CHAPTER 3

Organisation of Data

CBSE Class 11 · Economics (Statistics, Part A) · Chapter 3

CBSE · Economics · Class 11

WHAT THIS CHAPTER DOES

A Explain why raw data must be organised.

B Classify data on the four bases and tell discrete from continuous.

Boards prep that builds confidence, not anxiety.

TODAY'S MISSION

Today's mission

1. Explain why raw data must be organised.
2. Classify data on the four bases and tell discrete from continuous.
3. Build a frequency distribution from raw data using tally marks.
4. Master class limits vs boundaries and the inclusive vs exclusive method.

WHY THIS MATTERS

Why this chapter matters

1. It is the bridge between COLLECTING data (Ch 2) and PRESENTING it (Ch 4) — the tidying step.
2. High-scoring: the frequency-table numerical is mechanical once practised.
3. Real-world link: every survey result you see in the news was first organised exactly this way.

TOPIC

A

Raw Data & the Need to Organise

THEOREM · LOAD-BEARING RESULT

From raw data to organised data

> RAW (unclassified) data is the unprocessed mass of figures exactly as collected — disordered and hard to interpret. ORGANISATION is the process of arranging this raw data into a systematic form (classes and tables) so it becomes meaningful and ready for analysis.

STATEMENT

Classification of data is the process of arranging data into HOMOGENEOUS groups or classes according to some common characteristic. The organised data reveals points of similarity and dissimilarity,

WHY THIS MATTERS

- A list of, say, 200 marks tells us almost nothing at a glance
- Grouped into a frequency table it instantly shows where most students lie, the spread, and the extremes — the whole purpose of statistics.

WATCH OUT FOR

NOTE Organising data is NOT the same as collecting it (Ch 2) or presenting it as a graph (Ch 4). This chapter is the middle step — classification and tabulation into a frequency distribution.

TOPIC

The four bases of classification

GEOGRAPHICAL (SPATIAL)

Geographical or spatial classification arranges data according to PLACE or location — by country, state, district, region or city. For example, the population or rainfall of different Indian states listed side by side is a geographical classification. The classifying characteristic is WHERE the data

CHRONOLOGICAL (TEMPORAL)

Chronological or temporal classification arranges data according to TIME — by year, quarter, month or hour. A table showing India's GDP for 2019, 2020, 2021 and 2022 is a chronological classification; such time-ordered data is also called a TIME SERIES.

QUALITATIVE (BY ATTRIBUTE)

Qualitative classification groups data according to some ATTRIBUTE or quality that cannot be measured numerically — such as sex, literacy, religion, marital status or employment status. Because the attribute is either present or absent, the data splits into categories (male/female,

QUANTITATIVE (BY VARIABLE)

Quantitative classification groups data according to a measurable characteristic — a VARIABLE such as height, weight, income, marks or age — that is expressed in numbers. The variable may be DISCRETE (whole separate values) or CONTINUOUS (any value

TOPIC

B

Variables — Discrete vs Continuous

THEOREM · LOAD-BEARING RESULT

Discrete and continuous variables

> A VARIABLE is a measurable characteristic that takes different values. A DISCRETE variable can take only certain exact, separate values (with gaps between them). A CONTINUOUS variable can take ANY value within a range, including fractions.

STATEMENT

DISCRETE variable: assumes only whole/specific values and jumps from one to the next — e.g., number of children (0, 1, 2, 3; never 2.5), number of cars, members in a family. CONTINUOUS variable:

WHY THIS MATTERS

- The type of variable decides HOW you organise it — single-value array (discrete) versus class-interval table (continuous)
- Choosing wrongly produces a malformed table.

WATCH OUT FOR

NOTE Size is irrelevant. 'Number of buses in a city' may be in thousands yet is DISCRETE; 'time taken in seconds' may be tiny yet is CONTINUOUS. The only test is whether fractional in-between values are possible.

TOPIC

C

The Frequency Distribution

TOPIC

Key terms of a frequency distribution

FREQUENCY & CLASS

The FREQUENCY of a value or a class is the NUMBER OF TIMES it occurs — how many observations fall into that group. A CLASS (or class interval) is a group into which the range of the variable is divided, written with a lower and an upper value, for example 20–30. A FREQUENCY

CLASS LIMITS & BOUNDARIES

The CLASS LIMITS are the two end-values actually written in a class — the LOWER limit and the UPPER limit (20 and 29 in the inclusive class 20–29). The CLASS BOUNDARIES (true limits) are the values obtained after the adjustment that removes the gap between inclusive classes

