Chapter 8

Sampling

When Good Math Meets Bad Data

A researcher surveys 500 members of **insomnia.sleep-disorders.net** about their nightly sleep and builds a 95% CI: $(\mathbf{5.69},\ \mathbf{5.91})$ hours.

The true population mean (Statistics Canada) is approximately **7.0 hours**.

Connection: The math was correct. The **sample** was problematic.
Chapters 15–16 assume a **random sample** from the population.
A self-selected online community is not a random sample of all Canadians.
When this assumption fails, the interval is precise but meaningless.

Guiding Questions

1. Why do we use samples to learn about populations?
2. What makes a sample good or bad?
3. How do we choose our samples?
4. What can go wrong even with a good sample?

Sampling Is Everywhere

Example 8.1

We sample every day without thinking about it.

You taste a **spoonful** of soup, not the entire pot, to decide whether the seasoning is right.

That spoonful is a **sample**. The whole pot is the **population**.

Sampling Is Everywhere

More everyday examples

The same logic applies whenever you make a decision from partial information:

1. Reading the first few pages of a book before buying
2. Taking a few introductory classes before choosing a major

In each case you examine a **small part** to learn about the **whole**.

PART 1

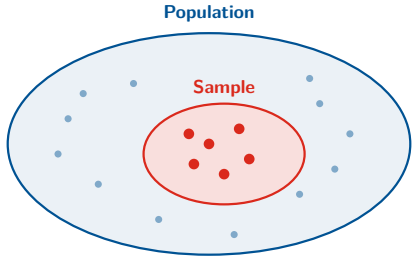
Populations, Samples, and Terminology

What makes a sample representative of a population?

Population vs. Sample

Population and Sample

- The **population** is the entire group of individuals we want information about.
- A **sample** is the part of the population we actually observe.
  We use it to draw conclusions about the whole.

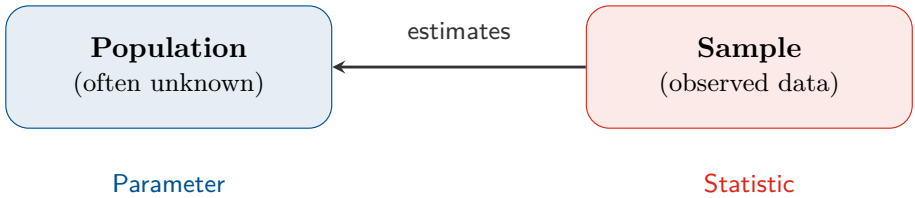


Example 8.2

Parameter vs. Statistic

Parameter and Statistic

A **parameter** is a numerical summary of a *population* characteristic.
A **statistic** is a numerical summary of a *sample* characteristic.



Principle: We use statistics (from the sample) to estimate parameters (of the population).

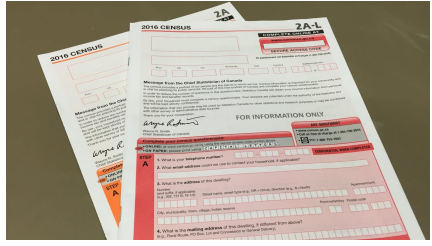
Census: The Ideal Case

Census

A **census** collects data from *every* individual in the population.

When a census works well:

- Population is small and accessible
- Cost and time are not prohibitive



Census: The Practical Limits

Why censuses often fall short:

- **Cost and time:** surveying an entire large population is expensive and slow
- **Infeasibility:** some populations are too large or inaccessible to enumerate fully

