Unit

# 6 BASIC STATISTICS

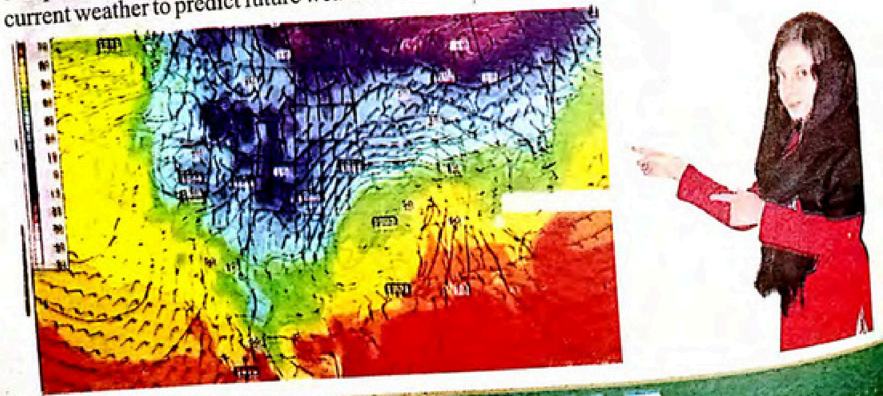
## In this unit the students will be able to

- Construct a grouped frequency table.
- Construct histograms with equal and unequal class intervals;
- Construct a frequency polygon;
- Construct a cumulative frequency table.
- Draw a cumulative frequency polygon;
- Calculate (for ungrouped and grouped data)
  - arithmetic mean by definition and using deviation from assumed mean,
  - median, mode, geometric mean, harmonic mean.
- Recognize the properties of arithmetic mean;
- Calculate weighted mean and moving averages.
- Estimate median, quartiles and mode graphically
- Measure range, variance and standard deviation.

### Why it's important

**Weather Forecasts**

Do you watch the weather forecast sometime during the day? How do you use that information? Have you ever heard the forecaster talk about weather models? These computer models are built using statistics that compare prior weather conditions with current weather to predict future weather.

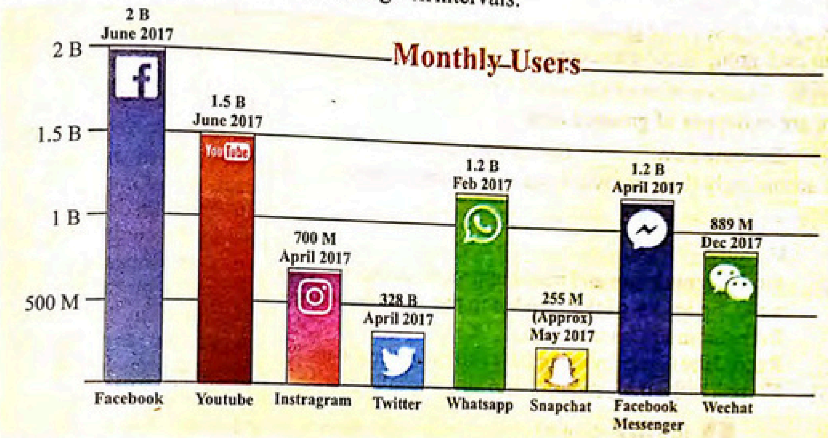








Statistics is the science of collecting, arranging, analyzing, and interpreting data to draw conclusions or answer questions. Data are pieces of information, which are often numerical. Large amounts of data can be organized in a frequency table. A frequency table shows the number of pieces of data that fall within given intervals.




### Activity

Scooters Riding is a fun to get exercise. The table shows prices for 35 types of nonmotorized scooters found at an online store.

(i). What is the cost of the least expensive scooter? the most expensive scooter?

(ii). How many scooters cost \$51 to \$75?

(iii). How could you reorganize the prices so that they are easier to find and read?

| Prices of Scooters(\$.) | | | | |
|---|---|---|---|---|
| 60 | 25 | 29 | 25 | 40 |
| 35 | 50 | 80 | 30 | 70 |
| 80 | 90 | 80 | 55 | 70 |
| 40 | 100 | 52 | 45 | 60 |
| 99 | 60 | 100 | 89 | 11 |
| 99 | 100 | 49 | 35 | 90 |
| 92 | 20 | 49 | 80 | 50 |

The scale allows you to record all of the data. It includes the least value, 11, and the greatest value, 100. The scale is 1 to 100.

| Price(\$) | Tally Marks | Frequency |
|---|---|---|
| 1 — 25 | IIII | 4 |
| 26 — 50 | 卌 卌 I | 11 |
| 51 — 75 | 卌 II | 7 |
| 76 — 100 | 卌 卌 III | 13 |

The interval separates the scale into equal parts. The interval is 25.

