified sampling, and find the number to be taken from each group. **(3)**

(b) State one advantage of stratified sampling over simple random sampling. **(1)**

(c) Explain how systematic sampling could be used to select a sample of 42 from the whole population of 840. **(2)**

(Total for Question 3 is 6 marks)

4 A library has 280 registered members. The librarian selects 56 members to ask about opening hours. The population is split as follows: Section 1: 120, Section 2: 100, Section 3: 60.

(a) Describe how the sample could be selected using stratified sampling, and find the number to be taken from each group. **(3)**

(b) State one advantage of stratified sampling over simple random sampling. **(1)**

(c) Explain how systematic sampling could be used to select a sample of 56 from the whole population of 280. **(2)**

(Total for Question 4 is 6 marks)

5 A factory produces 720 batteries each day. An inspector tests a sample of 36 of them. The population is split as follows: Group A: 360, Group B: 120, Group C: 240.

(a) Describe how the sample could be selected using stratified sampling, and find the number to be taken from each group. **(3)**

(b) State one advantage of stratified sampling over simple random sampling. **(1)**

(c) Explain how systematic sampling could be used to select a sample of 36 from the whole population of 720. **(2)**

(Total for Question 5 is 6 marks)

6 A teacher puts the names of all 800 students in a hat and draws out 40 names.

(a) Name the sampling method described. **(1)**

(b) State one advantage of this method in this situation. **(1)**

(c) State one disadvantage of this method in this situation. **(1)**

(Total for Question 6 is 3 marks)

7 For each statement below about using a sample rather than a census, state whether the statement describes an advantage or a disadvantage of sampling.

(a) "Results can usually be obtained more quickly." **(1)**

(b) "The results may be less accurate, since conclusions about the whole population are based on only part of it." **(1)**

(c) "Less time and money are usually needed to collect and process the data." **(1)**

(d) "No individual data is available for population members who were not chosen, so no prediction can be made about any one of them specifically." **(1)**

(Total for Question 7 is 4 marks)

8 An interviewer is told to question 11 people, of whom 8 must be women and 15 men, and she stops people in the street until she has the required numbers.

(a) Name the sampling method described. **(1)**

(b) State one advantage of this method in this situation. **(1)**

(c) State one disadvantage of this method in this situation. **(1)**

(d) The population is actually 160 women and 300 men. Show that the quotas are consistent with this population when the sample is 23. **(2)**

(Total for Question 8 is 5 marks)

9 The 370 members of a club are numbered 1 to 370. A systematic sample of 74 members is selected, starting with member 2.

(a) Write down the sampling interval and the numbers of the first four members selected. **(2)**

(b) Find the number of the last member selected. **(2)**

(c) The club's members who are numbered with a multiple of 5 are all of one type. Explain why the sample may be biased. **(2)**

(Total for Question 9 is 6 marks)

10 A market researcher stands outside a shopping centre and asks the first 40 adults who walk past to complete a questionnaire.

(a) Name the sampling method described. **(1)**

(b) State one advantage of this method in this situation. **(1)**

(c) State one disadvantage of this method in this situation. **(1)**

(Total for Question 10 is 3 marks)

11 An interviewer is told to question 45 people, of whom 29 must be women and 16 men, and she stops people in the street until she has the required numbers.

(a) Name the sampling method described. **(1)**

(b) State one advantage of this method in this situation. **(1)**

(c) State one disadvantage of this method in this situation. **(1)**

(d) The population is actually 580 women and 320 men. Show that the quotas are consistent with this population when the sample is 45. **(2)**

(Total for Question 11 is 5 marks)

12 The 140 members of a club are numbered 1 to 140. A systematic sample of 28 members is selected, starting with member 5.

(a) Write down the sampling interval and the numbers of the first four members selected. **(2)**

(b) Find the number of the last member selected. **(2)**

(c) The club's members who are numbered with a multiple of 5 are all of one type. Explain why the sample may be biased. **(2)**

(Total for Question 12 is 6 marks)

13 An interviewer is told to question 15 people, of whom 13 must be women and 15 men, and she stops people in the street until she has the required numbers.

(a) Name the sampling method described. **(1)**

(b) State one advantage of this method in this situation. **(1)**

(c) State one disadvantage of this method in this situation. **(1)**

(d) The population is actually 260 women and 300 men. Show that the quotas are consistent with this population when the sample is 28. **(2)**

(Total for Question 13 is 5 marks)

14 The 400 members of a club are numbered 1 to 400. A systematic sample of 16 members is selected, starting with member 17.