MID-VALUE & MAGNITUDE

The MID-VALUE (class mark) of a class is (lower limit + upper limit) ÷ 2 — for the class 20–30 it is (20+30)/2 = 25 — and it represents the whole class in later calculations such as the mean. The MAGNITUDE (class width or size) is (upper limit – lower limit) of a class — here 30 – 20 = 10. A good answer gives both

FREQUENCY ARRAY

A FREQUENCY ARRAY is the form used for DISCRETE data: instead of class intervals, each distinct VALUE the variable takes is listed once alongside the number of times it occurs. For example, the number of children per family (0, 1, 2, 3, 4) each paired with how many families have that

WORKED EXAMPLE

Constructing a frequency distribution (tally marks)

1. Raw data — marks of 30 students: 23, 45, 12, 34, 56, 28, 41, 19, 37, 50, 22, 48, 31, 15, 39, 44, 27, 33, 52, 18, 25, 47, 36, 29, 42, 11, 38, 49, 24, 30.
2. Step 1 — Range = 56 − 11 = 45. Choose 5 equal exclusive classes of magnitude 10: 10–20, 20–30, 30–40, 40–50, 50–60 (lower limit included, upper limit excluded).
3. Step 2 — Tally each value once. 10–20: 12,19,15,18,11 → 5. 20–30: 23,28,22,27,25,29,24 → 7. 30–40: 34,37,31,39,33,36,38,30 → 8 (note 30 goes to 30–40, not 20–30). 40–50: 45,41,48,44,47,42,49 → 7. 50–60: 56,50,52 → 3.
4. Step 3 — Frequencies: 5, 7, 8, 7, 3.
5. Step 4 — CHECK: 5 + 7 + 8 + 7 + 3 = 30 = N. The total matches the number of students, so the table is correct.

TOPIC

D

Inclusive vs Exclusive Method

THEOREM · LOAD-BEARING RESULT

Two ways to write class intervals

> In the EXCLUSIVE method the upper limit of a class is excluded and becomes the lower limit of the next class (10–20, 20–30) — no gap, suited to continuous data. In the INCLUSIVE method both limits are included (10–19, 20–29) and there is a gap between classes — convenient for discrete data.

STATEMENT

EXCLUSIVE: classes are 10–20, 20–30, 30–40... and the rule is lower limit included, upper limit excluded (a value of 20 goes to 20–30). INCLUSIVE: classes are 10–19, 20–29, 30–39... and both

WHY THIS MATTERS

- Statistical formulae (mean, median by interpolation, histograms) assume continuous, gap-free classes
- So an inclusive series is usually CONVERTED to exclusive boundaries before further work.

WATCH OUT FOR

NOTE In the exclusive method the upper limit is EXCLUDED — a value of exactly 30 goes to 30–40, NOT 20–30. Forgetting this single rule is the commonest tally error.

WORKED EXAMPLE

Converting an inclusive series to exclusive

1. Inclusive series: 20–29, 30–39, 40–49 with frequencies 5, 8, 4.
2. Step 1 — Find the GAP = lower limit of next class – upper limit of present class = 30 – 29 = 1.
3. Step 2 — Half the gap = 1 ÷ 2 = 0.5.
4. Step 3 — Subtract 0.5 from every lower limit, add 0.5 to every upper limit: 20–29 → 19.5–29.5; 30–39 → 29.5–39.5; 40–49 → 39.5–49.5.
5. Step 4 — Frequencies are UNCHANGED (5, 8, 4). The classes now touch with no gap — a proper exclusive (true-limit) series ready for calculation.

TOPIC

E

Tally, Loss of Information & Bivariate Data

TOPIC

Three remaining ideas

TALLY MARKS

TALLY MARKS are the counting device used while building a frequency table. As you read each observation, you put one vertical stroke against its class; the FIFTH stroke is drawn across the previous four to make a bundle of five, so frequencies can be counted quickly in fives

LOSS OF INFORMATION

Classification brings clarity but at a price: a LOSS OF INFORMATION. Once individual observations are placed into a class such as 20–30, their exact values (21, 27, 29...) are no longer visible — for all further calculation every item in the class is treated as the mid-value (25). So grouping

UNIVARIATE VS BIVARIATE

A UNIVARIATE frequency distribution involves only ONE variable — for example, students classified by marks alone. A BIVARIATE frequency distribution involves TWO variables measured together and shown in a two-way table — for example, families classified simultaneously by income AND

WHY ORGANISE AT ALL

Pulling the chapter together: organisation of data CONDENSES a large mass of figures, brings out points of SIMILARITY and DISSIMILARITY, makes COMPARISON possible, and prepares the data for tabulation, diagrams and analysis in the next chapters. The cost is the loss of individual detail, but the

TRY IT · SOLVE BEFORE YOU PEEK

Quick self-test (60 seconds)

Work it out before you flip the answer.

