Data Handling

Representation Of Data Using Tally Marks

Data is defined as a collection of numbers which give the required information. For example, marks scored by the students in a class, number of members in a family, number of books sold etc.

Data are of two types:

(i) **Primary data:** It is the data collected by the person directly for a specific purpose without referring any source. Primary data is collected through surveys, local sources etc.

(ii) **Secondary data:** It is the data collected through other sources like research organizations, financial institutions etc.

The original form of data is called **raw data**. But when the data is arranged in ascending or descending order, it is referred to as **array**.

Example: The marks obtained by 10 students in a test out of 30 are as follows:

10, 15, 25, 22, 12, 18, 28, 29, 17, 18

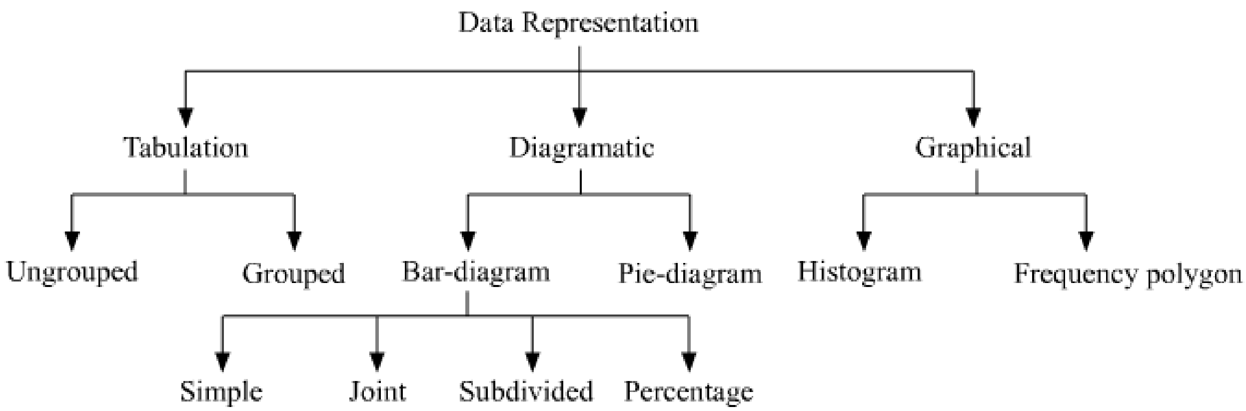
This data is in the form of raw data or ungrouped data.

Now, after arranging the given data in ascending order we get 10, 12, 15, 17, 18, 18, 22, 25, 28, 29 and on arranging them in descending order we get 29, 28, 25, 22, 18, 18, 17, 15, 12, 10.

Such arrangement of the data in ascending or descending order is called an array.

There are different ways of representing the data.

Below given chart explain the same.



Let us now discuss some examples based on representation of a data through tally marks.

Example 1:

A survey was conducted in a village of 80 people. The blood groups of 80 people are as follows.

A, AB, A, B, O, AB, A, AB, O, B, A, O, O, B, A, A, B, AB, O, AB, O, B, A, O, B, O, A, O, A, AB, A, B, AB, O, B, O, O, B, AB, O, A, A, B, AB, AB, O, B, O, O, B,

O, O, B, A, A, O, O, B, O, O, AB, O, B, O, AB, O, A, O, O, B, A, AB, AB, A, B, AB, O, B, O, O.

Represent this data using tally marks.

Solution:

The given data is represented by using tally marks as follows.

Blood Group	**Tally Marks**	**No. of people**
A	卌卌卌||	17
B	卌卌卌|||	18
AB	卌卌卌	15
O	卌卌卌卌卌卌	30
Total		80

Example 2:

Neha threw a dice 150 times and noted the number appearing each time. This data is represented by the following table containing tally marks. Fill in the blanks in this table.

Observation	Tally Marks	No. of Observations
1	_____	28
2	卌 卌 卌	_____
3	_____	42
4	卌 卌 卌 卌 ||	_____
5		13
6	_____	_____
Total		150

Solution:

We easily fill the first four blanks by counting the number of tallies.

Observation	Tally Marks	No. of Observations
1	卌 卌 卌 卌 卌 |||	28
2	卌 卌 卌	15
3	卌 卌 卌 卌 卌 卌 卌 卌 ||	42
4	卌 卌 卌 卌 ||	22
5	卌 卌 |||	13
6	_____	_____
Total		150

Now, let us find the last two blanks.

