Lesson 17

DATA HANDLING

17.1 We studied in our previous class that data are those numerical facts which are collected for definite purposes.

We studied types of data, collection of data, arranging the data and tabulation of data with the help of tally marks along with the reading and making of pictograms and bar diagrams. Collection of data, graphing and representation help us organize our experiences and draw inferences from them.

In this chapter, we will study reading and making of double bar graphs and central tendencies, arithmetic mean, median and mode of non-classified data.

In daily life, we face different data which we observe in newspapers, magazines, television or other sources. Let us look at some common forms of data that we come across.

Daily routine of Amar

Work	Time Spent
Food and routine	2 hours
School	6 hours
Games and entertainment	3 hours
Cooperation in house works	2 hours
Studies	3 hours
Sleeping	8 hours

Table 17.1

Number of patients on Monday in a Primary Health Centre

Name of disease	Number of patients
Fever	22
Cold and Cough	26
Eye diseases	08
Skin diseases	12
Accidental injury	07
Dental diseases	05

Table 17.2

Performance of Arti in first two tests

Subject	First Test	Second test
Hindi	5	8
English	6	8
Mathematics	3	9
Science	6	9
Social Science	5	8
Sanskrit	8	7

Table 17.3

What do these collections of data tell us? For example we can say that Amar spends 6 hours in school and 3 hours at home on studies from his daily routine. (Table 17.1)

Similarly Arti performed better in second test as compared to first test in all the subjects and improved the best in Mathematics.

Can we organize and present these data in a different way, so that their analysis and interpretation becomes better? We shall address such questions in this chapter.

We have seen that how can we arrange the collected informations in the form of frequency distribution table and then represent it in the form of pictographs or bar graphs. We can say that tallest bar is mode if bar represents frequency.

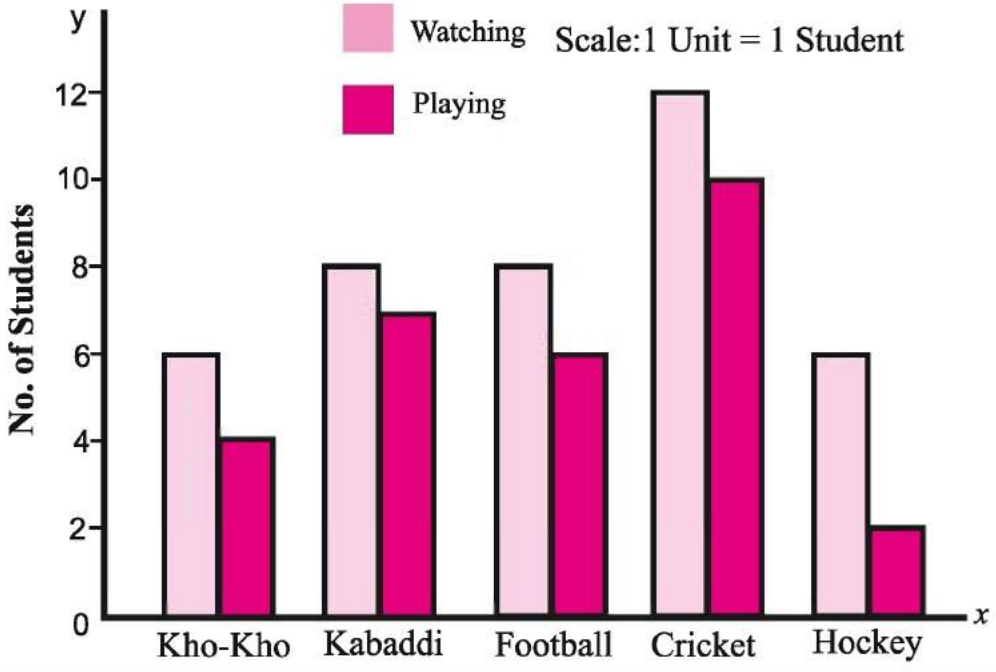
17.2 Draw double bar graphs

Consider the following data which were obtained from a survey executed in a class.

Favourite game	Kho-Kho	Kabaddi	Football	Cricket	Hockey
Watching	6	8	8	12	6
Playing	4	7	6	10	2

Above data show the interest of playing and watching a game by different students of a class. By seeing these data we can tell that which game is played by the maximum students and which game is viewed by the least students.