SOLUTION

1. Q1. Name the four bases of classification with one example each.
2. Q2. Give one discrete and one continuous variable.
3. Q3. In the exclusive series 20–30, 30–40, which class does the value 30 fall into?
4. Q4. Convert the inclusive class 10–19 to exclusive boundaries.
5. Cover the slide and answer aloud before checking the Recap below.

TOPIC

Class limits vs class boundaries

TRAP → TRUTH

× **MISTAKE** Class limits and class boundaries are the same thing.

✓ **CORRECT** CLASS LIMITS are the two end-values actually WRITTEN in a class (e.g., 20 and 29 in the inclusive class 20–29). CLASS BOUNDARIES (true limits) are the values obtained AFTER the adjustment that removes the gap, used for continuous treatment (19.5 and 29.5 for 20–29). In an EXCLUSIVE series the upper limit equals the next lower limit, so limits and boundaries coincide; in an INCLUSIVE series they differ by the half-gap adjustment.

TOPIC

Inclusive vs exclusive method

TRAP → TRUTH

× **MISTAKE** In the exclusive method (e.g., 20–30, 30–40) an item of value 30 is counted in the 20–30 class because 30 is written there.

✓ **CORRECT** In the EXCLUSIVE method the UPPER limit is EXCLUDED. A value of exactly 30 goes into the 30–40 class, NOT 20–30. The rule is 'lower limit included, upper limit excluded' ($20 \leqslant x < 30$). In the INCLUSIVE method (20–29, 30–39) BOTH limits are included and there is a gap between classes.

TOPIC

Discrete vs continuous variable

TRAP → TRUTH

× **MISTAKE** A variable is continuous if it takes large values, and discrete if it takes small values.

✓ **CORRECT** It has NOTHING to do with size. A DISCRETE variable takes only whole/exact separate values with gaps (number of children: 0,1,2 — never 1.5). A CONTINUOUS variable can take ANY value within a range, including fractions (height, weight, time — 162.4 cm is valid). The test is whether in-between (fractional) values are possible, not how big the numbers are.

TOPIC

The adjustment (inclusive to exclusive)

TRAP → TRUTH

× **MISTAKE** To convert an inclusive series to exclusive you just rewrite the limits.

✓ **CORRECT** You must apply the ADJUSTMENT. Find the gap = (lower limit of next class – upper limit of present class); halve it; SUBTRACT half from every lower limit and ADD half to every upper limit. For classes 20–29, 30–39 the gap is 1, half is 0.5, so the exclusive class boundaries become 19.5–29.5, 29.5–39.5. Frequencies do NOT change — only the limits do.

TOPIC

Mid-value (class mark)

TRAP → TRUTH

× **MISTAKE** The mid-value of a class is just any number inside the class.

✓ **CORRECT** The MID-VALUE (class mark) is a FIXED number: mid-value = (lower limit + upper limit) ÷ 2. For the class 20–30 it is (20+30)/2 = 25. It represents the whole class in later calculations (mean, etc.). It is computed from the LIMITS written in the table.

TOPIC

Frequency distribution vs frequency array

TRAP → TRUTH

× **MISTAKE** A frequency array and a frequency distribution mean exactly the same thing.

✓ **CORRECT** A FREQUENCY ARRAY is for DISCRETE data — each distinct VALUE is listed with its frequency (no class intervals). A (grouped) FREQUENCY DISTRIBUTION is for CONTINUOUS data — values are grouped into CLASS INTERVALS, each with a frequency. Array = single values; distribution = intervals.

TOPIC

Loss of information in classification

TRAP → TRUTH

× **MISTAKE** Classifying data into class intervals keeps every original observation intact.

✓ **CORRECT** Grouping causes a LOSS OF INFORMATION. Once items go into a class like 20–30, the INDIVIDUAL values (21, 27, 29...) are LOST — for further work all are treated as the mid-value (25). Classification gains clarity and comparability but at the cost of the original detail of each observation.

TOPPER TEMPLATE · MARK-BY-MARK

5-6 mark: 'Construct a frequency distribution from the given raw data using tally marks.'

