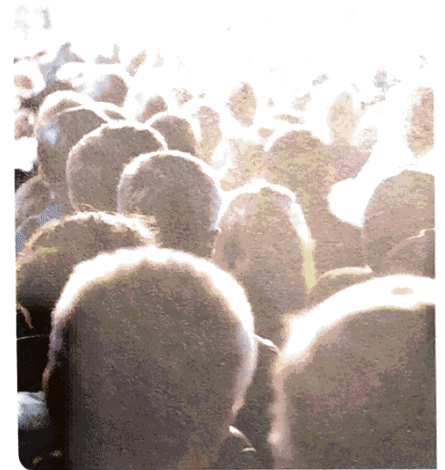


What is the difference between a population and a sample?

When we think of the term **population**, we usually think of people in our town, region, state or country.

→ In statistics, the term **population** includes all members of a defined group that we are studying for data driven decisions.

→ A part of the population is called a **sample**. It is a subset of the population, a selection of individuals from the population.



Random samples must have two characteristics:

- 1 Every individual has an equal opportunity of selection.
- 2 The sample has essentially the same characteristics as the population.

Exercise 8A

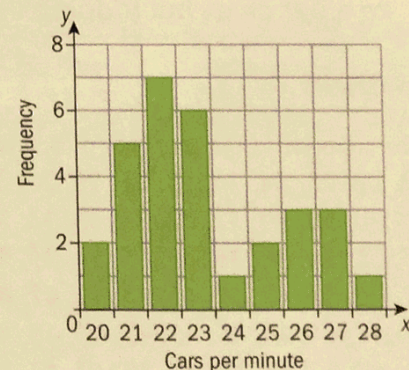
- 1 Classify each of the following as either discrete or continuous data.
 - a The number of fish caught by an angler.
 - b The length of the fish.
 - c The time taken to catch a fish.
 - d The number of friends that the angler took with him.
- 2 Are the test scores at the start of the chapter discrete or continuous data?

8.2 Presenting data

A **frequency table** is an easy way to view your data quickly and look for patterns.

You can also show discrete data in a **bar chart**.

A bar chart is sometimes called a column graph.



Example 1

A student counted how many cars passed his house in one-minute intervals for 30 minutes. His results were:

23, 22, 22, 22, 24, 22, 21, 21, 23, 23, 27, 21, 21, 22, 23, 25, 27, 26, 23, 23, 22, 27, 26, 25, 28, 26, 22, 20, 21, 20.

Display this data in a frequency table.

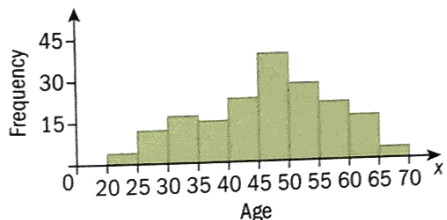
Draw a bar chart for this data.

▶ Continued on next page

Answer

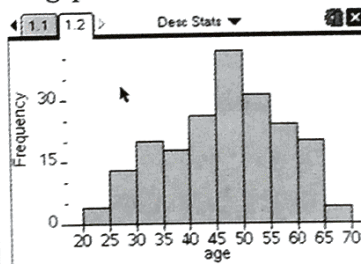
Age	Tally	Frequency
$20 \leq \text{age} < 25$		4
$25 \leq \text{age} < 30$	 	12
$30 \leq \text{age} < 35$	 	20
$35 \leq \text{age} < 40$	 	18
$40 \leq \text{age} < 45$	 	26
$45 \leq \text{age} < 50$	 	42
$50 \leq \text{age} < 55$	 	31
$55 \leq \text{age} < 60$	 	24
$60 \leq \text{age} < 65$	 	19
$65 \leq \text{age} < 70$		4

Equal class intervals of 5 years
25 is in the class $25 \leq \text{age} < 30$



Numbers on the edges of the bars or an x-axis like scale.

No gaps between the bars



You can use a GDC to draw histograms. See Chapter 17, Section 5.4.

GDC help on CD: Alternative demonstrations for the TI-84 Plus and Casio FX-9860GII GDCs are on the CD.



Exercise 8B



EXAM-STYLE QUESTION

- 1 All of the IB students in a school were asked how many minutes a day they studied mathematics. The results are given in the table.

Time spent studying mathematics (min)	$0 \leq t < 15$	$15 \leq t < 30$	$30 \leq t < 45$	$45 \leq t < 60$	$60 \leq t < 75$	$75 \leq t < 90$
Number of students	21	32	35	41	27	11

- Is this data continuous or discrete?
- Use your GDC to help you draw a fully labeled histogram to represent this data.

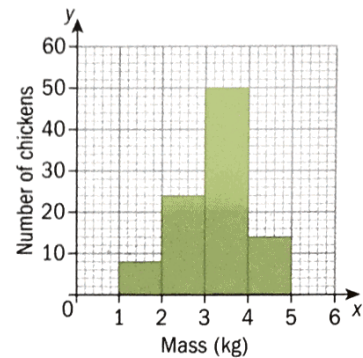


EXAM-STYLE QUESTION

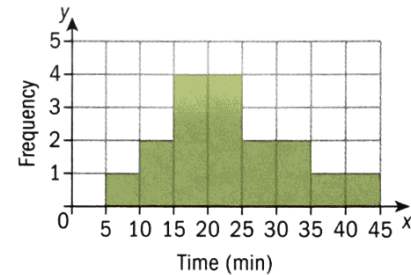
- 2 The following table shows the age distribution of mathematics teachers who work at Caring High School.
- Is the data discrete or continuous?
 - How many mathematics teachers work at Caring High School?
 - Use your GDC to help you draw a fully labeled histogram to represent this data.

Age	Number of teachers
$20 \leq x < 30$	5
$30 \leq x < 40$	4
$40 \leq x < 50$	3
$50 \leq x < 60$	2
$60 \leq x < 70$	3

- 3 The following histogram shows data about frozen chickens in a supermarket. The masses in kg are grouped such that $1 \leq w < 2$, $2 \leq w < 3$ and so on.
- Is the mass of frozen chickens discrete or continuous data?
 - Draw the grouped frequency table for this histogram.
 - How many frozen chickens are there in the supermarket?



- 4 The histogram on the right shows how many minutes it takes for students to return home after school.
- Is the data discrete or continuous?
 - Represent the data in a grouped frequency table.
 - What is the shortest time that a student could take to get home?



8.3 Measures of central tendency

A measure of central tendency tells us where the middle of a set of data lies. The three most common measures of central tendency are the **mode**, the **mean** and the **median**.

Another word is 'average'.

The mode

→ The mode is the value that occurs most frequently in a set of data.

The mode in a list of numbers is the number that occurs most often. Remember that **mode** starts with the same first two letters that **most** does.

There can be more than one mode. If no number occurs more than once in the set, then there is no mode for that set of numbers.