Name: _____

GCSE Statistics

Sampling Methods

Total marks available: 80

Total marks achieved: _____

Instructions
• Use black ink or ball-point pen.
• Fill in the boxes at the top of this page with your name, centre number and candidate number.
• Answer all questions.
• Answer the questions in the spaces provided – there may be more space than you need.
• Scientific calculators may be used.
• You must show all your working out with your answer clearly identified at the end of your solution.

Information
• The marks for each question are shown in brackets – use this as a guide as to how much time to spend on each question.

Advice
• Read each question carefully before you start to answer it.
• Try to answer every question.
• Check your answers if you have time at the end.

Q1. The table shows information about houses for sale in Oxford.

Number of bedrooms	1	2	3	4	5 or more	Total
Number of houses for sale	140	300	420	240	100	1200

(Source: *adapted from rightmove.co.uk*)

The estate agent wants to investigate the prices of these houses.

She takes a stratified sample of 60 houses according to the number of bedrooms.

(a) Work out the number of houses in her sample for each number of bedrooms.

Number of bedrooms	1	2	3	4	5 or more
Number of houses in the sample					

(3)

(b) Describe how to select the 60 houses in the sample.

.....

.....

.....

.....

.....

.....

.....

.....

(3)

(Total for question = 6 marks)

Q2. Tom and Samira want to collect data on the numbers of hours students at their school spend on homework.

There are 1100 students at their school.

Tom is planning to use a random sample of 50 students.

(a) Explain what is meant by a random sample.

.....

.....

.....

(1)

(b) Describe how Tom could use random numbers to take a random sample of the students at his school.

.....

.....

.....

.....

(3)

Samira is planning to use a stratified sample that is stratified by school year.

(c) Comment on whether Samira's plan is appropriate.

.....

.....

.....

.....

(2)

(Total for question = 6 marks)

Q3. Jaspreet wants to find out about the method of transport that students use to get to school. She plans to ask the first 20 students arriving at school one morning how they have travelled to school.

(a) Give one advantage and one disadvantage of Jaspreet's plan.

Advantage.

.....

.....

Disadvantage.

.....

.....

(2)

Rosemary thinks it would be better to take a random sample.
She is going to use a database of all of the students at the school to select her random sample.

(b) Explain how Rosemary can use the school database to select a random sample of 20 students.

.....

.....

.....

.....

.....

.....

(3)

(Total for question = 5 marks)

Q4. Stacey is the manager of a large company.

She wants to investigate the differences between the numbers of days of holiday taken by employees in different age groups.

She obtains the following information about the ages of the employees.

Age group	Number of employees	Number of males	Number of females
20 to 34	106	48	58
35 to 49	135	73	62
50 to 65	84	61	23

Assess the suitability of taking a sample of 20 employees stratified by age and by gender, for her investigation.

.....

.....

.....

.....

.....

.....

.....

(Total for question = 3 marks)

Q5. Hamish wants to take a sample of bus passengers.

Hamish plans to take a sample of 50 men and 50 women.

Name this sampling method and explain why this might be an appropriate method to use.

.....

.....

.....

.....

(2)

(Total for question = 2 marks)

Q6.

One member of parliament in the UK wants to investigate the ages of the people living in her constituency.

She suggests using the electoral register as a sample frame for her investigation.

(a) State one use of a sample frame in an investigation.

.....

.....

(1)

(b) Assess the suitability of using the electoral register as a sample frame for this investigation.

.....

.....

.....

.....

(2)

(Total for question = 3 marks)

Q7. The mayor of a town wants a new library built.

He wants to survey the residents of the town to find out where each resident wants the library to be built.

The mayor has to decide between three different sampling methods for his survey.

Method 1: Opportunity sampling by sampling the first 100 residents who are available.

Method 2: Cluster sampling by dividing the town into 8 districts, selecting one of the districts at random and surveying all the residents in that district.

Method 3: Stratified sampling by dividing the residents into groups by age and randomly selecting a sample of residents from each age group in proportion to the number of residents in that group.

(a) Give one **advantage** of Method 1

.....

.....

(1)

(b) Give one **disadvantage** of Method 2

.....

.....

(1)

There are 6200 residents in the town of which 930 are aged 60 or over.

A sample of 150 residents of the town is taken, stratified by age.

(c) Calculate the number of residents aged 60 or over who should be included in the sample.

.....

(2)

The mayor wants the sample to be representative of the residents of the town.

(d) Which of the three methods of sampling should the mayor use?
Give a reason for your answer.

.....

.....

.....

(2)

(Total for question = 6 marks)

Q8. There are 11 727 students at a university.
Their nationality is classified as UK, EU or International.

The table shows information about the nationality of these students.

Nationality	UK	EU	International	Total
Number of students	9393	979	1355	11 727

(Source: *www.ox.ac.uk*)

The manager of a book shop wants to carry out a survey into the books read by the students at this university.

