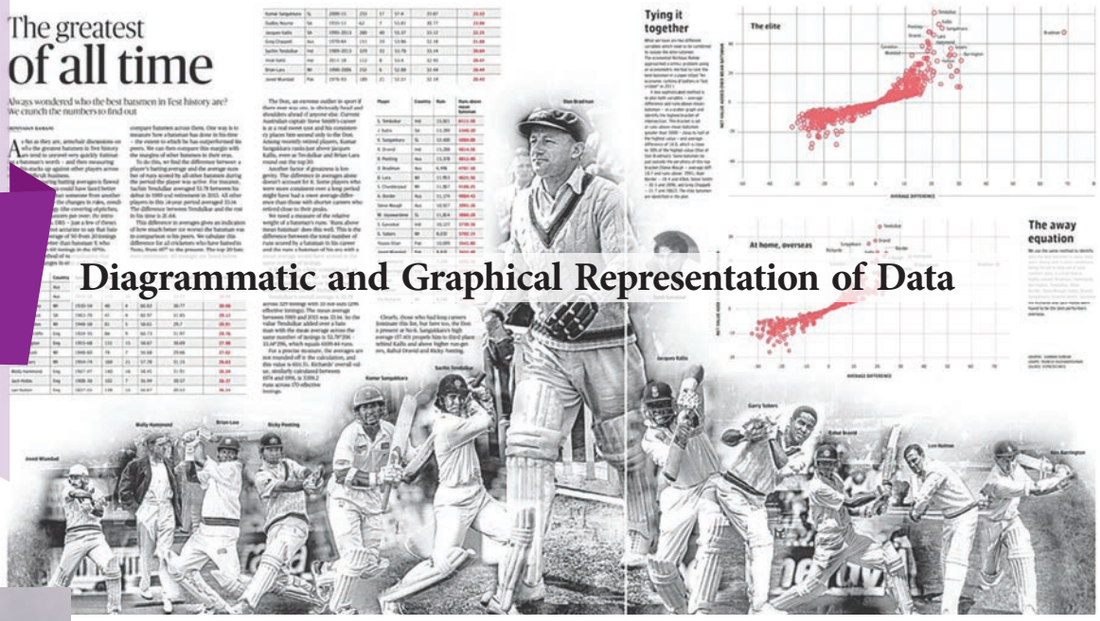
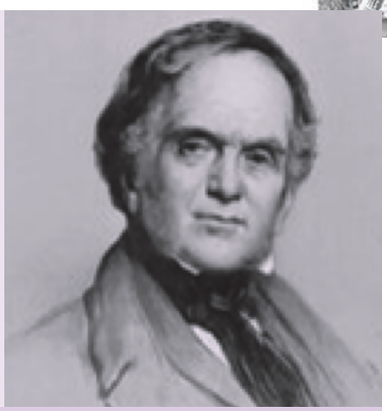
Chapter

4



Diagrammatic and Graphical Representation of Data



William Palyfair
(22 Sep, 1759- 11 Feb, 1823)

William Playfair was a Scottish engineer and political economist, the founder of graphical methods of statistics.

He invented several types of diagrams: In 1786, the line, area and bar chart of economic data, and in 1801 the pie chart and circle graph, used to show part-whole relations.

'Good statistical inference never strays very far from data'
- Brian S Yandell

Learning Objectives

- Presents the data in diagrams
- Understands the various types of diagrams
- Compares the tabular data with diagrammatic representation of data
- Represents the data in a graph
- Enumerates the unknown value using graphs
- Distinguishes diagrammatic and graphical representation of data

Introduction

In the preceding unit, we discussed the techniques of classification and tabulation that help in condensing and presenting the data in a tabular form. These ways of presentation of data in numbers is dull and uninteresting to the common man. For that,

the most convincing and appealing ways are highlighting the salient features of statistical data through visual/pictorial presentation using diagrams and graphs. It is an accepted fact that the pictorial representation is more appealing, attractive and has long last effect. Moreover, a layman who averse to numbers can understand the diagrams more easily. Hence, the newspapers, magazines, journals, advertisements etc., present their numerical facts through diagrams and graphs. It is imperative on the part of a student to understand and apply the pictorial representation in a real life situation.

4.1 Meaning and Significance of Diagrams and Graphs

Diagrams:

A diagram is a visual form for presenting statistical data for highlighting the basic facts and relationship which are inherent in the data. The diagrammatic presentation is more understandable and it is appreciated by everyone. It attracts the attention and it is a quicker way of grasping the results saving the time. It is very much required, particularly, in presenting qualitative data.

Graphs:

The quantitative data is usually represented by graphs.Though it is not quite attractive and understandable by a layman, the classification and tabulation techniques will reduce the complexity of presenting the data using graphs. Statisticians have understood the importance of graphical presentation to present the data in an interpretable way. The graphs are drawn manually on graph papers.

Significance of Diagrams and Graphs:

Diagrams and graphs are extremely useful due to the following reasons:

(i) They are attractive and impressive

(ii) They make data more simple and intelligible

(iii) They are amenable for comparison

(iv) They save time and labour and

(v) They have great memorizing effect.

Cricket is the game in which, graphs and charts are used predominantly to represent data than many other games.

4.2 Rules for Constructing Diagrams

While constructing diagrams for statistical data, the following guidelines are to be kept in mind:

- A diagram should be neatly drawn in an attractive manner
- Every diagram must have a precise and suitable heading
- Appropriate scale has to be defined to present the diagram as per the size of the paper
- The scale should be mentioned in the diagram
- Mention the values of the independent variable along the X-axis and the values of the dependent variable along the Y-axis
- False base line(s) may be used in X-axis and Y-axis, if required
- Legends should be given for X-axis, Y-axis and each category of the independent variable to show the difference
- Foot notes can be given at the bottom of the diagram, if necessary.

Legend means a brief verbal description about the shades/colours applied in the chart/graph. The legend is also known as chart key. It is most often located on the right hand side of the graph/chart and can sometimes be surrounded by a border

4.3 Types of Diagrams

In practice, varieties of diagrams are used to present the data. They are explained below.

4.3.1 Simple Bar Diagram

Simple bar diagram can be drawn either on horizontal or vertical base. But, bars on vertical base are more common. Bars are erected along the axis with uniform width and space between the bars must be equal. While constructing a simple bar diagram, the scale is determined as proportional to the highest value of the variable. The bars can be coloured to make the diagram attractive. This diagram is mostly drawn for categorical variable. It is more useful to present the data related to the fields of Business and Economics.

Example 4.1

The production cost of the company in lakhs of rupees is given below.

(i) Construct a simple bar diagram.

(ii) Find in which year the production cost of the company is (a) maximum (b) minimum (c) less than 40 lakhs.

(iii) What is the average production cost of the company?

(iv) What is the percentage increase from 2014 to 2015?

Year	Production Cost
2010	55
2011	40
2012	30
2013	25
2014	35
2015	70

Solution:

(i) We represent the above data by simple bar diagram in the following manner:

Step-1: Years are marked along the X-axis and labelled as 'Year'.

Step-2: Values of Production Cost are marked along the Y-axis and labelled as 'Production Cost (in lakhs of ₹).

Step-3: Vertical rectangular bars are erected on the years marked and whose height is proportional to the magnitude of the respective production cost.

Step-4: Vertical bars are filled with the same colours.

The simple bar diagram is presented in Fig.4.1.

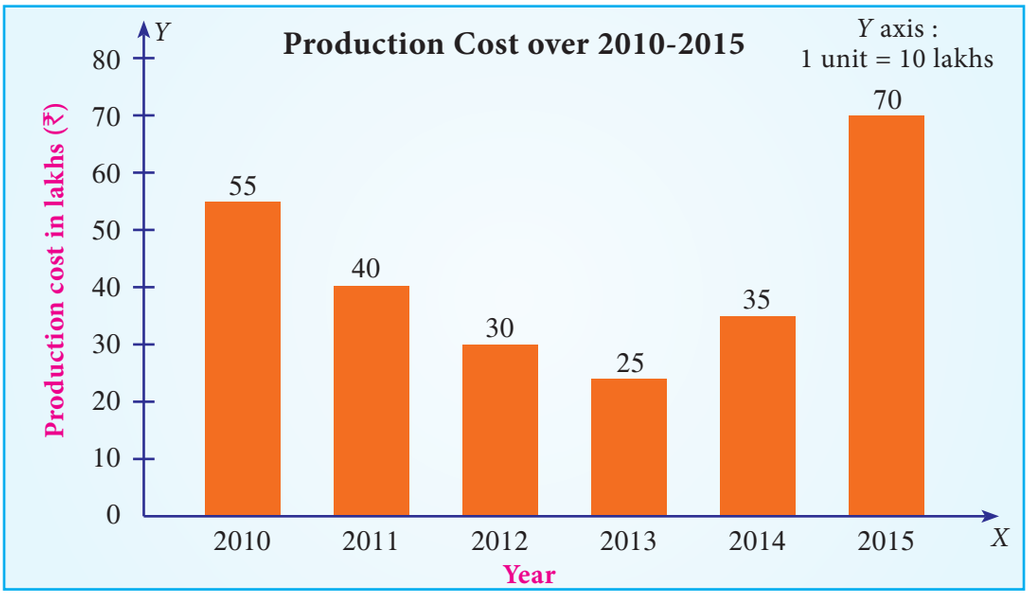


Fig 4.1

(ii) (a) The maximum production cost of the company was in the year 2015.

(b) The minimum production cost of the company was in the year 2013.

(c) The production cost of the company during the period 2012- 2014 is less than 40 lakhs.

