

# STATISTICS

## UNIT - 7

### Let us know the history of statistics.....

Statistics has been used in India since ancient days. As early as between 321 BC to 296 BC, we find in Arthshastra written by Kautilya the use of data in a variety of ways. This book describes in detail the agricultural data, figures for rural and urban population as well as economic data and the processes how they were collected during Maurya dynasty. This process of collecting data continued in the times of Mughal Emperor Akbar as well. Abul Fazal in his book 'Ain-e-Akbari', written around 1596 – 1597, describes this process of data collection and its use.

In the British period, East India Company needed to keep a record of its accounts as well as detailed information about the areas under its control. In 1807, the company got a survey done in its states. This survey included one crore 50 lakh people spread over 60,000 square miles area. This report included information on many significant aspects. The more important of these include the geographical description of each district, the religion as well as rites and rituals of the citizens, the natural wealth of the country, fisheries, agriculture situation and industrial situation. A government officer, A. Shakespeare, presented in 1848, the first census report. This was related to the area and the revenue of all districts of the Northwest province. The first effort to collect the detailed census data of India was made in the years 1867 to 1872. The first nationwide census took place in 1881. Since then nationwide census is being done every 10 years.

After independence the need for an appropriate statistical structure was felt for the economic and social development of the country. Prof P.C. Mahalanobis was the first statistical adviser to the Indian cabinet in 1949. His contribution to development of statistics in India is unforgettable. Prof Mahalanobis was the founder director of Indian Statistical Institute set up in Calcutta in the year 1932. This was declared as an institute of national importance in 1959. Besides this in 1949, Central Statistical Organisation was set up. From the 20<sup>th</sup> century to now, efforts to develop statistical methods, concepts and uses are continuing.

It appears that the word Statistics comes from the Latin word 'Status', which means political state or administration.

# Data Handling and Analysis

16

Knowingly or unknowingly we keep using data all the time. We organise the information from our prior experiences, analyse it and draw conclusions. For example in the month of July if the sky is cloudy and the wind blows from the east, we say that it would rain today. Similarly people who travel know that some trains generally come on time but there are some that are often late. While buying pulses, wheat grain, rice etc., we examine a small part of the material and decide if it is worth buying or not. In the cricket match we consider the rate at which runs are being scored and at what rate they are further required etc. In the newspaper every day we look at the maximum and minimum temperatures, average humidity, time for sunrise and sunset etc. Being able to see these data with comprehension and drawing conclusions from it helps us analyse better and to make better judgements.

Individuals, families, panchayats, State and Indian Government and all other institutions and organisations use data for taking decisions and planning. The better our ways of collecting and organising data and the sharper our analysis is, the better would be our decisions and their implementation.

## Data Collection and Representation

Suppose you have 30 students in a class and you are asked to collect the following data how will you go about it?

1. Information about the blood group of each student of the class.
2. The number of students of the class who come walking and those who use other means

Students of a class started collecting this data. They decided to do this in two groups. Each group went to every student and asked about their blood group and the means of transport to the school. Group 1 made the following table:-

**TABLE - 1**

A		B		AB		O	
Rh <sup>+</sup>	Rh <sup>-</sup>	Rh <sup>+</sup>	Rh <sup>-</sup>	Rh <sup>+</sup>	Rh <sup>-</sup>	Rh <sup>+</sup>	Rh <sup>-</sup>

