

Types of Data, Sampling and Bias

Solutions

For AQA GCSE Mathematics (8300) • Foundation tier

Each question is shown in grey, followed by the working, one step per line. The final answer is in a box. *Marks* shows how the marks are usually given in the exam, and **Tip** points out a common mistake.

1 (1 mark)

Which one of the following is continuous data?

Circle your answer.

Options: number of doors, number of goals, time to boil a kettle, number of coins

Continuous data is **measured**. It can take any value in a range.

Time is measured, so the time to boil a kettle is continuous.

The numbers of doors, goals and coins are **counted**, so they are discrete.

Answer: time to boil a kettle

Tip: Ask yourself: do I count it or measure it? Counted data is discrete. Measured data (length, mass, time) is continuous.

2 (3 marks)

Tick one box for each type of data.

	Categorical	Discrete	Continuous
Eye colour	✓		
Number of brothers and sisters		✓	
Time taken to run 100 metres			✓
Favourite sport	✓		
Mass of an apple			✓

Categorical data is words, not numbers (it is put into categories).

Discrete data is counted.

Continuous data is measured.

Eye colour: a word, so categorical.

Number of brothers and sisters: counted, so discrete.

Time taken to run 100 metres: measured, so continuous.

Favourite sport: a word, so categorical.

Mass of an apple: measured, so continuous.

Answer: See table

Marks: 3 marks for all five correct. 2 marks for four correct. 1 mark for three correct.

3 (1 mark)

Kwame records the number of emails he gets each day for 40 days.
The table shows his results.

Number of emails	Frequency
0 to 9	6
10 to 19	17
20 to 29	12
30 to 39	5

Kwame says,

‘This data is continuous because it has been put into groups.’

Is Kwame correct? Tick a box.

Give a reason for your answer.

The number of emails is **counted**. It can only be a whole number.

So the data is discrete.

Putting data into groups does not change whether it is discrete or continuous.

The data is grouped **and** discrete.

Answer: No. The number of emails is counted (whole numbers only), so the data is discrete, even though it is grouped.

Tip: “Grouped” and “continuous” do not mean the same thing. Counted data stays discrete when it is put into groups; measured data (such as time) is continuous whether or not it is grouped.

4 (2 marks)

Tick one box for each source of data.

	Primary	Secondary
Ali measures the heights of students in his class	✓	
Beth uses rainfall figures from a website		✓
Cara counts cars passing her house	✓	

Primary data is collected by the person themselves.

Secondary data has already been collected by someone else.

Ali measures the heights himself, so primary.

Beth uses figures from a website, so secondary.

Cara counts the cars herself, so primary.

Answer: Primary, Secondary, Primary

Marks: 2 marks for all three correct. 1 mark for two correct.

5 (2 marks)

Sam wants to find out how often people in his town use a gym.
He asks 20 people at his local gym.

(a) Give one reason why this sample is biased.

Sam only asks people who are at the gym.

These people are more likely to use a gym than other people in the town.

So his results will make gym use look higher than it really is.

Answer: Only asks people who already use a gym

Tip: Say **why** the sample is biased, linked to the question: the people asked are at a gym.
Just writing “the sample is too small” does not explain bias.

(b) Sam says,

‘If I ask 200 people at the gym, my sample will not be biased.’

Is Sam correct? Tick a box.

Give a reason for your answer.

Asking more people can make the results more reliable.

But all 200 people would still be at the gym.

They are still more likely to use a gym than other people in the town.

So the sample would still be biased.

Answer: No. All 200 people would still be gym users, so the sample is still biased.

Tip: A bigger sample does not remove bias. Bias comes from **who** is asked, not **how many** people are asked.

6 (2 marks)

Ella wants to estimate how much time students at her school spend on homework each week.

She asks 6 students in her maths class.

Give **two** reasons why Ella’s sample may not represent all the students in her school.

Reason 1: the sample is too small. She only asks 6 students.

Reason 2: all the students are from one maths class.

They are all the same age and may get the same homework, so they are not like the whole school.

Answer: Too small; all from one class (one year group)

Tip: Your two reasons must be different. Saying “too few students” and “not enough students” is the same reason twice.

Marks: 1 mark for each different reason.

7 (2 marks)

Maya claims most houses on her road have a garden.
Her road has 60 houses, numbered 1 to 60
She looks at houses 1 to 10 only.

- (a) Give a reason why this sample may not be useful in testing Maya's claim.

She only looks at houses 1 to 10.

These are all at one end of the road, and may be different from the other houses.

The sample is not random.

Answer: Only one part of the road / not random

- (b) Maya then looks at 30 houses on the road chosen at random.

21 of them have a garden.

Does this support Maya's claim? Tick a box.

Give a reason for your answer.

$$\frac{21}{30} = 70\%$$

70% is more than half, so most houses in the sample have a garden.

The sample is random, so this supports Maya's claim.

Answer: Yes, 21 out of 30 is more than half

8 (4 marks)

A school has 1200 students.
Jo wants to find out what the students think about school lunches.
She is going to ask a sample of 80 students.

- (a) Circle the population for this survey.

Options: the 80 students, all 1200 students, Year 11 students, the teachers

The population is the whole group you want to find out about.

Jo wants to know what all the students think.

The 80 students are the sample, not the population.

Answer: all 1200 students

- (b) In the sample, 52 of the 80 students say they like school lunches.
Estimate how many of the 1200 students like school lunches.

The school is $1200 \div 80 = 15$ times bigger than the sample.

So multiply the number in the sample by 15:

$$52 \times 15 = 780$$

Answer: 780

Marks: 1 mark for $\frac{52}{80} \times 1200$ or $1200 \div 80 = 15$. 1 mark for 780.

- (c) Jo says,

“If I asked 200 students instead of 80, my estimate would probably be more reliable.”

Is Jo correct? Tick a box.

Give a reason for your answer.

Yes.

A larger sample is more likely to represent the whole school.

Answer: Yes, a larger sample is more reliable

9 (4 marks)

A town has 18 000 adults.

The council wants to know if they support a new cycle lane.

It asks a sample of 200 adults.

The table shows the results.

Answer	Number of adults
Support	136
Against	52
Don't know	12

- (a) Work out an estimate for the number of adults in the town who support the new cycle lane.

In the sample, 136 out of 200 support the cycle lane.

The fraction who support it is $\frac{136}{200}$.

Use the same fraction of all 18 000 adults:

$$\frac{136}{200} \times 18\,000 = 12\,240$$

Answer: 12 240

Tip: Use the whole sample (200) as the denominator. Do not leave out the “Don’t know” answers and divide by $136 + 52 = 188$.

Marks: 1 mark for $\frac{136}{200} \times 18\,000$ or $18\,000 \div 200 = 90$. 1 mark for 12 240.

- (b) The sample was taken outside a bike shop.
Explain why your estimate in part (a) is likely to be too high.

People outside a bike shop are more likely to be cyclists.

Cyclists are more likely to support a cycle lane than other adults.

So the sample has too many people who support it, and the estimate is too high.

Answer: People outside a bike shop are more likely to cycle, so more likely to support a cycle lane than adults in the town in general.

- (c) Give one reason why the council used a sample instead of asking all 18 000 adults.

Asking 18 000 people would take a very long time and cost a lot.

A sample is quicker and cheaper.

Answer: It is quicker (or cheaper) to ask a sample.