(iii) Average production Cost of the company

$$= \frac{55 + 40 + 30 + 25 + 35 + 70}{6}$$

$$= ₹ 42.5 \text{ Lakhs}$$

(iv) Percentage increase in the production cost of the company is

$$= \frac{70 - 35}{35} \times 100$$

$$= 100\%$$

4.3.2 Pareto Diagram:

Vilfredo Pareto (1848-1923), born in Paris in an Italian aristocratic family, studied Engineering and Mathematics at the University of Turin. During his studies at the University of Lousane in Switzerland, Pareto derived a complicated mathematical formula to prove the distribution of income and wealth in society is not random. Approximately 80% of total wealth in a society lies with only 20% of the families. The famous law about the 'Vital few and trivial many' is widely known as 'Pareto Principle' in Economics.

Pareto diagram is similar to simple bar diagram. But, in Pareto diagram, the bars are arranged in the descending order of the heights of the bars. In addition, there will be a line representing the cumulative frequencies (in %) of the different categories of the variable. The line is more useful to find the vital categories among trivial categories

Example 4.2

Administration of a school wished to initiate suitable preventive measures against breakage of equipment in its Chemistry laboratory. Information collected about breakage of equipment occurred during the year 2017 in the laboratory are given below:

Equipment	No. of breakages
Burette	45
Conical flask	75
Test tube	150
Pipette	30

Draw Pareto Diagram for the above data. Which equipment requires more attention in order to reduce breakages?

Solution:

Since we have to find the vital few among the several, we draw Pareto diagram.

Step 1 : Arrange the equipment according to the descending order of the number of breakages.

Step 2 : Find the percentage of breakages for each equipment using the formula $= \frac{No.\ of\ Breakages}{Total\ No.\ of\ Breakages} \times 100$

Step 3 : Calculate cumulative percentage for each equipment.

Step 4 : Mark the equipment along the X-axis and the number of breakages along the Y-axis. Construct an attached simple bar diagram to this data. In an attached simple bar diagram, the vertical bars are erected adjacently.

Step 5 : Mark the cumulative no. of breakages for each equipment corresponding to the mid-point of the respective vertical bar.

Step 6 : Draw a free hand curve joining those plotted points.

Equipment	No. of Breakages (Frequency)	No. of Breakages in percentage	Cumulative No. of Breakages in percentage
Test tube	150	$\frac{150}{300} \times 100 = 50$	50
Conical flask	75	$\frac{75}{300} \times 100 = 25$	75
Burette	45	$\frac{45}{300} \times 100 = 15$	90
Pipette	30	$\frac{30}{300} \times 100 = 10$	100
Total	**300**	**100**	

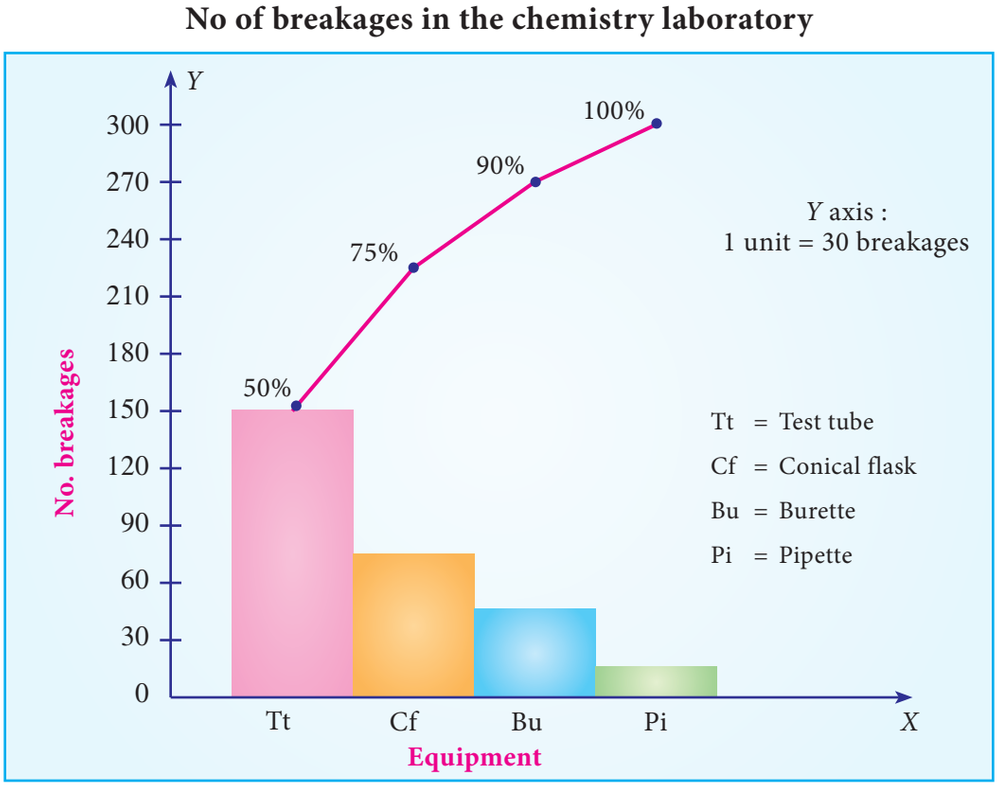


Fig 4.2: Pareto Diagram for No. of Breakages in the Chemistry Laboratory

From Fig 4.2, it can be found that 50% of breakages is due to Test tube, 25% due to Conical Flask. Therefore, the School Administration has to focus more attention on reducing the breakages of Test Tubes and Conical Flasks.

4.3.3 Multiple Bar Diagram

Multiple bar diagram is used for comparing two or more sets of statistical data. Bars with equal width are placed adjacently for each cluster of values of the variable. There should be equal space between clusters. In order to distinguish bars in each cluster, they may be either differently coloured or shaded. Legends should be provided.

Example 4.3

The table given below shows the profit obtained before and after tax payment(in lakhs of rupees) by a business man on selling cars from the year 2014 to 2017.

Year	Profit before tax	Profit after tax
2014	195	80
2015	200	87
2016	165	45
2017	140	32

(i) Construct a multiple bar diagram for the above data.

(ii) In which year, the company earned maximum profit before paying the tax?

(iii) In which year, the company earned minimum profit after paying the tax?

(iv) Find the difference between the average profit earned by the company before paying the tax and after paying the tax.

Solution:

Since we are comparing the profit earned before and after paying the tax by the same Company, the multiple bar diagram is drawn. The diagram is drawn following the procedure presented below:

Step 1 : Years are marked along the X-axis and labeled as "Year".

Step 2 : Values of Profit before and after paying the tax are marked along the Y-axis and labeled as "Profit (in lakhs of ₹)".

Step 3 : Vertical rectangular bars are erected on the years marked, whose heights are proportional to the respective profit. The vertical bars corresponding to the profit earned before and after paying the tax in each year are placed adjacently.

Step 4 : The vertical bars drawn corresponding to the profit earned before paying the tax are filled with one type of colour. The vertical bars drawn corresponding to the profit earned after paying the tax are filled with another type of colour. The colouring procedure should be applied to all the years uniformly.

Step 5 : Legends are displayed to describe the different colours applied to the bars drawn for profit earned before and after paying the tax.

The multiple bar diagram is presented in Fig 4.3.

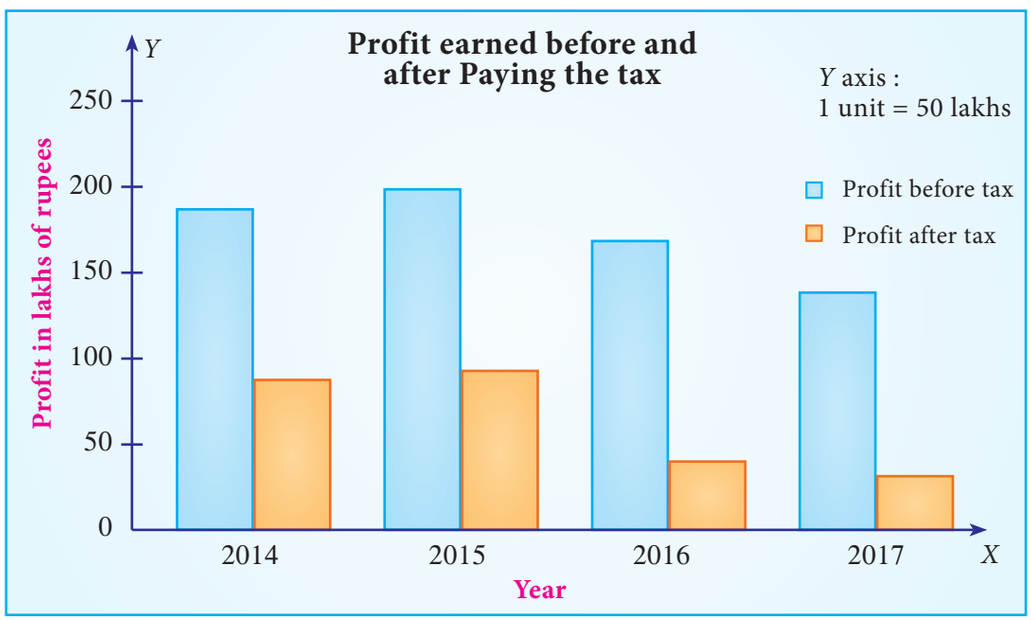


Fig 4.3 Multiple Bar Diagram for Profit by the Company earned before and after paying the Tax

(i) The company earned the maximum profit before paying the tax in the year 2015.

