MATHS FIELD GUIDE · GRADE 9

Sampling and statistical questions

OPENING PROBLEM

Who is missing from a cycling-club survey of school transport?

Question	Population
School travel	all pupils
Club travel	club members

02 Decide which population a claim describes
03 Rules and worked cases
04 Selection can introduce bias
05 From the givens to the answer
06 Ask what one observation represents
07 Check your reasoning
08 Check your reasoning

ILLUSTRATED PROBLEM SHEET

Decide which population a claim describes

A fair sample, a reliable measure and an honest display each answer a different concern.

1

Question	Population
School travel	all pupils
Club travel	club members

Who is missing from a cycling-club survey of school transport?

Pupils outside the club. They may use transport differently.

2

Group	Records	Cycle
A	20	8
B	10	6

Pool the observations. What fraction cycle?

14/30 = 7/15. Averaging 40% and 60% would give both groups equal weight.

3

Day	Monday	Tuesday
Rain?	yes	no
Walkers	8	20

Does the difference prove rain caused the change?

It is consistent with that idea, but other day-to-day differences may explain it.

4

Measure	Before	After
Total pupils	20	40
Cyclists	10	16

Which rose: the headcount or share cycling?

The count rose by six. The share fell from 50% to 40%.

Same method on more volunteers
may retain selection bias.

Sampling across every year group
can improve representation.

CHAPTER PROBLEM

A poll has 150 invited learners and 120 replies. Find the response rate as a percentage; enter the number only.

80. 120/150 × 100 = 80.

MATHS REFERENCE · GRADE 9

Sampling and statistical questions

Rules, conditions and worked cases

Recognise the case	Mathematical reasoning
Population and sample	To estimate school travel times, pupils at school form the population. Those actually surveyed form the sample.
Selection mechanism	Sampling every tenth name after a random start can reach beyond friends or volunteers. Check list ordering for bias.
Statistical question	“How long do pupils take to travel?” anticipates variation. A single named pupil’s recorded time is one observation.

Read the representations

A statistical question anticipates variation

A statistical question expects different possible observations and answers by examining data. “How long do learners in this year group take to travel to school?” anticipates varied times.

“How many minutes are in one hour?” has a fixed factual answer and does not need a study of variation.

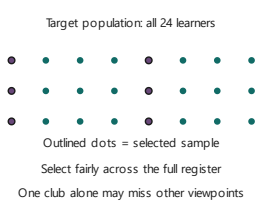
Specify the population, the variable, its units and the observation period. For travel time, define whether waiting and transfers are included.

Population is the group the claim is about

The population contains all cases addressed by the question. A sample is the subset actually observed.

A census observes every population member; a sample observes fewer. A sampling frame is the list or source from which selection is made.

The picture distinguishes selected cases from the full group. A sample should give relevant parts of that population a fair opportunity to be represented.



Class	0–10	10–20	20–30
Frequency	4	7	3

How many observations are at least 10 and below 30?

Seven plus three gives ten. Use the stated interval endpoints.

THEOREMS AND REPRESENTATIONS

Selection can introduce bias

Selection can introduce bias

Surveying a sports club about everyone's preferred activity may overrepresent sport. Voluntary online responses can overrepresent people with strong opinions or internet access.

A simple random sample gives every same-sized subset of the frame the same selection chance. It reduces deliberate selection bias but does not guarantee a perfectly representative result.

If comparing age groups, random sampling within each age group can ensure each is included. Record sample sizes, nonresponses and how selection was carried out.

Measurement and wording also shape the data

Ask a neutral question with clear response categories. "Do you agree our excellent new menu is better?" encourages a particular answer.

Use one clear reference period and compatible units. Do not compare one person's usual journey with another person's longest-ever journey as if measured alike.

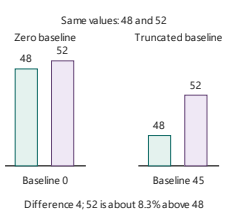
Missing responses are not automatically zero. Record them separately and state which observations are included in each calculation.

An axis can exaggerate a difference

Both panels below show 48 and 52. Starting bar heights at 45 makes the second height look much larger relative to the first.

The numerical difference remains 4, and the increase relative to 48 is $4/48 \times 100 \approx 8.3\%$.

For bar-length comparisons use a zero baseline. If a restricted axis is justified in another kind of plot, mark it prominently and avoid interpreting displayed heights as ratios.



Association does not establish cause

A study may show two variables changing together. That association does not by itself show that changing one causes the other.

For example, longer journeys and less sleep could also be associated with different start times, home routines or transport schedules.

Describe what was observed, consider alternative explanations and avoid causal claims that the study design cannot justify.

Class	0–10	10–30
Frequency	4	8
Freq/width	0.4	0.4

How can unequal-width histogram bars have equal height?

Frequency density is frequency divided by width. The wider bar has twice the area and twice the frequency.

WORKED EXAMPLES

From the givens to the answer

Limit a conclusion to the evidence

1. A voluntary lunch survey has 24 replies from one class, and 18 prefer option A. The response share is 18/24 = 75%.
2. This supports “75% of these respondents chose A.” It does not establish that 75% of the whole school prefer A.
3. To study the school, define the population, select fairly across relevant year groups and address nonresponse. A larger biased sample alone does not fix the selection method.

Will asking more of the same volunteers remove bias?

A learner proposes a much larger online club poll to prove what the entire school wants.

1. More responses can reduce some chance fluctuation but do not repair exclusion of nonmembers or nonusers.
2. The selection frame still misses parts of the target population.
3. Improve how cases are selected and measured, then report remaining limitations rather than treating size as a guarantee.

Check another case

Example 1

A class survey has 32 replies out of 40 invitations. How many invitations received no reply?

