

- A **categorical variable** is one which describes a particular quality or characteristic. It can be divided into **categories**. The information collected is called **categorical data**.

Examples of categorical variables are:

Getting to school: the categories could be train, bus, car and walking.

Colour of eyes: the categories could be blue, brown, hazel, green, grey.

- A **quantitative variable** is one which has a numerical value and is often called a **numerical variable**. The information collected is called **numerical data**.

Quantitative variables can be either discrete or continuous.

A **quantitative discrete variable** takes exact number values and is often a result of **counting**.

Examples of discrete quantitative variables are:

The number of people in a household: the variable could take the values 1, 2, 3,

The score out of 30 for a test: the variable could take the values 0, 1, 2, 3,, 30.

A **quantitative continuous variable** takes numerical values within a certain continuous range. It is usually a result of **measuring**.

Examples of quantitative continuous variables are:

The weight of newborn babies: the variable could take any value on the number line but is likely to be in the range 0.5 kg to 8 kg.

The heights of Year 10 students: the variable would be measured in centimetres. A student whose height is recorded as 145 cm could have exact height anywhere between 144.5 cm and 145.5 cm.

CENSUS OR SAMPLE

The two types of data collection are by census or sample.

A **census** is a method which involves collecting data about every individual in a *whole population*.

The individuals in a population may be people or objects. A census is detailed and accurate but is expensive, time consuming and often impractical.

A **sample** is a method which involves collecting data about a *part of the population* only.

A sample is cheaper and quicker than a census but is not as detailed or as accurate. Conclusions drawn from samples always involve some error.

A sample must truly reflect the characteristics of the whole population and so it must be **unbiased** and **large enough**.

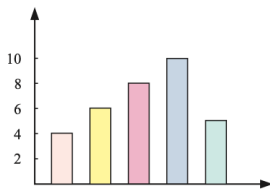
'Large enough' is discussed in year 11.



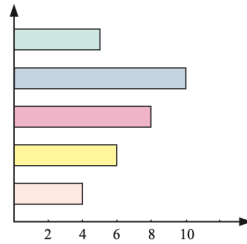
A **biased sample** is one in which the data has been unfairly influenced by the collection process and is not truly representative of the whole population.

Acceptable graphs to display categorical data are:

Vertical column graph



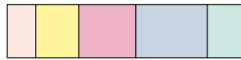
Horizontal bar chart



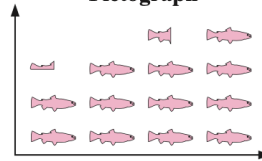
Pie chart



Segment bar chart



Pictograph



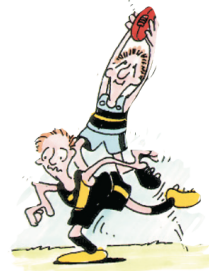
white red blue green

EXERCISE 10A

1 Classify the following variables as either categorical or numerical:

- a the time taken to travel to school
- b the number of cousins a person has
- c voting intention at the next election
- d the number of cars in a household
- e the speed of cars on a particular stretch of highway
- f favourite type of apple
- g province of New Zealand where born
- h the weight of three-year-old children

- 2** Write down the possible categories for the following categorical variables:
- a** gender
 - b** favourite football code
 - c** hair colour
 - d** type of fuel used in a car
- 3** For each of the following possible investigations, classify the variable as categorical, quantitative discrete or quantitative continuous:
- a** the number of goals scored each week by a hockey team
 - b** the weights of the members of a rugby team
 - c** the most popular TV station
 - d** the number of kittens in each litter
 - e** the number of bread rolls bought each week by a family
 - f** the pets owned by students in a year 10 class
 - g** the number of leaves on a rose plant stem
 - h** the amount of sunshine in a day
 - i** the number of people who die from heart attacks each year in New Zealand
 - j** the amount of rainfall in each month of the year
 - k** the countries of origin of immigrants
 - l** the stopping distances of cars doing 60 km/h
 - m** the number of cars passing through an intersection per hour
 - n** the pulse rates of a group of soccer players at rest
- 4** State whether a census or a sample would be used for these investigations:
- a** the reasons for people using taxis
 - b** the heights of the year 10 netballers at a particular school
 - c** finding the percentage of people in New Zealand who suffer from asthma
 - d** the resting pulse rates of members of your favourite sporting team
 - e** finding the country of origin of immigrants to New Zealand
 - f** the amount of daylight each month where you live
- 5** Comment on any possible bias in the following situations:
- a** Year 13 students only are interviewed about changes to the school uniform.
 - b** Motorists stopped in peak hour are interviewed about traffic problems.
 - c** Real estate agents are interviewed about the prices of houses.
 - d** A 'who will you vote for' survey at an expensive city restaurant.