1	**FIND RANGE & CHOOSE CLASSES** 1 m	First find the RANGE = highest value – lowest value. Decide the number of classes and the class MAGNITUDE (width) so that classes are equal and convenient, and write the class intervals using the EXCLUSIVE method (e.g., 10–20, 20–30, ...), stating that the lower limit is included and the upper limit excluded.
2	**TALLY EACH OBSERVATION** 2 m	Go through the raw data ONCE, placing one TALLY MARK against the correct class for each value. Group tallies in bundles of five (the fifth crosses the previous four) for easy counting. Be careful with boundary values: a value equal to an upper limit (e.g., 20) goes to the NEXT class (20–30), not the previous one.
3	**WRITE FREQUENCIES & TOTAL** 1 m	Count the tallies in each class to get the class FREQUENCY, write it in the frequency column, and total the frequencies. The total MUST equal the number of observations (N); show this check explicitly.
4	**PRESENT THE NEAT TABLE** 1 m	Present the final answer as a clean three-column table — Class Interval | Tally Marks | Frequency — with the total row. A tidy, correctly-totalled table secures full marks.

TOPPER TEMPLATE · MARK-BY-MARK

3-4 mark: 'Distinguish between class limits and class boundaries / find mid-value & magnitude.'

1	**CLASS LIMITS** 1 m	CLASS LIMITS are the two end-values actually WRITTEN in a class — the LOWER limit and the UPPER limit (e.g., in 20–29 they are 20 and 29). They are used to write and read the class.
2	**CLASS BOUNDARIES** 1 m	CLASS BOUNDARIES (true limits) are the values got AFTER the adjustment that removes the gap between inclusive classes — got by subtracting half the gap from the lower limit and adding it to the upper limit (20–29 becomes 19.5–29.5). They allow the data to be treated as continuous.
3	**MID-VALUE** 1 m	MID-VALUE (class mark) = (lower limit + upper limit) ÷ 2. For 20–30 it is (20+30)/2 = 25; it represents the whole class in later calculations.
4	**MAGNITUDE + WHEN THEY COINCIDE** 1 m	MAGNITUDE (width) = upper limit – lower limit of a class (e.g., 30 – 20 = 10). NOTE: in an EXCLUSIVE series, limits and boundaries are the SAME because there is no gap; they differ only in an INCLUSIVE series.

TOPPER TEMPLATE · MARK-BY-MARK

4-mark: 'Distinguish between the inclusive and exclusive method: convert the series to

1	**EXCLUSIVE METHOD** 1 m	In the EXCLUSIVE method classes are written so the UPPER limit of one class is the LOWER limit of the next (10–20, 20–30). The upper limit is EXCLUDED: a value of 20 goes into 20–30. There is NO gap between classes, so it suits continuous data.
2	**INCLUSIVE METHOD** 1 m	In the INCLUSIVE method BOTH limits are INCLUDED in the class (10–19, 20–29) and there is a GAP between the upper limit of one class (19) and the lower limit of the next (20). It is convenient for discrete data.
3	**THE ADJUSTMENT RULE** 1 m	To convert inclusive to exclusive, find the GAP = (lower limit of next class – upper limit of present class); take HALF of it; SUBTRACT half from each lower limit and ADD half to each upper limit. Frequencies are unchanged.
4	**WORKED CONVERSION** 1 m	For 20–29, 30–39: gap = 30 – 29 = 1, half = 0.5. So 20–29 → 19.5–29.5 and 30–39 → 29.5–39.5. Each class now joins the next with no gap — the exclusive (true-limit) series.

PYQ PATTERNS

Top PYQ patterns to drill

#1	Construct a frequency distribution / array from the given raw data using tally marks. (4-6 marks)	Most school papers + SQP
#2	Distinguish between class limits and class boundaries (and find mid-value & magnitude). (3-4 marks)	Annual
#3	Distinguish between the inclusive and exclusive method; convert an inclusive series to exclusive. (3-4 marks)	Annual
#4	Distinguish between a discrete and a continuous variable with examples. (3 marks)	Frequent
#5	Explain the bases / types of classification of data. (3-4 marks)	Unit tests

RECAP · MEMORISE THESE

Recap

1 Classify & types — Four bases: geographical (place), chronological (time), qualitative (attribute), quantitative (variable). Variables are discrete (separate values) or continuous (any value, fractions).

2 Frequency table — Frequency = count in a class; mid-value = (L+U)/2; magnitude = U–L. Build by tally marks in one pass; Σf must equal N. Discrete → array; continuous → grouped distribution.

3 Inclusive vs exclusive — Exclusive excludes the upper limit, no gap; inclusive includes both, has a gap. Convert by the half-gap adjustment (frequencies unchanged). Grouping causes loss of information.

WHAT'S NEXT

What's next

→ Chapter 4 — Presentation of Data (tables, bar diagrams, histograms, frequency polygons, ogives).

→ Sit the 15-MCQ Quick Drill below.

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

You can now turn any heap of raw data into a clean frequency table.

Classification, tally marks, class limits and the inclusive/exclusive method — now prove it.

readyforboards.com

Helpline: +91 70330 05444

Boards prep that builds confidence, not anxiety.