DATA HANDLING

3

CONTENTS

DATA

The collection of facts which are expressed numerically with the specific purpose is called **data.** or appropriate information and facts called data.

For example,

1.	Some people died in accident of two cars	Three people died in accident of two cars.
2.	Sachin made highest runs in today's match	Sachin made 190 runs in today's match

TYPES OF DATA

(i) Primary data

(ii) Secondary data

(iii) Raw (ungrouped) data

(iv) Grouped data

- **Primary data :** The data which is collected by the observer is called **Primary data.**
- **Secondary data :** The secondary data is not collected by observer. It is obtained from published or unpublished sources.

  Note: The Primary data is more reliable than the secondary data as the information collected by observer is more accurate in comparison to information collected by another person (other than observer).
- **Raw data :** The data which is collected for specific purpose and put as it is (without any arrangement) is called raw data. Each entry (number) in raw data is known as observation.
- **Grouped data :** If representation of data are with some group, is called grouped data.

  For example : 5, 7, 9, 11, 11, 15, 16, 19, 21, 22, 23, 24, 26, 29, 30.

  In grouped form : 0 – 10, 10 – 20, 20 – 30, etc. Each group is called class.

SOME TERMS

- **Array :** To understand the data in better way, we arrange the data either in ascending order or descending order which is called an array.
- **Range :** The difference between the highest and lowest values of the given data is called range. For example, 0, 3, 8, 10, 10, 12, 18, 19, 19, 20

  So, Range = 20 – 0 = 20

◈ **Frequency :** The number of times a particular observation occurs is called its frequency. It is denoted by (f_i).

◈ **Tally Marks :** The way of representation of frequency numbers is called tally marks.

One	|	Six	𝍸 |
Two	||	Seven	𝍸 ||
Three	|||	Eight	𝍸 |||
four	||||	Nine	𝍸 ||||
five	𝍸	Ten	𝍸 𝍸

➢ FREQUENCY DISTRIBUTION TABLE

If we rearrange above data in the tabular form, showing the frequency of each observation, the tabular form of representation of the data is called frequency distribution and the table is called frequency table.

❖ EXAMPLES ❖

Ex. 1 The number of children in 20 families of a locality are given below :

2, 2, 3, 2, 1, 3, 3, 2, 2, 1, 2, 2, 3, 1, 2, 1, 1, 3, 2, 2

Represent this information in a frequency distribution table.

Sol. Arranging the data in ascending order, we have

$\underbrace{1,1,1,1,1}_{5\text{ times}},\underbrace{2,2,2,2,2,2,2,2,2,2}_{10\text{ times}},\underbrace{3,3,3,3,3}_{5\text{ times}}$

Number of Children	Tally Marks	Number of Families
1	𝍸	5
2	𝍸 𝍸	5 + 5 = 10
3	𝍸	5
	Total	20

Ex. 2 A die is a cube where six faces are marked with numbers (or dots) from 1 to 6 one number on each face. The score obtained in 25 throws are 5, 4, 3, 2, 1, 1, 2, 5, 4, 6, 6, 6, 3, 2, 1, 4, 3, 2, 1, 5, 5, 6, 2, 1 and 3. Prepare a frequency table for the above scores.

Sol. Arranging the data in ascending order as follows :

$\underbrace{1,1,1,1,1}_{5\text{ times}},\ \underbrace{2,2,2,2,2}_{5\text{ times}},\ \underbrace{3,3,3,3}_{4\text{ times}},\ \underbrace{4,4,4}_{3\text{ times}}$

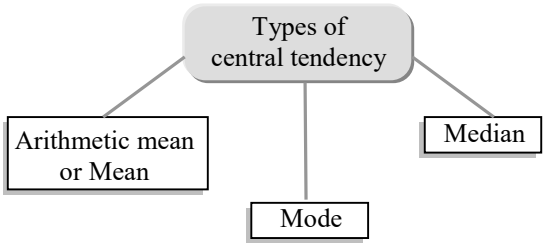
$\underbrace{5,5,5,5}_{4\text{ times}},\ \underbrace{6,6,6,6}_{4\text{ times}}$

The frequency distribution table is as under :

Scores	Tally Marks	Frequency
1	𝍸	5
2	𝍸	5
3	||||	4
4	|||	3
5	||||	4
6	||||	4
	Total	25

➢ CENTRAL TENDENCY AND ITS TYPES

Average is a number that represents or shows the central tendency of the given data.