But in order to find difference in the students who like watching and playing a particular game, we have to compare the number of students watching and playing. To do this, we'll learn to draw the graphs which are said to be double bar graphs. The comparison of both the interests is given by bar graphs side by side.



Example1 The number of C.F.L. tubes and L.E.D. bulbs sold by a seller every year from the year 2011 to2015 are given below.

Year	2011	2012	2013	2014	2015
CFL Tube	1200	1400	1100	900	600
LED Bulb	100	400	700	1000	1400

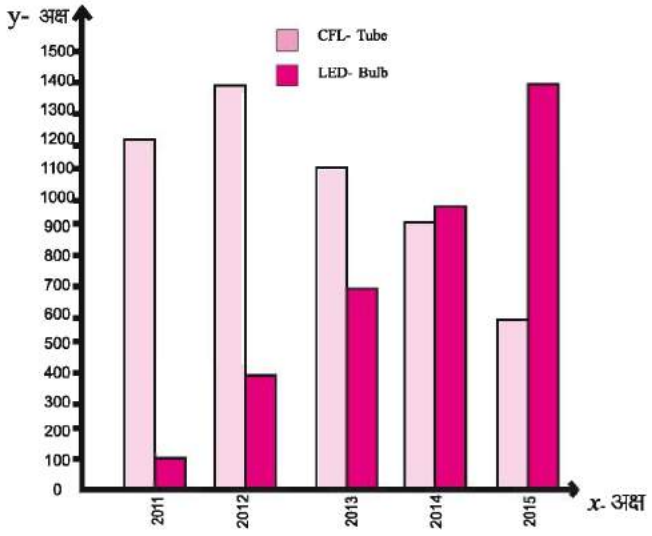
Draw a double bar graph and answer the following questions.

1. The sale of which type of light equipment's increased continuously?
2. Increased or decreased in light equipment's in 2015 as compared to 2011?
3. In which year the difference in the sale of light equipment's of both the types stood highest?

Solution

Steps of construction of double bar graph

1. Make x-axis (horizontal) and y-axis (vertical) on the graph paper. Both of these meet at the origin (0,0).
2. Write down the year 2011 to 2015 on the x-axis.
3. Mention the number of C.F.L. tubes and L.E.D. bulbs on y- axis.
4. Take proper scale on y-axis so that number of both the light equipment's can be written easily. We can take 1 cm. = 100 on y- axis.
5. Find the length of each column by dividing by 100 to the number.
6. Exhibit C.F.L. tube and L.E.D. bulb altogether by columns.



Scale : 1cm = 100

(1) It is obvious after seeing double bar graph that the sale of LED bulb increased continuously.

(2) There is an increase in the total light equipment's in the year 2015 as compared to 2011.

(3) It is obvious after seeing double bar graph that the difference in the sale of both the light equipment's in the year 2011 stood highest.

Do and learn

1. The marks of the subjects Mathematics and Science of five students of class 7 are given in the table. Exhibit these by the vertical double bar graph.

Name of Student	Mathematics	Science
Arti	65	75
Varsha	70	75
Simran	55	70
Radha	75	80
Jyoti	50	60

2. The details of different expenditures occurring in one month of two families are shown in the table. Make a double bar graph on the basis of this table and answer the following questions.

Expenditure Head	Family 1	Family 2
House rent	2000	2500
Electricity, Water, Telephone	800	600
Food Items	8000	7000
Children Education	2000	3000
Savings	2200	1900

(i) Under which head the expenditure is maximum?
(ii) Under which head the expenditure is minimum?
(iii) If monthly income of both the families is Rs.15000 then what percentage of expenditure is used on children education in both the families?

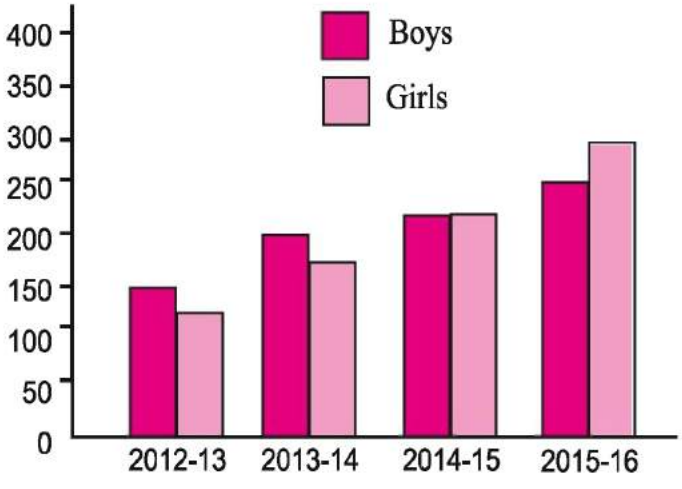
Exercise 17.1

1. In the following graph, number of students of a school are shown below according to session. Answer the questions on the basis of this graph.