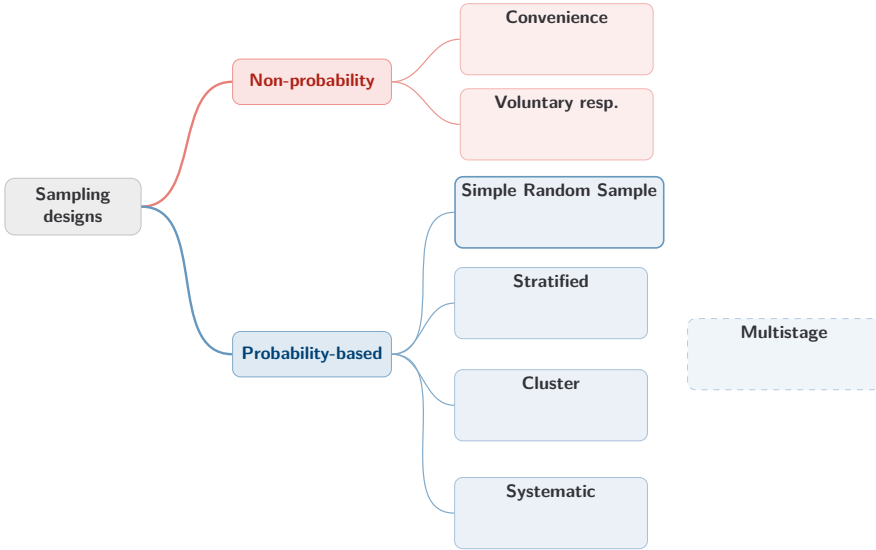
> ⚠ **Caution:** A well-designed sample is often more practical, and more accurate, than a flawed census.

Sampling Design: Roadmap

Sampling Design

A **sampling design** describes exactly how to choose a sample from the population.
As a census is usually impractical, we need a design that balances bias risk, cost, and precision.

Sampling Designs at a Glance



PART 2

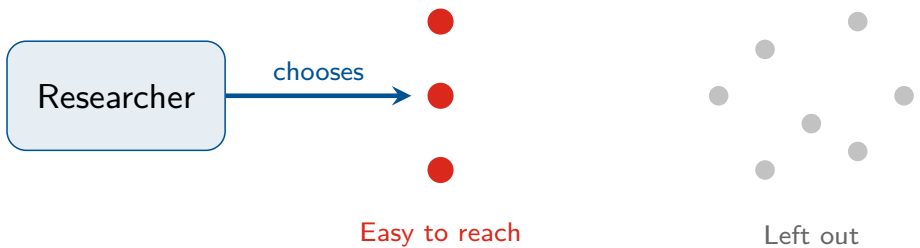
Biased Sampling Methods

Why do convenience and voluntary response samples mislead us?

Convenience Sampling

Convenience Sample

A **convenience sample** is selected by taking the members of the population that are easiest to reach.



Surveying at the Gym

Example 8.3

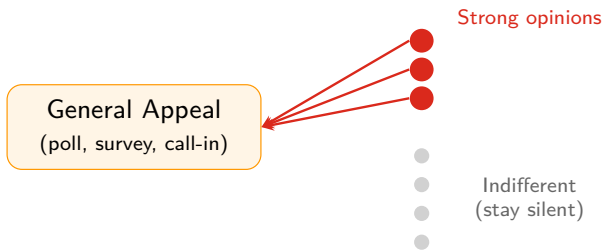
Context: Western University wants to estimate how many hours per week its students exercise. A research assistant surveys students as they leave the campus gym on a weekday evening.

Question: Why is this problematic?

Voluntary Response Sampling

Voluntary Response Sample

A **voluntary response (or volunteer) sample** consists of people who choose themselves by responding to a general appeal.



Notation: Bias from self-selection in this instance is sometimes called **voluntary response bias**.

Instagram Poll on Exercise

Example 8.4

Context: Western's student council posts an Instagram Story poll:
"How many hours a week do you exercise?"
Followers tap one of four ranges (0–2, 3–5, 6–8, 9+).

Find: Why is this problematic?

Convenience vs. Voluntary Response

Convenience

The *researcher* chooses who to survey based on accessibility.

Example: Surveying students at a campus coffee shop.

Voluntary Response

The *respondents* choose themselves whether to participate.

Example: A Twitter/X click-to-vote poll on AI.

Principle: Both are **non-random** methods. Both introduce **selection bias**. Additionally, we are **unable to quantify** the bias or its direction, so we cannot adjust for it. Both methods are unreliable for making inferences about the population.