(ii) The company earned the minimum profit after paying the tax in the year 2017.

(iii) The average profit earned before paying the tax $= \frac{700}{4}$ = ₹ 175 lakhs

The average profit earned after paying the tax $= \frac{244}{4}$ = ₹ 61 lakhs

Hence, difference between the average profit earned by the company before paying the tax and after paying the tax is

= 175 – 61 = ₹ 114 lakhs.

4.3.4 Component Bar Diagram(Sub-divided Bar Diagram)

A component bar diagram is used for comparing two or more sets of statistical data, as like multiple bar diagram. But, unlike multiple bar diagram, the bars are stacked in component bar diagrams. In the construction of sub-divided bar diagram, bars are drawn with equal width such that the heights of the bars are proportional to the magnitude of the total frequency. The bars are positioned with equal space. Each bar is sub-divided into various parts in proportion to the values of the components. The subdivisions are distinguished by different colours or shades. If the number of clusters and the categories in the clusters are large, the multiple bar diagram is not attractive due to more number of bars. In such situation, component bar diagram is preferred.

Example 4.4

Total expenditure incurred on various heads of two schools in an year are given below. Draw a suitable bar diagram.

Expenditure Head	Amount (in lakhs)	
	School I	School II
Construction/Repairs	80	90
Computers	35	50
Laboratory	30	25
Watering plants	45	40
Library books	40	30
Total	230	235

Which school had spent more amount for

(a) construction/repairs (b) Watering plants?

Solution :

Since we are comparing the amount spent by two schools in a year towards various expenditures with respect to their total expenditures, a component bar diagram is drawn.

Step 1 : Schools are marked along the *X*-axis and labeled as "School".

Step 2 : Expenditure Head are marked along the *Y*-axis and labeled as "Expenditure (₹ in lakhs)".

Step 3 : Vertical rectangular bars are erected for each school, whose heights are proportional to their respective total expenditure.

Step 4 : Each vertical bar is split into components in the order of the list of expenditure heads. Area of each rectangular box is proportional to the frequency of the respective expenditure head/component. Rectangular boxes for each school are coloured with different colours. Same colours are applied to the similar expenditure heads for each school.

Step 5 : Legends are displayed to describe the colours applied to the rectangular boxes drawn for various expenditure heads.

The component bar diagram is presented in Fig 4.4.

Expenditure Head	Amount (₹ in lakhs)			
	School I		School II	
	Amount Spent	Cumulative Amount Spent	Amount Spent	Cumulative Amount Spent
Construction/Repairs	80	80	90	90
Computers	35	115	50	140
Laboratory	30	145	25	165
Watering plants	45	190	40	205
Library books	40	230	30	235

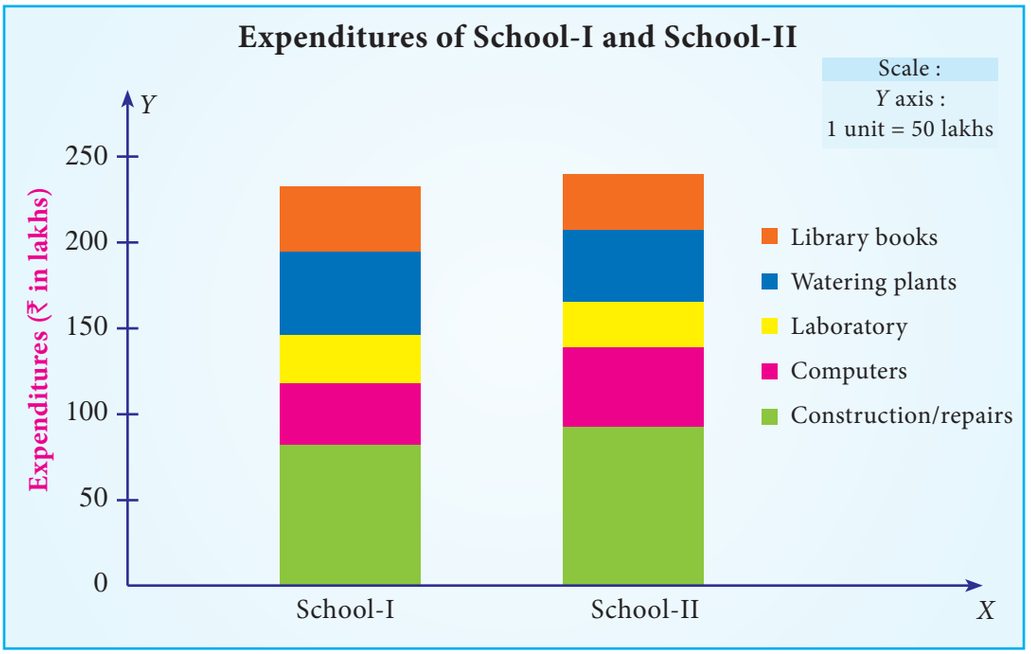


Fig 4.4 Component Bar diagram for expenditures of School I and School II

(i) School- II had spent more amount towards Construction/Repairs.

(ii) School- I had spent more amount towards Watering plants.

4.3.5 Percentage Bar Diagram

Percentage bar diagram is another form of component bar diagram. Here, the heights of the components do not represent the actual values, but percentages. The main difference between sub-divided bar diagram and percentage bar diagram is that, in the former, the height of the bars corresponds to the magnitude of the value. But, in the latter, it corresponds to the percentages. Thus, in the component bar diagram, heights of the bars are different, whereas in the percentage bar diagram, heights are equal corresponding to 100%. Hence, percentage bar diagram will be more appealing than sub-divided bar diagram. Also, comparison between components is much easier using percentage bar diagram.

Example 4.5

Draw the percentage sub-divided bar diagram to the data given in Example 4.4. Also find

(i) The percentage of amount spent for computers in School I

(ii) What are the expenditures in which School II spent more than School I.

Solution:

Since we are comparing the amount spent by two schools in a year towards various expenditures with respect to their total expenditures in percentages, a percentage bar diagram is drawn.

Step 1 : Schools are marked along the *X*-axis and labeled as "School".

Step 2 : Amount spent in percentages are marked along the Y-axis and labeled as "Percentage of Expenditure (₹ in lakhs)".

Step 3 : Vertical rectangular bars are erected for each school, whose heights are taken to be hundred.

Step 4 : Each vertical bar is split into components in the order of the list of percentage expenditure heads. Area of each rectangular box is proportional to the percentage of frequency of the respective expenditure head/component. Rectangular boxes for each school are coloured with different colours. Same colours are applied to the similar expenditure heads for each school.

Step 5 : Legends are displayed to describe the colours applied to the rectangular boxes drawn for various expenditure heads.

The percentage bar diagram is presented in Fig 4.5.

Expenditure Head	Amount (₹ in lakhs)					
	School I			School II		
	Amount spent	Percentage of Amount spent	Cumulative Percentage	Amount spent	Percentage of Amount spent	Cumulative Percentage
Construction / Repairs	80	35	35	90	38	38
Computers	35	15	50	50	21	59
Laboratory	30	13	63	25	11	70
Watering plants	45	20	83	40	17	87
Library books	40	17	100	30	13	100
Total	**230**	**100**		**235**	**100**	

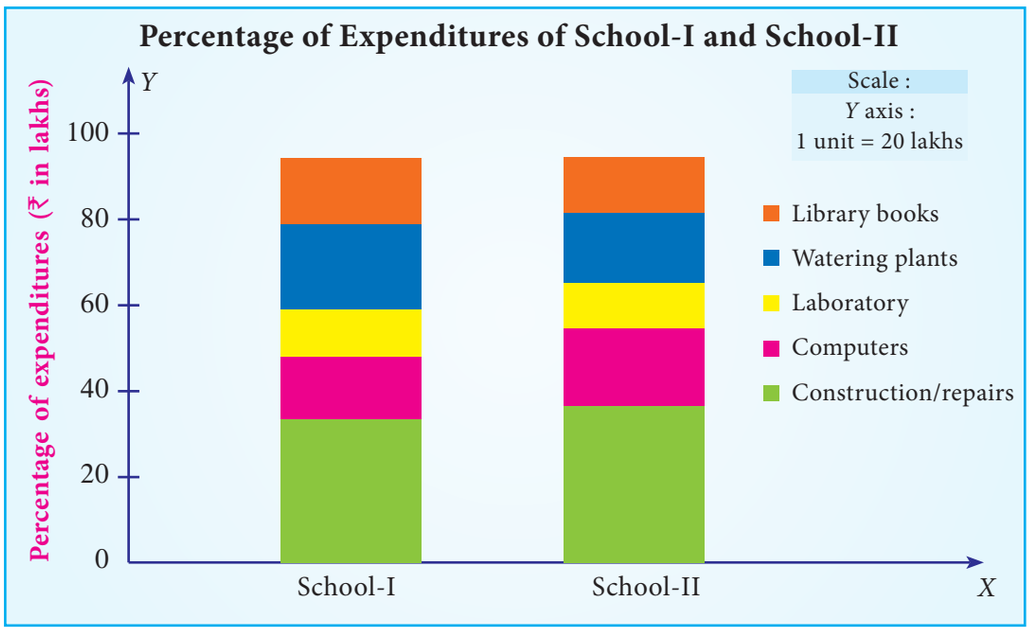


Fig 4.5 Percentage Bar diagram for expenditures of School I and School II

(i) 21% of the amount was spent for computers in School I

(ii) 38% of expenditure was spent for construction/Repairs by School II than School I.

4.3.6 Pie Diagram

The Pie diagram is a circular diagram. As the diagram looks like a pie, it is given this name. A circle which has 360° is divided into different sectors. Angles of the sectors, subtending at the center, are proportional to the magnitudes of the frequency of the components.