Group 2 made the following table:-

**TABLE - 2**

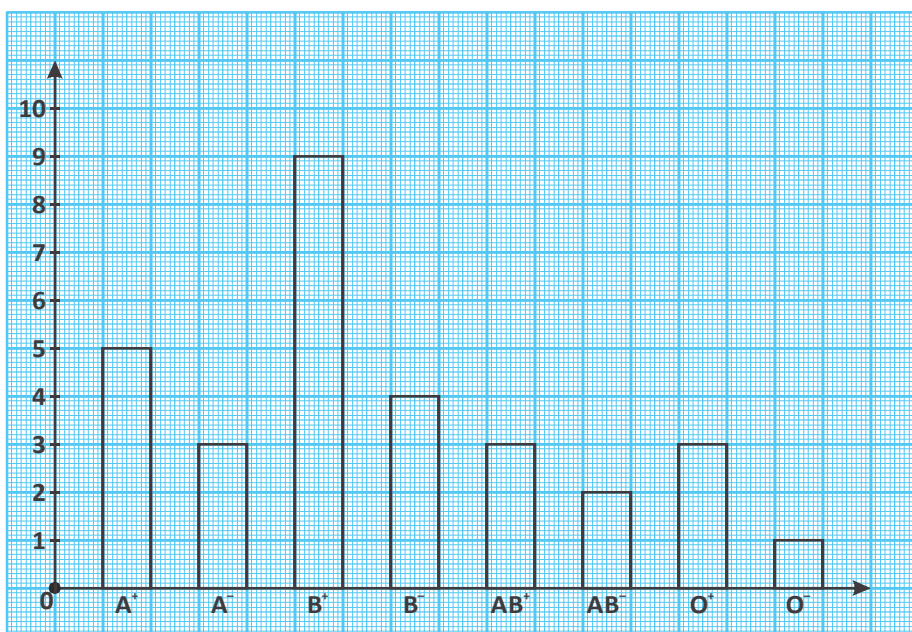
On Foot	Bicycle	Scooter	Bus	Others

## Frequency Table

In order to understand the collected data better Group 1 re-organised the data and made table-3:-

**TABLE - 3**

Blood Group	Tally Mark to Count	In Number
A <sup>+</sup>		5
A <sup>-</sup>		3
B <sup>+</sup>		9
B <sup>-</sup>		4
AB <sup>+</sup>		3
AB <sup>-</sup>		2
O <sup>+</sup>		3
O <sup>-</sup>		1
<b>Total</b>		<b>30</b>



In this table along with the tally marks the frequency count is also written as a number, for example 9 written opposite B<sup>+</sup> shows that there are nine people who have the blood group B<sup>+</sup>. In the same manner the other numbers show the frequency of other blood groups. Such a table is called frequency table. The group then made a bar diagram based on the table.

### Try This

- Write any five conclusions that can be drawn from the bar diagram.
- Similarly also represent the data of group -2 as a bar diagram.
- Make the attendance table for the students of your class in the month of January and answer the following:-
  - On which day was the attendance maximum?
  - When was the attendance minimum?
  - Write some more conclusions.
- From hockey, cricket, kabbadi, football and volleyball, which is the sport your class fellows like the most? Collect data for this and make a frequency table to find the answer to the following-
  - Which is the most popular sport?
  - Which sport is liked by less children?



## Place in Ascending and Descending Orders

Our data may have values which are repeated or may not. If the data is not large then we can draw conclusions by simply placing it in ascending-descending order. For example-

In a class the marks obtained in mathematics exam of 15 students out of 100 are as follows:

45, 35, 56, 22, 99, 71, 80, 63, 42, 36, 18, 77, 54, 82, 41

Writing these in ascending order-

18, 22, 35, 36, 41, 42, 45, 54, 56, 63, 71, 77, 80, 82, 99

Writing in descending order-

99, 82, 80, 77, 71, 63, 56, 54, 45, 42, 41, 36, 35, 22, 18

Now you can answer the following:-

- What is the lowest and the highest marks obtained.
- What is the difference between them?

Discuss with your friend and find out some more conclusions can be drawn from the data.

## Grouped Frequency Table

### 1. Inclusive Class

When the number of data points is large and they have a large range between their maximum and the minimum then frequency table will be very large. In such situations instead of finding

the frequency of one number and find the frequency of small groups. (We will call these groups classes)

**EXAMPLE-1.** In a 50 over match the number of runs made by a team in each over is given below.

7, 8, 2, 5, 7, 12, 6, 20, 18, 9, 11, 5, 19, 10, 3, 6, 12, 8, 16, 0, 12, 7, 8, 11, 15, 13, 4, 7, 1, 22, 2, 17, 1, 6, 21, 4, 9, 15, 0, 5, 1, 9, 26, 10, 14, 3, 16, 2, 6, 8

When we make a frequency table for this data we will need to find the number of overs in which no runs were scored, the number of overs where one run was scored etc. In this way we will have to go up to 26. This is because in one over 26 runs were scored. This would be a huge table.

Can we therefore reason in the following manner:-

How many overs in which runs from 1 to 6 were scored? How many overs in which runs from 7 to 12 were scored?

And in the same manner number of overs in which 13 to 18 or 19 to 24 and then 25 to 30 runs were scored. They call these groups classes. Depending upon our needs the groups can be smaller or bigger. In this example you can choose the groups to be 1 to 4, 5 to 8, 9 to 12 or from 1 to 5, 6 to 10, 11 to 15 etc. You could choose any other group size as well.

How do we find out the frequency of these groups?

From the above suggested grouping choose anyone. Look at the number of runs made in each over. Put a tally mark in the appropriate group. Do this for all the 50 overs. You will get the following frequency table:-

**TABLE - 4**

Number of runs scored	Tally mark	Number of overs (frequency)
0-4		12
5-9		18
10-14		09
15-19		07
20-24		03
25-29		01
	<b>Total</b>	<b>50</b>

While using such tables we use some terms, for example class interval, lower class limit, upper limit, centre point, inclusive class, non-inclusive class etc. Let us try to understand these.