Statistical Bias

Bias

The design of a sample is **biased** if it systematically favours certain outcomes.

Common misconception: Bias refers to the **method**, not the result.
A biased design may still produce an accurate sample.

Selection Bias

Selection Bias

Selection bias occurs when the sampling method systematically over- or under-represents certain groups.

Selection bias can arise from several sources:

- Researcher picks "convenient" individuals
- Only people with strong opinions respond
- Certain groups are harder to reach
- The list from which you actually draw your sample does not include the full population

Common misconception: A bigger biased sample is still biased.
More data cannot fix selection bias; only changing the sampling method can.

How Convenience Sampling Distorts Averages

Example 8.5

Context: Western is 40% first-year and 60% upper-year students.
A researcher surveys students at the campus gym and gets a sample that is 30% first-year and 70% upper-year.

The true mean weekly exercise hours by group: first-year $\mu_1 = 2$, upper-year $\mu_2 = 7$.

Calculate: Compare the population mean to the convenience sample estimate.

The Friendship Paradox

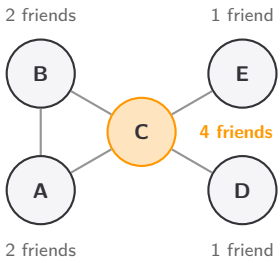
Example 8.6

Context: Researchers poll children at schools to estimate the average number of children per household.
The survey finds an average of 4.023 children per household, but the true population average is 1.6.

What explains this large discrepancy?

Friendship Paradox

The friendship paradox (Feld, 1991): Your friends have more friends than you do, on average.



Person	Friends	Friends' avg
A	2	3.0
B	2	3.0
C	4	1.5
D	1	4.0
E	1	4.0
Avg	**2.0**	**3.1**

4 of 5 people have *fewer* friends than their friends' average.

Intuition: Popular people show up in more friend lists, so they are over-sampled when you reach people through connections.

What Bias Does to Your Confidence Level

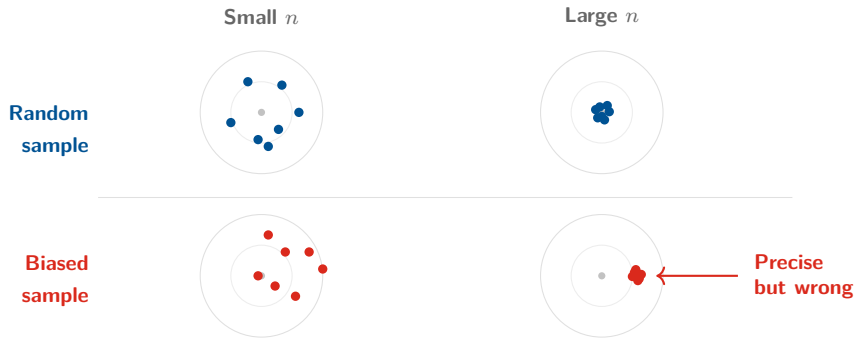
Suppose a biased method shifts the sample mean by $\delta = 0.5$ hours (e.g., surveying only gym users about exercise habits).
Population $\sigma = 6$ hours.

As n grows, the margin of error shrinks but the bias stays fixed:

n	Margin of error	Actual capture rate
100	± 1.18	87%
400	± 0.59	62%
1,000	± 0.37	25%
10,000	± 0.12	$\approx 0\%$

Caution: The capture rate only shrinks as n grows. More data makes the problem **worse**, not better.

Precision vs. Accuracy



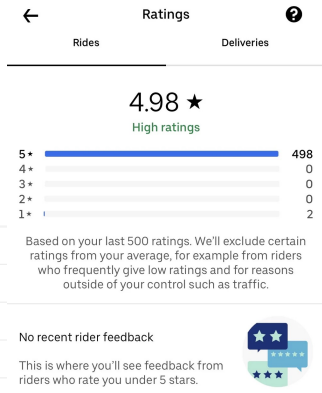
Principle: Increasing n narrows the precision but has no effect on the accuracy.