8

40 − 32 = 8.

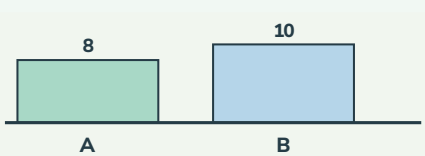
Example 2

List three checks before accepting a headline based on a graph.

Check the source, target population, sample selection and response coverage.

Read axes, units, baselines, denominators and any rounding.

Compare the headline with what the measured data support, including uncertainty and alternative explanations.



These bars share a zero baseline. How large is the ratio B:A?

10/8 = 1.25. Starting an axis at 7 instead could exaggerate the apparent difference.

ILLUSTRATED PROBLEM SHEET

Ask what one observation represents

Combine totals and counts before dividing. A mean of group means weights groups equally.

1

Group	Count	Mean
A	4	80
B	16	60

What is the combined mean?

(320 + 960)/20 = 64. The larger group pulls the mean towards 60.

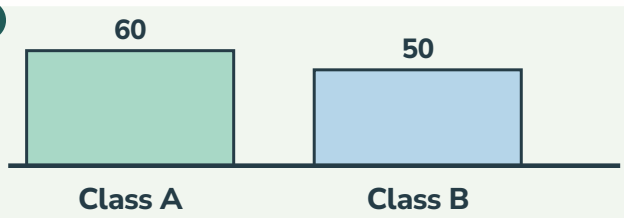
2

Value	2	5	8
Count	1	2	1

Find the mean without writing the full list.

Weighted total 2 + 10 + 8 = 20. Four observations give mean 5.

3



Bars show percent cycling. A has 20 pupils and B has 30. Which has more cyclists?

A has 12 cyclists and B has 15. A higher percentage need not mean a higher count.

4

Fresh sample: 20 pupils

↓

Only cycling-club members asked

Why might its travel survey overstate cycling?

The selection method favours cyclists. More responses from that same club do not remove the bias.

A group with zero observations
contributes zero weight.

All groups of equal size
group means receive equal weights.

CHAPTER PROBLEM

A bar chart shows counts 64 and 80. Find the increase from 64 as a percentage; enter the number only.

25. (80 – 64)/64 × 100 = 25.

QUESTION BANK · FOLD OR COVER THE RIGHT COLUMN

Check your reasoning

Try the problem	Solution and reasoning
01 Which is a statistical question? A) How do travel times vary among Year 9 learners this week? B) How many seconds are in a minute? C) What is 7 squared? D) What is the school's registered address?	How do travel times vary among Year 9 learners this week? Travel times vary across learners and require data.
02 A study asks about all city households but samples only apartment owners. What is the main problem? A) The population has no variation B) The sampling frame excludes other household types C) The sample is necessarily too random D) All responses must be false	The sampling frame excludes other household types Many households have no selection opportunity in that frame.
03 A survey asks only the first ten people leaving the gym about exercise. Which issue is most relevant? A) Selection may overrepresent people who exercise B) A census was completed C) Random selection is guaranteed D) The mode cannot be calculated for categories	Selection may overrepresent people who exercise Gym visitors may differ from the population about which a general exercise claim is made.
04 Which question is most neutral? A) Which lunch option do you prefer: A, B, neither, or no preference? B) Why is option A so much better? C) Everyone likes A; do you? D) Do you want our wonderful option A?	Which lunch option do you prefer: A, B, neither, or no preference? It offers balanced responses without praising an option.
05 Which is more representative of the whole school: every tenth name after a random start on a mixed pupil list, or the first ten volunteers?	The list method can cover all pupils. Volunteers select themselves and can differ systematically.
06 A value is missing from a travel-time record. Which action is justified? A) Delete the learner from the population definition B) Record it as missing and state how it is handled C) Replace it with zero without comment D) Assume it equals the maximum	Record it as missing and state how it is handled Missing means unknown, not a recorded zero-minute journey.

Group	Under 10	Total
A	6	10
B	10	25

Which group has the larger proportion under ten?

A has 60%; B has 40%. Compare shares when sample sizes differ.

QUESTION BANK · FOLD OR COVER THE RIGHT COLUMN

Check your reasoning

Try the problem	Solution and reasoning
07 Only volunteers answered a class poll. Which claim is directly supported? A) A stated share of respondents selected A B) The same share of the whole country selected A C) The poll proves A causes higher scores D) All nonrespondents selected B	A stated share of respondents selected A The evidence directly describes those who replied.
08 What if both group sizes are equal?	Then the combined mean equals the simple average of the two group means.
09 Plan a fair sample for comparing journey times in two year groups.	Define both populations, a common journey-time measure and the same observation period. Select randomly within each year-group register, explaining the chosen sizes. Record nonresponses and limitations, and report each group's size and distribution.
10 Learners who attend a club have higher average scores. Why does that alone not prove the club caused higher scores?	The comparison shows an association in the observed data. Prior attainment, motivation or other differences may explain part of it. A causal conclusion needs a design and evidence that address competing explanations.
11 What is a sample? A) Only the highest values B) The mean of the observations C) The observed subset of a target population D) Every possible population worldwide	The observed subset of a target population The sample consists of the cases actually selected and observed.

BANANALEARN LEARNING LIBRARY

Where will your maths take you next?

GRADE 9 MATHS

Choose your next problem.

Explore more illustrated guides and practise a related topic.

Browse Grade 9 guides →

Keep exploring

GRADE 9 · TOPIC GUIDE

Comparing stacked and percentage bar graphs

Open guide →

GRADE 9 · TOPIC GUIDE

Weighted averages and interpretation

Open guide →

Visit bananalearn.com →