(a) Write down the sampling interval and the numbers of the first four members selected. **(2)**

(b) Find the number of the last member selected. **(2)**

(c) The club's members who are numbered with a multiple of 25 are all of one type. Explain why the sample may be biased. **(2)**

(Total for Question 14 is 6 marks)

15 An interviewer is told to question 47 people, of whom 19 must be women and 28 men, and she stops people in the street until she has the required numbers.

(a) Name the sampling method described. **(1)**

(b) State one advantage of this method in this situation. **(1)**

(c) State one disadvantage of this method in this situation. **(1)**

(d) The population is actually 380 women and 560 men. Show that the quotas are consistent with this population when the sample is 47. **(2)**

(Total for Question 15 is 5 marks)

16 A factory produces 340 batteries each day. An inspector tests a sample of 34 of them. The population is split as follows: Group A: 140, Group B: 80, Group C: 120.

(a) Describe how the sample could be selected using stratified sampling, and find the number to be taken from each group. **(3)**

(b) State one advantage of stratified sampling over simple random sampling. **(1)**

(c) Explain how systematic sampling could be used to select a sample of 34 from the whole population of 340. **(2)**

(Total for Question 16 is 6 marks)

17 Priya is a student researcher who wants to survey students at her large sixth-form college about their travel-to-college habits. She obtains a printed list of all students enrolled on A-level and BTEC courses from the college's admissions database, dated 1 September. She plans to carry out her survey the following March.

(a) State what is meant by a sampling frame. **(1)**

(b) Explain one way in which the admissions database, printed on 1 September, might not be a suitable sampling frame for Priya's survey in March. **(2)**

(c) Suggest one improvement Priya could make to obtain a more suitable sampling frame. **(1)**

(Total for Question 17 is 4 marks)

18 A council wants to survey households. It lists all 500 households and chooses every 20th one after a random start.

(a) Name the sampling method described. **(1)**

(b) State one advantage of this method in this situation. **(1)**

(c) State one disadvantage of this method in this situation. **(1)**

(Total for Question 18 is 3 marks)

19 The 340 members of a club are numbered 1 to 340. A systematic sample of 68 members is selected, starting with member 4.

(a) Write down the sampling interval and the numbers of the first four members selected. **(2)**

(b) Find the number of the last member selected. **(2)**

(c) The club's members who are numbered with a multiple of 5 are all of one type. Explain why the sample may be biased. **(2)**

(Total for Question 19 is 6 marks)

20 A factory produces 120 batteries each day. An inspector tests a sample of 24 of them. The population is split as follows: Group A: 30, Group B: 30, Group C: 60.

(a) Describe how the sample could be selected using stratified sampling, and find the number to be taken from each group. **(3)**

(b) State one advantage of stratified sampling over simple random sampling. **(1)**

(c) Explain how systematic sampling could be used to select a sample of 24 from the whole population of 120. **(2)**

(Total for Question 20 is 6 marks)

21 A factory produces 580 batteries each day. An inspector tests a sample of 58 of them. The population is split as follows: Group A: 220, Group B: 220, Group C: 140.

(a) Describe how the sample could be selected using stratified sampling, and find the number to be taken from each group. **(3)**

(b) State one advantage of stratified sampling over simple random sampling. **(1)**

(c) Explain how systematic sampling could be used to select a sample of 58 from the whole population of 580. **(2)**

(Total for Question 21 is 6 marks)

22 The 292 members of a club are numbered 1 to 292. A systematic sample of 73 members is selected, starting with member 3.

(a) Write down the sampling interval and the numbers of the first four members selected. **(2)**

(b) Find the number of the last member selected. **(2)**

(c) The club's members who are numbered with a multiple of 4 are all of one type. Explain why the sample may be biased. **(2)**

(Total for Question 22 is 6 marks)

23 A supermarket had 850 customers pass through its tills between 9 a.m. and 5 p.m. last Saturday. The store manager wants to select a systematic sample of 25 of these customers to survey about checkout waiting times.