Uber/Lyft Star Ratings

Example 8.7

Context: You check an Uber driver's profile and see a 4.2-star average from 23 ratings.
Seven are 1-star reviews, complaining about rude behaviour or wrong routes.

Question: Should you conclude 30% of rides are terrible?



Quick Check: Biased or Not?

Example 8.8

Find: For each scenario, decide: is the sampling method **biased**?
If so, identify the type (convenience or voluntary response).

1. A radio station asks listeners to call in with their opinion on a new law.

2. A professor surveys every student in her class about study habits.

3. A company assigns each employee an ID and uses software to randomly select 50.

4. A journalist interviews people exiting a polling station about the election.

PART 3

Probability-Based Sampling Methods

How do we sample correctly when bias is not an option?

Why Probability-Based Sampling Methods?

A sample is **random** if it comes from a probability sampling design: one where every unit has a **known inclusion probability**.

If those probabilities are unknown or undefined, the sample is **non-random**.

Known inclusion probabilities mean we can:

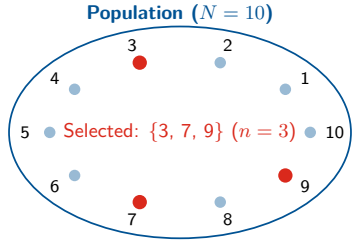
- Eliminate selection bias
- Generalise results to the population
- Measure the uncertainty of our estimates

Random selection is Condition 1 of the CLT (Chapter 15).
Without it, confidence intervals break down.

Simple Random Sampling (SRS)

Simple Random Sample (SRS)

A **simple random sample (SRS)** of size n consists of n individuals from the population, chosen in such a way that every set of n individuals has an equal chance of being the sample actually selected.

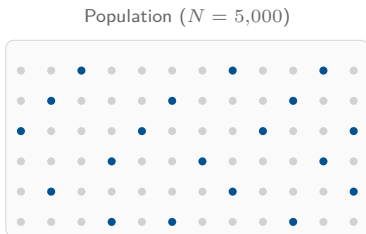


Intuition: Equivalently: every individual in the population has equal probability n/N of being selected, where n is the sample size and N is the population size.

Drawing an SRS

Example 8.9

Context: Western's registrar has an alphabetical list of all 5,000 first-year students. Researchers want to survey 200 of them about their transition to university.



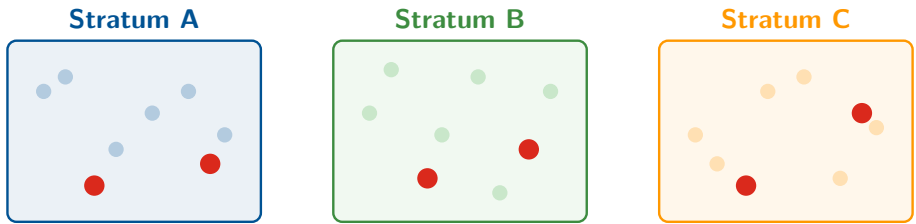
1. Assign each student an ID (1 to 5,000)
2. Randomly select 200 distinct IDs
3. Survey those 200 students

Intuition: Every set of 200 students is equally likely to be chosen.
No subgroup is systematically favoured or excluded.

Stratified Random Sampling

Stratified Random Sampling

Stratified random sampling divides the population into non-overlapping groups (**strata**), then takes an SRS from *each* stratum.



SRS from each stratum (red = selected)

When to Use Stratified Sampling

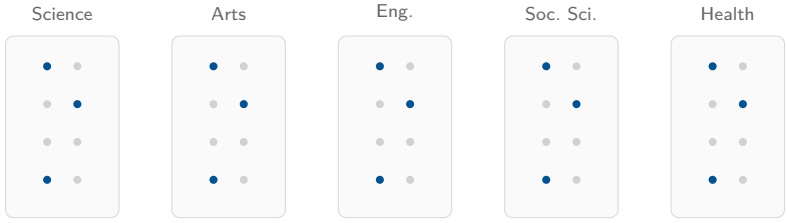
Stratified sampling works well when:

- The population has **distinct subgroups**
- You want to **ensure representation** from all subgroups
- The population differs more **between** groups than **within** groups (e.g., exercise habits differ more across faculties than within one faculty)

Stratified Sampling for Exercise Hours

Example 8.10

Context: Western has roughly 30,000 undergraduates across five faculties. Researchers want to survey 300 students about weekly exercise hours.