◈ **Arithmetic mean :**

Mean

$$= \frac{\text{Sum of the given observations } (\Sigma x)}{\text{Number of the given observations } (n)}$$

❖ EXAMPLES ❖

Ex.3 The maximum temperature recorded in a city during the last six months of a year are given below. Find the mean of the maximum temperature.

Month	July	Aug.	Sep.	Oct.	Nov.	Dec.
Maximum Temp. (in ºC)	42	40	40	37	29	22

Sol. Total temp. = 42 + 40 + 40 + 37 + 29 + 22 = 210
Number of months = 6

$$\therefore \quad \text{Arithmetic mean} = \frac{\text{Sum of temperature}}{\text{Number of months}}$$

$$= \frac{210}{6} = 35$$

Thus, average or mean temperature = 35ºC.

Ex.4 The average or mean weight of 50 students of a class is 32 kg. Find the total weight of the students.

Sol. $\because$ Mean $= \frac{\text{Sum of all the items}}{\text{Number of items}}$

or Mean × Number of items = Sum of all the items
For the given data, we have
Number of item (students) = 50
Average or Mean weight = 32 kg
From the formula we have
Sum of all weights = (32 × 50) kg
= 1600 kg

Ex.5 The runs scored by two batsman in their recent five matches are given as below :

Batsman-I	81	75	71	69	74
Batsman-II	12	81	5	105	117

Who performed better ?

Sol. Arithmetic mean of the scores of Batsman-I

$$= \frac{81+75+71+69+74}{5} \qquad = \frac{370}{5} = 74$$

Arithmetic mean of the scores of Batsman-II

$$= \frac{12+81+5+105+117}{5} \qquad = \frac{320}{5} = 64$$

The mean of the scores of Batsman-I is more than the mean of the scores of Batsman-II. So, the performance of Batsman-I is better.

Note : The average or mean is helpful for making comparisons.

➢ ARITHMETIC MEAN OF TABULATED DATA

Let $x_1, x_2, x_3, ..., x_n$ are n observations and $f_1, f_2, f_3, ..., f_n$ are their frequencies respectively.

Observation (x_i)	Frequency (f_i)	$x_i \times f_i$
x_1	f_1	x_1f_1
x_2	f_2	x_2f_2
⋮	⋮	⋮
x_n	f_n	x_nf_n
Total	$\Sigma f_i = N$	Σx_if_i

$$\text{Mean} = \frac{(f_1x_1 + f_2x_2 + f_3x_3 + ... + f_nx_n)}{(f_1 + f_2 + f_3 + ... + f_n)}$$

(Σ (called sigma) shows addition.)

$$\text{or} \quad \bar{x} = \frac{\sum_{i=1}^{n}(f_ix_i)}{(N)}$$

$$\text{Here } N = \sum_{i=1}^{n} f_i$$

❖ EXAMPLE ❖

Ex.6 The ages (in years) of 50 players are given below :

Age (in years)	24	25	26	27	28
Number of players	14	15	3	10	8

Find mean age.

Sol.

Age in years (x_i)	No. of players (f_i)	$x_i \times f_i$
24	14	336
25	15	375
26	3	78
27	10	270
28	8	224
Total	$\Sigma f_i = 50$	$\Sigma x_iy_i = 1283$

$$\text{Mean} = \frac{\Sigma(x_if_i)}{\Sigma f_i} = \frac{1283}{50}$$

Mean = 25.66

➢ MEDIAN

The another method for finding central tendency called median.

◈ **Definition :** The median is that value of the given data, which divides it into exactly two parts.

◈ **Method :** Arrange the data in ascending or descending order. Let the total number of observations be *n*.

Case I : When n is odd

$\Rightarrow$ Median = $\left(\frac{n+1}{2}\right)$ th term.

❖ EXAMPLES ❖

Ex.7 Find median of the data 34, 46, 56, 27, 28, 35, 45.

Sol. First arranging in ascending order, we get

27, 28, 34, $\boxed{35,}$ 45, 46, 56

Here, n = 7(which is odd)

$\therefore$ Median = $\left(\frac{7+1}{2}\right)$ th term

= $\left(\frac{8}{2}\right)$ th term = 4^{th} term

= 35 (4^{th} term)

Median = 35

Case II : When n is even

$$\Rightarrow \text{Median} = \frac{\left(\frac{n}{2}\right)\text{th term} + \left(\frac{n}{2}+1\right)\text{th term}}{2}$$

Ex.8 The weights of 8 students (in kg) are :

30, 42, 24, 37, 21, 25, 38, 31. Find median.

Sol. Arranging the weights in ascending order, we get 21, 24, 25, **30, 31**, 37, 38, 42

n = 8 (even)

$$\therefore \text{Median} = \frac{\left(\frac{n}{2}\right)\text{th term} + \left(\frac{n}{2}+1\right)\text{th term}}{2}$$

$$= \frac{\left(\frac{8}{2}\right)\text{th term} + \left(\frac{8}{2}+1\right)\text{th term}}{2}$$

$$= \frac{4\text{th term} + 5\text{th term}}{2}$$

$$= \frac{30+31}{2} = \frac{61}{2}$$

Median weight = 30.5 kg.

Ex.9 Find the median of the following data :

97, 93, 67, 100, 95, 98, 94

Sol. On arranging the data in ascending order, we get 67, 93, 94, 95, 97, 98, 100

Number of observations = 7 (odd)

Middle number is 4^{th}

$\therefore$ Median = 4^{th} number = 95

Working rule to find median.

Steps :

(i) Arrange the data in ascending or descending order.

(ii) Count the number of items.

(iii) If it is odd, observe one middle value and write it as median and if it is even, observe the two middle values and find median by the following formula :

$$\text{Median} = \frac{\text{Sum of the middle values}}{2}$$

Ex.10 Given the mean of 5 numbers is 25. If the four numbers are 16, 26, 20, 32, find the fifth number.

Sol. Let the missing number be x.

Then sum of the numbers

= 16 + 26 +20 +32 + x = 94 + x

$$\text{NowMean} = \frac{\text{Sum of the numbers}}{5}$$

$\therefore$ 5 × Mean = Sum of the numbers

i.e., 5 × 25 = 94 + x

or 125 = 94 + x $\Rightarrow$ x = 125 – 94 = 31

Thus, the fifth number is 31.

Ex.11 A data is given as below in which one number was missed by the observer :

18, 16, 37, 40, 28, 20, ..., 35

If the median of the data is 26, find the missing number.

Sol. On arranging the given numbers in ascending order, we get

16, 18, 20, 28, 35, 37, 40

There are 8 numbers including the missing number.

But median is 26 which is less than 28.

Also for 8 numbers, the middle numbers will be 4^{th} and 5^{th}. Thus, the missing number must be at 4^{th} place.

Now, $\text{Median} = \frac{4\text{th} + 5\text{th}}{2}$

or $26 = \frac{\text{Missing number} + 28}{2}$

or $26 \times 2 = \text{Missing number} + 28$

or $52 = \text{Missing number} + 28$

or $\text{Missing number} = 52 - 28 = 24$

Ex.12 Find the median of first five multiples of 6.

Sol. The first five multiples of 6 are :

6, 12, 18, 24, 30

which are in ascending order.

n = 5 (odd)

$\text{Median value} = \left(\frac{n+1}{2}\right)\text{th value}$

$= \left(\frac{5+1}{2}\right)\text{th value} = 3^{rd}\text{ value}$

Median value = 18

➢ MODE

This is also used for central tendency.

◈ **Definition :** The observation which has maximum frequency is called the mode.

❖ EXAMPLES ❖

Ex.13 Find mean, median and mode of the following data. Choose the best average(s) to represent the data. 3, 11, 7, 5, 17, 11

Sol. **Mean :** Sum of the numbers

$= 3 + 11 + 7 + 5 + 17 + 11 = 54$

$\text{Mean} = \frac{\text{Sum of the numbers}}{\text{Number of items}} = \frac{54}{6} = 9$

Mode : 11 as it occurs two times.

$\therefore$ Mode = 11

Median :

Step 1 : On arranging the data in ascending order, we get

3, 5, 7, 11, 11, 17

Step 2 : Two middle values are 7 and 11.

$\therefore$ $\text{Median} = \frac{7+11}{2} = 9$

We find that mean and median are the same and are in the centre of the data. The mode is not an appropriate central value for this data.

Ex.14 Find the mode of numbers

2, 2, 2, 2, 3, 3, 3, 4, 4, 5, 5, 5, 6, 6, 8, 8

Sol.

Number	Tally Marks	Frequency
2	||||	4
3	|||	3
4	||	2
5	|||	3
6	||	2
8	||	2

Number 2 has highest frequency (4), so mode is 2.

Ex.15 Find the mode of the following data :
2, 5, 4, 7, 5, 5, 8, 12, 8, 9

Sol.

Number	Tally Marks	Frequency
2	|	1
4	|	1
5	|||	3
7	|	1
8	||	2
9	|	1
12	|	1

The number 5 has the highest frequency 3. Hence mode of given numbers is 5.

Ex.16 The following table shows the marks obtained by 40 students of class VII :

Marks obtained	30	25	23	19	17	14
No. of Students	3	2	4	11	13	7

Find the mode.

Sol. As the marks obtained by maximum students is 17 so the mode of the given data is 17.

Ex.17 The scores in mathematics test (out of 30) of 15 students are as follows :
20, 19, 12, 25, 24, 23, 20, 25, 20, 16, 9, 20, 15, 10, 5
Find mean, median and mode. Are they equal?

Sol. As we know that

$$\text{Mean} = \frac{\text{Sum of all observations}}{\text{Total number of observations}}$$

$$= \frac{20+19+12+25+24+23+20+25+20+16+9+20+15+10+5}{15}$$

$$= \frac{263}{15} \qquad \text{Mean} = 17.54$$

Now arranging the given marks in ascending order, we get
5, 9, 10, 12, 15, 16, 19, **20**, 20, 20, 20, 23, 24, 25, 25
n = 15 (odd)

$$\text{Median} = \left(\frac{n+1}{2}\right)\text{th term}$$

$$= \left(\frac{15+1}{2}\right)\text{th term} \quad = 8^{th}\text{ term}$$

Median marks = 20

Mode :

Marks	Tally Marks	Frequency
5	|	1
9	|	1
10	|	1
12	|	1
15	|	1
16	|	1
19	|	1
20	||||	4
23	|	1
24	|	1
25	||	2

Number 20 has the highest frequency i.e., 4, so, mode marks = 20.

➢ BAR GRAPH

A bar graph is a pictorial representation of the numerical data by a number of bars (rectangles) of uniform width erected horizontally or vertically with equal spacing between them.

Steps :

(i) On a graph paper draw horizontal line and vertical line OX and OY respectively, which are called X-axis and Y-axis respectively.

(ii) Marks points at equal space (intervals) along the X-axis. Below these points write the names of the data whose values are to be plotted.

(iii) Assume (choose) a suitable scale. On that scale find the heights of the bars for the given numerical values.

(iv) Mark these heights parallel to Y-axis from the point taken in step (ii)

(v) Draw bars of equal width for heights marked in step (iv) on X-axis. The bars should be centred on the points marked on X-axis. These bars show the given numerical data.

❖ EXAMPLES ❖

Ex.18 Raju recorded the temperature at 2 p.m. for a week as follows :

Days	Mon	Tues	Wed	Thu	Fri	Sat	Sun
Temperature (°C)	30	25	35	30	20	38	28

Draw a bar graph for the above data.

Sol. We construct the bar graph using the following steps :

(i) Take a graph paper and draw two perpendicular lines OX and OY.

(ii) Along OX mark "Names of the days" and along OY "Temperature in °C".

(iii) Along OX, choose suitable width for each bar.

(iv) Along OY, choose an appropriate scale and mark the temperature in Celsius.

Here, 1 large division = 10°C

(v) Calculate the heights of various bars as follows:

(a) Height of bar for Monday

$= \frac{1}{10} \times 30 = 3$ large divisions.

(b) Height of bar for Tuesday

$= \frac{1}{10} \times 25 = 2.5$ large divisions.

(c) Height of bar for Wednesday

$= \frac{1}{10} \times 35 = 3.5$ large divisions.

(d) Height of bar for Thursday

$= \frac{1}{10} \times 30 = 3$ large divisions.

(e) Height of bar for Friday

$= \frac{1}{10} \times 20$

= 2 large divisions

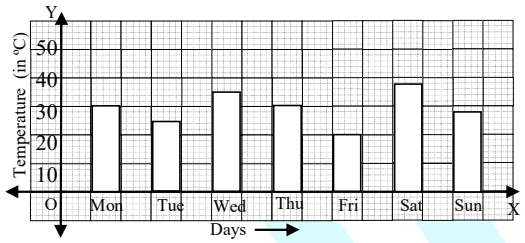
(f) Height of bar for Saturday

$= \frac{1}{10} \times 38$

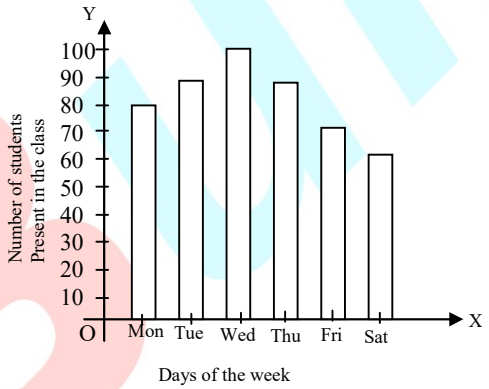
= 3.8 large divisions.

(g) Height of bar for Sunday

$= \frac{1}{10} \times 28$ = 2.8 large divisions.



Ex.19 Answer the following questions for the given bar graph.



(i) What does the bar graph represent ?

(ii) On which day minimum number of students attended the class ?

(iii) On which day was the highest attendance recorded ?

(iv) If all students attended the class on Wednesday, how many students are there in the class ?

Sol. (i) Along vertical line, the students who were present on days of a weak are shown. So the bar graph represents the attendance of a class during the days of a week.

(ii) The bar for Saturday is the shortest, therefore, the minimum attendance was on Saturday.

(iii) The bar for Wednesday is the highest, therefore, the maximum/highest attendance was on Wednesday.

(iv) Take a scale/ruler and put it parallel to the line OX at the highest point of the bar for Wednesday. We find that the scale touches the OY line at 100 marks. So the total number of students in the class are 100.

➢ DOUBLE BAR GRAPH

A double bar graph is a bar graph that represent two sets of data on the same graph. The two bars corresponding to each category are drawn next to each other.

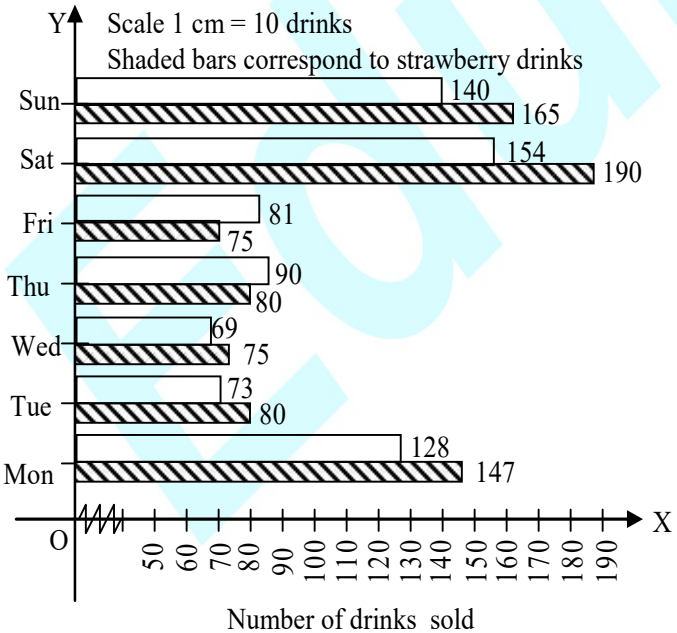
❖ EXAMPLES ❖

Ex.20 The number of two types of drinks sold on days of a week are given below. Construct a horizontal double bar graph for the data.