(a) Calculate the sampling interval, k, that the manager should use. **(2)**

(b) Describe fully how the manager should select the systematic sample of 25 customers, using the value of k found in part (a). **(2)**

(c) The manager considers only sampling customers who pass through the tills between 9 a.m. and 10 a.m. each day, to make identifying the sample quicker. Explain why this could introduce bias into the results. **(2)**

(Total for Question 23 is 6 marks)

24 State what is meant by a 'population' in the context of statistical sampling.

(Total for Question 24 is 1 mark)

25 State what is meant by a 'sample'.

(Total for Question 25 is 1 mark)

---

26 State what is meant by a 'sampling frame'.

(Total for Question 26 is 1 mark)

---

27 A local council wants to find out residents' opinions on a proposed new cycle lane. It posts a survey link on the council's official social media page, and asks residents to comment with their views.

(a) State the sampling method used. **(1)**

(b) Explain two separate ways in which this method could lead to a biased sample of residents' opinions. **(2)**

(c) Suggest a more appropriate sampling method the council could use, given that it holds the electoral register for the area. **(1)**

(Total for Question 27 is 4 marks)

28 A quality inspector picks a random starting point on a production line and then selects every 20th item after that. State the name of this sampling method.

(Total for Question 28 is 1 mark)

29 A researcher stands outside a supermarket and interviews the first 40 shoppers who agree to stop and answer her questions. State the name of this sampling method.

(Total for Question 29 is 1 mark)

30 A factory produces 464 identical components during a shift. A systematic sample of 8 components is required. Calculate the sampling interval that should be used.

(Total for Question 30 is 2 marks)

31 A college has 936 students enrolled. A systematic sample of 24 students is required. Calculate the sampling interval that should be used.

(Total for Question 31 is 2 marks)

---

32 A university researcher is investigating the exercise habits of the 18000 students at a large university, who live across 12 halls of residence.

(a) The researcher plans to survey the first 50 students who arrive at the university gym one Monday morning. State the name of this sampling method. **(1)**

(b) Explain one disadvantage of the method in part (a) for this particular study. **(2)**

(c) As an alternative, the researcher randomly selects 3 of the 12 halls of residence and surveys every student living in those 3 halls. State the name of this sampling method, and give one advantage of using it instead of the method in part (a). **(2)**

(Total for Question 32 is 5 marks)

33 A sixth-form college has 840 students: 460 in Year 12 and 380 in Year 13. The college wants to take a stratified sample of 60 students, by year group, to survey about wellbeing support.

(a) Show that the number of Year 13 students required in the sample is 27, to the nearest whole number. **(2)**

(b) Find the number of Year 12 students required in the sample. **(2)**

(c) Once the number of students required from Year 12 has been found, explain how the college should select which Year 12 students to include. **(1)**

(Total for Question 33 is 5 marks)

---

34 A hospital has 320 patients across three wards: Ward A has 140 patients, Ward B has 100 patients and Ward C has 80 patients. A stratified sample of 40 patients, proportional to ward size, is required. Calculate the number of patients that should be sampled from Ward C.

(Total for Question 34 is 2 marks)

35 A cinema chain has three branches with 900, 600 and 300 regular customers respectively. A stratified sample of 60 customers, proportional to branch size, is required. Calculate the number of customers that should be sampled from the branch with 600 regular customers.

(Total for Question 35 is 2 marks)

---

36 In a certain town, 45% of residents are male and 55% are female. A market researcher plans to interview a quota sample of 80 residents, in these proportions. Calculate the number of male residents she should aim to interview.

(Total for Question 36 is 2 marks)

37 A scientist records the daily maximum temperature at a weather station for every day of one calendar year (365 days): 90 days in Winter, 92 days in Spring, 92 days in Summer and 91 days in Autumn. She wants to select a sample of 20 days for detailed analysis, with the number of days chosen from each season in proportion to the number of days in that season.

(a) State the name of the sampling method described. **(1)**

(b) Calculate the number of days that should be sampled from Summer. **(2)**

(c) Calculate the number of days that should be sampled from Winter, giving your answer to the nearest whole number. **(2)**

(d) Explain why, when using stratified sampling with proportional allocation, rounding each stratum's calculated sample size independently can sometimes cause a problem, and state how this can be resolved. **(2)**

(Total for Question 37 is 7 marks)

38 A company has 640 staff across two sites: Site A has 400 staff and Site B has 240 staff. A stratified sample of 40 staff, proportional to site size, is required. Of the staff sampled from Site A, 60% work in Operations and the rest work in Support. Find the number of Operations staff at Site A that should be included in the sample.