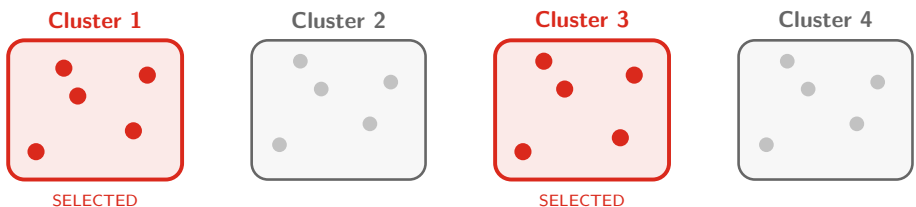
SRS of 60 from each stratum $\rightarrow n = 300$

Intuition: Every faculty is guaranteed representation.
Stratifying reduces variability when exercise habits differ across faculties.

Cluster Sampling

Cluster Sampling

Cluster sampling divides the population into groups (**clusters**), takes a random sample of clusters, and measures *all* individuals in the selected clusters.



Stratified vs. Cluster: Key Difference

Stratified

Sample from **ALL** strata.

Use *some* individuals from *each* group.

→ Ensures complete coverage.

Cluster

Use **SOME** clusters completely.

Include *all* individuals from selected groups.

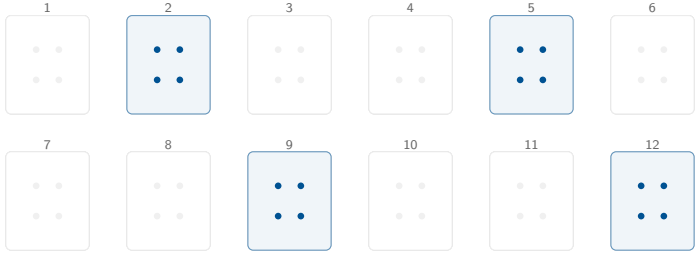
→ Reduces cost.

Intuition: Stratified = sample from **every** group. Cluster = use **some** groups entirely.

Cluster Sampling for Exercise Hours

Example 8.11

Context: Western has 12 first-year residence buildings, each housing 200–400 students. Surveying every building is too time-consuming.



Randomly select 4 buildings, survey ALL students inside

Intuition: More practical than an SRS across all residences.
Trade-off: students in the same building may have similar habits.

When to Use Cluster Sampling

Cluster sampling works well when:

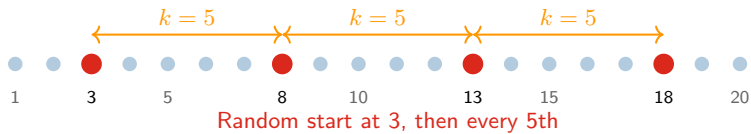
- Clusters are **geographically dispersed**
- It is expensive to sample **across many locations**
- Clusters are internally **diverse** (heterogeneous)

> ⚠ **Caution:** Cluster sampling is cheaper and more practical, but often less precise than SRS or stratified sampling, because members within a cluster tend to be similar.

Systematic Sampling

Systematic Sampling

Systematic sampling selects every kth element from a list of all elements in the population, after randomly choosing a starting point.



Caution: If there is a **periodic pattern** in the population (e.g., every 20th product comes from the same machine), systematic sampling can be severely biased.

Quality Control Inspection

Example 8.12

Context: A factory inspector checks every 20th product off the assembly line (after randomly choosing where to start in the first 20).
This spreads the sample evenly across production.

When it breaks: if the population has a repeating pattern with the same period as k, every selected unit comes from the same part of the cycle.

⚠ **Caution:** Systematic sampling fails when the sampling interval matches a hidden cycle in the data.