Tally marks are counters used to record items in a group.

From the frequency table, you can see that just over half the scooters cost between \$51 and \$100.

## 6.1 Frequency Distribution

The grouped frequency table is a statistical method to organize and simplify a large set of data into smaller "groups."

The main purpose of the grouped frequency table is to find out how often each value occurred within each group of the entire data.

### 6.1.1 Construction of Grouped Frequency Table

There are two types of grouped data.

(i) Discrete data (ii) Continuous data

Thus accordingly there are two types of frequency table.

**For your information**

The study of statistics began when an English man, John Graunt (1620-1674) collected and studied the death records in various cities of Britain and was fascinated by the patterns he found in the whole population even though people died randomly.

**(a) Construction of Discrete Frequency Table**

**Steps of construction:**

(i). Find the minimum and maximum value in the data and write the values of the variable in the variable column from minimum to maximum.

(ii). Record the values by using tally marks (vertical bars '|').

(iii). Complete the frequency column by using tally marks.

**Example 1** In a shoe store 40 customers bought shoes with following shoe size.

6, 6, 7, 6, 8, 7, 7, 8, 6, 10, 6, 8, 8, 10 ,7, 9, 7, 10, 6 ,10 ,10 ,9, 7 ,9, 6, 10, 10, 7, 11, 8, 8, 7, 6, 6 ,8 ,9, 7, 8, 7 ,9. Construct a frequency table of the data.

**Solution** Let $x$ = shoe size.

| $x$ | Tally Marks | Frequencies ($f$) |
|---|---|---|
| 6 | 卌 \|\|\|\| | 9 |
| 7 | 卌 卌 | 10 |
| 8 | 卌 \|\|\| | 8 |
| 9 | \|\|\|\| | 4 |
| 10 | 卌 \|\|\| | 8 |
| 11 | \| | 1 |

**Construction of Continuous Frequency Table**

**Steps of construction:**

(i). Find Range: Deduct lowest value from the highest value.

(ii). Determine the number of groups (k). Most of the data has between 5 to 15 groups. It is your decision to choose the number of groups for your data. Number of classes must be balanced that is not too large or not too small.

(iii). Determine the width (number of values per group) of group interval by dividing Range by k.

(iv). Create three columns titled "Groups." "Tally Marks." And "Frequency."

(v). Insert the data in the table .

(vi). Determine the frequencies for all the groupsby tallying the data.





**Related concepts**

**Class Limits**

Every class has two limits, the **upper-class limit** and the lower **class limit**. For each, class the two limits may be fixed such that the mid-point of each class falls on an integer rather than a fraction.

**Note** Choose the first class interval such that the smallest observation is included.

The mid-point of each class is calculated by the formula.

$$\text{Mid-point} = \frac{\text{Lower limit} + \text{Upper limit}}{2}$$

For example if in a data the lowest value is 3 and the highest value is 20, then the better way is to make the classes as 3 — 7, 8 — 12, 13 —17, 18 — 22, because the mid-point of each class is an integer. The mid-points usually denoted by ($x_i$) are called the **class marks**.

**Class Width**

The class width is the difference between consecutive lower-class limits. It is found by dividing the range by the number of groups formed.

**Class Boundaries**

For computing classboundaries, the upper limit of the preceding class is subtracted from the lower limit of the following class and the difference is then divided by '2'. The quantity which is obtained is then subtracted from the lower limit and added to the upper limit of each class.

For example, consider the following table.

| Classes-Limits | Class-Boundaries |
|---|---|
| 1 — 4 | 0.5 — 4.5 |
| 5 — 8 | 4.5 — 8.5 |
| 9 — 12 | 8.5 — 12.5 |
| 13 — 16 | 12.5 — 16.5 |
| 17 — 20 | 16.5 — 20.5 |

Formula for calculating class-boundaries (CB)

$$\frac{5-4}{2} = \frac{1}{2} = 0.5$$

So 0.5 is subtracted from the lower limit and added to the upper limit of each class.

