This activity is about collecting data.

Information sheet   Data and sampling

Definitions

Qualitative data	is descriptive data such as colour, gender, type of employee
Quantitative data	is numerical; it can be discrete or continuous
Discrete data	takes only particular, exact values such as number of students in a class or shoe size
Continuous data	can take any value in a range, for example measurements such as height, weight or temperature
Primary data	is collected directly from the population
Secondary data	has been collected previously
Population	all the members of the group being considered
Census	a survey involving all members of the population
Sample	a sub-group of the population
Random sample	a sample taken in such a way that every member of the population has an equal chance of being included
Representative sample	a sample reflecting the population in the proportion of individuals with particular characteristics relevant to the investigation (for example: age, gender)
Biased sample	a sample which is not representative

How to select a representative sample (stratified sample)

- Decide what characteristics of the population are likely to have some bearing on the investigation (for example: age, gender, ethnicity, income).

- Determine how many of the population lie in each of the sub-groups defined by these characteristics.

- Divide the sample in the same proportion.

- Use random sampling to select individuals from each sub-group.

How to use random numbers to select a sample

- List and number each member of the group.

- Decide how many random digits you need to use to enable all members of the group to have an equal chance of being selected.

- Use a calculator or computer to generate random numbers (or use a table of random numbers).

- Use as many of the digits as you need, treating them as whole numbers.

- If a random number is higher than the largest number in your list, simply ignore it and try again. Also ignore any repeat values that occur.

Practical considerations

- What will the data collection cost?

- How long will it take?

- How convenient will it be?
  (for both researchers and members of the sample)

- What is the response rate likely to be?

Designing a questionnaire

- Think carefully about what you want to find out.

- Questions should be short, simple and easy to understand, and easy to answer.

- Avoid ambiguous, leading or personal questions.

- Think about how you will analyse the results.

Try these

1 Describe how you would use random numbers to select a random sample of 20 items from a population of 750.

2 A GP would like to select a random sample of her patients to include in a survey to find out how satisfied patients are with the service they receive from her practice.

Describe how she could select a random sample, and what decisions and problems may be involved.

3 A botanist wants to study the distribution of plants in a random sample of 10 square metres of land, from a rectangular site that is 30 metres long and 20 metres wide.
How could he use random numbers to select a sample?

4 How would you select a representative sample to carry out a survey in the following cases?

a A sample of 500 students from a college to answer questions about canteen facilities.

b A sample of 1000 people from a city to answer questions about car parking.

c A sample of 5000 people from a region to answer questions about a plan for a new international airport in the region.

d A sample of 20 000 people from England to answer questions about regional devolution.

5 A survey is to be carried out in a town to find out what parents of children under five think of the provision of nursery schools in the town.

Describe two possible ways of selecting a sample of 400 such parents, and list the advantages and disadvantages of these methods.

6 The table gives the number of UK employees in different age groups.

Number of employees (000s)

Age (years)	Men	Women
16 – 17	152	192
18 – 24	1736	1590
25 – 34	3569	2955
35 – 49	5691	5012
50 – 64	3930	3453
65+	499	322

Source: Labour Market Status by age group (August 2011) from www.ons.gov.uk

A representative sample of 5000 workers is to be selected to receive a questionnaire by post about health & safety at work.

a Calculate how many workers from each sub-group are required.

b What other characteristics of the workforce are likely to be relevant in this situation?

c What problems are likely to arise in the gathering of information by this method?

Reflect on your work

- What is meant by a **representative** sample?

- What are the main steps in selecting a **stratified** sample?

- What must you bear in mind when designing a **questionnaire**?

- List 3 different ways of collecting views from the general public.
  What are their advantages and disadvantages?

Parking permits?

Parking at college is a problem.
The principal thinks that a permit system might help.
A student's permit could state the times when the student
is expected to be in college.
Unauthorised parking could be punished by ticketing or by
towing vehicles away. These could both result in fines.

Design a questionnaire to find out:

• how students currently use the car park

• whether there is support amongst students for a permit system

• what students think about the suggested punishments for unauthorised parking.

Describe how you would carry out such a survey.

Think about…

• How would you distribute the questionnaire to students?

• How would you encourage students to respond?

• Describe any problems there are likely to be.

The principal suggests that it is only necessary to consult a representative sample of the students.

Describe how you would select a representative sample of 500 students.

Think about…

• What characteristics of the student population do you think are relevant in this situation?

• How would you select the representative sample?