Name: _____ Date: _____

UNIT 5

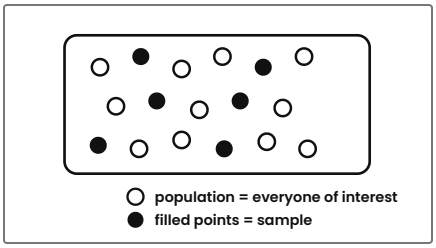
5.5.8 — Sampling and Response Bias

Population • Sampling Methods • Representativeness • Bias • Non-response • Fair Conclusions

A. POPULATION, SAMPLE AND REPRESENTATIVENESS

Ask who was included, who was missed and how responses were collected.

1. Define population and sample using the source terminology.
2. What does it mean for a sample to be representative?
3. Define bias as used in the source deck.
4. Explain why a large sample can still be biased.



B. RECOGNISE SAMPLING METHODS

The method should match the population and the question.

1. Match each method to its description: simple random, systematic, stratified, convenience and voluntary.
2. Which methods in the source are highlighted as having a greater risk of being unrepresentative?
3. Why does systematic sampling need a random start?
4. When might stratified sampling be useful?

simple random	chance
systematic	every kth after random start
stratified	sample important subgroups
convenience	easiest to reach
voluntary	people choose to respond

C. SOURCE PRACTICE: THE SPORTS-CENTRE SURVEY

The source identifies this as a convenience sample and suggests random selection from the school register or stratification by year group.

1. A school asks students leaving the sports centre, “Should the school spend more on sports equipment?” Identify the sampling method.
2. Identify one group that is less likely to be included.
3. Explain how the selection could affect the estimated support for more sports spending.
4. Give one improved sampling method from the source and explain why it is broader.

Name: _____ Date: _____

UNIT 5

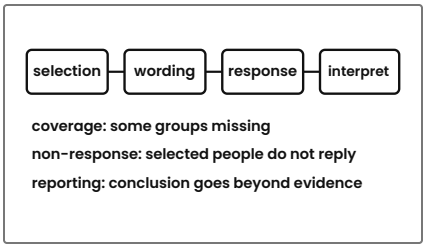
5.5.8 — Sampling and Response Bias

Population • Sampling Methods • Representativeness • Bias • Non-response • Fair Conclusions

D. WHERE BIAS CAN ENTER

Fair data collection begins before the first response.

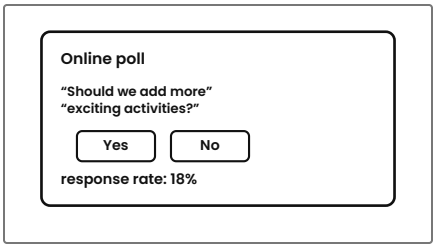
1. Put these stages in order: interpretation, selection, response, wording.
2. Explain coverage bias.
3. Explain non-response bias.
4. Explain how a conclusion can be biased even if the data were collected carefully.



E. SOURCE PRACTICE: VOLUNTARY ONLINE POLL

A response rate is evidence about possible bias.

1. A website asks, “Should we add more exciting activities?” Why is the wording potentially leading?
2. Only 18% respond. What source of bias can arise because visitors choose whether to vote?
3. Why might people with strong opinions be overrepresented?
4. Explain why the result may not represent all website visitors.



F. BUILD A FAIR STUDENT-WELLBEING SURVEY

The source sample uses a stratified random sample and the neutral question: “How often have you felt stressed by schoolwork this month?”

1. Define a suitable population for a school survey about student wellbeing.
2. Choose a sampling method that supports representation across year groups.
3. Write one neutral, specific survey question.
4. State one limitation that should still be reported.

Name: _____ Date: _____

UNIT 5

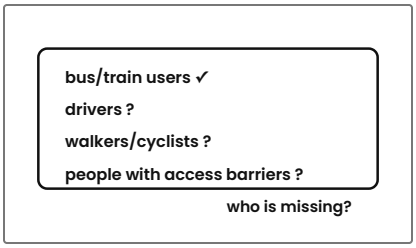
5.5.8 — Sampling and Response Bias

Population • Sampling Methods • Representativeness • Bias • Non-response • Fair Conclusions

G. REAL-WORLD REASONING: PROPOSED BUS ROUTE

The source specifically notes that drivers, walkers, cyclists and people who cannot easily access public transport may be excluded.

1. A city surveys only people using public transport about a proposed bus route. Which viewpoints might be missing?
2. Identify the main coverage problem.
3. Suggest a broader sampling approach.
4. Why should the final report name who was sampled before generalising?



H. SOURCE CRITIQUE: LARGE SAMPLES AND STRONG CLAIMS

1. A report says, “95% of respondents support the plan, so 95% of residents support it.” What information is missing before this conclusion can be justified?
2. A voluntary online poll receives 2,000 responses. Why can the large number of responses still leave the survey biased?
3. Would a smaller random sample necessarily be worse evidence? Explain using the source ideas.
4. Rewrite the 95% conclusion more carefully.