(Total for Question 38 is 3 marks)

---

39 A school has 1200 students across three year groups: Year 9 has 440 students, Year 10 has 400 students and Year 11 has 360 students. A stratified sample of 60 students, proportional to year group size, is required. Calculate the number of Year 10 students that should be included in the sample, and describe how the individual Year 10 students should then be chosen so that they form a simple random sample within that year group.

(Total for Question 39 is 3 marks)

40 A factory production line manufactures 1350 identical components during a shift. The quality control team wants a systematic sample of 40 components.

(a) Calculate the sampling interval, giving your answer to the nearest whole number. **(2)**

(b) Using an interval of 34, explain why selecting every 34th component from a random start may result in fewer than 40 components being sampled during the shift, and determine the smallest number of components that could be sampled. **(3)**

(c) Suggest one way the quality control team could adjust their method to obtain the full sample of exactly 40 components. **(1)**

(Total for Question 40 is 6 marks)

41 A firm employs 300 staff across three departments: Sales (120 staff: 72 male, 48 female), Production (150 staff: 90 male, 60 female) and Admin (30 staff: 12 male, 18 female). Human Resources wants a stratified sample of 40 staff, proportional to department size.

(a) Show that the number of staff sampled from the Production department is 20. **(2)**

(b) Find the number of staff that should be sampled from the Sales department and from the Admin department. **(3)**

(c) Within the Sales department, 72 of the 120 staff are male and 48 are female. Using the Sales sample size found in part (b), calculate how many male and how many female staff should be included in the Sales sample, rounding appropriately so that your two values sum to 16. **(3)**

(Total for Question 41 is 8 marks)

42 A charity has 730 volunteers across four regions: North (210), South (180), East (160) and West (180). A stratified sample of 50 volunteers, proportional to region size, is required, with each region's value rounded to the nearest whole number. Calculate the number of volunteers that should be sampled from each region, and explain why the four rounded values might not sum to exactly 50.

(Total for Question 42 is 4 marks)

Mark scheme · 2.1 Statistics: Sampling

Question 1

(a) 160 / 500 x 25 **M1**

(a) cao 8 **A1**

(a) Answer: 8 Premium members.

(b) sum of the 8 scores divided by 8 **M1**

(b) cao 8 **A1**

(b) Answer: Mean = 8.

(c) 12 x 6.5 + 8 x 8 + 5 x 7.6 **M1**

(c) = 78 + 64 + 38 = 180 **M1**

(c) cao 180 / 25 = 7.2 **A1**

(c) Answer: Overall mean = 7.2.

Question 2

(a) sampling fraction 48/240 = 1/5, applied to each group **M1**

(a) numbers from each group: Year 7 16, Year 8 10, Year 9 22 **A1**

(a) within each group the items are selected at random (for example by random number generator) **B1**

(a) Answer: Year 7 16, Year 8 10, Year 9 22

(b) every group is represented in proportion, so the sample is more representative of the population **B1**

(b) Answer: more representative of each group

(c) sampling interval 240/48 = 5 **M1**

(c) choose a random start between 1 and 5 and then select every 5th item from the numbered list **A1**

(c) Answer: every 5th item after a random start from 1 to 5

Question 3

(a) sampling fraction 42/840 = 1/20, applied to each group **M1**

(a) numbers from each group: Year 7 18, Year 8 6, Year 9 18 **A1**

(a) within each group the items are selected at random (for example by random number generator) **B1**

(a) Answer: Year 7 18, Year 8 6, Year 9 18

(b) every group is represented in proportion, so the sample is more representative of the population **B1**

(b) Answer: more representative of each group

(c) sampling interval 840/42 = 20 **M1**

(c) choose a random start between 1 and 20 and then select every 20th item from the numbered list **A1**

(c) Answer: every 20th item after a random start from 1 to 20

Question 4

(a) sampling fraction 56/280 = 1/5, applied to each group **M1**

(a) numbers from each group: Section 1 24, Section 2 20, Section 3 12 **A1**

(a) within each group the items are selected at random (for example by random number generator) **B1**

(a) Answer: Section 1 24, Section 2 20, Section 3 12

(b) every group is represented in proportion, so the sample is more representative of the population **B1**