Quality Control Inspection

Example 8.12 (continued)

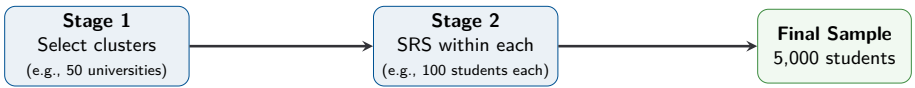
Examples of periodic patterns that break systematic sampling:

- **Day-of-week:** Sales records sorted by date; selecting every 7th record always lands on Sunday, underrepresenting weekday traffic.
- **Apartment layout:** Every 6th unit is a corner apartment; sampling every 6th unit overrepresents corner (larger, higher-rent) units.

Multistage Sampling

Multistage Sampling

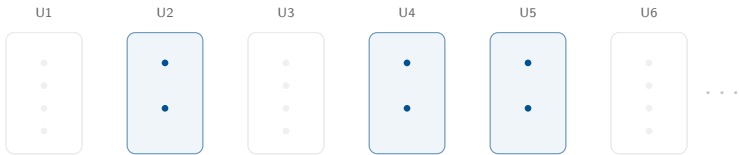
Multistage sampling is a sampling plan that uses a combination of sampling methods in various stages.



Multistage Sampling for Exercise Hours

Example 8.13

Context: Researchers want to estimate average weekly exercise hours for university students across Ontario.
No single list of all university students exists, so they sample in stages.



Stage 1: Randomly select 10 universities

Stage 2: SRS of 100 from each selected university

Intuition: Unlike stratified sampling (which samples from **all** groups), multistage uses only **some** groups from the first stage.

Summary of Sampling Designs

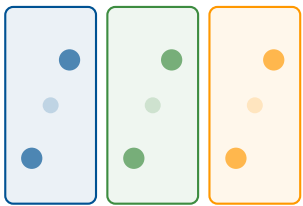
Task: Complete the table: fill in the key feature for each design.

Design	Key Feature
Census	
SRS	
Stratified	
Cluster	
Systematic	
Multistage	

Applies when: Convenience and voluntary response samples are **not** on this list because they are biased and should not be used for inference.

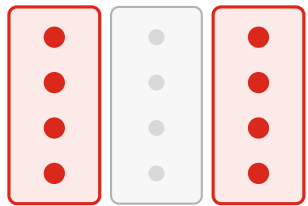
Visual Comparison

Stratified



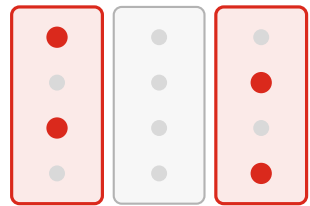
SRS from
EVERY group

Cluster



ALL members from
SOME groups

Multistage



SOME members from
SOME groups

PART 6

Practical Challenges in Sampling

Even a well-designed sample can go wrong: what else should we watch for?

Cautions About Sample Surveys

A good sampling *design* is necessary but not sufficient.
Four additional sources of error can undermine even a well-chosen sample:

- **Under-coverage**: some groups never get a chance to be sampled
- **Nonresponse**: chosen individuals cannot be reached or refuse
- **Response bias**: systematic pattern of incorrect answers
- **Wording effects**: question phrasing influences responses

Caution: These errors affect **accuracy**: systematic bias that no amount of data can fix.

Sampling frame

Sampling frame

A **sampling frame** is the list or source used to choose the sample.

Goal: The sampling frame should match the population as closely as possible.

Example: If the population is all students at a university, the registrar's student list could be the sampling frame.

Under-coverage

Under-coverage

Under-coverage occurs when some groups in the population have no chance to be selected.

Key idea: Even with a good random sample, under-coverage occurs if the *sampling frame* leaves out part of the population.

Example: A random sample from a voter list under-covers adults who are not registered to vote.

> ⚠ **Caution:** If a group is missing from the frame, it cannot appear in the sample.