Total number of observations = 150

Number of observations excluding the 6^{th} observation = 28 + 15 + 42 + 22 + 13 = 120

Hence, number of observations for the 6^{th} observation 6 = 150 – 120 = 30

Observation	Tally Marks	No. of Observations
1	卌 卌 卌 卌 卌 |||	28

2		15
3		42
4		22
5		13
6		30
Total		150

Mean of Data Sets

Application of Mean in Real Life

The runs scored by the two opening batsmen of a team in ten successive matches of a cricket series are listed in the table.

Player A	24	50	34	24	20	96	105	50	13	27
Player B	26	22	30	10	42	98	40	54	10	122

Using this data, we can compare the performances of the players for each individual game. For example, player B performed better than player A in the first match, player A then performed better than player B in the second match, etc.

This method, however, is not useful in trying to determine the overall performances of the two players and comparing them. For this we need to calculate the average or mean score of each player. The player having the better average or mean score has the better overall performance.

In this lesson, we will learn how to find the mean of a data set.

Did You Know?

1. Arithmetic mean (AM), mean or average are all the same.

2. Mean is used in calculating average temperature, average mark, average score, average age, etc. It is also used by the government to find the average individual expense and income.

3. Mean cannot be determined graphically.

4. Mean is supposed to be the best measure of central tendency of a given data.

5. Mean can be determined for almost every kind of data.

Properties of Mean

1. Sum of the deviations taken from the arithmetic mean is zero.

If the mean of *n* observations $x_1, x_2, x_3 \ldots . x_n$ is $\bar{x}$ then
$(x_1 - \bar{x}) + (x_2 - \bar{x}) + (x_3 - \bar{x}) + \ldots + (x_n - \bar{x}) = 0$.

2. If each observation is increased by *p* then the mean of the new observations is also increased by *p*.

If the mean of *n* observations $x_1, x_2, x_3 \ldots . x_n$ is $\bar{x}$ then the mean of
$(x_1 + p),\ (x_2 + p),\ (x_3 + p),\ \ldots,\ (x_n + p)$ is ($\bar{x}$ + *p*).

3. If each observation is decreased by *p* then the mean of the new observations is also decreased by *p*.

If the mean of *n* observations $x_1, x_2, x_3 \ldots . x_n$ is $\bar{x}$ then the mean of
$(x_1 - p),\ (x_2 - p),\ (x_3 - p),\ \ldots,\ (x_n - p)$ is ($\bar{x}$ - *p*).

4. If each observation is multipled by *p*(where p≠0p≠0) then the mean of the new observations is also multiplied by *p*.

If the mean of *n* observations $x_1, x_2, x_3 \ldots . x_n$ is $\bar{x}$ then
the mean of $px_1,\ px_2,\ px_3,\ \ldots,\ px_n$ is $p\bar{x}$.

5. If each observation is divided by *p*(where p ≠ 0) then the mean of the new observations is also divided by *p*.

If the mean of *n* observations $x_1, x_2, x_3 \ldots . x_n$ is $\bar{x}$ then
the mean of $\frac{x_1}{p}, \frac{x_2}{p}, \frac{x_3}{p}, \ldots, \frac{x_n}{p}$ is $\frac{x}{p}$.

Solved Examples

Easy

Example 1:

The amounts of money spent by Sajan during a particular week are listed in the table.

Day	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
Money spent (in rupees)	270	255	195	230	285	225	115

Find the average amount of money spent by him per day.

Solution:

$$\begin{aligned}\text{Average amount of money spent by Sajan per day} &= \frac{\text{Total money spent}}{\text{Total number of days}}\\ &= \text{Rs}\,\frac{270+255+195+230+285+225+115}{7}\\ &= \text{Rs}\,\frac{1575}{7}\\ &= \text{Rs}\,225\end{aligned}$$

Example 2:

The average weight of the students in a class is 42 kg. If the total weight of the students is 1554 kg, then find the total number of students in the class.

Solution:

Let the total number of students in the class be *x*.

$$\text{Average weight of the students} = \frac{\text{Total weight of the students}}{\text{Total number of students}}$$

$$\Rightarrow \text{Total number of students} = \frac{\text{Total weight of the students}}{\text{Average weight of the students}}$$

$$\begin{aligned}\Rightarrow \therefore x &= \frac{1554}{42}\\ &= 37\end{aligned}$$

Thus, there are 37 students in the class.

Medium

Example 1:

For what value of *x* is the mean of the data 28, 32, 41, *x*, *x*, 5, 40 equal to 31?

