CHAPTER 2

Collection of Data

CBSE Class 11 · Economics (Statistics, Part A) · Chapter 2

CBSE · Economics · Class 11

WHAT THIS CHAPTER DOES

A Distinguish the two SOURCES of data — primary and secondary.

B Name and explain the methods of collecting primary data.

Boards prep that builds confidence, not anxiety.

TODAY'S MISSION

Today's mission

1. Distinguish the two SOURCES of data — primary and secondary.
2. Name and explain the methods of collecting primary data.
3. Compare the CENSUS method with the SAMPLE method, and sampling with non-sampling errors.
4. Know the Census of India and the NSSO as India's key secondary sources.

WHY THIS MATTERS

Why this chapter matters

1. Collection is the FIRST step of the statistical method (collect → organise → present → analyse → interpret).

2. Comparison-heavy and high-scoring: clean two-column 'distinguish' answers win full marks.

3. Real-world link: every Census headline, every NSSO employment figure, every opinion poll lives in this chapter.

TOPIC

A

Sources of Data — Primary vs Secondary

THEOREM · LOAD-BEARING RESULT

The two sources of data

> Data come from TWO sources. PRIMARY data are collected first-hand by the investigator for a specific enquiry; SECONDARY data are already collected by someone else and reused. The difference is about the SOURCE, not the data itself.

STATEMENT

PRIMARY data: original data gathered directly from the field by the investigator for the first time, for a definite purpose. SECONDARY data: data that have already been collected, processed and

WHY THIS MATTERS

- The choice of source decides cost, time, accuracy and suitability
- Primary data fit the enquiry exactly but cost more
- secondary data are cheap and quick but may not match the enquiry perfectly and must be checked for reliability.

WATCH OUT FOR

NOTE Never say primary and secondary are 'different data'. Census data are PRIMARY for the Census office but SECONDARY for a student quoting them — same data, different source.

TOPIC

How primary data are collected

DIRECT PERSONAL INVESTIGATION

In direct personal investigation the investigator personally goes to the field and collects the data face-to-face from the informants. It gives the most ORIGINAL, accurate and reliable information because the investigator can observe, cross-check and adapt questions on

INDIRECT ORAL INVESTIGATION

When the informants cannot or will not give information directly, indirect oral investigation is used: the investigator questions OTHER persons (called witnesses or third parties) who are likely to know about the units under study. It is used in large enquiries and by official

QUESTIONNAIRE (MAILED)

In the mailed-questionnaire method a printed list of questions (the QUESTIONNAIRE) is sent by post or email to the respondents, who fill it in THEMSELVES and return it. Its merits are low cost, wide geographical coverage and no investigator bias; its serious demerits are a LOW response rate, the

SCHEDULE (ENUMERATOR-FILLED)

A SCHEDULE contains the same kind of questions as a questionnaire, but it is filled in by trained ENUMERATORS (investigators) who go to the respondents, read out the questions and record the answers. Its merits are that it works even with ILLITERATE

WORKED EXAMPLE

Primary or secondary? Questionnaire or schedule?

1. A student surveys 100 classmates with her own form ⇒ PRIMARY data (collected first-hand).
2. She later quotes literacy figures from the Census report ⇒ SECONDARY data for her (primary for the Census office).
3. A printed form respondents fill in themselves and post back ⇒ QUESTIONNAIRE.
4. The same questions read out and recorded by a trained field worker ⇒ SCHEDULE.
5. Rule of thumb: ask 'WHO collected it first?' (source) and 'WHO fills the form?' (questionnaire vs schedule).

TOPIC

B

Designing the Instrument — Questionnaire & Pilot Survey

THEOREM · LOAD-BEARING RESULT

A good questionnaire and the pilot survey

> "The quality of primary data depends on the quality of the questionnaire. A good questionnaire is short, simple, clear and unbiased, and it should be tested by a PILOT survey — a small trial run — before the main survey.

STATEMENT

Qualities of a GOOD questionnaire: it should be as SHORT as possible; questions should be SIMPLE, CLEAR and UNAMBIGUOUS; arranged in a LOGICAL order (general to specific); avoid LEADING,

WHY THIS MATTERS

- Poorly worded or leading questions produce non-sampling errors at the very source
- The pilot survey catches these flaws cheaply, BEFORE money is spent on the main survey, so the final data are more reliable.