**Example 2** The heights of 30 students of 10th class in cm are as follows. Construct group frequency table.

| Heights in cm |
|---|
| 162, 165, 170, 170, 162, 159, 162, 163, 175, 166, 171, 174, 155, 160, 173, |
| 140, 145, 140, 146, 150, 172, 158, 155, 163, 165, 171, 153, 158, 149, 153 |

**Solution**

The width of each group is 5 because $\frac{175-140}{7} = 5$. The width of class boundary is 5.

| Groups | Class boundaries | Heights in cm | Frequencies ($f$) |
|---|---|---|---|
| 139 – 144 | 138.5 – 144.5 | 140, 140 | 2 |
| 145 – 150 | 144.5 – 150.5 | 146, 150, 149, 145 | 4 |
| 151 – 156 | 150.5 – 156.5 | 155, 155, 153, 153, | 4 |
| 157 – 162 | 156.5 – 162.5 | 162, 162, 159, 162, 160, 158,158 | 7 |
| 163 – 168 | 162.5 – 168.5 | 165, 163,165, 163, 166 | 5 |
| 169 – 174 | 168.5 – 174.5 | 170, 170, 171, 174,173, 172, 171 | 7 |
| 175 – 180 | 174.5 – 180.5 | 175 | 1 |

**Example 3** Construct a frequency table of the weights (kg) of 30 students from the following data by using 5 as a class-interval. Find the class-boundaries and class-marks also. 25, 30, 40, 21, 24, 25, 36, 30, 45, 50, 22, 25, 36, 46, 35, 38, 40, 28, 34, 45, 42, 46, 38, 48, 28, 29, 31, 33, 30, 26.

**Solution**

| Class-Limits | Tally Marks | Frequency | Class-Boundaries | Class Marks |
|---|---|---|---|---|
| 21 – 25 | ~~IIII~~ I | 6 | 20.5 — 25.5 | 23 |
| 26 – 30 | ~~IIII~~ II | 7 | 25.5 — 30.5 | 28 |
| 31 – 35 | IIII | 4 | 30.5 — 35.5 | 33 |
| 36 – 40 | ~~IIII~~ I | 6 | 35.5 — 40.5 | 38 |
| 41 – 45 | III | 3 | 40.5 — 45.5 | 43 |
| 46 – 50 | IIII | 4 | 45.5 — 50.5 | 48 |






## 6.1.2 Histogram

A histogram is a vertical bar graph with no space between the bars. The area of each bar is proportional to the frequency it represents.

A histogram has advantages over the other methods that it can be used to represent data with both equal and unequal class intervals.

Whenever we construct a histogram representing a set of figures which have been rounded off from the original measurements, we have to make the class intervals continuous by using class boundaries. This is to avoid having gaps in between the bars.

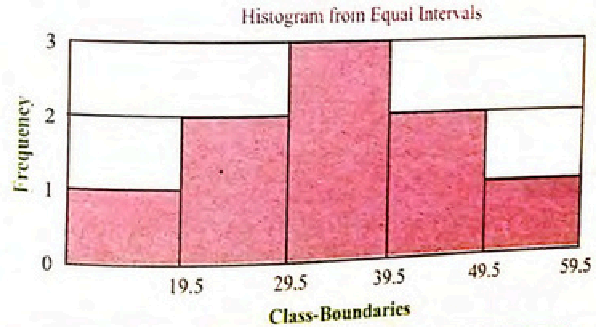
**Example 4** Construct a histogram from the following frequency distribution.

| Class-limits | 20 — 29 | 30 — 39 | 40 — 49 | 50 — 59 | 60 — 69 |
|---|---|---|---|---|---|
| Frequency | 1 | 2 | 3 | 2 | 1 |

**Solution**

To draw a histogram class-boundaries are marked along x-axis and frequencies of each class are marked along Y-axis as shown in the figure.

| Class-Limits | Frequency | Class-Boundaries |
|---|---|---|
| 20 — 29 | 1 | 19.5 — 29.5 |
| 30 — 39 | 2 | 29.5 — 39.5 |
| 40 — 49 | 3 | 39.5 — 49.5 |
| 50 — 59 | 2 | 49.5 — 59.5 |
| 60 — 69 | 1 | 59.5 — 69.9 |



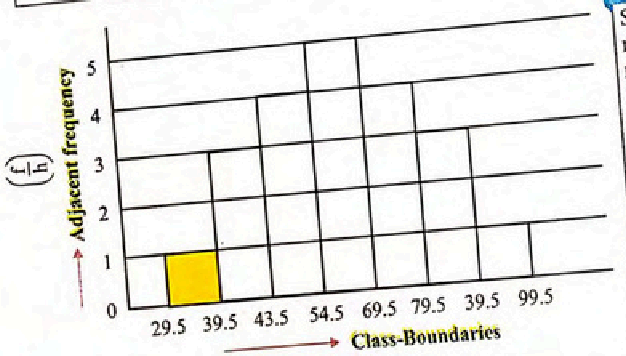
**Example 5** Draw a histogram for the following data.