(b) Answer: more representative of each group

(c) sampling interval 280/56 = 5 **M1**

(c) choose a random start between 1 and 5 and then select every 5th item from the numbered list **A1**

(c) Answer: every 5th item after a random start from 1 to 5

Question 5

(a) sampling fraction 36/720 = 1/20, applied to each group **M1**

(a) numbers from each group: Group A 18, Group B 6, Group C 12 **A1**

(a) within each group the items are selected at random (for example by random number generator) **B1**

(a) Answer: Group A 18, Group B 6, Group C 12

(b) every group is represented in proportion, so the sample is more representative of the population **B1**

(b) Answer: more representative of each group

(c) sampling interval 720/36 = 20 **M1**

(c) choose a random start between 1 and 20 and then select every 20th item from the numbered list **A1**

(c) Answer: every 20th item after a random start from 1 to 20

Question 6

(a) simple random sampling **B1**

(a) Answer: simple random sampling

(b) every student has an equal chance of being selected, so the sample is unbiased **B1**

(b) Answer: every student has an equal chance of being selected, so the sample is unbiased

(c) needs a complete list of the population and may be slow for a large population **B1**

(c) Answer: needs a complete list of the population and may be slow for a large population

Question 7

(a) advantage **B1**

(a) Answer: Advantage.

(b) disadvantage **B1**

(b) Answer: Disadvantage.

(c) advantage **B1**

(c) Answer: Advantage.

(d) disadvantage **B1**

(d) Answer: Disadvantage.

Question 8

(a) quota sampling **B1**

(a) Answer: quota sampling

(b) the sample reflects the known proportions in the population, and no sampling frame is needed **B1**

(b) Answer: the sample reflects the known proportions in the population, and no sampling frame is needed

(c) the choice of individuals is not random, so interviewer bias can occur **B1**

(c) Answer: the choice of individuals is not random, so interviewer bias can occur

(d) proportion of women = 160/460 **M1**

(d) 160/460 x 23 = 8 women, matching the quota (cso) **A1**

(d) Answer: 8 women and 15 men

Question 9

(a) interval = 370/74 = 5 **B1**

(a) 2, 7, 12, 17 **B1**

(a) Answer: interval 5; 2, 7, 12, 17

(b) 2 + (74 - 1) x 5 **M1**

(b) 367 **A1**

(b) Answer: 367

(c) if the numbering follows a pattern with the same period as the sampling interval, the sample may contain only that type **B1**

(c) so each member does not have an equal chance of selection when the list is ordered in a periodic way **B1**

(c) Answer: possible bias from a periodic list

Question 10

(a) opportunity (convenience) sampling **B1**

(a) Answer: opportunity (convenience) sampling

(b) quick, cheap and needs no sampling frame **B1**

(b) Answer: quick, cheap and needs no sampling frame

(c) not random, so the sample may not be representative of the population **B1**

(c) Answer: not random, so the sample may not be representative of the population

Question 11

(a) quota sampling **B1**

(a) Answer: quota sampling

(b) the sample reflects the known proportions in the population, and no sampling frame is needed **B1**

(b) Answer: the sample reflects the known proportions in the population, and no sampling frame is needed

(c) the choice of individuals is not random, so interviewer bias can occur **B1**

(c) Answer: the choice of individuals is not random, so interviewer bias can occur

(d) proportion of women = 580/900 **M1**

(d) 580/900 x 45 = 29 women, matching the quota (cso) **A1**

(d) Answer: 29 women and 16 men

Question 12

(a) interval = 140/28 = 5 **B1**

(a) 5, 10, 15, 20 **B1**

(a) Answer: interval 5; 5, 10, 15, 20

(b) 5 + (28 - 1) x 5 **M1**

(b) 140 **A1**

(b) Answer: 140

(c) if the numbering follows a pattern with the same period as the sampling interval, the sample may contain only that type **B1**

(c) so each member does not have an equal chance of selection when the list is ordered in a periodic way **B1**

(c) Answer: possible bias from a periodic list

Question 13

(a) quota sampling **B1**

(a) Answer: quota sampling

(b) the sample reflects the known proportions in the population, and no sampling frame is needed **B1**

(b) Answer: the sample reflects the known proportions in the population, and no sampling frame is needed

(c) the choice of individuals is not random, so interviewer bias can occur **B1**