In the above example 0-4, 5-9, 10-14 etc., are all classes.

Look at any two classes of this frequency table. You will find that the lower limit of the next group starts where the upper limit of the first group ends. This means that the upper limit of any group is not the same as the lower limit of the next group. These groups/ classes are called inclusive because the lower and upper limits are both included in that group. The group 0 to 4 includes overs in which 0, 1, 2, 3, 4 runs were scored. There are five such situations and hence the class interval is also 5. In the same manner the group of 5 to 9, includes overs in which 5, 6, 7, 8, 9 runs were scored, the class interval here is also 5. For the first group 0 to 4 the lower limit is 0 and the upper limit 4. Similarly, the other groups the lower limits are 5, 10, 15, ..... and the upper limits are 9, 14, 19... In each group the difference between the lower and upper limit is 4.

$$\text{The midpoint of the class 0 to 4 is } = \frac{0+4}{2} = 2$$

$$\text{and of class 5 to 9 } = \frac{5+9}{2} = \frac{14}{2} = 7$$

We can similarly find the midpoints of other classes which are known as class marks.

Look at the frequency table given below. The heights of a group of people is recorded in inclusive classes:-

**TABLE - 5**

<b>Class Interval</b> (Height in Centimetre)	141-150	151-160	161-170	171-180	<b>Total</b>
<b>Frequency</b>	9	11	15	10	<b>45</b>

Now discuss the following questions with your friends:

1. How many classes are there in this frequency table?
2. Which group has 180 as its upper limit?
3. What are the lower and upper limits of the class 151 to 160?
4. Which class has the highest frequency?
5. What is the frequency of the first class?
6. What is the meaning of the statement that the class 171 to 180 has a frequency 10?
7. Are these classes inclusive? Yes or no, give reason for your choice?

This table helps to show the data in a simple and concise form and we can see the main features of the data at a glance. Such a table is called grouped frequency distribution table.

## 2. Exclusive Class

In table 5 above you saw that the number of people having heights between 141 to 150 *cm* is nine. The number of people with heights between 151 to 160 *cm* is 11. If there is a person who has a height between 150 and 151 *cm* which group would you place that person in?

Similarly if the height of a person is 160.4 or 160.6 *cm* which group would you place that person in?

For this we will have to examine the way we are building the classes. Can we do something such that the upper limit of one group is the lower limit of the next group so that there is no gap in between? For example:-

**EXAMPLE-2.** The weights of class IX students were measured, the data is shown in the following frequency table:-

<b>Weight</b> (In kilograms)	30-33	33-36	36-39	39-42	42-45
<b>Frequency</b> (No. of children)	4	9	12	7	3

**EXAMPLE-3.** The monthly income of all families of a village is given below:-

<b>Income</b> (In rupees)	0-1000	1000-2000	2000-3000	3000-4000
<b>Frequency</b> (No. of families)	12	30	13	5

These are examples of exclusive classes. Such an arrangement sometimes creates a problem. For example if a family has an income of Rs.2000 per month, then in which group it will be placed, in group 2 or in 3? Similarly in example 1, if the weight of a child is exactly 39 kg then which group would she go in?

In such cases it is assumed that whenever the value is equal to the value of the upper limit of a group, it will be placed in the next group. On this basis we can say that the family with income Rs.2,000 would go in class 3 (2000 – 3000). And the 39 kg child would be counted in class 4 (39 - 42).

## Changing Inclusive Classes to Exclusive Classes

When inclusive classes are changed to exclusive classes then lower limits of all class are decreased by half the class interval between classes and the upper limit is increased by the same amount.

**TABLE - 6**

Inclusive Classes		Exclusive Classes	
Class Interval	Frequency	Class Interval	Frequency
6 - 10	8	5.5 - 10.5	8
11 - 15	11	10.5 - 15.5	11
16 - 20	10	15.5 - 20.5	10
21 - 25	15	20.5 - 25.5	15
26 - 30	6	25.5 - 30.5	6

As can be seen in this example the difference between the lower limit of one group and upper limit of the next group is 1 (The second group has a lower limit of 11, which is one more than the upper limit of the lower group.) Thus half of this that is 0.5 is subtracted from all lower limits and 0.5 added to the upper limit values. That makes the lower limit in the first group to be 5.5 and the upper limit to be 10.5. Similarly for the last group the limit are 25.5 to 30.5. The class interval remains 5.

### Try This

Change the inclusive classes of Table 5 to exclusive classes and add two people with heights 150.5 cm and 160.5 cm to the data.



**EXAMPLE-4.** The class marks of a distribution are 104, 114, 124, 134, 144, 154 and 164. Find the class size and the class limits.