Procedure:

The following procedure can be followed to draw a Pie diagram for a given data:

(i) Calculate total frequency, say, N.

(ii) Compute angles for each component using the formula. $\frac{class\ frequemcy}{N} \times 360$

(iii) Draw a circle with radius of sufficient length as a horizontal line.

(iv) Draw the first sector in the anti-clockwise direction at an angle calculated for the first component.

(v) Draw the second sector adjacent to the first sector at an angle corresponding to the second component.

(vi) This process may be continued for all the components.

(vii) Shade/colour each sector with different shades/colours.

(viii) Write legends to each component.

Example 4.6

Draw a pie diagram for the following data (in hundreds) of house hold expenditure of a family.

Items	Expenditure
Food	87
Clothing	24
Recreation	11
Education	13
Rent	25
Miscellaneous	20

Also find

(i) The central angle of the sector corresponding to the expenditure incurred on Education

(ii) By how much percentage the recreation cost is less than the Rent.

Solution :

The following procedure is followed to draw a Pie diagram for a given data:

(i) Calculate the total expenditure, say, N.

(ii) Compute angles for each component food, clothing, recreation, education, rent and miscellaneous using the formula $\frac{class\ frequency}{N} \times 360$

Item	Expenditure	Angle of the circle
Food	87	$\frac{87}{180} \times 360 = 174$
Clothing	24	$\frac{24}{180} \times 360 = 48$
Recreation	11	$\frac{11}{180} \times 360 = 22$
Education	13	$\frac{13}{180} \times 360 = 26$
Rent	25	$\frac{25}{180} \times 360 = 50$

Miscellaneous	20	$\frac{20}{180} \times 360 = 40$
Total	N=180	360

(iii) Draw a circle with radius of sufficient length as a horizontal line.

(iv) Draw the first sector in the anti-clockwise direction at an angle calculated for the first component food.

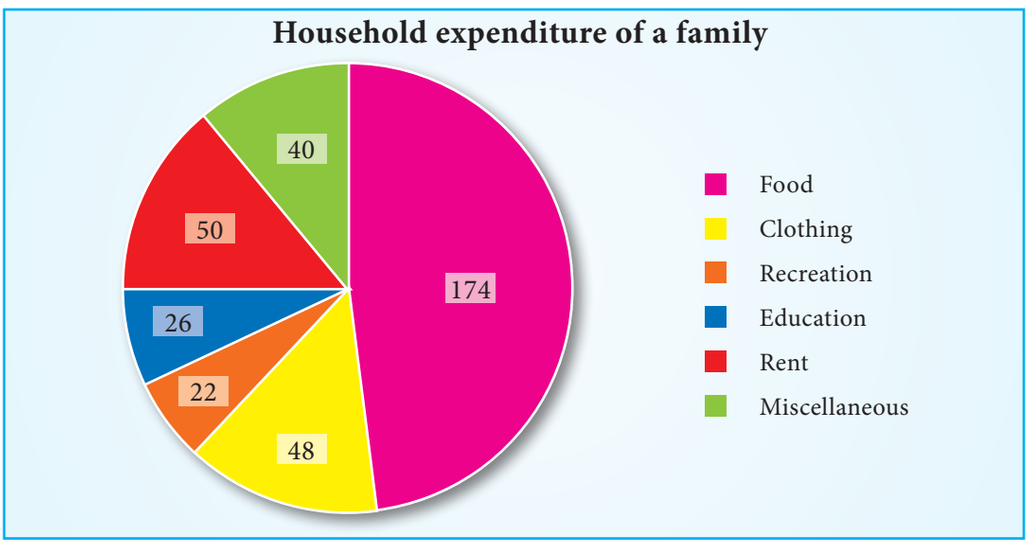
(v) Draw the second sector adjacent to the first sector at an angle corresponding to the second component clothing.

(vi) This process is continued for all the components namely recreation, education, rent and miscellaneous.

(vii) Shade/colour each sector with different shades/colours.

(viii) Write legends to each component.

The pie diagram is presented in Fig 4.6.



The central angle of the sector corresponding to the expenditure incurred on Education is 26°

Recreation cost is less than rent by 28°

4.3.7 Pictogram

Pictograms are diagrammatic representation of statistical data using pictures of resemblance. These are very useful in attracting attention. They are easily understood. For the purpose of propaganda, the pictorial presentations of facts are quite popular

and they also find places in exhibitions. They are extensively used by the government organizations as well as by private institutions. If needed, scales can be fixed.

Despite its visual advantages, pictogram has limited application due to the usage of pictures resembling the data. It can express an approximate value than the given actual numerical value..

Example 4.7

The following table gives the sugarcane production in tonnes per acre for various years.

Year	2013	2014	2015	2016	2017
Sugar Cane (in tonnes per acre)	10	13	9	15	18

Represent the above data by pictogram.

Solution :

The above data is represented by pictogram in the following manner:

2013	
2014	
2015	
2016	
2017	

 = 1 tonne

Example 4.8

The Pictogram given below shows the number of persons who have traveled by train from Chennai to Rameshwaram on each day of a week

Sunday	
Monday	
Tuesday	
Wednesday	
Thursday	
Friday	
Saturday	

= 100 persons

From the Pictogram find:

(i) Number of travelers travelled during the week

(ii) On which day there was a maximum rush in the train

(iii) The difference between the maximum and minimum number of travelers.

Solution :

(i) Here total number of is 48, and each represents 100 persons.

Hence number of travelers travelled during the week is 48×100=4800 .

(ii) The maximum rush in the train is on Thursday.

(iii) Maximum number of persons travelled on Thursday = 10

Hence the number of persons travelled on Thursday is 10×100 = 1000

Minimum number of persons travelled on Wednesday = 4

Hence the number of persons travelled on Wednesday is 4×100=400

Therefore difference between maximum and minimum number of travelers is 1000-400=600 persons.

4.4 Types of Graphs

Graphical representation can be advantageous to bring out the statistical nature of the frequency distribution of quantitative variable, which may be discrete or continuous.

The most commonly used graphs are

(i) Histogram

(ii) Frequency Polygon

(iii) Frequency Curve

(iv) Cumulative Frequency Curves (Ogives)

4.4.1 Histogram

A histogram is an attached bar chart or graph displaying the distribution of a frequency distribution in visual form. Take classes along the X-axis and the frequencies along the Y-axis. Corresponding to each class interval, a vertical bar is drawn whose height is proportional to the class frequency.

Limitations:

We cannot construct a histogram for distribution with open-ended classes. The histogram is also quite misleading, if the distribution has unequal intervals.

Example 4.9

Draw the histogram for the 50 students in a class whose heights (in cms) are given below.

Height	111 – 120	121 – 130	131 – 140	141 – 150	151 – 160	161 – 170
Number of students	4	11	15	9	8	3

Find the range, whose height of students are maximum.

Solution:

Since we are displaying the distribution of Height and Number of students in visual form, the histogram is drawn.

Step 1 : Heights are marked along the X-axis and labeled as "Height(in cms)".

Step 2 : Number of students are marked along the Y-axis and labeled as "No. of students".

Step 3 : Corresponding to each Heights, a vertical attached bar is drawn whose height is proportional to the number of students.

The Histogram is presented in Fig 4.7.

For drawing a histogram, the frequency distribution should be continuous. If it is not continuous, then make it continuous as follows.

Height (in Cm)	No.of Students
110.5 - 120.5	4
120.5 - 130.5	11
130.5 - 140.5	15
140.5 - 150.5	9
150.5 - 160.5	8
160.5 - 170.5	3

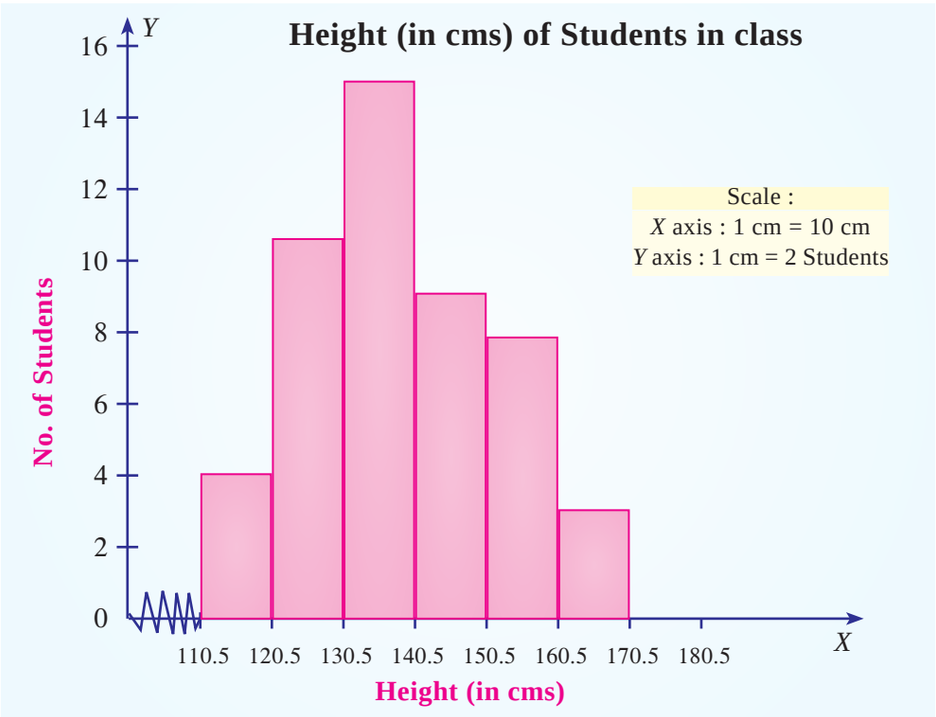


Fig 4.7 Histogram for heights of students in a class

The tallest bar shows that maximum number of students height are in the range 130.5 to 140.5 cm

Example 4.10

The following table shows the time taken (in minutes) by 100 students to travel to school on a particular day

Time	0-5	5-10	10-15	15-20	20-25
No. of Students	5	25	40	17	13

Draw the histogram. Also find:

(i) The number of students who travel to school within 15 minutes.