(i) In which session, number of girls were more than number of boys in the school?

(ii) In which session, number of both girls and boys in the school were equal?

(iii) What was the total number of students in the school in the session 2015-16?



2. Under the free textbook distribution scheme from the year 2011 to 2015, the distribution of books of the subjects Mathematics and Hindi of class 7 in a district were as follows:

Subject/Year	2011	2012	2013	2014	2015
Mathematics	8000	8500	9500	11000	13000
Hindi	9000	10000	10500	11500	14000

Draw a double bar graph and answer the following questions-

(i) The demand of book of which subject is maximum always?

(ii) In which year the difference in the demand of both the books is minimum?

(iii) In which year the difference in the demand of both the books is maximum?

3. Estimated distance of following cities of Rajasthan from Udaipur by road and by train are given in the following table. Draw a double bar graph on the basis of table and answer the following questions:-

Distance from Udaipur	By Road (in km.)	By train (in km.)
Ajmer	290	310
Jaipur	410	440
Bikaner	530	580
Jodhpur	270	300
Kota	360	570

(i) Which city is at a maximum distance from Udaipur by road?

(ii) Which city has least difference in the distance by road and by train?

(iii) Which city has highest difference in the distance by road and by train?

17.3 Collection of Data

In our daily life, expression of facts through data plays very important role For example a better way to express, "India's population is enough" is to say, "the population of India according to census 2011 is 1 Arab, 21 crore 8 lakh." Similarly it would be appropriate to say that number of students in our school is 867 rather than number of students in our school is enough. Hence we can say that we can express our thoughts very clearly via data. As we need to collect stone, lime, cement, bricks etc. before constructing a building, similarly it is very much necessary to collect data to draw conclusion and analyse it. We can be able to find the solution logically to understand the complex situations by the proper use of these data. But it is very much necessary that the data taken should be pure, general and authenticate.

We can classify data into two parts on the basis of its sources of collection of data.

(a) Primary Data (b) Secondary Data

(a) Primary data - The data which is collected first time by self or by help are known as primary data. For example, if you have to study the family position of students of your class then you'll have to collect information regarding monthly income-expenditure, number of brothers-sisters, source of income etc. These data shall be classified as primary data.

(b) Secondary Data - The data which has already been collected by any person or institute, which can be in either published or unpublished form are called secondary data. For example one can obtain the data related with census or literacy from the Census Department of India authorized by Government of India.

17.4 Organisation of Data

When we collect data, we have to record and organize it. Why do we need to do that?

Consider the following example:-

The heights of 8 students under health test in a school were as follows:-

Vijay – 140 cm Kishor – 138 cm Vidhya – 130 cm
Tabassum – 135 cm Ramesh – 145 cm Sarika – 125 cm
Divyanshi – 131 cm Mohit – 144 cm

It was not an easy task to draw conclusions from these data in this form. Sarika wrote these heights in ascending order in the tabular form.

Name of student	Height (in m.)	Name of student	Height (in cm)
Sarika	125	Kishor	138
Vidhya	130	Vijay	140
Divyanshi	131	Mohit	144
Tabassum	135	Ramesh	145

Answer following questions:-

1. What is the name of the tallest student?
2. What is the name of the shortest student?
3. What is the difference in heights of Kishor and Tabassum?

When we put data in a proper tabular form it becomes easy to understand and interpret it. Many kinds of data we come across are put in tabular form. Our school rolls, progress report, temperature record, daily attendance record and many other are all in tabular form.

Can you think of a few more data that you come across in tabular form?

Do and Learn

Determine the weight of students of your class by weighing machine. Tabulate these by arranging data. Write these data in ascending or descending order. Then answer the following questions:-

1. The weight of which student is maximum in the class?
2. How many students have weight more than 25 kg.in the class?
3. How many students have weight in between 20 to 30 kg.?

17.5 Measures of Central tendency

You must have studied or heard following statements in your daily life.