**SOLUTION :** The class size is the difference between the adjoining class values.

Thus the size of the class is  $114 - 104 = 10$

We need classes of size 10 whose mid points are respectively 104, 114, 124, 134, 144, 154 and 164.

Therefore the lower limit of the first group is  $= \left( 104 - \frac{10}{2} \right) = 99$

The upper limit of the first group is  $= \left( 104 + \frac{10}{2} \right) = 109$

The other class interval values would be

99-109, 109-119, 119-129, 129-139, 139-149, 149-159, 159-169



### Exercise - 16.1

1. Explain the following:-  
Class interval, size of a class, class mark, class frequency, class limits
2. Give the difference between inclusive and exclusive classes.
3. The weather department has given the following as the data of maximum temperatures in August for Delhi. Make a frequency table of this data.  
32.5, 33.3, 33.8, 31.0, 28.6, 33.9, 33.3, 32.4, 30.4, 32.6, 34.7, 34.9,  
31.6, 35.2, 33.3, 33.3, 36.4, 36.6, 37.0, 34.5, 32.5, 31.4, 34.4, 33.6,  
37.3, 37.5, 36.9, 37.0, 36.3, 36.9, 36.9





4. The following are the distances of the work places of 40 teachers from their homes:-  
 7, 9, 5, 3, 7, 8, 10, 20, 3, 5, 11, 25, 15, 12, 7, 13, 18, 12, 11, 3  
 12, 6, 12, 14, 7, 2, 9, 15, 6, 15, 17, 2, 16, 32, 19, 10, 12, 17, 18, 11  
 Make a frequency distribution table of class size 5 for this data.
5. The value of Pi up to 50 places of decimal is given below-  
 3.14159265358979323846264338327950288419716939937510
- Make a frequency distribution table of the digits from 0 to 9 occurring in this expansion.
  - Which is the number that occurs the least?
  - Which is the number that occurs the most? What can we conclude from this?
6. The per hectare production of rice in 40 fields of a village in quintals is given below, make a frequency distribution for this data.  
 31, 20, 25, 18, 28, 20, 18, 26, 15, 12, 25, 16, 30, 20, 22, 24, 45, 28, 30, 16,  
 30, 40, 20, 30, 20, 30, 28, 47, 40, 35, 28, 45, 20, 35, 32, 18, 20, 26, 23, 16  
 Write the conclusions that you draw from this frequency distribution table.
7. 40 children were asked about the number of hours they watched TV in the previous week, their responses were:  
 1, 5, 6, 2, 7, 4, 10, 12, 5, 8, 10, 12, 36, 22, 6, 15, 3, 1, 2, 4, 21, 16, 17,  
 13, 14, 2, 7, 9, 23, 26, 31, 33, 5, 35, 25, 26, 29, 30, 9, 31
- Make a frequency distribution table of class size 5.
  - What is the lower limit of the first class?
  - Give the limits of the fourth class
  - What is the class mark of the 7<sup>th</sup> class
  - How many children watched television for 20 or more hours in the week?

## The Pictorial Depiction of Data

There is a saying that “a picture is better than thousand words” The comparis of different sets of data can be represented with the help of graph. We will discuss the following diagrams here:

- Histograms
- Frequency polygons
- Cumulative frequency curve or ogive

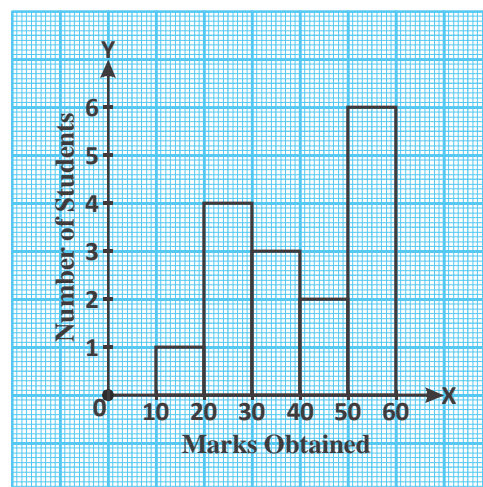
## Histogram

This is a simple and elegant method of displaying frequency distribution. While constructing this class interval (the independent variable) is taken on the X axis and the frequencies (dependent variable) is shown on the Y axis. In this we make rectangles with the class interval as base with height in proportion to the frequency of that class. Thus we see a continuous series of rectangles with equal bases. The areas of these rectangles are proportional to their corresponding frequencies.

**EXAMPLE-5.** The marks obtained by 16 students of a class in an examination are given below:-

Marks Obtained	Frequency
10-20	1
20-30	4
30-40	3
40-50	2
50-60	6

Plot a histogram for this data.