Days	Strawberry Drinks	Chocolate Drinks
Monday	147	128
Tuesday	80	73
Wednesday	75	69
Thursday	80	90
Friday	75	81
Saturday	190	154
Sunday	165	140

Also find the averages in both types of drinks and tell for which drink the sale is more.

Sol. Here also, a zig-zag is marked along OX which shows that marking is skipped and begins with 50.



Now, Average (mean) sale of strawberry drinks

$$= \frac{147+80+75+80+75+190+165}{7}$$

$$= \frac{812}{7} = 116 \text{ strawberry drinks}$$

Average (mean) sale of chocolate drinks

$$= \frac{128+73+69+90+81+154+140}{7}$$

$$= \frac{735}{7} = 105 \text{ chocolate drinks}$$

Therefore, the average sale of strawberry drinks is more than the average sale of chocolate drinks.

Ex.21 The minimum and maximum temperature of different cities on a particular day are given below :

Cities	Temperature	
	Minimum	Maximum
Ambala	26ºC	37ºC
Chennai	27ºC	36ºC
Bangalore	24ºC	28ºC
Delhi	28ºC	38ºC
Jaipur	29ºC	39ºC
Jammu	26ºC	41ºC
Bhopal	25ºC	35ºC

Construct a vertical double bar graph. Also answer the following questions on the basis of the bar graph :

(i) Which city has the greatest difference in the minimum and maximum temperatures on the given day ?

(ii) Which is the hottest city and which is the coldest city ?

(iii) Name the city which has the least difference between the minimum and maximum temperatures.

(iv) Name the two cities where the maximum temperature of one was less than the minimum temperature of the other.

Sol. (i) Difference in temperatures for various cities :

Ambala $(37 - 26)^\circ C = 11^\circ C$

Chennai $(36 - 27)^\circ C = 9^\circ C$

Bangalore $(28 - 24)^\circ C = 4^\circ C$

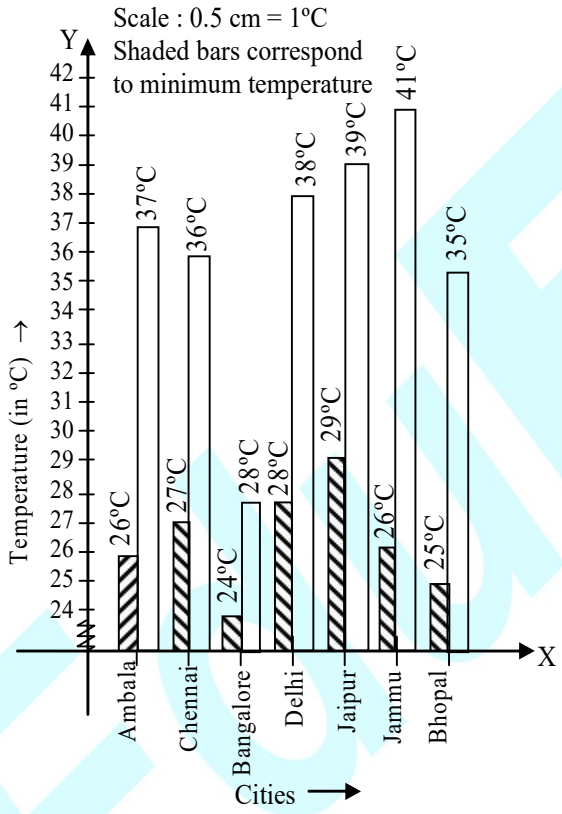
Delhi $(38 - 28)^\circ C = 10^\circ C$

Jaipur $(39 - 29)^\circ C = 10^\circ C$

Jammu $(41 - 26)^\circ C = 15^\circ C$

Bhopal $(35 - 25)^\circ C = 10^\circ C$

$\therefore$ Greatest difference is in Jammu.



(ii) Hottest city is the city which has the highest maximum temperature i.e., Jammu.

Coldest city is the city which has the lowest minimum temperature i.e., Bangalore

(iii) City with least difference in minimum and maximum temperatures is Bangalore.