| Class-Limits | 30 — 39 | 40 — 43 | 44 — 54 | 55 — 69 | 70 — 79 | 80 — 89 | 90 — 99 |
|---|---|---|---|---|---|---|---|
| Frequency | 10 | 12 | 44 | 75 | 40 | 30 | 10 |

**Solution** Histogram with unequal class intervals.

The class intervals are not equal. In constructing the histogram, we must ensure that the areas of the rectangle are proportional to the class frequencies, as the frequency in a histogram is represented by the area of each rectangle.

| Class-Limits | Class-Boundaries | Class-Intervals ($h$) | Frequency ($f$) | Adjusted Frequency $\frac{f}{h}$ |
|---|---|---|---|---|
| 30 — 39 | 29.5 — 39.5 | 10 | 10 | 1 |
| 40 — 43 | 39.5 — 43.5 | 4 | 12 | 3 |
| 44 — 54 | 43.5 — 54.5 | 11 | 44 | 4 |
| 55 — 69 | 54.5 — 69.5 | 15 | 75 | 5 |
| 70 — 79 | 69.5 — 79.5 | 10 | 40 | 4 |
| 80 — 89 | 79.5 — 89.5 | 10 | 30 | 3 |
| 90 — 99 | 89.5 — 99.5 | 10 | 10 | 1 |




**Note**

Since the areas of the rectangles in a histogram must be proportional to the class frequencies, height of rectangle $x$ class width = class frequency.

OR

height of rectangle = class frequency/class width Thus, the height of a rectangle may be found by dividing the class frequency by the class width.

**For your information**

In a histogram, the frequency is represented by (height of bar) $x$ (number of standard intervals)






## 6.1.3 Frequency polygon

A frequency polygon is drawn by joining all the midpoints at the top of each rectangle. The midpoints at both ends are joined to the horizontal axis to accommodate the end points of the polygon. This will make the graph nearer with the end points falling off to zero on the horizontal axis.

We can draw the frequency polygon of a distribution without first drawing the histogram. We plot each class frequency against the mid value of the class to obtain points which we join by straight lines to form the frequency polygon.

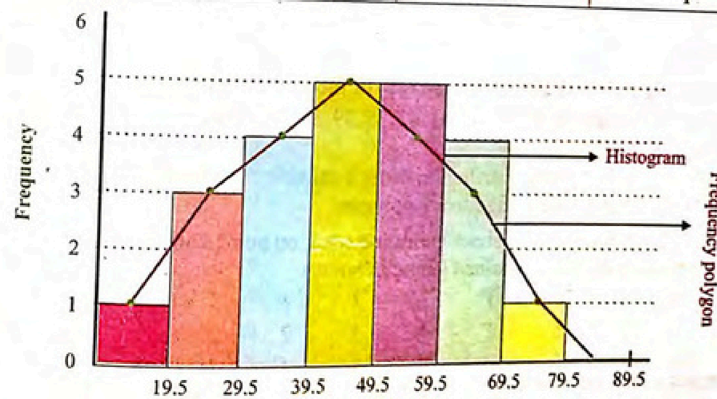
Frequency polygon are specially useful to compare two sets of data e.g. prices of commodities (vegetables, pulses, gold etc) over a given period of time. Now a days it is commonly used in cricket to show the progress of a batting team or comparision of scores of chasing a team.

**Example 6** Construct a frequency polygon for the following frequency distribution.

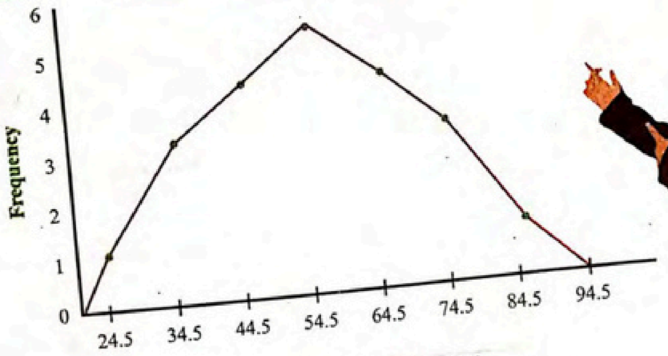
| Class-Limits | 20 — 29 | 30 — 39 | 40 — 49 | 50 — 59 | 60 — 69 | 70 — 79 | 80 — 89 |
|---|---|---|---|---|---|---|---|
| Frequency | 1 | 3 | 4 | 5 | 4 | 2 | 1 |

**Solution**

| Classes | Class-Boundaries | Class Marks (Mid-Points) | Frequency |
|---|---|---|---|
| 20 — 29 | 19.5 — 29.5 | 24.5 | 1 |
| 30 — 39 | 29.5 — 39.5 | 34.5 | 3 |
| 40 — 49 | 39.5 — 49.5 | 44.5 | 4 |
| 50 — 59 | 49.5 — 59.5 | 54.5 | 5 |
| 60 — 69 | 59.5 — 69.5 | 64.5 | 4 |
| 70 — 79 | 69.5 — 79.5 | 74.5 | 2 |
| 80 — 89 | 79.5 — 89.5 | 84.5 | 1 |




Frequency polygon




## Exercise 6.1

1. Construct a frequency distribution of the marks of 30 students during a quiz with 100 points by taking 10 as the class-interval. Indicate the class-boundaries and class-marks.
40, 60, 65, 70, 35, 50, 56, 74, 72, 49, 85, 76, 82, 83, 68, 90, 67, 66, 58, 46, 74, 88, 76, 69, 57, 63, 66, 47, 82, 90.

2. Following are the mistakes made by a group of students of class 10th in a test of essay writing. Using an appropriate size of class interval, make a frequency distribution and also indicate the number of class intervals.
4, 7, 12, 9, 21, 16, 3, 19, 17, 24, 14, 15, 8, 13, 11, 16, 15, 6, 5, 8, 11, 20, 18, 22, 6.

3. Draw a Histogram for the following data.

| Class-Limit | 20 — 24 | 25 — 29 | 30 — 34 | 35 — 39 | 40 — 44 | 45 — 49 | 50 — 54 |
|---|---|---|---|---|---|---|---|
| Frequency | 1 | 3 | 4 | 5 | 4 | 2 | 1 |

4. The following data give the weights in (kg) of the students in the 10th class.
25, 30, 32, 29, 24, 40, 36, 37, 28, 27, 41, 42, 35, 39, 31, 32, 34, 42, 40, 43, 36, 26, 22, 23, 42, 39, 35, 41, 39, 29.
   (i). Prepare a frequency distribution using a suitable class interval.
   (ii). Draw histogram and frequency polygon.

5. A teacher asked the students about their time spent on homework completion time. The following set of data was obtained (Time in hours).

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 4 | 4 | 6 | 3 | 1 | 2 | 2 | 3 | 1 | 4 |
| 1 | 2 | 5 | 3 | 4 | 5 | 2 | 2 | 3 | 1 |
| 3 | 1 | 2 | 2 | 3 | 1 | 4 | 2 | 6 | 2 |

Construct a frequency table and draw a histogram showing the results.






**Activity**

"Getting Ready for School in the morning"

Gather your Data:

Ask your school-mates the question, how much time (minutes) do you take in getting ready for your school? Write the data in a notebook like this Ali Haidar 20 minutes. Collect responses from approximately 20 students in the form of minutes only. Like 30 min, 10 min, 12 min, 15 min. and so on. Make a grouped frequency table for your data.


## 6.2 Cumulative Frequency Distribution

A cumulative frequency table provides information about the sum of a variable against the other value. For example, how many workers of a factory have salary less than Rs.7000, the number of students scoring below 50% marks, number of maiz cob measuring more than 6 inches etc.

When the same data is presented on a graph paper the freehand curve formed is called an Ogive.

### 6.2.1 Construct a cumulative frequency table

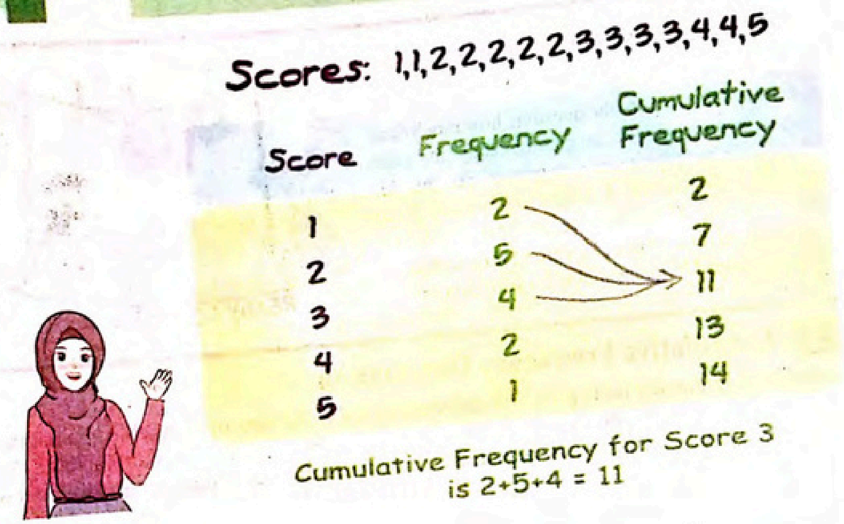
The frequency of the last class is the total frequency of a set of data.