WATCH OUT FOR

NOTE A pilot survey is NOT the final survey — it is a small trial. And a 'leading' question (one that hints at the expected answer) is a classic flaw examiners want you to flag.

TOPIC

C

Census Method vs Sample Method

TOPIC

Studying the whole vs studying a part

CENSUS (COMPLETE ENUMERATION)

In the CENSUS or complete-enumeration method, data are collected from EVERY single unit of the population (the 'universe'). Its great merit is COMPLETENESS and detail — no unit is left out, so it gives intensive, reliable information and is the natural choice

SAMPLE (PARTIAL ENUMERATION)

In the SAMPLE method, data are collected from only a representative PART (the sample) of the population, and conclusions about the whole are drawn from it. Its merits are that it is far CHEAPER, FASTER and needs fewer trained workers, and it is the ONLY feasible option for

WHEN TO USE WHICH

The choice between census and sample depends on the OBJECT of enquiry, the SIZE of the population, the FUNDS and TIME available, and the ACCURACY required. Use the CENSUS when the population is small, when complete detail is essential, or when even small units matter; use

RANDOM VS NON-RANDOM SAMPLING

Within the sample method, sampling can be RANDOM or NON-RANDOM. In RANDOM (probability) sampling every unit has a KNOWN, equal chance of being selected — e.g., the lottery or random-number method — which removes investigator bias and lets us measure the

TOPIC

D

Errors in Data Collection

THEOREM · LOAD-BEARING RESULT

Sampling vs non-sampling errors

> "Two kinds of error can creep into data. A SAMPLING error arises because we study only a part instead of the whole; a NON-SAMPLING error arises from mistakes in the actual collection and processing of data.

STATEMENT

SAMPLING error: the difference between the sample result and the true population value, caused by studying only a SAMPLE. It exists ONLY in sample surveys and DECREASES as

WHY THIS MATTERS

- Knowing the source of an error tells you how to reduce it
- A sampling error is cut by enlarging or improving the sample
- a non-sampling error is cut by better questionnaires, training and processing — not by a bigger sample.

WATCH OUT FOR

NOTE Do NOT say both errors fall as the sample grows. Only the SAMPLING error falls; the NON-SAMPLING error can actually RISE with a larger study because there is more data to mis-measure and mis-process. Also: the census has NO sampling error but can have large non-sampling errors.

WORKED EXAMPLE

Classify the error

1. A sample mean of household income differs from the true mean ⇒ SAMPLING error (only a part was studied).
2. Some households refuse to answer the survey ⇒ NON-SAMPLING error (non-response).
3. An enumerator misreads and records 45 as 54 ⇒ NON-SAMPLING error (recording mistake).
4. A leading question pushes respondents toward 'yes' ⇒ NON-SAMPLING error (measurement bias).
5. Lesson: only the FIRST is reduced by a bigger sample; the rest need better instruments, training and processing.

TOPIC

E

Important Sources of Secondary Data in India

TOPIC

Census of India & the NSSO

CENSUS OF INDIA

The CENSUS OF INDIA is the most important source of demographic and socio-economic secondary data in the country. It is conducted ONCE EVERY TEN YEARS (a decennial census) by the Office of the Registrar General and Census Commissioner of India, under the Ministry of Home Affairs. It uses the CENSUS (complete enumeration) method — every person in the country is counted — and provides data on population size, age, sex,

NSSO (NATIONAL SAMPLE SURVEY OFFICE)

The NSSO — National Sample Survey Office — was set up by the Government of India to conduct nationwide SAMPLE surveys on a regular basis covering socio-economic subjects such as EMPLOYMENT and unemployment, household CONSUMER expenditure, LITERACY, health, agriculture and the use of public services. Unlike the Census, it uses the SAMPLE method, so it can produce more frequent, detailed estimates at lower

TRY IT · SOLVE BEFORE YOU PEEK

Quick self-test (60 seconds)

Work it out before you flip the answer.