Solution:

$$\text{Mean of the given data set} = \frac{\text{Sum of all observations}}{\text{Number of observations}}$$
$$\Rightarrow 31 = \frac{28+32+41+x+x+5+40}{7}$$
$$\Rightarrow 217 = 2x+146$$
$$\Rightarrow 2x = 71$$
$$\Rightarrow \therefore x = 35.5$$

Thus, for x = 35.5, the mean of the data 28, 32, 41, x, x, 5, 40 is 31.

Example 2:

The numbers of children in five families are 0, 2, 1, 3 and 4. Find the average number of children. If two families having 6 and 5 children are included in this data set, then what is the new mean or average?

Solution:

$$\text{Mean of the given data set} = \frac{\text{Sum of all observations}}{\text{Number of observations}}$$
$$\therefore \text{Mean of the initial data set} = \frac{0+2+1+3+4}{5} = \frac{10}{5} = 2$$

Thus, the average number of children for the five families in the initial data set is 2.

Two families are added to the initial set of families.

$$\therefore \text{Mean of the new data set} = \frac{0+2+1+3+4+6+5}{7} = \frac{21}{7} = 3$$

Thus, the average number of children for the seven families in the new data set is 3.

Example 3:

The mean of fifteen numbers is 7. If 3 is added to every number, then what will be the new mean?

Solution:

Let $x_1, x_2, x_3, \ldots, x_{15}$ be the fifteen numbers having the mean as 7 and x¯ be the mean.

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \ldots + x_n}{n}$$
$$\Rightarrow 7 = \frac{x_1 + x_2 + x_3 + \ldots + x_{15}}{15}$$
$$\Rightarrow x_1 + x_2 + x_3 + \ldots + x_{15} = 15 \times 7$$
$$\Rightarrow x_1 + x_2 + x_3 + \ldots + x_{15} = 105 \qquad \ldots(1)$$

The new numbers are $x_1 + 3,\ x_2 + 3,\ x_3 + 3,\ \ldots,\ x_{15} + 3$.

Let $\overline{X}$ be the mean of the new numbers.

$$\overline{X} = \frac{(x_1 + 3) + (x_2 + 3) + \ldots + (x_{15} + 3)}{15}$$
$$\Rightarrow \overline{X} = \frac{(x_1 + x_2 + \ldots + x_{15}) + 3 \times 15}{15}$$
$$\Rightarrow \overline{X} = \frac{105 + 45}{15} \qquad \text{(By equation 1)}$$
$$\Rightarrow \overline{X} = \frac{150}{15}$$
$$\Rightarrow \therefore \overline{X} = 10$$

Thus, the mean of the new numbers is 10.

Hard

Example 1:

The average salary of five workers in a company is Rs 2500. When a new worker joins the company, the average salary is increased by Rs 100. What is the salary of the new worker?

Solution:

Let the salary of the new worker be Rs *x*.

Before the joining of the new worker, we have:

$$\text{Mean salary of the five workers} = \frac{\text{Sum of the salaries of the five workers}}{5}$$

$$\Rightarrow 2500 = \frac{\text{Sum of the salaries of the five workers}}{5}$$

$$\Rightarrow \therefore \text{Sum of the salaries of the five workers} = 2500 \times 5 = 12500 \qquad \text{...(1)}$$

After the joining of the new worker, we have:

Number of workers = 5 + 1 = 6

Average salary = Rs (2500 + 100) = Rs 2600

$$\text{Mean salary of the six workers} = \frac{\text{Sum of the salaries of the six workers}}{6}$$

$$\Rightarrow 2600 = \frac{\text{Sum of the salaries of the five workers} + \text{Salary of the new worker}}{6}$$

$$\Rightarrow 2600 = \frac{12500 + x}{6} \qquad (\text{By equation 1})$$

$$\Rightarrow 15600 = 12500 + x$$

$$\Rightarrow \therefore x = 15600 - 12500 = 3100$$

Thus, the salary of the new worker is Rs 3100.

Example 2:

Find two numbers that lie between $\frac{2}{5}$ and $\frac{1}{2}$.

Solution:

The given numbers are $\frac{2}{5}$ and $\frac{1}{2}$.

Mean of the two numbers = $\dfrac{\frac{2}{5} + \frac{1}{2}}{2} = \dfrac{\frac{4+5}{10}}{2} = \dfrac{9}{10 \times 2} = \dfrac{9}{20}$

Now, we know that the mean of any two numbers lies between the numbers.