(c) Answer: the choice of individuals is not random, so interviewer bias can occur

(d) proportion of women = 260/560 **M1**

(d) 260/560 x 28 = 13 women, matching the quota (cso) **A1**

(d) Answer: 13 women and 15 men

Question 14

(a) interval = 400/16 = 25 **B1**

(a) 17, 42, 67, 92 **B1**

(a) Answer: interval 25; 17, 42, 67, 92

(b) 17 + (16 - 1) x 25 **M1**

(b) 392 **A1**

(b) Answer: 392

(c) if the numbering follows a pattern with the same period as the sampling interval, the sample may contain only that type **B1**

(c) so each member does not have an equal chance of selection when the list is ordered in a periodic way **B1**

(c) Answer: possible bias from a periodic list

Question 15

(a) quota sampling **B1**

(a) Answer: quota sampling

(b) the sample reflects the known proportions in the population, and no sampling frame is needed **B1**

(b) Answer: the sample reflects the known proportions in the population, and no sampling frame is needed

(c) the choice of individuals is not random, so interviewer bias can occur **B1**

(c) Answer: the choice of individuals is not random, so interviewer bias can occur

(d) proportion of women = 380/940 **M1**

(d) 380/940 x 47 = 19 women, matching the quota (cso) **A1**

(d) Answer: 19 women and 28 men

Question 16

(a) sampling fraction 34/340 = 1/10, applied to each group **M1**

(a) numbers from each group: Group A 14, Group B 8, Group C 12 **A1**

(a) within each group the items are selected at random (for example by random number generator) **B1**

(a) Answer: Group A 14, Group B 8, Group C 12

(b) every group is represented in proportion, so the sample is more representative of the population **B1**

(b) Answer: more representative of each group

(c) sampling interval 340/34 = 10 **M1**

(c) choose a random start between 1 and 10 and then select every 10th item from the numbered list **A1**

(c) Answer: every 10th item after a random start from 1 to 10

Question 17

(a) a list of all the members of the population, from which the sample is selected, oe **B1**

(a) Answer: A list of all the members of the population, used to select the sample.

(b) identifies that some students will have left their course and new students will have joined between September and March, so the list is out of date **B1**

(b) explains the consequence: some students currently at the college have no chance of selection, and some listed students may no longer attend, so the frame does not accurately represent the population being studied in March, oe **B1**

(b) Answer: The list is six months out of date, so some students who have since left cannot actually be surveyed, and new students who joined after September have no chance of being selected, meaning the frame does not accurately represent the population in March.

(c) request an updated/current enrolment list from the college, printed close to the date of the survey, oe **B1**

(c) Answer: Ask the college for an up to date enrolment list, printed close to March when the survey will actually take place.

Question 18

(a) systematic sampling **B1**

(a) Answer: systematic sampling

(b) simple and quick to carry out, and it spreads the sample evenly across the list **B1**

(b) Answer: simple and quick to carry out, and it spreads the sample evenly across the list

(c) may be biased if the list has a periodic pattern **B1**

(c) Answer: may be biased if the list has a periodic pattern

Question 19

(a) interval = 340/68 = 5 **B1**

(a) 4, 9, 14, 19 **B1**

(a) Answer: interval 5; 4, 9, 14, 19

(b) 4 + (68 - 1) x 5 **M1**

(b) 339 **A1**

(b) Answer: 339

(c) if the numbering follows a pattern with the same period as the sampling interval, the sample may contain only that type **B1**

(c) so each member does not have an equal chance of selection when the list is ordered in a periodic way **B1**

(c) Answer: possible bias from a periodic list

Question 20

(a) sampling fraction 24/120 = 1/5, applied to each group **M1**

(a) numbers from each group: Group A 6, Group B 6, Group C 12 **A1**

(a) within each group the items are selected at random (for example by random number generator) **B1**

(a) Answer: Group A 6, Group B 6, Group C 12

(b) every group is represented in proportion, so the sample is more representative of the population **B1**

(b) Answer: more representative of each group

(c) sampling interval 120/24 = 5 **M1**

(c) choose a random start between 1 and 5 and then select every 5th item from the numbered list **A1**

(c) Answer: every 5th item after a random start from 1 to 5

Question 21

(a) sampling fraction 58/580 = 1/10, applied to each group **M1**

(a) numbers from each group: Group A 22, Group B 22, Group C 14 **A1**

(a) within each group the items are selected at random (for example by random number generator) **B1**