She is going to take a sample of 600 of these students.

The manager plans to sample 200 UK students, 200 EU students and 200 International students.

(a) Write down the name of this method of sampling.

.....

(1)

(b) Give a reason why this method of sampling might **not** be appropriate.

.....

.....

(1)

A shop assistant suggests that it would be better to select a sample of 600 students, stratified by nationality.

(c) Work out how many EU students there should be in this sample.

.....

(Total for question = 2 marks)

Q9. At a university, 70% of students are undergraduates and 30% of students are postgraduates.

Amy and Robert want to do a survey.

Amy decides to use simple random sampling to collect a sample of 100 students.

She uses the university database as a sample frame and she numbers each student on the database. She then generates exactly 100 random numbers and uses these random numbers to select her sample.

(a) Give **two** reasons why Amy's method may **not** produce a sample of 100 students.

.....

.....

.....

.....

(2)

Robert decides to use quota sampling to collect a sample of 100 students.
He plans to stand outside the main building until he has interviewed 70 undergraduates and 30 postgraduates.

(b) Give **two** advantages of using quota sampling.

.....

.....

.....

.....

(2)

(c) Explain why this quota sample is **not** a random sample.

.....

.....

(1)

(Total for question = 5 marks)

Q10. The table gives information about the numbers of students from different types of schools who applied to Cambridge University in 2016

Type of school	Applications in 2016		
	Gender		Total
	Male	Female	
Maintained	3674	2899	6573
Independent	1510	1268	2778
Other and Overseas	300	312	612
Total	5484	4479	9963

(Source: *www.cam.ac.uk*)

Richard is going to take a sample of 200 of these students stratified by gender.

(a) Work out how many female students there should be in this sample.

.....

(2)

(b) Describe a situation when it would **not** be appropriate to take a sample stratified by gender.

.....

.....

(1)

Richard could have used a different category for his stratified sample.

(c) What is this different category?

.....

(1)

(Total for question = 4 marks)

Q11. Sanjit writes this hypothesis.

At my school, younger boys are faster at running 400 metres than older girls.

He decides to collect some data from the students at his school to see if his hypothesis is correct.

He plans to ask each student their name, their gender, their best time to run 400 m and their age.

One of these variables is an extraneous variable.

(a) Which one?

.....

(1)

Sanjit decides to take a stratified sample of the students at his school.

He decides to stratify by gender.

(b) By considering Sanjit's hypothesis, explain whether or not this is the best way to stratify.

.....

.....

.....

.....

(2)

(Total for question = 3 marks)

Q12. Samira wrote a questionnaire to investigate mobile phone use by the students at her school.

She decided to select a simple random sample from the 850 students at her school.

(a) Explain what is meant by 'random' in this case.

.....

.....

(1)

(b) State the population for Samira's investigation.

.....

(1)

Samira obtained a list of all the students at her school, numbered 001 to 850, to use as a sampling frame.

Samira decided to select her sample using random numbers generated by her calculator.

Here are her first 16 random numbers.

Terri says these 16 random numbers will not give Samira 16 students for her sample.

(c) Explain why Terri is correct.

Give two reasons.

.....

.....

.....

.....

(2)

Terri suggests that Samira should select her sample by writing all the students' names on pieces of paper, then picking them from a box without looking.

(d) Comment on whether Terri's method is appropriate.

.....

.....

.....

.....

(2)

(Total for question = 6 marks)

Q13.

Reeta read a report that said that these days more 18 to 24 year olds find out what is happening in the news from social media than find out by watching TV.

She decided to investigate how the 12 000 students at her university find out what is happening in the news.

Reeta planned to use a sample of the students at her university stratified by gender and by age.

(a) Explain why this method of sampling would be appropriate.

.....

.....

(1)

Reeta found that she could not get a list of all the students at her university to use for her stratified sample.

Instead she decided she would question students in the university cafeteria.

Here is the data collection sheet Reeta plans to use.

Access to news	Male	Female
Newspapers		
Online subscription news service		
Television news		
Social media		

(b) Discuss whether Reeta's data collection sheet is appropriate for her to use.

You should consider how Reeta might use the responses and describe any problems she may have when she uses the data collection sheet.

.....

.....

.....

.....

.....

.....

(3)

(Total for question = 4 marks)

Q14. John and Susan are investigating films.

John wants to find out the favourite type of film of the students at his school.

Susan wants to find out which type of film made the most money in UK cinemas last year.

(a) State the population for

(i) John's investigation,

.....

.....

(1)

(ii) Susan's investigation.

.....

.....

(1)

Susan plans to collect the information for her investigation from the internet.

(b) Give advice to Susan so that she can ensure that the information she collects is reliable.

.....

.....

(1)

John is deciding between two different sampling methods for his investigation.

Method A: Quota sampling by sampling 20 students from each year group.

Method B: Opportunity sampling by selecting the first 50 students he sees one day.

(c) For each method, give one reason why the method is **not** random.

Method A: .....

.....

Method B: .....

.....

(2)

(d) Explain which of the two methods John should use in order to minimise bias.
Give a reason for your answer.

.....

.....

.....

.....

(2)

(Total for question = 7 marks)

Q15.

Bethany is the union representative at the large company where she works.
She is investigating the salaries of the employees at the company.

Bethany takes a simple random sample of 8 employees at the company and asks each of them to tell her their salary.

Describe what you understand by the term 'simple random sample'.

.....

.....

(Total for question = 1 mark)

Q16.

A firm produces pottery.
The firm produces 1000 plates each week.

The manager of the firm wants to take a systematic sample of 5% of the plates to check for faults.

(a) Describe how the manager should take this sample.

.....

.....

.....

.....

(2)

(b) Give one disadvantage of using systematic sampling.

.....

.....

(1)

(Total for question = 3 marks)

Q17.

Reeta read a report that said that these days more 18 to 24 year olds find out what is happening in the news from social media than find out by watching TV.

She decided to investigate how the 12 000 students at her university find out what is happening in the news.

Reeta wrote the following as a hypothesis:

Do more students get their news from social media than from other sources?

(a) Comment on whether it is appropriate to use this as a hypothesis.

.....

.....

(1)

Reeta planned to use a sample of the students at her university stratified by gender and by age.

(b) Explain why this method of sampling would be appropriate.

.....

.....

(1)

(Total for question = 2 marks)

Q18.

Mehmet wants to conduct an investigation about school meals at his school.
He wants to take a sample of students from his school.

(a) Write down a sampling frame he could use.

.....

.....

(1)

Mehmet has two hypotheses:

- Younger students have school meals, rather than packed lunches, more often than older students do.
- More female students than male students are vegetarian.

Here are sections A, B and C of his plan for his investigation.

A. Sampling method:

Take a sample of 10 students from each of the year groups at school.

B. Questions:

1. What is your gender? Please tick (✓)
   Male ☐ Female ☐
2. What is your year group? Please tick (✓)
   Year 7 ☐ Year 8 ☐ Year 9 ☐ Year 10 ☐ Year 11 ☐ Year 12 ☐ Year 13 ☐
3. What is your favourite meal?.....
4. Do you think eating meat is cruel to animals? Please tick (✓)
   Yes ☐ No ☐
5. How often do you eat school meals? Please tick (✓)
   Once ☐ Two times ☐ Three times ☐ Four times ☐ Five times ☐

C. Presenting data:

Draw a bar chart with age on horizontal axis and total number of school meals on vertical axis (to see if younger students have taller bars).

Draw a pictogram showing number of vegetarian females and number of vegetarian males.

(b) Name the sampling method that Mehmet plans to use.

.....

(1)

(Total for question = 2 marks)

Q19. Emily and Jess work in the head office of a chain of restaurants.

The menu in all of the restaurants has recently been changed.
Emily and Jess have been asked to find out the effect of these changes.

They are going to start by finding out what the kitchen staff think of the new menu.

Emily thinks that they should use a cluster sample of kitchen staff in the restaurants.
The restaurants would be the clusters.

Jess thinks that they should use judgement sampling.
The head chef of the company would be asked to select suitably experienced kitchen staff for the judgement sample.

(a) Discuss whether these two sampling methods are suitable for Emily and Jess to use.

.....

.....

.....

.....

.....

(Total for question = 3 marks)

Q20.

Sue is interested to know how much time children in the UK spend doing chores around the house each week.

She plans to ask a sample of 20 students at her school to record the number of hours they each spend doing chores in a week.

(a) Explain why it is more appropriate for Sue to take a sample rather than a census in this case.

.....

.....

(1)

She will ask each of the 20 students to write down on a piece of paper the time they each spent doing chores around the house in one week.

(b) Describe one problem Sue might encounter during the statistical enquiry process with regard to non-response **or** unexpected outcomes.

You should explain how Sue could try to overcome the problem.

.....

.....

.....

(2)

(Total for question = 3 marks)

Q21

A presenter wants to get feedback on a lecture that she gave.

There were 467 students at the lecture.

The presenter plans to give a questionnaire to a sample of 50 of these students.

One of the questions on the questionnaire is

> To what extent do you agree with the statement, the presenter was knowledgeable?
>
> Use a scale from 0 to 5, where 0 means **I strongly disagree** and 5 means **I strongly agree**

(a) Circle the word from the list that best describes the type of data that this question collects.

ordinal **bivariate** **continuous** **grouped**

(1)

(b) Describe how the presenter can use a list of random numbers to select a simple random sample of 50 students.

.....

.....

.....

.....

.....

.....

(3)

(Total for question = 4 marks)