1. The average age of students of class 7 is 13 years.
2. The lunch taken by each student is 150 grams in mid-day meal.
3. The average temperature is 30^0 C of last 10 days.
4. Lakshya studies 5 hours every day.

Consider the following statements:-

Can you say that age of each and every student of class 7 is 13 years according to first statement or each and every student takes exactly 150 gm. food according to the second statement?

Clearly the answers of these questions are "no".

Then what do these statements mean?

We understand from "average" that age of maximum students of class 7 is near about13 years. The age of some students may be below 13 years or some may be above 13 years.

Similarly the average temperature of last few days 1s 32° which means that the temperature remained around 32°. The temperature may be below 32° or may be above 32°.

Similarly we can say that "average" is a number which shows or represent central tendency of a group of data or observations. As average is measure of central tendency of a group of highest and lowest data. There is a necessity of central values or different representatives to interpret different types of data.

Out of these one value of representation is algebraic or arithmetic mean.

17.6 Arithmetic Mean

For a group of data the commonly used value of representation is arithmetic mean. In brief, it is known as mean. Consider the following example:

Example 2 The net income of per day in a week of a fruit seller is Rs. 500, Rs. 650, Rs. 400, Rs.425, Rs. 450, Rs.600 and Rs.475 respectively. Find the average income of fruit seller?

Solution

$$\text{Average income of fruit seller} = \frac{\text{Total income of the week}}{\text{Number of days in the week}}$$

$$= \frac{500 + 650 + 400 + 425 + 450 + 600 + 475}{7}$$

$$= \frac{3500}{7} = \text{Rs. } 500$$

Average income of fruit seller would be Rs. 500 daily.

Example 3 Find arithmetic mean of first six even numbers.

Solution We know that first six even numbers are 2, 4, 6, 8, 10, 12. To find arithmetic mean, we have to add all the observations and then divide it by total number of observations. Hence in this case

$$\text{Arithmetic mean} = \frac{\text{Sum of all the observations}}{\text{Number of observations}}$$

$$= \frac{2 + 4 + 6 + 8 + 10 + 12}{6}$$

$$= \frac{42}{6} = 7$$

Thus arithmetic mean of first six even numbers is 7.

17.7 Range

Consider the following example:

Example 4 The salaries of five teachers working in a school are Rs. 25000, Rs. 18000, Rs. 20000, Rs. 22000 and Rs. 23000 per month.

1. What is the salary of the teacher who gets highest salary?
2. What is the difference of salaries of teachers who gets highest and lowest salary?
3. Find the mean of salaries of these teachers.

Solution Arranging salaries of teachers in ascending order we have 18000, 20000, 22000, 23000, and 25000.

The information's we draw from this are:

1. The salary of the teacher who gets highest salary is Rs. 25000.
2. Highest and lowest salaries are respectively Rs. 25000 and Rs. 18000. Difference between these two = 25000-18000 = Rs. 7000
3. Mean of salaries of teachers $= \frac{18000 + 20000 + 22000 + 23000 + 25000}{5}$

$$= \frac{108000}{5} = \text{Rs. } 21600$$

It is obvious from above example that we can estimate the expansion of observations in the form of difference in between highest and lowest observations. We call this result as **range** of observations of data.

Do and Learn

1. Find the mean heights of your family members.
2. Find the mean age of your family members.

Exercise 17.2

1. Following are the number of students of a school from class 6 to 12. 78,72,67,59,54,49,48 Then find:
   (i) In which class the number of students is maximum?
   (ii) In which class the number of students is minimum?
   (iii) What is the range of these data?
   (iv) Find the mean of these data.
2. Find the mean of first 10 whole numbers.
3. A cricketer scored the runs in 6 innings as follows:
   68, 03,17,78,12,104 Find the arithmetic mean of runs scored.
4. The number of passengers travelled in the bus running from Bikaner to Udaipur from Monday to Friday are as follows:
   45, 48,32,40,30 What is the mean of passengers in each day?
5. Following crops were raised up to five years in a village. The profit (in rupees) per acre on the crop were as follows:-

Crop	2011	2012	2013	2014	2015
Barley	6000	8000	5000	6500	8500
Gwar	7000	8000	12000	9000	8500
Groundnut	9000	7000	10000	8000	13000

   Answer the following questions on the basis of above table.
   (i) Find the mean profit of each crop in five years.
   (ii) Which crop should be raised next year on the basis of above answer?
6. If arithmetic mean of digits 3, 4, 8, 5, x, 3 is 4 then find the value of x.
7. The number of books given to students in 10 days from a library are as follows:-
   40, 57, 32, 59, 72, 66, 40, 62, 72, 60
   Find the mean of books given to students every day?
8. The average of five numbers is 18. If four numbers are 22, 20, 14 and 13 then find the fifth number.