(ii) Number of students whose travelling time is more than 20 minutes.

Solution:

Since we are displaying the distribution of time taken (in minutes) by 100 students to travel to school on a particular day in visual form, the histogram is drawn.

Step 1 : Time taken are marked along the X-axis and labeled as "Time (in minutes)".

Step 2 : Number of students are marked along the Y-axis and labeled as "No. of students".

Step 3 : Corresponding to each time taken, a vertical attached bar is drawn whose height is proportional to the number of students.

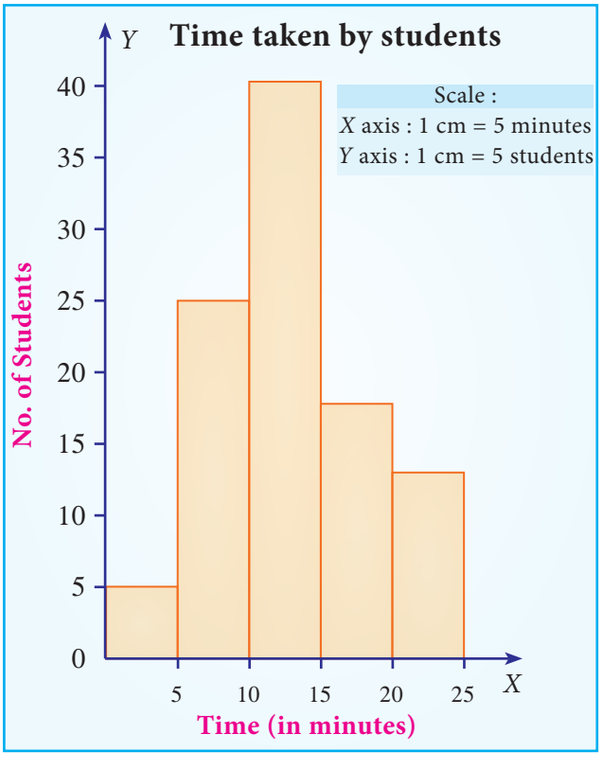


Fig 4.8 Histogram for time taken by students to travel to school

The Histogram is presented in Fig 4.8.

(i) 5+25+40=70 students travel to school within 15 minutes

(ii) 13 students travelling time is more than 20 minutes

Example 4.11

Draw a histogram for the following 100 persons whose daily wages (in ₹) are given below.

Daily wages	0 – 50	50 – 100	100 – 200	200 – 250	250 – 450	450 – 500
Number of persons	5	10	16	7	48	14

Also find:

(i) Number of persons who gets daily wages less than ₹ 200?

(ii) Number of persons whose daily wages are more than ₹ 250?

Solution:

Since we are displaying the distribution of 100 persons whose daily wages in rupees in visual form, the histogram is drawn.

Step 1 : Daily wages are marked along the *X*-axis and labeled as "Daily Wages (in ₹)".

Step 2 : Number of Persons are marked along the Y-axis and labeled as "No. of Persons".

Step 3 : Corresponding to each daily wages, a vertical attached bar is drawn whose height is proportional to the number of persons.

The Histogram is presented in Fig 4.9.

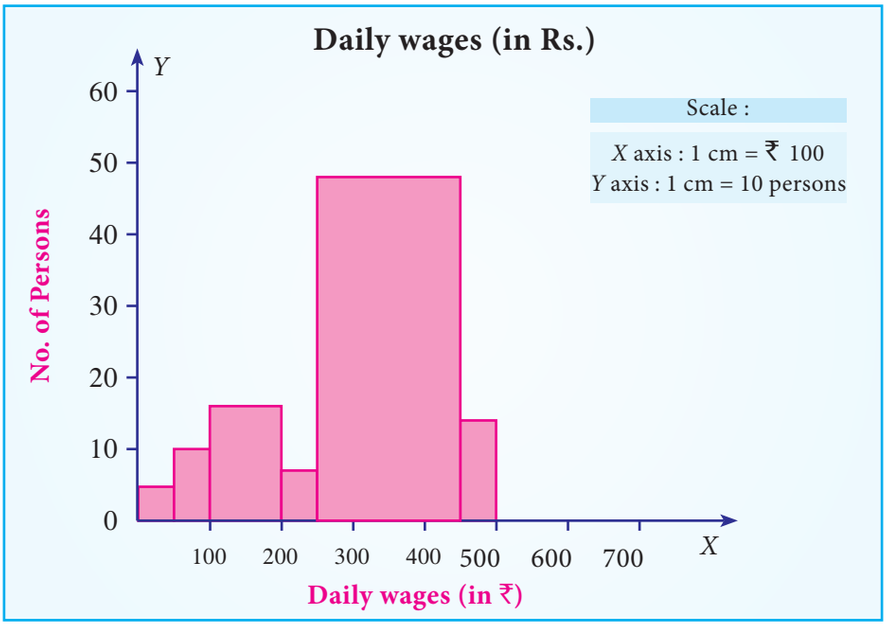


Fig 4.9 Histogram of daily wages (in ₹) for persons

(i) 5+10+16=31 persons get daily wages less than ₹ 200.

(ii) 48+14=62 persons get more than Rs 250.

4.4.2 Frequency Polygon

Frequency polygon is drawn after drawing histogram for a given frequency distribution. The area covered under the polygon is equal to the area of the histogram. Vertices of the polygon represent the class frequencies. Frequency polygon helps to determine the classes with higher frequencies. It displays the tendency of the data. The following procedure can be followed to draw frequency polygon:

(i) Mark the midpoints at the top of each vertical bar in the histogram representing the classes.

(ii) Connect the midpoints by line segments.

Example 4.12

A firm reported that its Net Worth in the years 2011-2016 are as follows:

Year	2011 -2012	2012 – 2013	2013 – 2014	2014 – 2015	2015 - 2016
Net Worth (₹ in lakhs)	100	112	120	133	117

Draw the frequency polygon for the above data

Solution:

Since we are displaying the distribution of Net worth in the years 2011-2016, the Frequency polygon is drawn to determine the classes with higher frequencies. It displays the tendency of the data.

The following procedure can be followed to draw frequency polygon:

Step 1 : Year are marked along the *X*-axis and labeled as 'Year'.

Step 2 : Net worth are marked along the *Y*-axis and labeled as 'Net Worth (in lakhs of ₹)'.

Step 3 : Mark the midpoints at the top of each vertical bar in the histogram representing the year.

Step 4 : Connect the midpoints by line segments.

The Frequency polygon is presented in Fig 4.10.

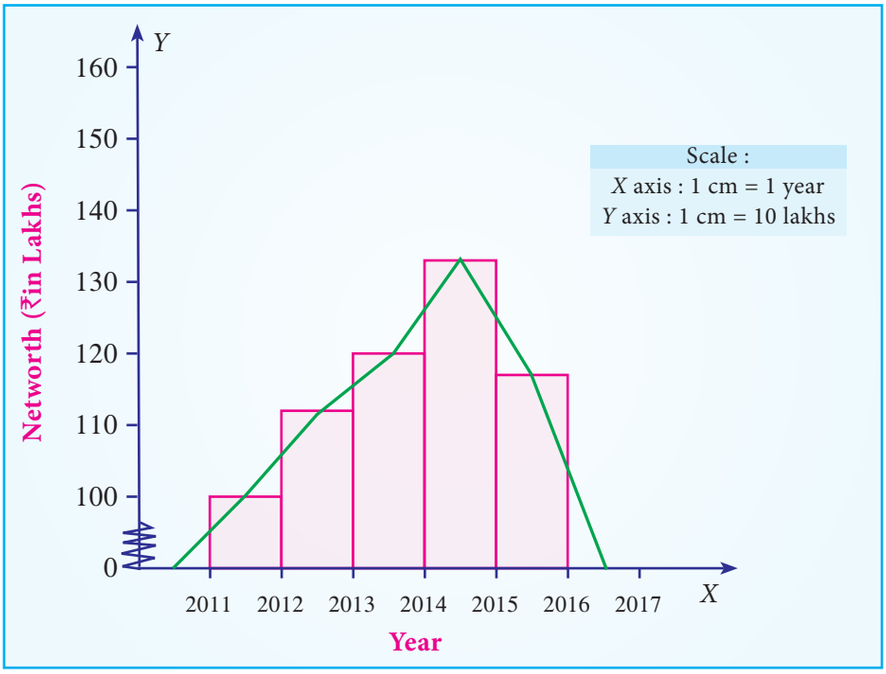


Fig 4.10 Frequency polygon for Net Worth in the years 2011-2016

4.4.3 Frequency Curve

Frequency curve is a smooth and free-hand curve drawn to represent a frequency distribution. Frequency curve is drawn by smoothing the vertices of the frequency

polygon. Frequency curve provides better understanding about the properties of the data than frequency polygon and histogram.

Example 4.13

The ages of group of pensioners are given in the table below. Draw the Frequency curve to the following data.

