

UNIT: Statistics & society — Data collection & sampling

(i) Types of Data

Classification of data

Data is divided into two broad classifications called **quantitative** and **categorical**.

'Type of car' and 'favourite footy team' are examples of **categorical** data because they are represented by categories, usually in words or symbols.

'Number of people in a family' and 'persons shoe size' are examples of **quantitative** data because they are represented by quantities or numbers.

Quantitative data

Quantitative data is **numerical** data. For example, if we asked each student in the class their height we would expect to get a variety of answers. However, each answer is a number. Quantitative data is further classified as **discrete** or **continuous**.

Discrete data – data that can only take **exact numerical values**. For example, the number of students in a class is an example of discrete data because it is obtained by counting. The values are **distinct**, separate numbers. You can't have 33.6 students.

Continuous data – data that can take any numerical value (sometimes within specified interval). For example, the height of a person is an example of continuous data because it is obtained by measuring. We can have heights of 143.1 cm, 143.15 cm or 143.02 cm.

Types of quantitative data

Discrete data	Continuous data
<ul style="list-style-type: none"> Examples: number of children, age next birthday, shoe size Separate, distinct, 'counted' numbers. 'In-between' values are not possible. 	<ul style="list-style-type: none"> Examples: home prices, a person's weight, water temperature of a pool. 'Measured' numbers on a smooth scale. 'In-between' values are possible.

- Classify each of the following types of data as categorical (C) or quantitative (Q):

a types of religion	b numbers of computers in homes
c types of driver's licences held	d populations of cities
e numbers of CDs bought online	f marital status
g methods of payment for shopping	h movie classifications
i patients' blood pressures	j swimmers' finishing places in a race
k amounts of time spent on the Internet per week	l towns or suburbs where people live
- What is the difference between categorical data and quantitative data? Give an example of each type.
- Classify each of the following types of quantitative data as discrete (D) or continuous (C):

a amount of rainfall each day
b masses of athletes
c lengths of mobile phone calls
d running times for a 100 m race
e star ratings of hotels
f heights of the world's tallest buildings
g numbers of seats on buses
h loudness of lawn mowers
i shirt sizes
j numbers of rooms in homes
k distances travelled by trucks in a year
l speeds of cyclists



Answers

- 1 a C b Q c C d Q e Q f C
 g C h C i Q j Q k Q l C
- 2 categorical (words, symbols)
 quantitative (quantities, numbers)
- 3 a C b C c C d C e D f C
 g D h C i D j D k C l C

The 'capture-recapture' technique

The 'capture-recapture' technique is used to **estimate** population size. The technique works by capturing a random sample of the population. This first sample is tagged and then released. At a later time a second sample is captured and the recaptured or tagged members recorded.

The 'capture-recapture' technique

- 1 Use p to represent the population size.
- 2 First sample. Write the number captured as the numerator and p as the denominator of a fraction.
- 3 Second sample. Write the number recaptured or tagged as the numerator and the number captured in the second sample as the denominator.
- 4 Equate the two fractions in steps 2 and 3.
- 5 Solve the equation for p .

Example 1:

Adam is a biologist who is estimating the population of fish in a lake. He randomly captures and tags 120 fish. Two months later he samples 80 fish and finds he has recaptured 5 fish, as they are tagged. Estimate the number of fish in the lake.

ANS: 1920

Example 2:

From a forest, 30 deer were caught, tagged and released. Two months later, a sample of 160 deer was taken and 14 were found to be tagged. Estimate the deer population of the forest.

ANS: 343

Example 3:

Tim is scientist who is concerned about the number of bats in the park. He catches and tags 60 bats and then releases them. Two weeks later he catches 85 bats in the park and finds that 25 of them are tagged. Estimate the number of bats in the park for Tim.

ANS: 204

- 4 From a lake, 100 rock lobsters were caught, tagged and returned. Six months later, a random sample of 80 lobsters was taken and 11 were found to be tagged. Which of the following values is the best estimate for the number of rock lobsters in the lake? Select A, B, C or D.
A 138 **B** 727 **C** 880 **D** 1375
- 5 From a forest, 30 deer were caught, tagged and released. Two months later, a sample of 160 deer was taken and 14 were found to be tagged. Estimate the deer population of the forest.
- 6 Tania randomly caught 200 blue starfish in a Fiji lagoon, then labelled and released them. Later that month, she collected 50 blue starfish from the same lagoon and eight were found to be labelled. How many blue starfish would you expect to find in this lagoon?
- 7 To estimate the kangaroo population in a forest, Casey tagged and released 160 kangaroos. Three weeks later, he took a sample of 120 kangaroos and found that 17 were tagged. Estimate the kangaroo population in the forest.
- 8 The fisheries department took 2000 freshwater salmon from a Tasmanian lake, tagged them and returned them. When a sample of 600 salmon was captured later, 12 were found to be tagged. Calculate the lake's salmon population.
- 9 To find the possum population in New Zealand's Otago province, 800 possums were caught and fitted with a microchip, then released. Some months later, 45 out of 640 possums were found to have a microchip. Estimate Otago's possum population correct to the nearest hundred.
- 10 Paula caught 15 grasshoppers in her backyard, then marked and released them. One week later she caught 65 grasshoppers and found that 20% were marked. Which of the following values is the best estimate for the number of grasshoppers in Paula's garden? Select A, B, C or D.
A 30 **B** 49 **C** 56 **D** 75
- 11 In the upper Cotter River region of the ACT, 1500 blackfish were caught, tagged and released. When 250 fish were later taken from the same region, 20 were found to be tagged. What was the estimated population of blackfish in the region?
- 12 To estimate the population of moose in Blackcomb Forest, a sample of 150 moose was captured, tagged and let loose. Two weeks later, another sample of moose was taken and 12% had a tag. Estimate the moose population of Blackcomb Forest:
a using the fraction $\frac{12}{100}$ for the recaptured sample ← **12/100 is the same as 12%.**
b using the unitary method. ← **The unitary method
12% of the population = 150. Find 100%.**
- 13 An entomologist released a large number of butterflies into a forest after he had tagged 300 of them. When he later took a sample of 200, he found 15 of them had a tag. Approximately how many butterflies are there in the forest?
- 14 Fifteen wild rabbits were caught on a farm, marked and released. When 48 rabbits were recaptured, seven were found to be marked. Estimate the total number of rabbits on the farm.

- 15 Tim caught a large number of trout in Lake Eucumbene, and 15% were found to be tagged. If he had previously captured 4410 trout from the lake, and tagged and released them, estimate the number of trout in the lake:
a using equivalent fractions
b using the unitary method.
- 16 Two hundred tiger snakes were captured on a property. They were marked with blue dye, and then released. When a new random sample of snakes was taken, 51 out of the 110 snakes captured had the blue mark. The local snake catcher estimated that there were about 400 tiger snakes on the property. Is he correct?
- 17 Steve caught 20 crocodiles from a river and 12 had a yellow tag. If 150 crocodiles were originally tagged, approximately how many crocodiles were in the river?
- 18 To estimate the number of water buffalo in a Northern Territory region, a sample of 600 buffalo was herded and marked with a red dye. One month later, another group of 150 buffalo was rounded up and 24 showed a red mark. How many buffalo were there in the region?

Answers

- 4 B
5 343
6 1250
7 1129
8 100 000
9 11 378
10 D
11 18 750
12 **a, b** 1250
13 4000
14 103
15 **a, b** 29 400
16 Yes (estimate is 431)
17 250
18 3750