9. The temperature of a particular week in a city is noted as below:

Day	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Temperature (in °C)	37	37.5	40	36.5	37.5	35	35.5

(i) Find the range of temperature by using the data.
(ii) Find the mean temperature of this week.
(iii) How many days did the temperature remain more than average?

10. In a singing competition in a school, three judges gave marks out of 100 to four singer contestants as follows:

Name of contestant	Judge I	Judge II	Judge III
Rashi	78	75	72
Suman	82	75	83
Poonam	68	64	69
Khushboo	49	56	51

1. What is the range of marks given by the judges?
2. Find the mean of total marks.
3. Point out the name of winner.
4. What is the difference in the means of winner and the contestant placed at the fourth place.

17.8 Mode

Second type of representative value is mode. Consider the example.

Example 5 The shoes of different sizes are available on the shoe shop. The shopkeeper recorded the sale of shoes to find weekly demand of shoes given in the table below.

Shoe number	5	6	7	8	9	10
Sale	12	27	40	45	26	18

If we find the mean of shoes sold by the shopkeeper, then

Solution

$$\text{Mean} = \frac{\text{Total Number of shoes sold}}{\text{Total types of number of shoes}} = \frac{168}{6} = 28$$

Then should the shopkeeper keep 28 pairs of shoes per week of each size? Definitely the shopkeeper will keep a greater stock of the shoes of number 7,8 as compared to other sizes on the basis of above record. Because the sale of shoes of size 7, 8 is comparatively more.

Of all, the sale of shoes of size 8 is the highest. This is another value of representation of data. This value of representation is said to be mode.

> The term which repeats maximum number of times in a given data is called as mode. It means the term which has maximum frequency is called as mode.

Example 6 Find the mode of the following numbers.
5, 4, 4, 2, 5, 7, 5, 6, 5, 4, 3, 5

Solution On arranging the numbers in ascending order
2, 3, 4, 4, 4, 5, 5, 5, 5, 5, 6, 7
It is obvious from the observations that number 5 repeats maximum number of times.
Hence mode will be 5.

17.8.1 Mode of large and unclassified data

If number of data is large then it is not easy to count after writing them either in ascending or descending order. In this situation, we tabulate the data with the help of tally marks. We had learnt how to tabulate the data in previous classes.

Example 7 30 athletes participated in a 100 meter race. The time taken (in seconds) by them in completing the race is as follows:
14, 12, 13, 12, 10, 12, 14, 13, 12, 11, 12, 13, 14, 12, 14, 12, 13, 14, 14, 11, 10, 11, 12, 14, 13, 12, 12, 11, 12, 14 Find the mode of these data.

Solution On tabulating the data

Time (in seconds)	Tally Marks	Number of athletes
10	||	2
11	||||	4
12	𝍸 𝍸 |	11
13	𝍸 |	6
14	𝍸 ||	7
	Total	30

By seeing this table we can immediately say that the mode of these data is 12 because maximum number of athletes completed their race in 12 seconds.

Think and Discuss

- Is it possible to have 2 modes in a group of numbers?
- Can mode be found out just by observation?

Do and Learn

1. 40 students of class 7 wrote their number of family members together. This number is shown below:
   4, 3, 5, 4, 7, 3, 5, 6, 4, 4, 4, 7, 6, 4, 5, 4, 3, 4, 5, 6, 7, 4, 4, 5, 3, 4, 6, 4, 5, 5, 4, 3, 4, 7, 6, 4, 3, 5, 4, 5 Find the mode of these data.
2. Find mode of following data.
   21, 22, 25, 24, 22, 23, 23, 24, 25, 24, 22, 24, 23, 24, 23, 24, 22, 21, 25, 23

We saw that mean provides us average of all the observations of the data and mode shows the observations which occurs maximum number of times.
Think over following examples.

1. You have to calculate the daily consumption of electricity in your house.
2. The seller of ready made garments has to fulfil supplies of his stall.
3. We have to find the height of the door for our home.
4. Have to select a sweet in the form of a favourite sweet for students of a class. Then which sweet could be selected?