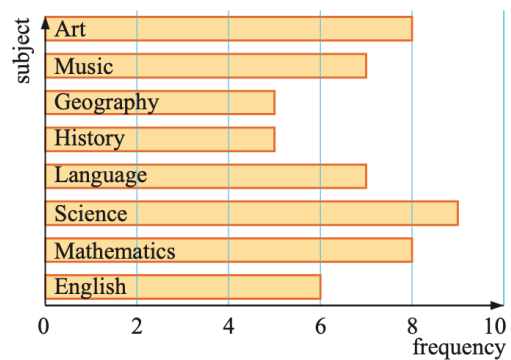
- 6** At a particular primary school, children were randomly chosen and asked to nominate their favourite fruit. The following data was collected:
- a** What are the variables in this investigation?
 - b** What is the dependent variable?
 - c** If we are trying to find out the favourite fruit of New Zealand children, is the sample unbiased?
 - d** If we are only interested in the favourite fruit of 368 children within the school:

<i>Type of fruit</i>	<i>Frequency</i>
Apple	20
Banana	24
Grapes	3
Orange	11
Mandarin	10
Nectarine	7
Peach	2
Pear	3

- i** is the sample unbiased? **ii** What is the sample size?

- e What is the mode?
- f Construct a vertical column graph to illustrate the data. Use a spreadsheet!

7 55 randomly selected year ten students were asked to nominate their favourite subject studied at school. The results of the survey are displayed in the bar chart shown.

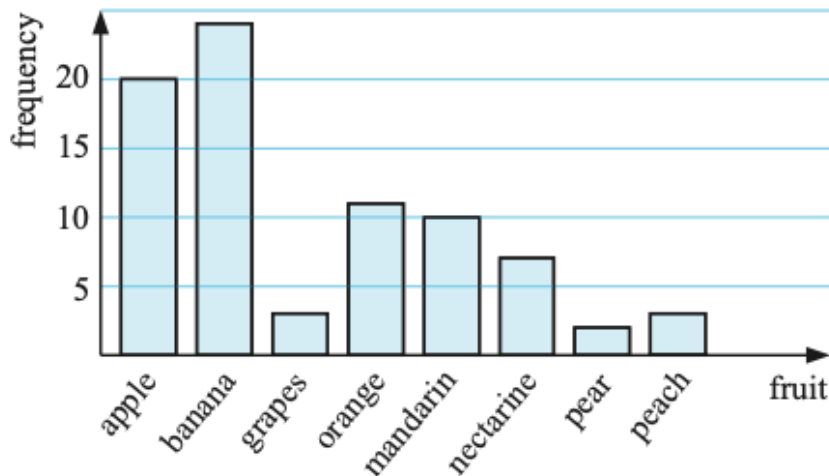


- a What are the variables in this investigation?
- b What are the dependent and independent variables?
- c What is the mode?
- d What given information indicates that the sample was unbiased?
- e If there are 173 year 10 students at the school, is the sample size sufficient?
- f Construct a pie chart for the data. If possible, use a spreadsheet.

EXERCISE 10A

- 1 a numerical b numerical c categorical
d numerical e numerical f categorical
g categorical h numerical
- 2 a male, female
b Australian Rules, Soccer, Gridiron, Rugby
c black, blonde, brown, grey, brunette
d super, unleaded, premium unleaded, diesel, LPG
- 3 a quantitative discrete
b quantitative continuous c categorical
d quantitative discrete e quantitative discrete
f categorical g quantitative discrete
h quantitative continuous i quantitative discrete
j quantitative continuous k categorical
l quantitative continuous
m quantitative discrete n quantitative discrete
- 4 a sample b census c sample d census
e census f sample
- 6 a type of fruit, frequency b type of fruit
c no, sample is from students at one place
d i no ii 80 e banana

f



- 7 a favourite subject, frequency
b dependent: subject; independent: frequency
c Science d randomly selected e yes