**Example 7** Find the cumulative frequency of the following data.

| $x$ | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|
| $f$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 4 | 3 | 8 |

**Solution**

| $x$ | $f$ | Method of finding (c.f) | c.f |
|---|---|---|---|
| 3 | 1 | 1 | 1 |
| 4 | 2 | $1+2=3$ | 3 |
| 5 | 3 | $3+3=6$ | 6 |
| 6 | 4 | $6+4=10$ | 10 |
| 7 | 5 | $10+5=15$ | 15 |
| 8 | 6 | $15+6=21$ | 21 |
| 9 | 7 | $21+7=28$ | 28 |
| 10 | 4 | $28+4=32$ | 32 |
| 11 | 3 | $32+3=35$ | 35 |
| 12 | 8 | $35+8=43$ | 43 |




**Example 8**

The consumption of petrol of 1000cc cars of a particular brand was surveyed and the following data was obtained. Construct a cumulative frequency distribution.

| Distance in Km/ Litre (Milage) | 10 — 12 | 13 — 15 | 16 — 18 | 19 — 21 | 22 — 24 |
|---|---|---|---|---|---|
| Frquencey | 16 | 20 | 36 | 21 | 7 |

**Solution** The cumulative frequency disturbution is consturcted below.

| Milage (Km/Litre) | Upper Class Boundaries | Frequency | Cumulative Frequency |
|---|---|---|---|
| 10 — 12 | 12.5 | 16 | 16 |
| 13 — 15 | 15.5 | 20 | 16 + 20 = 36 |
| 16 — 18 | 18.5 | 36 | 36 + 36 = 72 |
| 19 — 21 | 21.5 | 21 | 72 + 21 = 93 |
| 22 — 24 | 24.5 | 7 | 93 + 7 = 100 |

### 6.2.2 Cumulative Frequncy Polygon

A polygon in which cumulative frequencies are used for plotting the curve is called a cumulative frequency polygon. The curve is also called a Ogive. The construction of an Ogive is described in the following example.






**Example 9** In the following data marks of students are given during first pre-boa exam in the subject of maths.

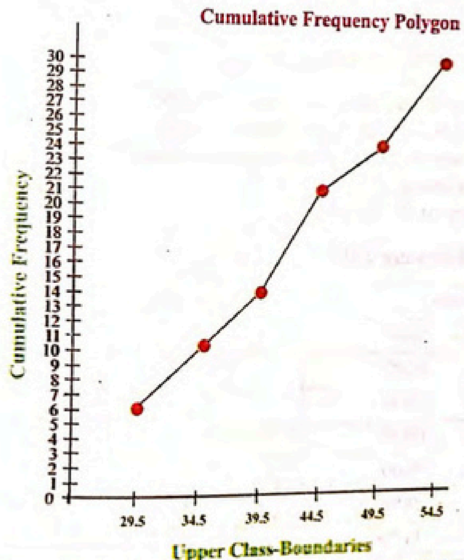
25, 30, 27, 28, 35, 36, 40, 41, 42, 45, 50, 44, 29, 26, 36, 31, 43, 46, 52, 53, 51, 42, 37, 27, 3. 46, 44, 34, 51, 54

By taking a suitable class-interval, prepare a frequency distribution, find less than cumulativ frequency and draw a cumulative frequency polygon (Ogive).

**Solution**

To construct the cumulative frequency distribution, we take the class-interval as 5.

| Class-limits | Frequency | Class-boundaries | Cumulative Frequency |
|---|---|---|---|
| 25 — 29 | 6 | 24.5 — 29.5 | 6 |
| 30 — 34 | 4 | 29.5 — 34.5 | 10 |
| 35 — 39 | 4 | 34.5 — 39.5 | 14 |
| 40 — 44 | 7 | 39.5 — 44.5 | 21 |
| 45 — 49 | 3 | 44.5 — 49.5 | 24 |
| 50 — 54 | 6 | 49.5 — 54.5 | 30 |




## Exercise 6.2

1. The following data give the wages (in Rs) of workers.
60, 75, 80, 85, 90, 84, 70, 73, 76, 84, 95, 100, 150, 66, 58, 90, 98, 120, 77, 90.
By taking 10 as the class-interval, prepare.
(i). Cumulative frequency distribution (ii). Cumulative frequency polygon

