Unit 20 · Statistics

STUDY NOTES

Key terms

Data: facts or numbers collected about something.

Raw data: data written down as they were collected, in no order.

Frequency: how many times an answer comes in the data.

Tally marks (tally bars): one stroke for each answer; every fifth goes across the four before it, so they are counted in fives.

Frequency table: every answer in the raw data with its tally marks and its frequency.

Simple bar diagram: the frequencies as rectangular bars, one for each item, each as tall as its frequency.

Making a frequency table

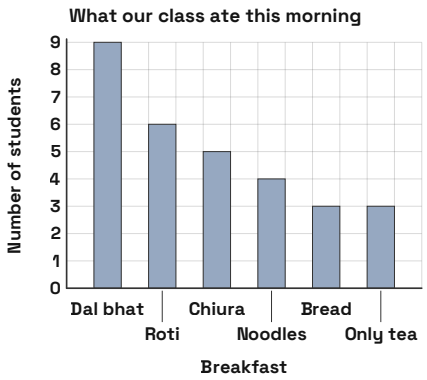
1. one row for each answer; numbers in order, smallest first
2. go through the raw data **once**, first to last, one tally mark each
3. count the tallies into the frequency column
4. **check**: the frequencies add up to the number of answers

- vegetables: $4 + 5 + 8 + 6 + 5 + 3 = 31$ answers; cauliflower most (8), cabbage least (3)
- milk (ml): 500, 700, 1000, 1500, 2000 → 10, 6, 7, 4, 3; total 30 families
- frequencies 7, 4, 9, 5 → 25 answers in the raw data
- "3, 5, 3, 4, 3, 5": frequency of 3 is 3 (not 6, not 9)

Drawing a simple bar diagram

1. **X-axis** along the bottom for the items, **Y-axis** up the side for the numbers; name both
2. a **title** that says what it shows
3. every bar the **same width**
4. the **same gap** between every two bars
5. a **scale** from **0**, in equal steps, so the tallest bar fits

- equal widths: a wider bar would look bigger than its number
- choose what one square stands for: 30, 40, 35, 20, 25 are multiples of 5 → 1 square = 5 students; 40 is 8 squares
- reading: 9 squares at 5 students a square = $9 \times 5 = 45$ (not $9 + 5$)



The same breakfasts as bars: dal bhat (9 of 30) is the tallest.

Reading a bar diagram

- tallest bar = most; shortest = least: yellow 50, red 20
- total = add every bar: $20 + 30 + 50 + 40 + 35 = 175$ students
- growth each year = this year − last year: city 35, 40, 50, 65, 90, 100 lakh → least growth 5 lakh (2071)
- frequency table from a bar diagram: each bar's label and height in a row, the total at the end
- a part as a percent of the whole: food 50 of 145 thousand → $50/145 \times 100\% \approx 34.48\%$

Why the scale starts at 0

- Bishal 48 votes, Nikita 52: a scale from 44 shows only 4 and 8, so her bar looks twice as tall
- she got only 4 votes more; look where a newspaper chart's scale starts

Where it is used

- census: raw data from every family, published as frequency tables; Florence Nightingale's diagrams of soldiers' deaths

COMMON MISTAKES

✗ hunting the list for all the cauliflowers first
✓ one pass, first to last, one mark each.

✗ frequencies not checked
✓ they must add up to the number of answers.

✗ scale starting at 44
✓ start at 0: from 44, 4 more votes looked twice as tall.