**SOLUTION :** We follow the following steps for plotting the histogram:-

- STEP-1** Take a graph paper, draw two perpendicular axis on it and show them as the X and the Y axis.
- STEP-2** Along the horizontal axis we will show the class values (marks obtained). Here we have used 1 *square* = 10 marks.
- STEP-3** On the perpendicular axis we shall display the frequency (the number of students with those marks). Here 1 *square* = 10 student.
- In this way we get the desired histogram.

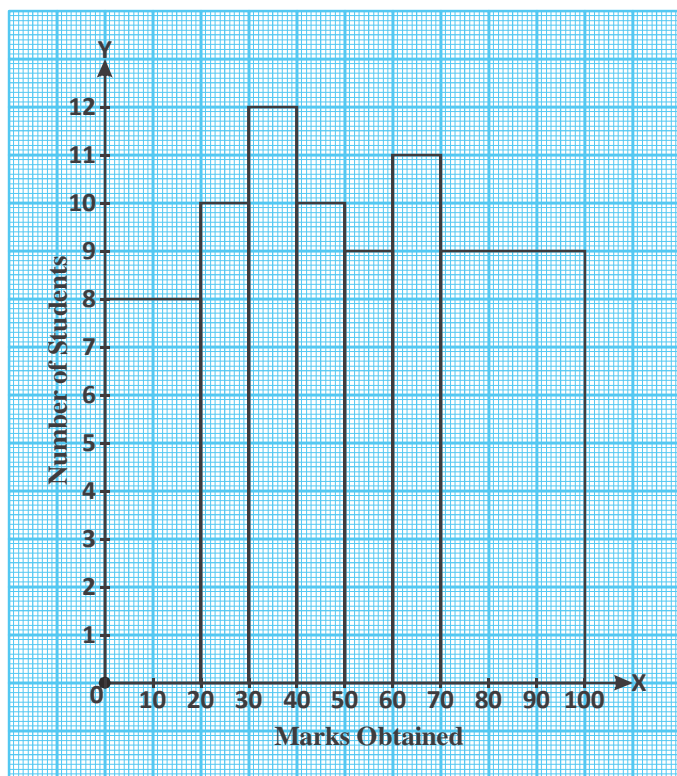
## Histogram for Unequal Class Interval

Now we consider this different situation. The marks obtained out of 100 by students of a class in science are the following:-

**TABLE - 7**

Mark Obtained	0-20	20-30	30-40	40-50	50-60	60-70	70 or more than 70
Number of Students	08	10	12	10	09	11	09



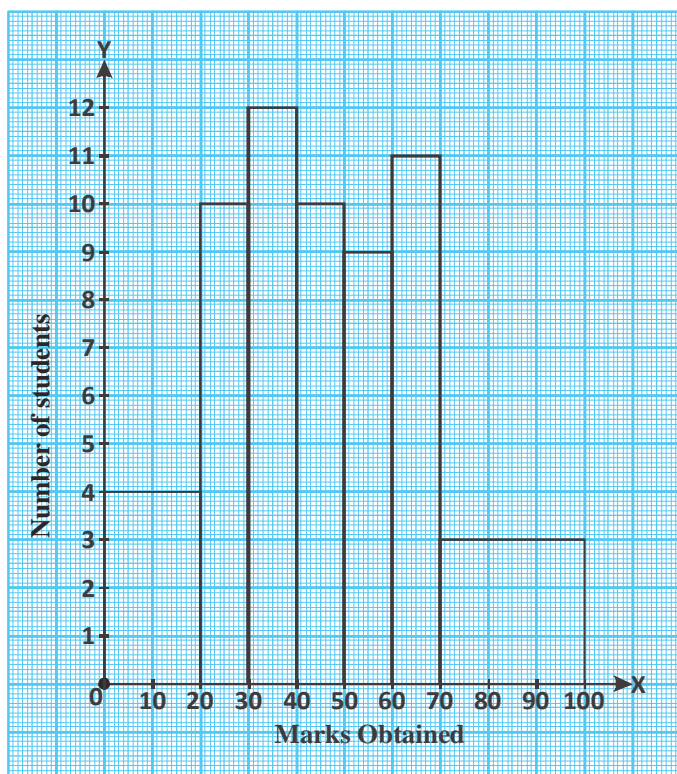


It is clear from the following table that number of students getting less than 20 is 8 and of those getting more than 70 is 9. The data is presented in unequal class intervals. The first interval has a size 20, the last has a size 30 and the rest have class size of 10. A student makes the histogram of the above data as shown in the adjoining diagram. Is this depiction correct?