So, two numbers that lie between $\frac{2}{5}$ and $\frac{1}{2}$ are $\frac{9}{20}$ and $\frac{19}{40}$.

Example 3:

If $\bar{x}$ is the mean of the *n* observations $x_1, x_2, x_3, \ldots, x_n$, then prove that

$$\frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)}{n} = 0$$

Solution:

It is given that $\bar{x}$ is the mean of the *n* observations $x_1, x_2, x_3, \ldots, x_n$.

Thus,

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \ldots + x_n}{n}$$

$$\Rightarrow x_1 + x_2 + x_3 + \ldots + x_n = n\bar{x} \qquad \ldots(1)$$

Now,

$$\begin{aligned}
\frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)}{n} &= \frac{\left(x_1 - \bar{x}\right) + \left(x_2 - \bar{x}\right) + \left(x_3 - \bar{x}\right) + \ldots + \left(x_n - \bar{x}\right)}{n} \\
&= \frac{\left(x_1 + x_2 + x_3 + \ldots + x_n\right) - \left(\bar{x} + \bar{x} + \bar{x} + \ldots + \bar{x}\right)}{n} \\
&= \frac{n\bar{x} - n\bar{x}}{n} \qquad \text{(By equation 1)} \\
&= \frac{0}{n} \\
&= 0
\end{aligned}$$

Thus, the given result is proved.

Median Of A Group Of Data

Example 1:

Find the median of the following data set.

54, 65, 20, 78, 101, 55, 16, 27, 89, 75, 92, 99, 45, 66, 77

Solution:

The given group of data can be arranged in ascending order as

16, 20, 27, 45, 54, 55, 65, 66, 75, 77, 78, 89, 92, 99, 101

This group of data contains 15 terms. The middle term of this data set is the 8^{th} term, which is 66.

Thus, the median of the given group of data is 66.

Example 2:

Find the median of the following group of data and compare it with the mean.

23, 12, 16, 27, 5, 10, 15, 7, 11, 17, 22

Solution:

The given group of data can be arranged in ascending order as

5, 7, 10, 11, 12, 15, 16, 17, 22, 23, 27

This group of data contains 11 terms. The middle term of this data set is the 6^{th} term, which is 15.

Thus, the median of the given group of data is 15.

Mean $= \frac{\text{Sum of all observations}}{\text{Number of observations}} = \frac{23+12+16+27+5+10+15+7+11+17+22}{11}$

$= \frac{165}{11} = 15$

Thus, the values of the mean and the median are same for the given group of data.

Mode Of A Data Set

Bhangra's is a very popular shop that sells watches of foreign brands in Delhi's posh Connaught Place market. The owner of the shop decided to stock the brand whose watches were selling the most. He decided to look at the previous month's sales, which is listed as

Brand	Number of watches sold
Alpha	17
Townzen	23
Tag Heuim	7
Twatch	13

Based on this information, the owner decided to stop keeping Tag Heuim watches because very few people buy them. He also decided to keep more varieties of Townzen watches because most people were buying this brand.

In this data set, the highest occurring event (23) corresponds to Townzen watches and is known as the **mode** of this data set. Just like mean and range, mode is another measure of central tendency of a group of data. It can be defined as

The value of a set of data that occurs most often is called the mode of the data.

Here, the data set did not contain too many terms and could thus be easily arranged in ascending order. However, in case of very large data, it is not always easy to arrange it in ascending or descending order. Therefore, in such cases, it is better to arrange the data set in the form of a table with tally marks.

A collection of data can have more than one mode. The data sets having one mode or two modes or more than 2 modes are said to have **uni-mode** or **bi-mode** or **multi-mode**.

Let us now look at some more examples to understand this concept better.

Example 1:

Find the mode of the following numbers.

2, 6, 7, 5, 4, 2, 6, 7, 9, 7, 8, 3, 2, 11

Solution:

The given set of numbers can be arranged in ascending order as

2, 2, 2, 3, 4, 5, 6, 6, 7, 7, 7, 8, 9, 11

Here, 7 and 2 occur most often (3 times). Therefore, both 7 and 2 are the modes of the given set of numbers.

Example 2:

Find the mode of the following data set.

1000, 200, 700, 500, 600, 160, 270, 300, 360, 950

Solution:

The increasing order of the given numbers is

160, 200, 270, 300, 360, 500, 600, 700, 950, 1000

Here, every number is occurring only once.

Thus, the given data has no mode.

Note: The above example shows that the mode of a data may or may not be unique. Also, there are some data sets which do not have any mode.

