Planning a statistical investigation

Leaving Certificate · Ordinary Level

Understand the idea

How to start

A population is the whole group you want to learn about. A sample is the smaller group you collect data from.

1. State a question that data can help answer.
2. Choose a sample that can represent the population fairly.
3. Describe what the results show and what the sample cannot tell you.

Key terms & definitions

Population
The full group the investigation is about.

Sample
The part of the population actually studied.

Bias
A systematic tendency to favour some outcomes or groups.

Discrete data
Counted values, usually separate whole numbers.

Key ideas

• Begin with a question that needs data and expects variation, rather than one with a single known answer.

• Define the population, decide how to sample fairly and choose what to measure.

• Distinguish categorical data from numerical data; numerical data may be discrete or continuous.

• Use neutral questions, consistent units and sensible recording methods. Do not collect unnecessary personal information.

• Consider sampling bias, sample size, non-response and measurement error before generalising.

• Describe conclusions in context and state limits. An association does not establish that one variable causes the other.

Quick reminder

• Define the population and question.

• Choose a sample that avoids systematic bias.

• State what the data can and cannot show.

Worked examples

Example 01 · Improve a sample

A student asks only their basketball team how much exercise first years do. Is the sample suitable?

Why this method?

Decide what the question is trying to find out. Then check whether the data and method can support that conclusion.

Population: all first-year students.	**Step 1** The question is about the whole year group, so identify that group first.
Sample used: the basketball team.	**Step 2** This is only one part of the year group.
The sample may overestimate exercise.	**Step 3** Team members are likely to do more organised exercise than many other students.
Choose students randomly across the first-year classes.	**Step 4** This gives students outside the team a chance of selection. Use the same question and time period for everyone.
A sample still has uncertainty.	**Step 5** Students may estimate their exercise inaccurately, and another random sample could give a different result.

Example 02 · Classify variables

Classify favourite sport, number of siblings and journey time to school.

Why this method?

Decide what the question is trying to find out. Then check whether the data and method can support that conclusion.

Categorical: values are categories.	**Step 1** Favourite sport
Numerical and discrete: it is a count.	**Step 2** Number of siblings
Numerical and continuous: in principle it can take any value in a range, even if recorded to whole minutes.	**Step 3** Journey time

Example 03 · A statistical question

Improve: “Do you agree our brilliant new lunch menu is healthier?”

Why this method?

Decide what the question is trying to find out. Then check whether the data and method can support that conclusion.

Use neutral wording rather than brilliant or healthier.	**Step 1** Remove leading language
For example: “How many days in the last school week did you buy a school lunch?”	**Step 2** Define what is measured
Survey a suitable sample and report the distribution. This measures lunch purchases, not health itself.	**Step 3** Collect and interpret

Example 04 · Choose a simple random sample

A school has 240 students. Select 30 for a survey about journey time.

Number the students 1 to 240.	**Step 1** Start with a complete list of the population.

Draw 30 distinct numbers randomly.	**Step 2** Select without replacement so each student has the same chance.
State the limit of the conclusion.	**Final step** A sample from one school does not represent every school.

Common mistake

Assuming a larger sample automatically removes bias. Asking many members of one sports club about favourite sports can still misrepresent the whole school. Use a sampling method suited to the population.

Exam technique

Explain how your sampling method reaches different groups, not just your friends.

Try it yourself

Work on paper before checking the answers at the end. These questions are for self-checking, not submitted homework.

Question 1

A student surveys only the school choir about how much music all students listen to. Explain one problem and suggest an improvement.

Question 2

Classify eye colour, number of pets and measured arm span as categorical, discrete numerical or continuous numerical.

Watch an explanation

Corbettmaths — Sampling: random

Corbettmaths — Questionnaires

Open the website chapter to load the embedded players. Use the rates and rules in your question for financial examples.

Self-check answers

Question 1

The sample may over-represent students interested in music. Select students from across the school using a suitable random sampling method.

A large sample from one club can still be biased. Match the sampling method to the population.

Question 2

Eye colour: categorical. Number of pets: discrete numerical. Arm span: continuous numerical.

Counts take separate whole-number values. A measured length can take any value within a range.

Open the full chapter online