When we think about first statement then in order to find out consumption of electricity daily, we will find out the weekly electricity consumption from electric meter, then by taking its mean, we can find out daily electric consumption.

Can we use the same method for second statement?

From the example of shoes we can see that for supply of clothes, mean is not a suitable representative value. Mode is suitable for it.

Similarly for third statement both median and mode are not suitable values. Here the height of door has to be decided according to the tallest member of the family. Similarly examine the fourth statement by thinking and find suitable representative value for it.

17.9 Median

We have seen that in some cases arithmetic mean is a proper measure of central tendency and in some cases mode is a proper measure of central tendency.

Think over another example.

The salaries of 9 employees in a factory are as follows:

3300, 4200, 5000, 3500, 4300, 3500, 4400, 3500, 5500

If we want to distribute the employees into two groups according to the payment, then what would be better representative in this case either arithmetic mean or mode? By arranging these data into ascending order

3300, 3500, 3500, 3500, 4200, 4300, 4400, 5000, 5500

In the above data we see that 4200 is such a number which has groups of 4-4 numbers on both of its side. It means that the payment of 4 employees is less than 4200 and payment of 4 employees is more than 4200. In this way on arranging numbers in ascending or descending order the number lying exactly in the middle is known as median.

> If data are arranged in either ascending or descending order then the term lying in the middle is known as median.

Example 8 Find median of following data.
0,47,35,20,30,40,50

Solution On arranging data in ascending order we get
0,20,30,35,40,47,50

In the above data there are 7 terms. To find middle term of the data it is divided by 2 after adding 1 to the number of terms. (When number of terms are odd).

This means that median term of above data is the fourth term which is 35. Hence the median of above data is 35.

Similarly if number of terms are even then after arranging in ascending order the mean of two middle terms is median.

$$\text{Median Term} = \frac{7+1}{2} = \frac{8}{2} = 4 \text{ fourth term.}$$

Do and Learn

1. The data arranged in ascending order are as follows:
   8, 11, 12, 16, 16 + x, 20, 25, 30
   If median is 18 then find x.
2. Jyoti scored the following marks (out of 10) in different subjects
   5,7,0,3,5,8
   Jyoti calculated mean, median and mode from the remaining numbers ignoring 0. Did she do it correctly?

Exercise 17.3

1. Find the mode from the following data:
   7, 6 ,4, 5, 6, 4, 6, 3, 2, 7, 8, 6, 4, 6, 5
2. Vandana took a dice. She tossed the dice 20 times and noted the obtained number each time:
   3, 4, 6, 3, 5, 2, 2, 3, 5, 4
   5, 6, 6, 1, 5, 6, 3, 5, 2, 4
   Find the median and mode with the help of above data.
3. The weights (in kgs.) of 30 labourors working in a factory is as follows:
   60, 65, 70, 65, 60, 70, 65, 70, 75, 80, 75, 60, 65, 70, 65, 65
   70, 65, 60, 70, 65, 75, 80, 75, 80, 65, 60, 65, 70, 80
   Find the median and mode with the help of above data.

4. Find the median of following variables:
   37, 31, 42, 43, 46, 25, 39, 45, 32
5. The heights of 21 persons of a class are as follows:
   147, 149, 150, 152, 148, 151, 148, 150, 151, 149
   152, 151, 152, 151, 150, 148, 149, 152, 153, 151, 152
   (i) Find the median and mode of the above data.
   (ii) Are there more than one mode in the above data?
6. The runs scored by the players in a cricket match are as follows:
   105, 47, 0, 36, 50, 16, 7, 70, 65, 36, 52
   Find the mean, median and mode from the above data. Are they all equal?

We Learnt

1. The data collected can be shown in the form of bar graphs with the help of frequency distribution table.
2. With the help of a double bar graph we can compare the two groups of observations in single inspection.
3. The collection, recording and presentation of data help us organize our experiences and draw inferences from them.
4. Before collecting data we need to know what we will use it for.
5. The data that is collected needs to be organized in a proper table, so that it becomes easy to understand and interpret.
6. Arithmetic mean is one of the representative values of data.
7. The mean can be obtained by dividing the sum of group of data by the number of data, which lies in between the lowest and highest value.
8. The term that occurs most often in a group of data is known as mode. A set of data can have more than one mode.
9. If data are arranged either in ascending or descending order, then the term which lies exactly in the middle is known as median.