SOLUTION

1. Q1. Give one difference between primary and secondary data, with an example of each.
2. Q2. What is the difference between a questionnaire and a schedule?
3. Q3. State two differences between the census method and the sample method.
4. Q4. How does a sampling error differ from a non-sampling error?
5. Q5. How often is the Census of India held, and what method does the NSSO use?
6. Cover the slide and answer aloud before checking the Recap below.

TOPIC

Primary vs secondary data

TRAP → TRUTH

× **MISTAKE** Primary data and secondary data are two completely different kinds of data.

✓ **CORRECT** Primary and secondary describe the SOURCE, not the data itself. The SAME data is PRIMARY for the agency that collects it first-hand for its own purpose, and becomes SECONDARY when anyone else later uses those already-collected figures. Census data is primary for the Census office but secondary for a student who quotes it.

TOPIC

Questionnaire vs schedule

TRAP → TRUTH

× **MISTAKE** A questionnaire and a schedule are the same thing.

✓ **CORRECT** A QUESTIONNAIRE is filled in by the RESPONDENT himself (often mailed and returned). A SCHEDULE is the SAME list of questions but filled in by a trained ENUMERATOR/ investigator who reads the questions to the respondent. The difference is WHO fills the form, not the questions.

TOPIC

Census vs sample method

TRAP → TRUTH

× **MISTAKE** The census method is always better because it covers everyone.

✓ **CORRECT** The census method covers EVERY unit, so it is complete but COSTLY, slow and labour-heavy, and is impossible where testing destroys the unit. The SAMPLE method studies only a representative PART, is cheaper and faster, and — with good sampling — gives reliable results. Neither is universally 'better'; the choice depends on cost, time, accuracy needed and population size.

TOPIC

Sampling vs non-sampling errors

TRAP → TRUTH

× **MISTAKE** Sampling and non-sampling errors are the same and both rise as the sample gets larger.

✓ **CORRECT** A SAMPLING error arises because we study only a PART instead of the whole — it DECREASES as the sample size increases and exists ONLY in sample surveys. NON-SAMPLING errors (wrong measurement, non-response, bias, processing mistakes) can occur in BOTH census and sample studies and may actually GROW with a larger study. They are distinct in cause.

TOPIC

Random sampling

TRAP → TRUTH

× **MISTAKE** Random sampling means picking units carelessly / haphazardly.

✓ **CORRECT** RANDOM (probability) sampling means every unit has a KNOWN, equal chance of selection (e.g., lottery or random-number method) — it removes the investigator's bias. 'Haphazard' picking is NOT random; it is a form of NON-random, convenience selection that can introduce bias.

TOPIC

Census of India frequency

TRAP → TRUTH

× **MISTAKE** The Census of India is conducted every year by the NSSO.

✓ **CORRECT** The CENSUS OF INDIA is conducted ONCE EVERY TEN YEARS (decennial) by the Office of the Registrar General and Census Commissioner — NOT by the NSSO. The NSSO (National Sample Survey Office) is a SEPARATE body that conducts SAMPLE surveys on socio-economic topics (employment, consumption, literacy) on a continuing basis.

TOPIC

Pilot survey

TRAP → TRUTH

× **MISTAKE** A pilot survey is the final, full survey of the whole population.

✓ **CORRECT** A PILOT survey is a SMALL TRIAL run of the questionnaire on a few respondents BEFORE the main survey. Its purpose is to test the questionnaire — to spot ambiguous, confusing or missing questions and estimate cost and time — so the instrument can be improved before the real data collection begins.

TOPPER TEMPLATE · MARK-BY-MARK

3-mark: 'Distinguish between primary data and secondary data.'

1	**DEFINE PRIMARY DATA** 1 m	PRIMARY data are data collected by the investigator HIMSELF for the FIRST TIME, directly from the field for a specific enquiry — they are original and have not passed through any other hands (e.g., a researcher surveying 200 households about their spending).
2	**DEFINE SECONDARY DATA** 1 m	SECONDARY data are data that have ALREADY been collected and processed by someone else and are then used by the investigator for a NEW purpose — taken from published or unpublished sources such as government reports, the Census, NSSO reports, books and websites.
3	**TWO-COLUMN BASES OF DIFFERENCE** 1 m	Bases of difference: ORIGINALITY — primary are original, secondary are second-hand; COLLECTION — primary are freshly collected, secondary are already available; COST & TIME — primary cost more time and money, secondary are cheaper and quicker; SUITABILITY — primary fit the enquiry exactly, secondary may need adjustment. (Key idea: the SAME data are primary for the collector and secondary for a later user.)

TOPPER TEMPLATE · MARK-BY-MARK

4-mark: 'Distinguish between the census method and the sample method.'

1	**COVERAGE** 1 m	COVERAGE: the CENSUS (complete enumeration) method studies EVERY unit of the population, whereas the SAMPLE method studies only a representative PART (sample) of the population and infers about the whole from it.
2	**COST, TIME AND LABOUR** 1 m	COST/TIME/LABOUR: the census method is EXPENSIVE, time-consuming and needs a large workforce; the sample method is far CHEAPER, FASTER and needs fewer enumerators because only a part is studied.
3	**ACCURACY AND SUITABILITY** 1 m	ACCURACY/SUITABILITY: the census gives complete, detailed information and suits a SMALL or highly varied population; the sample suits a LARGE population and is the only option when testing DESTROYS the unit (e.g., testing matchsticks). A well-designed sample can be highly reliable.
4	**ERRORS + EXAMPLE** 1 m	ERRORS/EXAMPLE: sampling error exists only in the SAMPLE method (and falls as the sample grows); the census avoids sampling error but is prone to large non-sampling errors due to its size. Example: the Census of India uses the census method (every person counted, once in 10 years); the NSSO uses the sample method.

TOPPER TEMPLATE · MARK-BY-MARK

3-mark: 'Distinguish between sampling and non-sampling errors.'

1	**SAMPLING ERROR** 1 m	A SAMPLING error arises because only a PART (sample) of the population is studied instead of the whole, so the sample estimate differs from the true population value. It exists ONLY in sample surveys and DECREASES as the sample size increases.
2	**NON-SAMPLING ERROR** 1 m	NON-SAMPLING errors arise from faults in the actual process of collecting and handling data — wrong measurement, biased or ambiguous questions, non-response, wrong recording, and mistakes in processing/tabulation. They are not caused by sampling.
3	**KEY CONTRAST** 1 m	Key contrast: sampling error occurs ONLY in sample studies and shrinks with a bigger sample; non-sampling error can occur in BOTH census and sample studies and may even GROW with a larger study because more data means more chances of mistakes.

PYQ PATTERNS

Top PYQ patterns to drill

#1	Distinguish between primary data and secondary data. (3-4 marks)	Most school papers + SQP
#2	Explain any two methods of collecting primary data. (3-4 marks)	Annual
#3	Distinguish between the census method and the sample method of collecting data. (3-4 marks)	Annual
#4	Distinguish between sampling and non-sampling errors. (3-4 marks)	Frequent
#5	State the qualities of a good questionnaire / explain a pilot survey. (3 marks)	Unit tests

RECAP · MEMORISE THESE

Recap

1 Two sources — PRIMARY = collected first-hand by the investigator; SECONDARY = already collected by someone else and reused. Same data: primary for collector, secondary for later user.

2 Methods & instrument — Primary methods: direct/ indirect investigation, telephone, mailed QUESTIONNAIRE (respondent fills) and enumerator-filled SCHEDULE. A good questionnaire is short, clear, unbiased and pilot-tested.

3 Census vs sample & errors — Census studies EVERY unit (complete, costly); sample studies a PART (cheap, fast). Sampling error: only in samples, falls with size. Non-sampling error: in both, from faulty collection/ processing.

WHAT'S NEXT

What's next

→ Chapter 3 — Organisation of Data (classification, variables, frequency distribution).

→ Sit the 15-MCQ Quick Drill below.

→ Then attempt the full School-Pattern Paper — 30 marks.

You now know exactly where economic data comes from.

Primary vs secondary, census vs sample, sampling errors and the NSSO — prove it.

readyforboards.com

Helpline: +91 70330 05444

Boards prep that builds confidence, not anxiety.