Example 3:

Determine whether the data 35, 30, 32, 35, 40, 30, 25, 30, 22, 30 has uni-mode, bi-mode or multi-mode.

Solution:
From the data, we observe that 30 repeats maximum times (4 times). Thus, the mode of this data is 30. Since there is only one mode of this data, so the given data has uni-mode.

Example 4:
Find the type of mode of the data 25, 23, 23, 25, 27, 26, 23, 24, 23, 25, 28, 25.

Solution:
From the data, we observe that 23 and 25 repeats maximum times (4 times). Thus, the mode of this data is 23 and 25. Since there are two modes of this data, so this data has bi-mode.

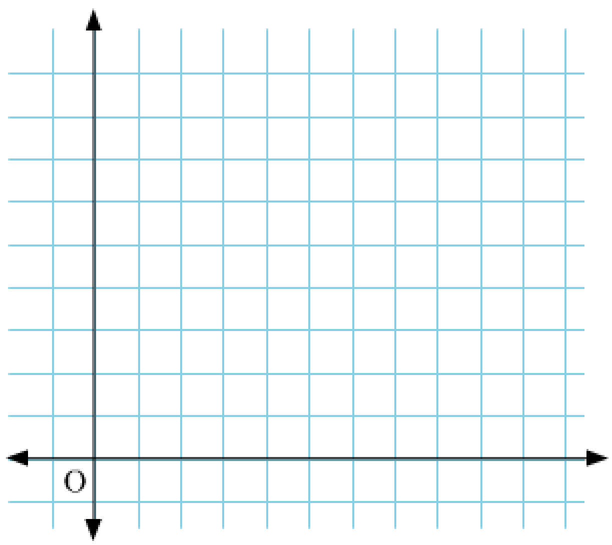
Example 5:
Which type of mode is represented by the data as shown below?

Data	Frequency
0	5
8	6
16	9
24	9
32	9
40	8

Solution:
From the frequency table, we observe that 16, 24 and 32 repeats maximum times (9 times). Thus, the mode of this data is 16, 24, and 32. Since there are three modes of this data, so this data has multi-mode.

Construction of Bar Graphs

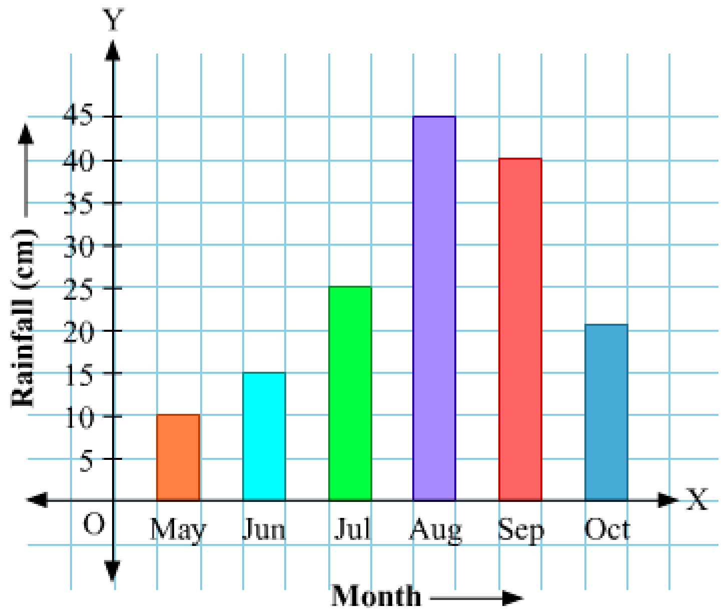
Look at the following figure.



This is a graph paper. A **graph paper** has various vertical and horizontal lines intersecting each other. Also, distance between every two lines is always equal. Thus, we get many small squares on the graph paper.

Graph paper is used to draw bar diagrams. A bar diagram drawn on such a graph paper is known as **bar graph**.

Look at the following graph given by meteorological department.



The above graph shows the rainfall in a place from the month of May to October of a year.

Here, the rainfall is shown by vertical bars of uniform width and with equal spaces between them. This type of representation of a data is known as **bar graph.**

Example:

The following information represents the amount of money earned by a trader in different months.

Month	Amount of money earned (Rs)
January	**9000**
February	**5000**
March	**7000**
April	**11000**

May	**10500**
June	**7000**

Represent this information with the help of a bar graph.

Solution:

We will represent the months on the horizontal line and the amount of money earned on the vertical line.

Here, we take the scale as 1 unit = Rs 1500

The bar graph of the given information is as follows: