

## Chapter Review 1

- 1 a Sampling frame: first 15 days in January 2019.

Allocate each date a number from 1 to 15.

Use the random number function on calculator to generate 5 numbers between 1 and 15.

$$\text{b } \bar{x} = \frac{26+16+25+13+17}{5}$$

$$= 19.4^{\circ}\text{C}$$

Your answer is likely different as it depends on the random numbers generated.

$$\text{c } \bar{x} = \frac{28+26+27+16+16+21+24+24+18+13+17+22+25+16+13}{15}$$

$$= 21.1^{\circ}\text{C}$$

The second, larger sample, will give a more reliable estimate of the mean temperature.

- 2 a i Advantage: very accurate.

Disadvantages: expensive, time consuming and difficult to process.

- ii Advantages: easier data collection, quick and cheap.

Disadvantages: less accurate, less representative and possibly biased.

- b Assign unique 3-digit identifiers 000, 001, ..., 499 to each member of the population. Use random number tables, a computer or a calculator to generate 3-digit numbers. If these correspond to an identifier then include the corresponding member in the sample, being careful to ignore repeats and numbers greater than 499. Repeat this process until the sample contains 100 members.

- 3 a i Collection of individual items.

- ii List of sampling units, with each unit given an identifying name or number.

- b i List of registered owners from the DVLA.

- ii List of people visiting a doctor's clinic in Oxford in July 2017.

- 4 a Advantages:

The results are the most representative of the population since the structure of the sample reflects the structure of the population.

It guarantees proportional representation of groups within a population.

Disadvantages:

You need to know the structure of the population before you can take a stratified sample.

Classification into mutually exclusive strata may be difficult to implement.

The sampling within each strata may suffer from the disadvantages of simple random sampling.

**4 b** Advantages:

Quick.  
Cheap.  
All units have an equal chance of selection.

## Disadvantages:

Can introduce bias (e.g. if the sample, by chance, only includes very tall people in an investigation into heights of students).  
A sampling frame is needed first.

**5 a** Cabin crew are not represented.

**b i** Get a list of the 300 employees at the airline.  $\frac{300}{30} = 10$  so choose one of the first ten workers on the list at random and every subsequent 10th worker on the list, e.g. if person 7 is chosen, then the sample includes workers 7, 17, 27, ..., 287 and 297.

**ii** The sample should contain  $\frac{1}{3} \times 30 = 10$  pilots and  $\frac{2}{3} \times 30 = 20$  cabin crew, as those are the corresponding proportions in the whole population. The 10 pilots in the sample should be selected by a simple random sample of the 100 pilots. The 20 cabin crew should be selected by a simple random sample of the 200 cabin crew.

**iii** Decide the categories e.g. age, gender, pilots/cabin crew and set a quota for each in proportion to their numbers in the population. Interview employees until quotas are full.

**6 a** Allocate a number between 1 and 120 (the total number of pupils) to each pupil. Use random number tables, a computer or a calculator to select 15 different whole numbers between 1 and 120.

Pupils corresponding to these numbers become the sample.

**b** Allocate numbers 1–64 to girls and 65–120 to boys.

Select  $\frac{64}{120} \times 15 = 8$  different random numbers between 1 and 64 for girls.

Select the remaining 7 sampling units using random numbers between 65 and 120 for boys.

Include the corresponding boys and girls in the sample.

**7 a** Stratified sampling.

**b** This method uses naturally occurring groupings (strata). The results are more likely to represent the views of the whole population since the sample reflects its structure.