Data Collection and Organisation

Data collection involves gathering numerical information through surveys, observations, or experiments. The collected raw data is then systematically arranged to facilitate analysis.

Concept Explanation

Raw data is unprocessed numerical information collected from a primary source. Organising data means arranging it in ascending or descending order and tabulating it using tally marks for clarity.

Formula Derivation

Tally marks are used to count data efficiently. Every group of five is represented by four vertical lines with a diagonal line crossing them, enabling quick counting.

Worked Illustration

Example: Organise the following data of students joining different sports: Hockey - 10, Cricket - 15, Basketball - 3, Tennis - 8.

Step 1: Arrange data in ascending order: 3, 8, 10, 15.

Step 2: Represent using tally marks:

- 3: |||
- 8: |||| \ ||| (one group of 5 and 3 single lines)
- 10: |||| \ |||| (two groups of 5)
- 15: |||| \ |||| \ |||| (three groups of 5)

Solved Example

Given the data above, create a tally mark table:

Sport	Tally Marks	Number of Students
Basketball	|||	3
Tennis	|||| \ |||	8
Hockey	|||| \ ||||	10
Cricket	|||| \ |||| \ ||||	15

Practice Set

- **Level 1 – Easy:** Organise the data: 4, 7, 2, 9, 5 using tally marks.
- **Level 2 – Moderate:** Given the number of books read by students: 12, 15, 9, 7, 10, organise and represent using tally marks.
- **Level 3 – Challenging:** A survey recorded the number of cars passing through a street in 5 hours: 23, 17, 35, 29, 40. Organise and represent the data using tally marks.

Answer Key

- Level 1: 2: ||, 4: ||||, 5: |||| \ |, 7: |||| \ |||, 9: |||| \ ||||
- Level 2: 7: |||| \ |||, 9: |||| \ ||||, 10: |||| \ ||||, 12: |||| \ |||| \ ||, 15: |||| \ |||| \ ||||
- Level 3: 17: |||| \ |||| \ |||, 23: |||| \ |||| \ |||| \ ||, 29: |||| \ |||| \ |||| \ |||| \ |, 35: |||| \ |||| \ |||| \ |||| \ |||| \ |||, 40: |||| \ |||| \ |||| \ |||| \ |||| \ ||||

Quick Reference

- Tally marks group in fives: four vertical lines crossed by a diagonal line.
- Arrange data in ascending or descending order before tallying.

Glossary

- **Raw Data:** Unprocessed numerical information collected from primary sources.
- **Tally Marks:** A counting method using groups of five lines for easy data representation.
- **Organisation of Data:** Arranging raw data systematically for analysis.

Data Representation

Data representation involves displaying organised data visually using pictographs and bar graphs to facilitate comparison and interpretation.

Concept Explanation

Pictographs use symbols to represent data quantities, while bar graphs use bars of varying heights to show numerical values.

Formula Derivation

In pictographs, one symbol represents a fixed number of units. In bar graphs, the height of each bar is proportional to the data value it represents.

Worked Illustration

Example: Lakhanpal collected data on students absent in classes 1 to 8:

- Class 1: 3 students
- Class 2: 5 students
- Class 3: 2 students
- Class 4: 6 students
- Class 5: 0 students
- Class 6: 4 students
- Class 7: 1 student
- Class 8: 7 students

He used a pictograph where one smiley face 😊 represents one student.

Solved Example

Construct a pictograph for the data above:

- Class 1: 😊😊😊
- Class 2: 😊😊😊😊😊
- Class 3: 😊😊
- Class 4: 😊😊😊😊😊😊
- Class 5: (no symbol)
- Class 6: 😊😊😊😊
- Class 7: 😊
- Class 8: 😊😊😊😊😊😊😊

Similarly, a bar graph can be drawn with classes on the x-axis and number of students absent on the y-axis, with bars of corresponding heights.

Practice Set

- **Level 1 – Easy:** Represent the data: 2, 4, 3, 5 using pictographs where one symbol represents one unit.
- **Level 2 – Moderate:** Create a bar graph for the data: 7, 3, 9, 6, 2.
- **Level 3 – Challenging:** Given data on monthly rainfall (in cm): 12, 20, 30, 35, 41, 10, 37, represent it using a bar graph with appropriate scale.

Answer Key

- Level 1: Pictograph with 2, 4, 3, 5 symbols respectively.
- Level 2: Bar graph with bars of heights 7, 3, 9, 6, 2 units.
- Level 3: Bar graph with y-axis scaled to cover 0 to 45 cm, bars drawn accordingly.

Quick Reference

- Pictograph: One symbol = fixed number of units.
- Bar Graph: Height of bar proportional to data value.

Glossary

- **Pictograph:** A chart that uses symbols to represent data quantities.
- **Bar Graph:** A graph that uses bars to show data values.

Statistical Measures

Statistical measures summarise data using mean, median, and mode to describe central tendency.

Concept Explanation

Mean is the average of data values, median is the middle value when data is ordered, and mode is the most frequently occurring value.