The class interval is unequal here. To make a bar diagram we would have to change them to equal class intervals, for example in the first class

If the class interval is 20 then the height of the rectangle/bar is 8

∴ The area under the curve therefore is

$$= \frac{8}{20} \times 10 = 4$$


Similarly, for the last class, the interval is 30 and hence the length of the rectangle/bar would be

$= \frac{9}{30} \times 10 = 3$  and since the other class intervals are 10 only we do not have to make any changes in them. We can therefore change the heights of the rectangles in the following way.

Length of the rectangle

$$= \frac{\text{Frequency}}{\text{Width of that class}} \times \text{Minimum class width in the data}$$

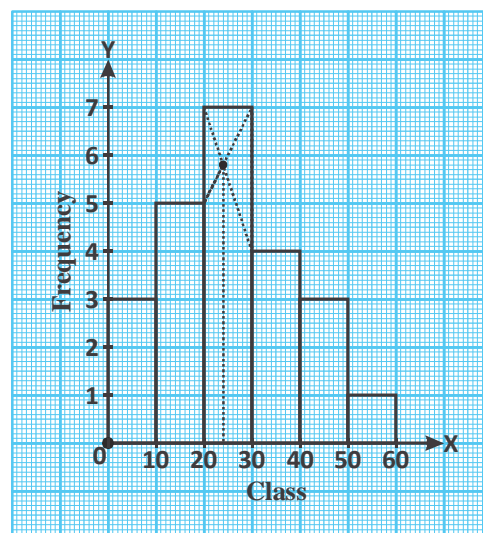
In this way for each class we find the height of the rectangle with a class size of 10. Therefore, the corrected histogram with modified lengths of rectangles would be as given below.

## Graphical Method to Locate Mode

The histogram of data with exclusive classes can be used to find out mode as well, for example:-

**TABLE - 8**

Class	Frequency
0 - 10	3
10 - 20	5
20 - 30	7
30 - 40	4
40 - 50	3
50 - 60	1



**Step-1** Make the histogram for the above data.

**Step-2** The rectangle with the maximum height is taken to be the modal class for the data. Join the upper right corner of the modal class to the right edge of previous rectangle and join the upper left corner of the modal class to the left edge of the next rectangle.

**Step-3** Draw a perpendicular line to the X axis from the intersecting point.

**Step-4** The point at which the perpendicular line meets the X axis is the mode of the data and it is 24, hence the mode of the data is 24.

## Frequency Polygon

Another way to depict a classified frequency distribution is to make a frequency polygon. To construct a frequency polygon we make a bar/rectangle on each class interval and the middle points of the upper sides of the rectangles are joined using straight lines. This diagram has many sides and hence is called frequency polygon. Frequency polygons are made in two ways.

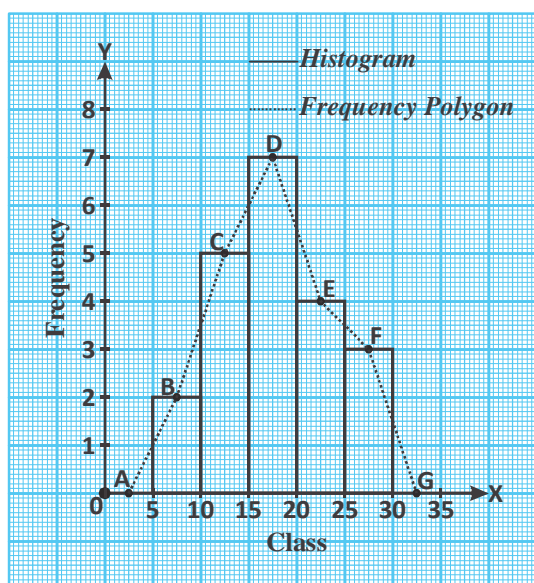
1. Using histograms
2. By direct method



### 1. Constructing Frequency Polygon using a Histogram

**TABLE - 9**

Class	5-10	10-15	15-20	20-25	25-30
Frequency	2	5	7	4	3



There are the following steps in this method:

**Step-1** Make a histogram from the frequency descriptions.

**Step-2** Mark the mid points B, C, D, E, F on the upper edge of each rectangle. Join these mid points sequentially using straight lines.

**Step-3** Now take the class intervals before and after the data distribution. Namely here for example, the intervals 0-5 and 30-35. Mark the mid points of these intervals on the X axis. These would be at 2.5 and 32.5 respectively. Call these A and G respectively. Join B to A and F to G. The required polygon is ABCDEFG.