(a) Answer: Group A 22, Group B 22, Group C 14

(b) every group is represented in proportion, so the sample is more representative of the population **B1**

(b) Answer: more representative of each group

(c) sampling interval 580/58 = 10 **M1**

(c) choose a random start between 1 and 10 and then select every 10th item from the numbered list **A1**

(c) Answer: every 10th item after a random start from 1 to 10

Question 22

(a) interval = 292/73 = 4 **B1**

(a) 3, 7, 11, 15 **B1**

(a) Answer: interval 4; 3, 7, 11, 15

(b) 3 + (73 - 1) x 4 **M1**

(b) 291 **A1**

(b) Answer: 291

(c) if the numbering follows a pattern with the same period as the sampling interval, the sample may contain only that type **B1**

(c) so each member does not have an equal chance of selection when the list is ordered in a periodic way **B1**

(c) Answer: possible bias from a periodic list

Question 23

(a) 850 / 25 **M1**

(a) cao k = 34 **A1**

(a) Answer: k = 34

(b) choose a random starting point between 1 and k (i.e. between 1 and 34) **B1**

(b) then select every 34th customer after that point (e.g. by adding 34 repeatedly), until 25 customers have been chosen **B1**

(b) Answer: Pick a random start between 1 and 34, then select every 34th customer after that until 25 customers are chosen.

(c) customers shopping between 9 and 10 a.m. may differ systematically from customers shopping at other times, e.g. more likely to be doing a quick shop before work **B1**

(c) so the sample would not represent all types of customer, and conclusions about waiting times could not be generalised to the whole day, oe **B1**

(c) Answer: Early-morning shoppers may differ from customers at other times of day (for example, doing a quick shop before work), so the sample would not represent all types of customer and results could not be generalised to the whole day.

Question 24

the whole set of individuals or items that information is wanted about, oe **B1**

Answer: The whole set of individuals or items about which information is wanted.

Question 25

a subset of the population, selected for study, oe **B1**

Answer: A subset of the population that is actually selected and studied.

Question 26

a list of all the members of the population, from which the sample is selected, oe **B1**

Answer: A list of all the members of the population, from which the sample is selected.

Question 27

(a) self-selecting sample (volunteer sample) **B1**

(a) Answer: Self-selecting sample (volunteer sample).

(b) only residents who use/follow that particular social media platform can respond, so residents without an account or who are less active online, e.g. some older residents, are excluded **B1**

(b) residents with strong opinions, for or against the cycle lane, are more likely to respond than those who feel neutral, so the sample may over-represent extreme views, oe **B1**

(b) Answer: Residents who do not use that social media platform, such as some older residents, cannot respond at all; and residents with strong feelings, for or against the cycle lane, are more likely to comment than those who are neutral, so the results may over-represent extreme opinions.

(c) take a simple random sample or a stratified sample of residents from the electoral register, oe **B1**

(c) Answer: Take a simple random sample (or a stratified sample) of residents directly from the electoral register.

Question 28

systematic sampling **B1**

Answer: Systematic sampling.

Question 29

opportunity sampling (convenience sampling) **B1**

Answer: Opportunity sampling (convenience sampling).

Question 30

464 / 8 **M1**

58 cao **A1**

Answer: 58

Question 31

936 / 24 **M1**

39 cao **A1**

Answer: 39

Question 32

(a) opportunity sampling (convenience sampling) **B1**

(a) Answer: Opportunity sampling (convenience sampling).

(b) only students who already choose to use the gym can be included in the sample **B1**

(b) so students with different exercise habits, e.g. those who never exercise or who exercise elsewhere, have no chance of selection, which is likely to bias the results, oe **B1**

(b) Answer: Only students who already attend the gym can be sampled, so students who never exercise or who exercise elsewhere have no chance of selection, biasing the results towards more active students.

(c) cluster sampling **B1**

(c) advantage, e.g. quicker/cheaper to carry out since students are grouped geographically, and every type of student (not just gym-users) has a chance of selection, oe **B1**

(c) Answer: Cluster sampling. It is quicker and cheaper to carry out, since students are grouped geographically in halls, and it includes all types of student, not just those who already use the gym.