Age	65 - 70	70 - 75	75 - 80	80 - 85	85 - 90
No.of pensioners	38	45	24	10	8

Solution:

Since we are displaying the distribution of Age and Number of Pensioners, the Frequency curve is drawn, to provide better understanding about the age and number of pensioners than frequency polygon.

The following procedure can be followed to draw frequency curve:

Step 1 : Age are marked along the *X*-axis and labeled as 'Age'.

Step 2 : Number of pensioners are marked along the *Y*-axis and labeled as 'No. of Pensioners'.

Step3 : Mark the midpoints at the top of each vertical bar in the histogram representing the age.

Step 4 : Connect the midpoints by line segments by smoothing the vertices of the frequency polygon

The Frequency curve is presented in Fig 4.11.

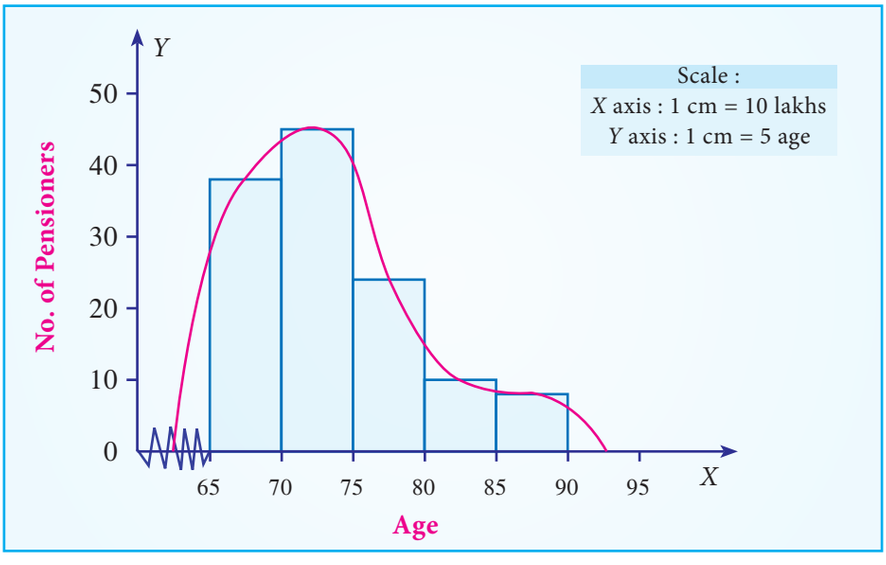


Fig 4.11 Frequency curve for Age and No. of pensioners

4.4.4 Cumulative frequency curve (Ogive)

Cumulative frequency curve (Ogive) is drawn to represent the cumulative frequency distribution. There are two types of Ogives such as 'less thanOgive curve' and 'more thanOgive curve'. To draw these curves, we have to calculate the 'less than' cumulative frequencies and 'more than' cumulative frequencies. The following procedure can be followed to draw the ogive curves:

Less than Ogive: Less than cumulative frequency of each class is marked against the corresponding upper limit of the respective class. All the points are joined by a free-hand curve to draw **the less than ogive** curve.

More than Ogive: More than cumulative frequency of each class is marked against the corresponding lower limit of the respective class. All the points are joined by a free-hand curve to draw the **more than ogive** curve.

Both the curves can be drawn separately or in the same graph. If both the curves are drawn in the same graph, then the value of abscissa (x-coordinate) in the point of intersection is the median.

Median is a measure of central tendency, which divides the given data/distribution into two equal parts. It is discussed much in detail in Unit V

If the curves are drawn separately, median can be calculated as follows:

Draw a line perpendicular to Y-axis at $y=N/2$. Let it meet the Ogive at C. Then, draw a perpendicular line to X-axis from the point C. Let it meet the X-axis at M. The abscissa of M is the median of the data.

Example 4.14

Draw the less than Ogive curve for the following data:

Daily Wages (in Rs.)	70- 80	80- 90	90-100	100-110	110-120	120-130	130-140	140-150
No. of workers	12	18	35	42	50	45	20	8

Also, find

(i) The Median

(ii) The number of workers whose daily wages are less than ₹ 125.

Solution:

Since we are displaying the distribution of Daily Wages and No. of workers, the Ogive curve is drawn, to provide better understanding about the wages and No. of workers.

The following procedure can be followed to draw Less than Ogive curve:

Step 1 : Daily wages are marked along the X-axis and labeled as "Wages(in ₹)".

Step 2 : No. of Workers are marked along the Y-axis and labeled as "No. of workers".

Step 3 : Find the less than cumulative frequency, by taking the upper class-limit of daily wages. The cumulative frequency corresponding to any upper class-limit of daily wages is the sum of all the frequencies less than the limit of daily wages.

Step 4 : The less than cumulative frequency of Number of workers are plotted as points against the daily wages (upper-limit). These points are joined to form less than ogive curve.

The Less than Ogive curve is presented in Fig 4.12.

Daily wages (less than)	No of workers
80	12
90	30
100	65
110	107
120	157
130	202
140	222
150	230

Daily Wages of Workers

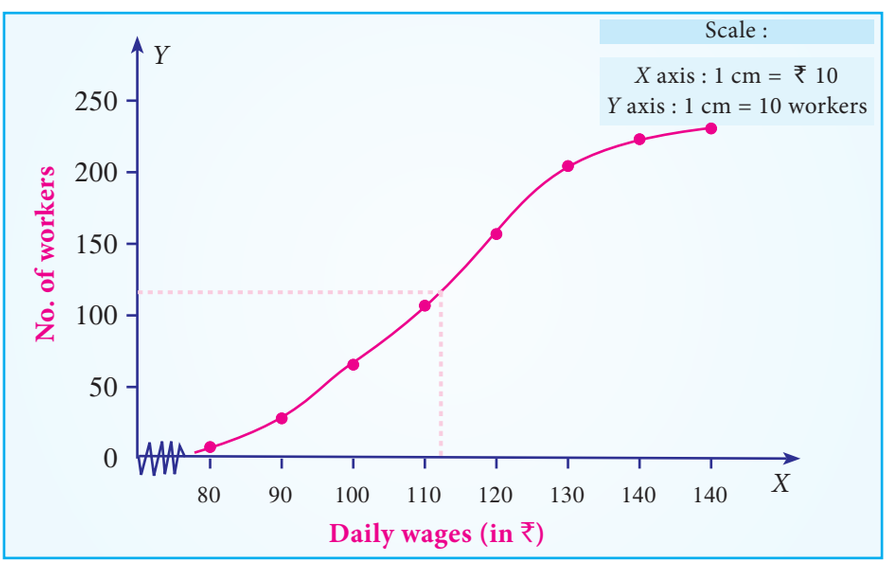


Fig 4.12 Less than Ogive curve for daily wages and number of workers

(i) Median = ₹ 113

(ii) 183 workers get daily wages less than ₹ 125

Example 4.15

The following table shows the marks obtained by 120 students of class IX in a cycle test-I . Draw the more than Ogive curve for the following data :

Marks	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-100
No. of students	2	6	8	20	30	22	18	8	4	2

Also, find

(i) The Median

(ii) The Number of students who get more than 75 marks.

Solution:

Since we are displaying the distribution marks and No. of students, the more than Ogive curve is drawn, to provide better understanding about the marks of the students and No. of students.

The following procedure can be followed to draw More than Ogive curve:

Step 1 : Marks of the students are marked along the *X*-axis and labeled as 'Marks'.

Step 2 : No. of students are marked along the *Y*-axis and labeled as 'No. of students'.

Step 3 : Find the more than cumulative frequency, by taking the lower class-limit of marks. The cumulative frequency corresponding to any lower class-limit of marks is the sum of all the frequencies above the limit of marks.

Step 4 : The more than cumulative frequency of number of students are plotted as points against the marks (lower-limit). These points are joined to form more than ogive curve.

The More than Ogive curve is presented in Fig 4.13.

Marks More than	**No of Students**
0	120
10	118

20	112
30	104
40	84
50	54
60	32
70	14
80	6
90	2

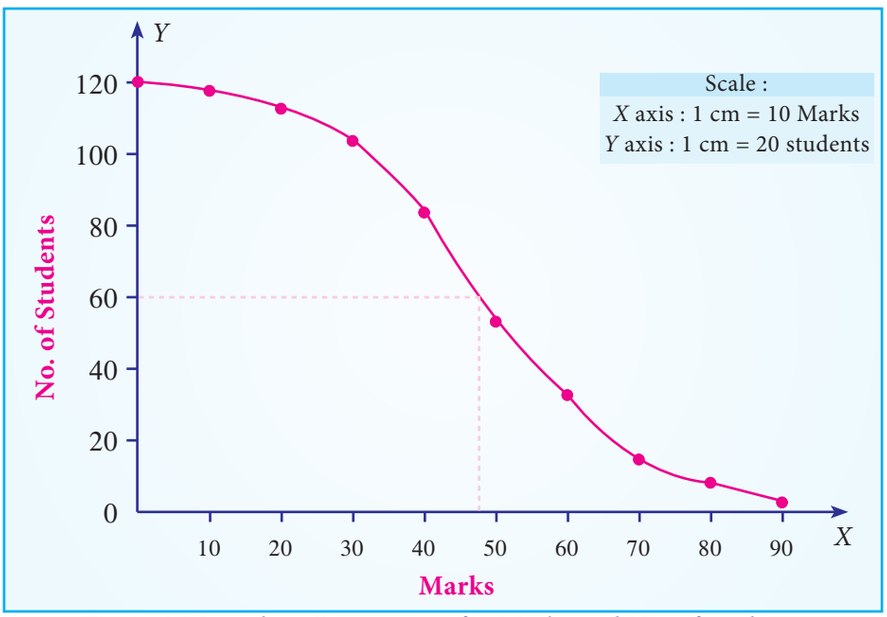


Fig 4.13 More than Ogive curve for Marks and No. of students

(i) Median = 47 students

(ii) 7 students get more than 75 marks.