(iv) The maximum temperature of Bangalore is less than the minimum temperature of Jaipur.

Ex.22 Sales of English and Hindi books in the years 1995, 1996, 1997 and 1998 are given below :

Years	1995	1996	1997	1998
English	350	400	450	620
Hindi	500	525	600	650

Draw a double bar graph and answer the following questions :

(i) In which year was the difference in the sale of the two language books least ?

(ii) Can you say that the demand for English books rose faster ?

Sol. We construct the bar graph using the following steps :

(a) Take a graph paper and draw two perpendicular lines OX and OY.

(b) Along OX, mark the years and along OY, mark the subjects.

(c) We choose a suitable scale to determine the heights of bars. Here, we choose the scale as 1 large division is equal to 100.

(d) Calculate the height of the bar for the different years for both subjects.

For year 1995,

Height of bar for English

$$= \frac{350}{100}$$

$$= 3.5 \text{ large divisions}$$

Height of bar for Hindi

$$= \frac{500}{100} = 5.0 \text{ large divisions.}$$

For year 1996,

Height of bar for English

$$= \frac{400}{100} = 4 \text{ large divisions}$$

Height of bar for Hindi

$$= \frac{525}{100} = 5.25 \text{ large divisions}$$

For year 1997,

Height of bar for English

$$= \frac{450}{100}$$

= 4.5 large divisions

Height of bar for Hindi

$$= \frac{600}{100}$$

= 6.0 large divisions

For year 1998,

Height of bar for English

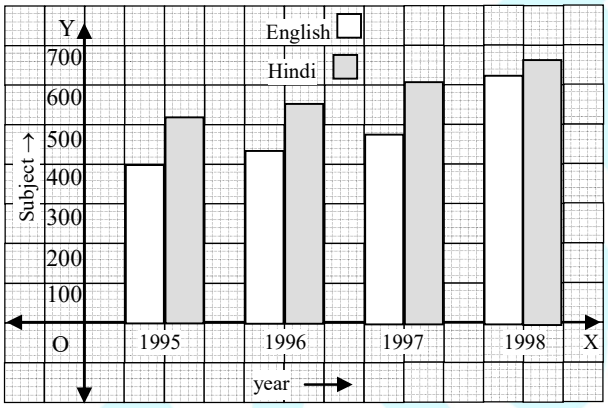
$$= \frac{620}{100}$$

= 6.2 large divisions

Height of bar for Hindi

$$= \frac{650}{100}$$

= 6.5 large divisions



(i) In year 1998, the difference in the sale of the two language books is least.

(ii) Yes, the demand for English books rose faster.

➢ PROBABILITY

Definition : The measure of the chance of happening something is called 'probability'.

Measure of Probability :

Probability of a sure or certain happening = 1

Probability of an impossible situation = 0

Probability of an event that may or may not happen = between 0 and 1.

◈ Basic terms

Experiments : The activities which we perform to achieve something. Eg. Throw a die, Toss a coin etc.

Random Experiment : The activities or experiment whose possible outcomes are known to us but it is not possible to predict the outcome when the experiment is performed. The activities of tossing a coin or throwing a die are random experiments.

Outcome : When a random experiment is performed, it ends into some happening. Such a happening is called an outcome of the experiment.

For example, let us list out the outcomes of following experiments.

Experiment	Outcomes
Tossing a coin	Head or tail
Throwing a die	1 or 2 or 3 or 4 or 5 or 6 dots on upper face.
Two coins tossed together	Two heads or two tails or one head and one tail
Watching traffic on a busy road	There may be an accident or all will go smooth

Equally Likely Outcomes :

In a random experiment when all the outcomes have equal chance to occur or they have equal probability then the outcomes are called equally likely.

For example :

(i) Getting head or tail on tossing a coin are equally likely outcomes.

So each have half-half chance to occur.

or Probability of head = Probability of tail = $\frac{1}{2}$.

(ii) Getting 1, 2, 3, 4, 5 or 6 dots on top of a die are equally likely outcomes. Each outcome has equal chance or probability that is one-sixth or $\frac{1}{6}$.