### Think and Discuss



The areas of the frequency polygon and of the histogram are the same, Why?  
(Hint : Use the property of congruency of triangles)

### 2. Constructing Frequency Polygon by Direct Method

**TABLE - 10**

Class	5-10	10-15	15-20	20-25	25-30
Frequency	2	5	7	4	3

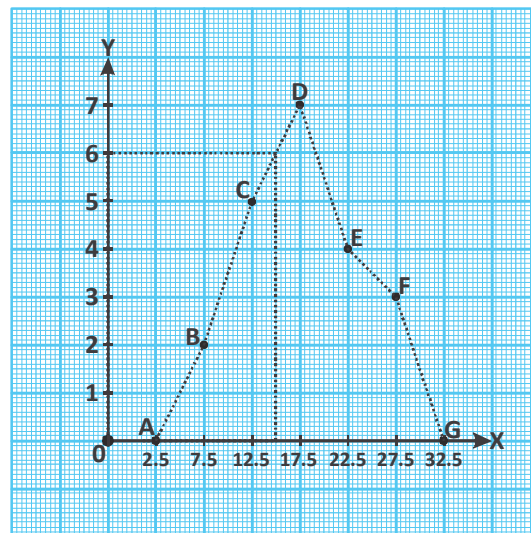




This method has the following steps:-

**Step-1** First find the class marks of each class.

Class	Frequency	Mid-Point
5-10	2	7.5
10-15	5	12.5
15-20	7	17.5
20-25	4	22.5
25-30	3	27.5



**Step-2** We will plot the mid points on the X-axis and the frequencies on the Y-axis. Mark the mid points on the X-axis.

**Step-3** Mark the points B, C, D, E, F using the corresponding frequency.

**Step-4** Mark on the X-axis the midpoint 2.5 (point A) of the class 0-5, that lies before the first class and the midpoint 32.5 (point G) of the class 30-35, which is after the first class. We have marked them on the X-axis because their frequencies are zero.

**Step-5** Join all the marked mid points sequentially.

**Step-6** The figure ABCDEFG obtained is the frequency polygon.

**Note :** Frequency polygon shows the increase or fall in the value of the frequency. Using this we can estimate the value of the frequency for any particular point. For example for 15 on the X-axis the corresponding frequency would be 6.

## Difference between a Histogram and a Frequency Polygon

Histogram	Frequency Polygon
1. Shows the frequency using bars.	1. More useful since it indicates the increase or fall in frequency. Frequency is shown as a polygon in this.
2. Frequencies are assumed to be distributed over the entire interval.	2. Frequencies of the class are assumed to be located at the midpoint.
3. It can be constructed for equal or unequal class widths.	3. This can be made only for equal class widths.

## Ogive or Cumulative Frequency Curve

The graphical depiction of the class and its associated cumulative frequency is called cumulative frequency curve or Ogive. The graph of the cumulative frequency curve is plotted in a manner similar to the way frequency polygon is made.

There are two ways to plot the cumulative frequency curve:

1. "Less than" method
2. "More than" method

1. **"Less than" method:** The cumulative frequency of a class is the sum of the frequency of that class and frequencies of all the classes before it.

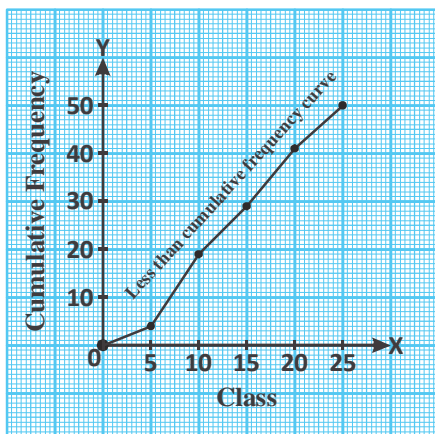
Based on following table let us try to plot the "less than" cumulative curve:-

**TABLE - 11**

Class	0-5	5-10	10-15	15-20	20-25
Frequency	4	15	10	12	09

**SOLUTION :** Cumulative frequency for "less than"

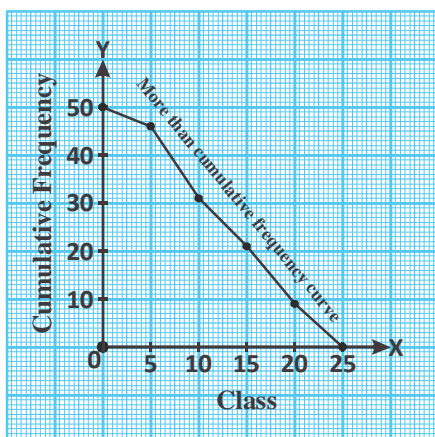
Class	Frequency	Cumulative Frequency
less than 0	0	0
less than 5	4	4
less than 10	15	$19=(15+4)$
less than 15	10	$29=(10+15+4)$
less than 20	12	$41=(12+10+15+4)$
less than 25	9	$50=(9+12+10+15+4)$



2. **"More than" method :** The cumulative frequency of the class is the sum of the frequency of that class and frequencies of all the classes following it.