Question 33

(a) 380 / 840 x 60 **M1**

(a) cso: 27.14... which rounds to 27 (answer given, exact working must be shown) **A1**

(a) Answer: 27 (given)

(b) 460 / 840 x 60, or 60 - 27 **M1**

(b) cao 33 **A1**

(b) Answer: 33 students.

(c) use simple random sampling (or systematic sampling) to choose 33 students at random from the Year 12 group, so each Year 12 student has an equal chance of selection, oe **B1**

(c) Answer: Use simple random sampling to choose 33 students at random from within Year 12, so every Year 12 student has an equal chance of being selected.

Question 34

80 / 320 x 40 **M1**

10 cao **A1**

Answer: 10

Question 35

600 / 1800 x 60 **M1**

20 cao **A1**

Answer: 20

Question 36

0.45 x 80 **M1**

36 cao **A1**

Answer: 36

Question 37

(a) stratified sampling **B1**

(a) Answer: Stratified sampling.

(b) 92 / 365 x 20 **M1**

(b) awrt 5 days **A1**

(b) Answer: 5 days (to the nearest whole number).

(c) 90 / 365 x 20 **M1**

(c) awrt 5 days **A1**

(c) Answer: 5 days.

(d) the individually rounded values may not sum to the required overall sample size, e.g. they could sum to one more or one fewer than needed **B1**

(d) one (or more) of the values must then be adjusted, typically the value(s) closest to rounding the other way, so that the strata sizes sum to the required total, oe **B1**

(d) Answer: Rounding each stratum's value separately may mean the totals do not add up to the required sample size, e.g. 20; this is resolved by adjusting one of the values (usually the one closest to rounding the other way) until the totals sum correctly.

Question 38

400 / 640 x 40 (= 25, the Site A sample size) **M1**

0.6 x 25 (method to find the Operations count) **M1**

15 cao **A1**

Answer: 15

Question 39

400 / 1200 x 60 **M1**

20 cao **A1**

use simple random sampling (e.g. a random number generator) to choose 20 students from within Year 10, so every Year 10 student has an equal chance of selection **B1**

Answer: 20 students; select them using simple random sampling (e.g. a random number generator) from within Year 10, so every Year 10 student has an equal chance of being chosen.

Question 40

(a) 1350 / 40 (= 33.75) **M1**

(a) cao 34 **A1**

(a) Answer: Interval = 34.

(b) 34 x 40 = 1360, which is greater than 1350, so for a late random start there are not enough components in the shift to make 40 selections at this interval **B1**

(b) 1350 / 34 **M1**

(b) cao: 1350/34 = 39.7..., so the smallest possible sample is 39 components **A1**

(b) Answer: The smallest possible sample is 39 components.

(c) reduce the sampling interval to 33 (since 1350/40 = 33.75 rounded down), or treat the production line as circular and wrap around to the start once the end is reached, oe **B1**

(c) Answer: Use a smaller interval of 33 instead of 34 (or treat the list of components as circular, wrapping back to the start once the end of the shift is reached).

Question 41

(a) 150 / 300 x 40 **M1**

(a) cso 20 (answer given, exact working must be shown) **A1**

(a) Answer: 20 (given)

(b) 120 / 300 x 40 (or equivalent method for Admin) **M1**

(b) Sales = 16 **A1**

(b) Admin = 4, cao (check: 16 + 20 + 4 = 40) **B1**

(b) Answer: Sales: 16, Admin: 4.

(c) 16 x 72/120 (= 9.6) **M1**

(c) 16 x 48/120 (= 6.4) **M1**

(c) cao: male = 10, female = 6 **A1**

(c) Answer: Male: 10, Female: 6.

Question 42

calculates at least one stratum using (region size / 730) x 50 **M1**

North 14, South 12, East 11, West 12 cao **A1**

identifies that these four values sum to 49, not 50 **B1**

explains that rounding each stratum independently does not guarantee the rounded values sum to the required total, so one value (the one whose exact figure sits closest to the next rounding boundary, here North at 14.38) must be adjusted upward to restore a total of 50, oe **B1**

Answer: North 14, South 12, East 11, West 12 (total 49, not 50); since each value was rounded independently, the totals need not sum to 50 - the value closest to the next rounding boundary (North, 14.38) should be adjusted up to 15 to restore the total, giving final sample sizes North 15, South 12, East 11, West 12.