2. Make cumulative frequency table for the following data

| Age in Years | 20–24 | 25–29 | 30–39 | 40–44 | 45–49 | 50–54 | 55–59 |
|---|---|---|---|---|---|---|---|
| Number of persons ($f$) | 1 | 2 | 26 | 22 | 20 | 15 | 14 |

3. In a city during the first week of August rainfall recorded is as fallows. Construct a cumulative frequency graph.

| Day | Sun | Mon | Tue | Wed | Thu | Fri | Sat |
|---|---|---|---|---|---|---|---|
| Rainfall in ml | 70 | 40 | 30 | 35 | 50 | 55 | 80 |

4. Draw less than and more than cumulative frequency polygon for the data given.

| Marks | Number of Students |
|---|---|
| 40 — 49 | 1 |
| 50 — 59 | 2 |
| 60 — 69 | 3 |
| 70 — 79 | 4 |
| 80 — 89 | 5 |
| 90 — 99 | 6 |

5. Determine from the data of Q4, the following.
(i). Number of students who obtained more than 50 marks.
(ii). Number of students who obtained less than 70 marks.
(iii). Number of students who secured marks between 50 and 70.
(iv). Class interval of all classes.
(v). Lower-class boundary of $5^{th}$ class.

6. Construct an ogive for the following table

| Salary | Workers |
|---|---|
| 4000 — 5000 | 3 |
| 5001 — 6000 | 5 |
| 6001 — 7000 | 12 |
| 7001 — 8000 | 9 |
| 8001 — 9000 | 5 |
| 9001 — 10000 | 4 |
| 10001 — 11000 | 2 |



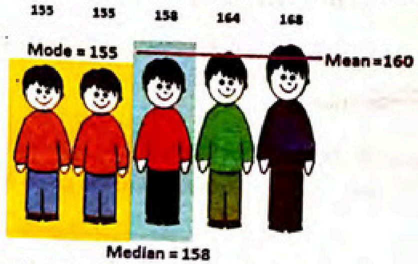



## 6.3 Measures of Central Tendency

Central Tendency of a data is the value that represents the whole data or the stage at which the largest number of item tends to concentrate and so it is called central tendency. Central Tendency or Averages are also sometimes called measures of location, because they locate the centre of a distribution.

### Types of Central Tendency Averages

(i). Arithmetic Mean (A.M)
(ii). Median
(iii). Mode
(iv). Geometric Mean (G.M)
(v). Harmonic Mean (H.M)
(vi). Quartiles




### 6.3.1 Calculate for ungrouped and grouped data

**(i) Arithmetic Mean**

**(a) Arithmetic Mean (for ungrouped data)**

Arithmetic Mean or simply Mean is calculated by adding all values of the data divided by the number of items (values). If $x_1, x_2, x_3, \ldots, x_n$ are $n$ values; then the Arithmetic Mean denoted by $(\bar{x})$ is given as

$$\bar{x} = \frac{x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + \ldots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n}$$

$\bar{x}$ = Arithmetic Mean, $\sum$ = sigma (Notation used for the summation)

$x_i = x_1, x_2, x_3, \ldots x_n (i = 1, 2, 3 \ldots n)$

$n$ = Total number of items in the data.

By short-cut method

Formula for short-cut method is,

$$\text{Mean} = \frac{\text{Sum of data}}{\text{Number of data}}$$

$$\bar{x} = a + \frac{\sum D_x}{n}$$

Where

$\bar{x}$ = Arithmetic Mean

$a$ = Provisional Mean (P.M)

$D_x = (x-a)$ (Deviation from Provisional Mean)

$\sum D_x$ = Sum of Deviations from P.M

$n$ = Total number of values in the data.

**Example 10**

Find the A.M of the values 2, 3, 4, 5, 6, 7, 8, 9, 10.

(i) By direct method

(ii) By short cut method

**Solution**

(i) By direct method

Let $\bar{x}$ be the A.M of the given items. Then by using formula

$$\bar{x} = \frac{x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + \dots + x_n}{n}$$

$$= \frac{2+3+4+5+6+7+8+9+10}{9} = \frac{54}{9} = 6$$

$\therefore \quad \bar{x} = 6$

(ii) By short cut method

| $x$ | $D_x\ (x-a),\ a=6$ |
|---|---|
| 2 | $2-6=-4$ |
| 3 | $3-6=-3$ |
| 4 | $4-6=-2$ |
| 5 | $5-6=-1$ |
| 6 | $6-6=0$ |
| 7 | $7-6=1$ |
| 8 | $8-6=2$ |
| 9 | $9-6=3$ |
| 10 | $10-6=4$ |
| | $\sum D_x = 0$ |