Example 4.16

The yield of mangoes were recorded (in kg)are given below:

Graphically,

(i) find the number of trees which yield mangoes of less than 55 kg.

(ii) find the number of trees from which mangoes of more than 75 kg.

(iii) find the median.

Draw the Less than and More than Ogive curves. Also, find the median using the Ogive curves

Yield (in kg)	No. of trees
40 – 50	10
50 – 60	15
60 – 70	17
70 – 80	14
80 – 90	12
90 – 100	2
Total	70

Solution:

Since we are displaying the distribution of Yield and No. of trees, the Ogive curve is drawn, to provide better understanding about the Yield and No. of trees

The following procedure can be followed to draw Ogive curve:

Step 1 : Yield of mangoes are marked along the *X*-axis and labeled as 'Yield (in Kg.)'.

Step 2 : No. of trees are marked along the *Y*-axis and labeled as 'No. of trees'.

Step 3 : Find the less than cumulative frequency, by taking the upper class-limit of Yield of mangoes. The cumulative frequency corresponding to any upper class-limit of Mangoes is the sum of all the frequencies less than the limit of mangoes.

Step 4 : Find the more than cumulative frequency, by taking the lower class-limit of Yield of mangoes. The cumulative frequency corresponding to any lower class-limit of Mangoes is the sum of all the frequencies above the limit of mangoes.

Step 5 : The less than cumulative frequency of Number of trees are plotted as points against the yield of mangoes (upper-limit). These points are joined to form less than ogive curve.

Step 6 : The more than cumulative frequency of Number of trees are plotted as points against the yield of mangoes (lower-limit). These points are joined to form more than O give curve.

Less than Ogive		More than Ogive	
Yield less than	No. of trees	Yield greater than	No. of trees
50	10	40	70
60	25	50	60
70	42	60	45
80	56	70	28
90	68	80	14
100	70	90	2

The Ogive curve is presented in Fig 4.14.

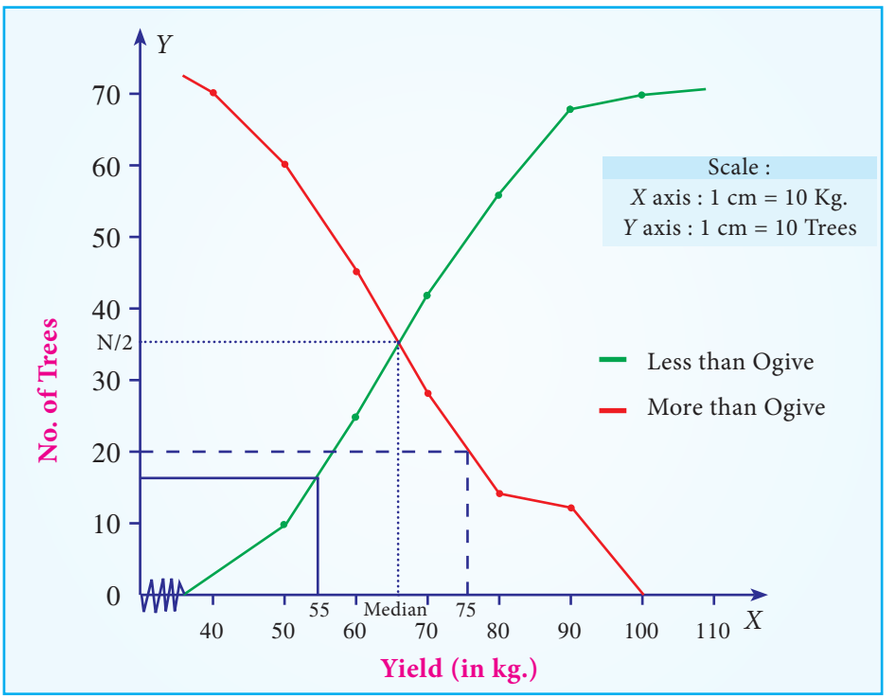


Fig 4.14 Ogive curve for Yield of mangoes and number of trees

(i) 16 trees yield less than 55 kg

(ii) 20 trees yield more than 75 kg

(iii) Median =66 kg

4.5 Comparison of Tables, Diagrams and Graphs

Data may be presented in the form of tables as well as using diagrams and graphs. Tables can be compared with graphs and diagrams on the basis of various characteristics as follows:

(i) Table contains precise and accurate information, whereas graphs and diagrams give only an approximate idea.

(ii) More information can be presented in tables than in graphs and diagrams.

(iii) Tables require careful reading and are difficult to interpret, whereas diagrams and graphs are easily interpretable.

(iv) For common men, graphs and diagrams are attractive and more appealing than tables.

(v) Diagrams and graphs exhibit the inherent trends in the distribution easily on comparable mode than the tables.

(vi) Graphs and diagrams can be easily misinterpreted than tables.

Comparison between diagrams and graphs

(i) Diagrams can be drawn on plain papers, whereas graphs require graph papers.

(ii) Diagrams are appropriate and effective to present information about one or more variables. Normally, it is difficult to draw graphs for more than one variable in the same graph.

(iii) Graphs can be used for interpolation and/or extrapolation, but diagrams cannot be used for this purpose.

(iv) Median can be determined using graphs, but not using diagrams.

(v) Diagrams can be used for comparison of data/variables, whereas graphs can be used for determining the relationship between variables.

Points to Remember

- A diagram is a Visual aid for presenting statistical data.
- Simple bar diagram can be drawn either on horizontal or vertical base.It is used in Business and Economics.
- Pareto diagram is similar to simple bar diagram. Here the bars are arranged in the descending order of the heights of the bars. Also ,there will be a line representing the cumulative frequencies (in %) of the different categories of the variable.
- Multiple bar diagram is used for comparing two or more sets of statistical data.
- A component bar diagram is used for comparing two or more sets of statistical data, as like multiple bar diagram. But, unlike multiple bar diagram, the bars are stacked in component bar diagrams.

- Percentage bar diagram is another form of component bar diagram. Here, the heights of the components do not represent the actual values, but percentages.
- The Pie diagram is a circular diagram. As the diagram looks like a pie, it is given this name. A circle which has 360° is divided into different sectors. Angles of the sectors, subtending at the center, are proportional to the magnitudes of the frequency of the components.
- Pictograms are diagrammatic representation of statistical data using pictures of resemblance. These are very useful in attracting attention.
- Histogram is an attached bar chart or graph displaying the distribution of a frequency distribution in visual form.
- Frequency curve is a smooth and free-hand curve drawn to represent a frequency distribution.
- Cumulative frequency curve (Ogive) is drawn to represent the cumulative frequency distribution. There are two types of Ogives such as '*less than*Ogive curve' and '*more than*Ogive curve'.
- If both the curves are drawn in the same graph, then the value of x-coordinate in the point of intersection is the median.

EXERCISE 4

AX7V4

I. Choose the best answer:

1. Which one of the following diagrams displays the class frequencies at the same height?

   (a) Simple bar diagram (b) percentage bar diagram
   (c)Sub-divided bar diagram (d) multiple bar diagram

2. Which one of the following diagrams use pictures to present the data?

   (a) Pictogram (b) Pareto Diagram (c) Pie diagram (d) Histogram

3. In which one of the following diagrams, data is transformed into angles?

   (a) Pictogram (b) Pareto Diagram (c) Pie diagram (d) Histogram

4. In which one of the following diagrams, bars are arranged in the descending order of their heights?

   (a) Pictogram (b) Pareto Diagram (c) Pie diagram (d) Histogram

5. The bars are _____ in multiple bar diagrams.

   (a) Sub-divided (b) placed adjacently
   (c) placed adjacently and sub-divided (d) sub-divided and are of equal height

6. In which one of the following diagrams, the heights of the bars are proportional to the magnitude of the total frequency?

(a) Simple bar diagram (b) percentage bar diagram

(c) Sub-divided bar diagram (d) multiple bar diagram

II. Fill in the blanks:

7. _____ is useful for interpolation and extrapolation.
8. Frequency curve is a _____ curve drawn from the histogram.
9. _____ diagram uses pictures to present the data.
10. Circular diagram is known as _____
11. While drawing less than Ogive, cumulative frequencies are marked against the _____ of the respective classes.
12. In _____ curve, cumulative frequencies are marked against the lower limit of the respective classes.
13. Intersection of less than Ogive and more than Ogive gives _____
14. Frequency curve is drawn from frequency polygon by _____ the vertices.
15. Frequency polygon is drawn from ____
16. Area under the frequency polygon and the area of the histogram are ____

III. Answer shortly :

17. What is a diagram?
18. What is a graph?
19. List various types of diagrams.
20. List various types of graphs.
21. What is the use of simple bar diagram?
22. Distinguish the simple bar diagram and the Pareto diagram.
23. What are the different types of Ogives?
24. Write down the formula used for computing the angles of the components in Pie diagram?
25. How do you present the data using pictograms?
26. Distinguish multiple bar diagrams and component bar diagrams.