Based on following table let us try to plot the "more than" cumulative curve:-

Clas	Cumulative Frequency
More than 0	$50 = (4+15+10+12+9+0)$
More than 5	$46 = (15+10+12+9+0)$
More than 10	$31 = (10+12+9+0)$
More than 15	$21 = (12+9+0)$
More than 20	$9 = (9+0)$
More than 25	0



## Importance of Cumulative Frequency Curve or Ogive

Cumulative Frequency Curve or Ogive is used in study of data in many ways.

For example:-

1. When you want to extrapolate the data to one below or one above the given data.
2. Ogive is used for comparative study.
3. Ogive is used to find the measures of central tendency like- median, first quartile, third quartile etc.
4. We can also find out the value of the variable which is included in the special cumulative frequency.

### Exercise - 16.2

1. Choose the correct alternative from the following:-

(i) In an inclusive class:-

- a) Both the limits are included in different classes
- b) Both the limits are included in the same class
- c) It is not decided in which class the limits are included
- d) None of the above

(ii) Method to make the cumulative frequency table is:-

- a) Less than
- b) More than
- c) Both a and b
- d) None of the above

(iii) Using the histogram we can find out

- a) Mode
- b) Median
- c) Both a and b
- d) None of the above

(iv) Using the cumulative frequency curve we can find out:-

- a) Mode
- b) Median
- c) Both a and b
- d) None of the above

(v) The width of the rectangle in a histogram depends on:-

- a) Class interval of the class
- b) Class mark
- c) On the class frequency
- d) On all of the above





2. The time taken (in second) by 25 students in doing a question is as follows:-

16, 20, 26, 27, 28, 30, 33, 37, 38, 40, 42, 43, 46,  
46, 46, 48, 49, 50, 53, 58, 59, 60, 64, 52, 20

- Make frequency table of this data using a class interval of 10 second.
- Make a histogram depicting this frequency distribution.
- Use the histogram to make a frequency polygon.

3. The length of 30 leaves of a plant is given in milimetre below:-

Length of leaves (in mm.)	111-120	121-130	131-140	141-150	151-160	161-170
Number of leaves	3	5	7	9	4	2

- Make a histogram to depict the frequency distribution.  
(*Hint:* Make the class intervals continuous)
- Make a frequency polygon using the direct method.
- In which class the number of leaves is the largest?

4. The number of runs made by two teams A and B in a 10 over (60 ball) match is given below:-

Number of Balls	Team A	Team B
1-6	3	6
7-12	2	6
13-18	7	1
19-24	8	10
25-30	3	6
31-36	8	5
37-42	5	3
43-48	11	4
49-54	6	7
55-60	2	11



Use the above data make a frequency polygon on one graph paper.

(*Hint:* First make the class intervals continuous)

5. The frequency distribution of marks obtained by 100 students is as follows:-

Marks obtained	0-10	10-20	20-30	30-40	40-50	50-60	Total
Number of students	7	10	23	51	6	3	100

Make a cumulative frequency curve using the above data.

6. Make the frequency table and the cumulative frequency curve for the following data:-

Marks Obtained	Number of Students
less than 10	3
less than 20	8
less than 30	12
less than 40	19
less than 50	31
less than 60	42
less than 70	60



7. The marks obtained by two groups of students in a test are given below:-

Class Interval	Group A	Group B
50-52	4	8
47-49	10	9
44-46	15	10
41-43	18	14
38-40	20	12
35-37	12	17
32-34	13	22
<b>Total</b>	<b>92</b>	<b>92</b>



Make frequency polygons for each of these groups on one graph paper.

## What Have We Learnt

1. Observation collected for a specific purpose is known as data.
2. When the number of observation is more we use tally marks to find the frequency.
3. The number of times a particular observation occurs in a data is called the frequency of that observation.



4. The table showing the frequencies of different observations in the data is called frequency distribution table or the frequency table.
  5. When the number of observations is very large then we organise data in groups. These groups are called classes and the data is referred to as classified data.
  6. The histogram is a graphical representation of classified data. In this rectangle is drawn for each group with X-axis showing the class width and heights on the Y-axis corresponding to the frequencies.
  7. If we mark points using the mid-points of the classes as the X-coordinates and the corresponding frequencies of the classes as the Y-coordinates, then the polygon made joining these points is called the frequency polygon.
  8. When we find the total number of observations up to the lower limit of all the class intervals, we obtained the ascending cumulative frequency.
  9. If in a classified frequency distribution the class intervals are different, that the rectangles in the bar graphs would have to be constructed using frequency density.  
$$\text{Frequency density} = \frac{\text{Class frequency}}{\text{Class width}} \times \text{Lower limit of the data}$$
  10. For the same data, the area of frequency polygon and cumulative frequency graph are equal.
- 