I. EXTENDED IB-STYLE: PROPORTIONAL STRATIFIED SAMPLING

Extended material: proportional allocation and response-rate calculation build on the source concepts but are not explicitly calculated in the source deck.

1. For proportional stratified sampling, use $n_h = \frac{N_h}{N} n$, where N_h is a subgroup size, N the population size and n the total sample size.
2. A school has 240 Year 9, 210 Year 10 and 150 Year 11 students. For a sample of 60, calculate the proportional number from each year.
3. A survey invites 250 people and receives 170 responses. Use $\text{response rate} = \frac{\text{responses}}{\text{people invited}} \times 100\%$ to calculate the response rate.
4. Explain why a high response rate does not by itself prove that the original sampling frame was representative.

ANSWER KEY

Name: _____ Date: _____

UNIT 5

5.5.8 — Sampling and Response Bias

Population • Sampling Methods • Representativeness • Bias • Non-response • Fair Conclusions

A. POPULATION, SAMPLE AND REPRESENTATIVENESS

1. Population: everyone or everything of interest. Sample: a smaller group selected from that population.
2. Representative means similar to the population in relevant ways.
3. Bias is a systematic tendency to favour some results.
4. A large sample can still systematically exclude groups, so size alone cannot remove selection or coverage bias.

B. RECOGNISE SAMPLING METHODS

1. Simple random: select by chance. Systematic: every kth member after a random start. Stratified: sample from important subgroups. Convenience: use the easiest people to reach. Voluntary: people choose whether to respond.
2. The source especially flags convenience and voluntary samples as potentially unrepresentative.
3. A random start avoids deliberately favouring a particular position in the ordered list.
4. Stratified sampling is useful when important subgroups, such as year groups, should be represented.

C. SOURCE PRACTICE: THE SPORTS-CENTRE SURVEY

1. It is a convenience sample because students are selected at one location.
2. Students who do not use the sports centre are less likely to be included.
3. The results may overestimate support for sports equipment because sports-centre users may be more supportive.
4. Select students randomly from the full school register, or use a stratified sample to represent year groups.

Name: _____ Date: _____

UNIT 5

5.5.8 — Sampling and Response Bias

Population • Sampling Methods • Representativeness • Bias • Non-response • Fair Conclusions

D. WHERE BIAS CAN ENTER

1. Selection → wording → response → interpretation.
2. Coverage bias occurs when some groups are missing from the sampling frame.
3. Non-response bias can occur when selected people do not reply and respondents differ from non-respondents.
4. Reporting can still overreach if the conclusion generalises beyond what the sample and response process support.

E. SOURCE PRACTICE: VOLUNTARY ONLINE POLL

1. The word “exciting” encourages a positive response and therefore makes the wording leading.
2. Voluntary-response bias can occur because visitors choose whether to vote.
3. People with strong opinions may be more motivated to respond.
4. The 18% who respond may differ systematically from all visitors, so the result may not represent the full population.

F. BUILD A FAIR STUDENT-WELLBEING SURVEY

1. Example population: all students in the school.
2. Use a stratified random sample from each year group.
3. Source example: “How often have you felt stressed by schoolwork this month?”
4. Possible limitations include non-response, imperfect coverage, or the possibility that some students interpret the question differently.

Name: _____ Date: _____

UNIT 5

5.5.8 — Sampling and Response Bias

Population • Sampling Methods • Representativeness • Bias • Non-response • Fair Conclusions

G. REAL-WORLD REASONING: PROPOSED BUS ROUTE

1. Possible missing viewpoints include drivers, walkers, cyclists and people who cannot easily access public transport.
2. The sampling frame undercovers people who are not current public-transport users.
3. Use a broader city sampling frame and select randomly, or stratify across relevant resident groups or areas.
4. The identity of the sampled group determines how far the conclusion can reasonably be generalised.

H. SOURCE CRITIQUE: LARGE SAMPLES AND STRONG CLAIMS

1. The sampling method, response rate and population coverage are missing.
2. Voluntary response can overrepresent people who choose to participate, so 2,000 responses can still be systematically unrepresentative.
3. No. A smaller random sample can provide a fairer basis for generalisation than a much larger biased sample.
4. Example: “Among the people who responded, 95% supported the plan; the extent to which this represents all residents depends on the sampling method, coverage and response rate.”

I. EXTENDED IB-STYLE: PROPORTIONAL STRATIFIED SAMPLING

1. $n_h = \frac{N_h}{N} n$.
2. Total population = 600. Year 9: 240/600 × 60 = 24; Year 10: 210/600 × 60 = 21; Year 11: 150/600 × 60 = 15.
3. $\text{response rate} = \frac{\text{responses}}{\text{people invited}} \times 100\%$ = 170/250 × 100% = 68%.
4. A high response rate only describes participation among those invited. If the sampling frame omitted important groups, the responding sample can still be biased.