IV. Answer in brief:

27. Write down the method of constructing Pie diagram.

28. Write down the significance of diagrams and graphs.

29. What are the features of percentage bar diagram?

30. How will you construct histogram for a given grouped data?

31. Mention any two features of tabulated data distinguishing it from diagrams and graphs.

32. Distinguish the two types of Ogives.

V. Answer in detail:

33. What are the general rules to be followed for constructing diagrams?

34. Profit (in ₹ '000) earned by a company during the period 2011-2016 are given below. Draw simple bar diagram to this data showing profit .

Year	Profit (₹'000)
2011	15000
2012	18000
2013	20000
2014	16000
2015	13000
2016	17000

Also,

(a) find the year in which the company earned maximum profit.

(b)find the year in which the company earned minimum profit.

(c) find the difference between the profit earned in the years 2012 and 2013.

35. Draw multiple bar diagram for the following data showing the monthly expenditure (in ₹ '000)of Shop A and Shop B.

Expenditure	Amount spent by	
	Shop A	Shop B
Rent	10	18
Investment	70	90
Salary	20	35
Electricity bill	10	15
Miscellaneous expenditure	5	7

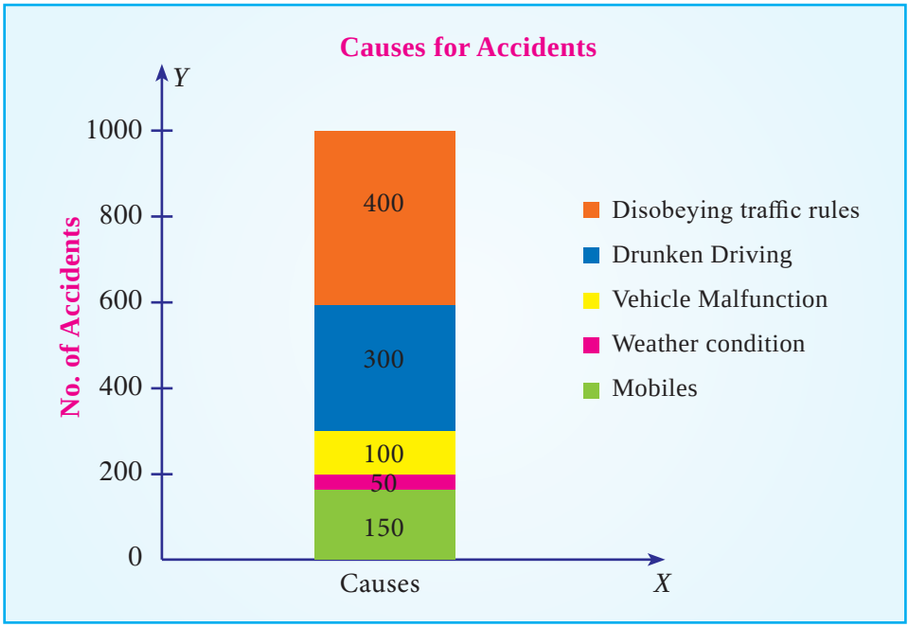
36. Draw component bar diagram to the data given in Exercise V (35).

37. Represent the data given in Exercise V(35) by percentage bar diagram.

38. The following table shows the number of students studying in three schools, preferred Walking/Cycling to go to their schools. Draw multiple bar diagram to the data.

Preference of Students	School A	School B	School C
Walking	400	550	150
Cycling	450	250	350

39. Component bar diagram showing the information about various causes and the number of road accidents occurred in a state during a period of one year are given below.



Using the diagram answer the following:

(i) ---------- caused more number of road accidents.

(ii) Minimum of road accidents is due to ----------

40. The number of students studying various undergraduate degree programmes in three Colleges are given in the following table.

Year of Study	College		
	A	B	C
First year	450	350	400

Second year	250	250	350
Third year	225	200	300

Draw sub – divided bar diagram.

41. The following table shows the details about the expenditures (in percentages) of Indian Hotel industries under various components.

Component of Expenditure	Percentage of Expenditure
Administrative	30
Salary	20
Maintenance	14
Food and Bevarages	12
Electricity	16
Savings	8

Draw pie diagram to represent the data. Also, find

(a) the angle of the sector corresponding to Salary?

(b) the difference between the angles of the sectors corresponding to Electricity and Maintenance.

42. The pictogram given below shows the number of mails received by the Directorate of Schools over a period of 6 months.

June	✉✉✉✉✉
July	✉✉✉✉
August	✉✉✉
September	✉✉✉✉✉✉
October	✉✉
November	✉✉✉✉

✉ =100 mails.

From the above pictogram find

(a) The total no. of mails received from June to November.

(b) During which month i) maximum ii) minimum no. of mails were received.

(c) What is the percentage of decrease in the receipt of mails from September toOctober?

(d) The average number of mails received per month.

43. Draw histogram for the following data.

Age	0-20	20-40	40-60	60-80	80-100
Number of persons	10	45	36	28	5

From the histogram say whether there is symmetry in the data or not.

44. A factory produces bolts. In the Quality Control test conducted on 500 bolts, the weights (grams) of the bolts were recorded. Draw frequency polygon and frequency curve to the data.

Weight(gram)	40-50	50-60	60-70	70-80	80-90
Number of bolts	30	90	130	210	40

45. Compare the uses of tables with diagrams and graphs for presenting statistical data.

46. Compare diagrams with graphs.

47. Details about Stipend (in Rs.) given to the apprentices in an organization are given below. Draw less than Ogive to the data.

Stipend in (Rs)	No. of apprentices
2000 – 3000	4
3000 – 4000	6
4000 – 5000	13
5000 – 6000	25
6000 – 7000	32
7000 – 8000	19
8000 – 9000	8
- 10000	3

(a) Estimate, from the curve, the number of apprentices whose stipend is less than Rs.5500.

(b) Find the median of the data.

48. Draw more than Ogive curve for the following data showing the marks secured by the students of Class XI in a school.

Marks	No. of students
0 – 10	2
10 – 20	4
20 – 30	9

30 – 40	10
40 – 50	8
50 – 60	5
60 – 70	3
70 - 80	2

(a) Estimate the total number of students who secured marks more than 33.

(b) Find the median of the data.

49. The lifetime (in hours) of 100 bulbs observed in a Quality Control testis given below.

Lifetime (in hours)	600-650	650-700	700-750	750-800	800-850
No. of Bulbs	6	14	40	34	6

(a) Draw less than Ogive and more than Ogive curves?

(b) Find the median lifetime of bulbs graphically.

ANSWERS

I. 1. (b) 2. (a) 3. (c) 4. (b) 5. (b) 6. (c)

II. 7. Graphs 8. Smooth and free hand 9. Pictogram

10. Pie diagram 11. Upper limit 12. More than Ogive

13. Median 14. Smoothening 15. Histogram 16. equal

V 34. (a) 2013 (b) 2015 (c) 2000000

39. (1) Disobeying traffic rules (2) Weather conditions.

41. (a) 72° (b) 8° 42. (a) 2400

(b) (i) September (ii) October (c) 66.7% (d) 400

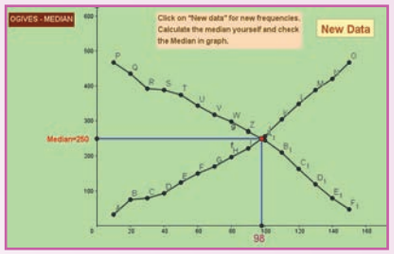
47. (a) 38 apprentices (b) 6300 48. (a) 26 students (b) 36

49. (ii) 737.5 hrs

ICT CORNER

OGIVES-MEDIAN

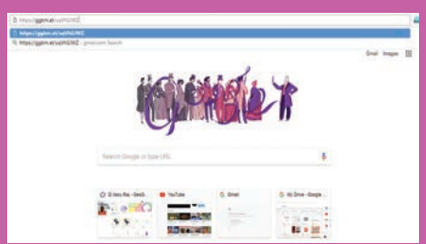
This activity is to find median of a statistical data using the Ogive curve. Also helps the students to learn how to draw Ogives.



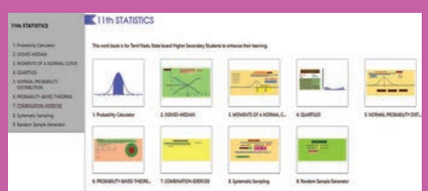
Steps:

- Open the browser and type the URL given (or) scan the QR code.
- GeoGebra work book called "**11th Standard Statistics**" will appear. In this several work sheets for statistics are given, Open the worksheet named "**Ogives-Median**"
- Ogives work sheet will open. On the right-hand side x value and frequencies are given in first two columns. 3rd and 4th columns are less than cumulative frequency and more than cumulative frequency respectively. Observe the data and compare with the curves on the left-hand side. You can press new data to get new problems as many times you want to work.
- For each problem you can see the median frequency and the median value on left side and bottom. If necessary, you can create your own problem by typing new frequencies in 2nd column.

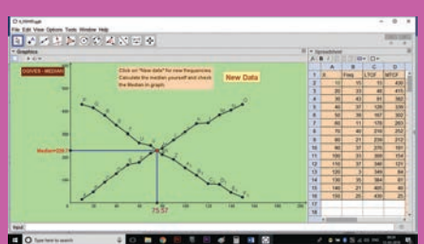
Step-1



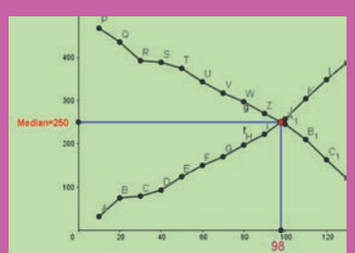
Step-2



Step-3



Step-4



Pictures are indicatives only*

URL:

https://ggbm.at/uqVhSJWZ