Nonresponse

Nonresponse

Nonresponse occurs when an individual chosen for the sample cannot be contacted or refuses to participate. The person was *selected*; they just did not respond.

Notation: **Nonresponse bias** is the systematic error that arises when non-respondents differ from respondents in ways relevant to the survey topic.

Common misconception: Nonresponse is not the same as under-coverage.
Under-coverage: a group never had a chance to be selected.
Nonresponse: a selected person did not reply.

Response Bias

Response Bias

A systematic pattern of incorrect responses in a sample survey leads to **response bias**.

Common sources:

- **Social desirability**: answering to appear favourable rather than honestly
- **Interviewer influence**: tone, appearance, or presence of the interviewer shapes answers
- **Sensitive questions**: respondents avoid or misreport uncomfortable topics

Wording Effects

Wording Effects

Wording effects are the most important influence on the answers given to a sample survey.

Quick example:

"Should we *reduce* wasteful government spending?"
vs.
"Should we *cut* important social programmes?"

Same policy, opposite framing. Very different response rates.

Intuition: Respondents react to the words used, not just the underlying question. Small changes in phrasing can swing results dramatically.

Quick Check: Name the Error

Example 8.14

Find: Identify the error type: under-coverage, nonresponse, response bias, or wording effect.

1. A health survey on dietary habits samples only registered gym members.

2. A company emails 500 employees a satisfaction survey; 42 respond.

3. "Do you support cutting wasteful spending on foreign aid?"

When Response Rates Are Low

Even the best surveys cannot contact everyone, and not everyone contacted will respond.

- Response rates should be **reported** in any research summary
- Lower response rate $\Rightarrow$ results are harder to generalise
- Reminders and follow-up calls can reduce nonresponse

> ⚠ **Caution:** A low response rate introduces a form of voluntary response bias: the people who *do* respond may differ systematically from those who do not.

Identify the Source of Error

Example 8.15

Find: For each scenario, identify the source of error: under-coverage, nonresponse, response bias, or wording effect.

1. A survey on alcohol consumption is conducted face-to-face by uniformed police officers.

2. A mailed survey on neighbourhood satisfaction gets only 18% of forms returned.

3. A survey on reading habits uses a list of public library cardholders to draw its sample.

4. "Do you agree that the city should stop wasting money on bike lanes?"

Chapter 8: Concept Card: Biased Methods

Design	Key Feature	Watch Out For
Convenience	Researcher picks easy-to-reach individuals	Selection bias
Voluntary resp.	Subjects self-select	Strong-opinion bias

Common misconception: Neither method should be used for inference.
A bigger biased sample is still biased.
More data cannot fix selection bias.

Chapter 8: Concept Card: Probability Methods

Design	Key Feature	Watch Out For
SRS	Every sample equally likely	Impractical for large N
Stratified	SRS from ALL strata	Need known strata
Cluster	ALL from SOME groups	Less precise
Systematic	Every kth after random start	Repeating-pattern bias
Multistage	Combine methods in stages	Complex to analyse

Principle: A well-designed sample of 1,000 beats a poorly designed sample of 10,000. Sample size controls precision (Chapter 16), but only the sampling method controls accuracy.

Chapter 8: Concept Card: Survey Errors

Error Type	What it is	Example
Under-coverage	Groups left out of the sampling frame	Landline-only phone survey
Nonresponse	Selected units do not respond	Mailed survey with 20% return
Response bias	Systematic incorrect answers	Interviewer influences respondent
Wording effects	Question phrasing shifts answers	"cut" vs. "invest in"

Principle: A well-designed sample can still produce biased results if nonresponse, under-coverage, or poor wording distort the data. Design and execution both matter.

CHAPTER 8

Summary

Key Takeaways

- **Random selection is essential.** Probability-based methods eliminate selection bias and allow valid inference.
- **Different methods serve different purposes.** Choose the design that matches the population structure and practical constraints.
- **Avoid non-probability methods.** Convenience and voluntary response samples introduce bias.
- **Even good designs can go wrong.** Under-coverage, nonresponse, response bias, and wording effects can undermine any well-chosen sample.