Formula Derivation

Mean:

the mean

$\bar{x}$

of

n

data points

$x_1, x_2, \ldots, x_n$

is given by

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

Median: Arrange data in ascending order. If n is odd, median is the middle value $x_{(\frac{n+1}{2})}$. If n is even, median is average of two middle values $\frac{x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}}{2}$.

Mode: The data value that appears most frequently.

Worked Illustration

Data: 3, 7, 7, 2, 9, 10, 7

- Arrange in ascending order: 2, 3, 7, 7, 7, 9, 10
- Mean: $\frac{2+3+7+7+7+9+10}{7} = \frac{45}{7} \approx 6.43$
- Median: Middle value (4th value) = 7
- Mode: 7 (appears 3 times)

Solved Example

Find mean, median, and mode of the data: 5, 8, 12, 5, 6, 8, 5

- Arrange: 5, 5, 5, 6, 8, 8, 12
- Mean: $\frac{5+5+5+6+8+8+12}{7} = \frac{49}{7} = 7$
- Median: 4th value = 6
- Mode: 5 (appears 3 times)

Practice Set

- **Level 1 – Easy:** Find mean, median, and mode of 4, 6, 8, 6, 4.
- **Level 2 – Moderate:** Data: 10, 15, 10, 20, 25, 20, 10, find mean, median, mode.
- **Level 3 – Challenging:** Data: 12, 15, 12, 18, 20, 15, 12, 22, 25, find mean, median, mode.

Answer Key

- Level 1: Mean = 5.6, Median = 6, Mode = 4 and 6 (bimodal)
- Level 2: Mean = 15, Median = 15, Mode = 10
- Level 3: Mean = 16.1, Median = 15, Mode = 12

Quick Reference

- Mean = Sum of data values / Number of data points

- Median = Middle value in ordered data
- Mode = Most frequent data value

Glossary

- **Mean:** Average value of data.
- **Median:** Middle value in ordered data.
- **Mode:** Most frequent value in data.

Probability

Probability quantifies the chance of an event occurring based on favourable and total outcomes.

Concept Explanation

Probability of an event E is the ratio of the number of favourable outcomes to the total number of possible outcomes.

Formula Derivation

Let $n(E)$ be the number of favourable outcomes and $n(S)$ be the total number of outcomes in sample space S. Then,

$$P(E) = \frac{n(E)}{n(S)}$$

Worked Illustration

Example: Probability of getting a 4 when rolling a fair six-sided die.

- Total outcomes $n(S) = 6$
- Favourable outcomes $n(E) = 1$ (only one face with 4)
- Probability $P(E) = \frac{1}{6}$

Solved Example

Find the probability of drawing an ace from a standard deck of 52 cards.

- Total outcomes $n(S) = 52$
- Favourable outcomes $n(E) = 4$ (4 aces)
- Probability $P(E) = \frac{4}{52} = \frac{1}{13}$

Practice Set

- **Level 1 – Easy:** Probability of getting heads in a coin toss.
- **Level 2 – Moderate:** Probability of drawing a red card from a deck of 52 cards.
- **Level 3 – Challenging:** Probability of rolling an even number on a six-sided die.

Answer Key

- Level 1: $P = \frac{1}{2}$
- Level 2: $P = \frac{26}{52} = \frac{1}{2}$
- Level 3: $P = \frac{3}{6} = \frac{1}{2}$ (even numbers are 2, 4, 6)

Quick Reference

- Probability $P(E) = \frac{\text{Number of favourable outcomes}}{\text{Total number of outcomes}}$
- Probability values range from 0 to 1.

Glossary

- **Event:** An outcome or set of outcomes.
- **Favourable Outcomes:** Outcomes that satisfy the event.
- **Sample Space:** Set of all possible outcomes.
- **Probability:** Measure of likelihood of an event.

Practical Applications

Data handling and probability are applied in real-life decision making through surveys, case studies, and data analysis.

Concept Explanation

Using collected data and statistical measures, one can interpret trends, make predictions, and support decisions in various fields such as education, business, and health.

Worked Illustration

Example: A survey shows that 60% of students prefer online classes. Using this data, a school can decide to continue or modify online teaching methods.

Solved Example

Case Study: A company analyses customer feedback data to improve product quality. They calculate mean satisfaction scores and identify common complaints (mode) to focus improvements.

Practice Set

- **Level 1 – Easy:** Interpret a pictograph showing daily sales of a shop.
- **Level 2 – Moderate:** Use mean and mode to analyse test scores of a class.

- **Level 3 – Challenging:** Given probability data from a survey, predict outcomes for a larger population.

Answer Key

- Level 1: Identify highest and lowest sales days from pictograph.
- Level 2: Calculate mean and mode from given scores and interpret.
- Level 3: Use probability to estimate expected number of occurrences in larger sample.

Quick Reference

- Use statistical measures to summarise data.
- Apply probability to predict outcomes.
- Interpret data visualisations for decision making.

Glossary

- **Survey:** Collection of data from a sample.
- **Case Study:** Detailed analysis of a particular instance.
- **Data Analysis:** Process of inspecting and interpreting data.