By using formula,

$$\bar{x} = a + \frac{\sum D_x}{n} = 6 + \frac{0}{9} = 6+0$$

$\therefore \quad \bar{x} = 6$





**(b) Arithmetic Mean (for grouped data)**

If $x_1, x_2, x_3, \dots, x_n$ are $n$ values with $f_1, f_2, f_3, \dots f_n$ as frequencies of the respective values, hence the Arithmetic Mean $\bar{x}$ is given as

$$\bar{x} = \frac{f_1x_1 + f_2x_2 + \dots + f_nx_n}{f_1 + f_2 + f_3 + \dots + f_n} = \frac{\sum fx}{\sum f}$$

**Example 11**

In a coaching class of 13 students, a test was conducted, and the following marks were obtained by the students.

10, 12, 12, 14, 9, 18, 9, 13, 16, 9, 17, 16, 14

Make frequency table and find Arithmetic mean.

**Solution**

| Score | 9 | 10 | 12 | 13 | 14 | 16 | 17 | 18 |
|---|---|---|---|---|---|---|---|---|
| Frequency ($f$) | 3 | 1 | 2 | 1 | 2 | 2 | 1 | 1 |

| ($x$) Score | ($f$) Frequency | $fx$ |
|---|---|---|
| 9 | 3 | 27 |
| 10 | 1 | 10 |
| 12 | 2 | 24 |
| 13 | 1 | 13 |
| 14 | 2 | 28 |
| 16 | 2 | 32 |
| 17 | 1 | 17 |
| 18 | 1 | 18 |
| | $\sum f = 13$ | $\sum fx = 169$ |

$$\text{Arithmetic Mean} = \frac{\sum fx}{\sum f}$$

$$= \frac{169}{13}$$

$$= 13$$

**Example 12**

The prices of 2 KW generators are given below along with their frequencies. Find the mean price.

(i) By direct method

(ii) By short-cut method

| Price (In 100s) | 90–94 | 95–99 | 100–104 | 105–109 | 110–114 | 115–119 | 120–124 |
|---|---|---|---|---|---|---|---|
| Frequency (f) | 4 | 11 | 15 | 24 | 18 | 9 | 3 |

**Solution**

(i) By direct method

| Class Interval | (x) Mid Value | (f) Frequency | fx |
|---|---|---|---|
| 90 — 94 | 92 | 4 | 368 |
| 95 — 99 | 97 | 11 | 1067 |
| 100 — 104 | 102 | 15 | 1530 |
| 105 — 109 | 107 | 24 | 2568 |
| 110 — 114 | 112 | 19 | 2016 |
| 115 —119 | 117 | 9 | 1053 |
| 120 — 124 | 122 | 3 | 366 |
| | | $\sum f = 85$ | $\sum fx = 8968$ |

The mean price of the generators $= \bar{x}$

$$= \frac{\sum fx}{\sum f}$$

$$= \frac{8968}{85}$$

$= 105.5$ hundred rupees

$= 10550/-$







(ii) Short–cut method

Let the assumed mean a = 107.

| Class Interval | (x) Mid -point | (f) Frequency | $D_x = x - a$ | $fD_x$ |
|---|---|---|---|---|
| 90 — 94 | 92 | 4 | 92 – 108 = –15 | –60 |
| 95 — 99 | 97 | 11 | 97 – 107 = –10 | –110 |
| 100 — 104 | 102 | 15 | 102 – 107 = –5 | – 75 |
| 105 — 109 | 107 | 24 | 107 – 107 = 0 | 0 |
| 110 — 114 | 112 | 19 | 112 – 107 = 5 | 95 |
| 115 — 119 | 117 | 9 | 117 – 107 = 10 | 90 |
| 120 — 124 | 122 | 3 | 122 – 107 = 15 | 45 |
| | | $\sum f = 85$ | | $\sum fD_x = -15$ |

$$\bar{x} = a + \frac{\sum fD_x}{\sum f}$$

$$= 107 + \left(-\frac{15}{85}\right)$$

$$= \frac{9095 - 15}{85}$$

$$= \frac{9080}{85}$$

$= 106.83$

$= 10683/-$

**(ii) Median**

It is a value which is in the center of observations when all the observations are arranged in ascending or descending order i.e. Median divide the data in two equal parts. Median is calculated for the following data:

(a) ungrouped data (b) discrete data (c) continuous data

**(a) Median for ungrouped data**

If in a data there are $n$ values, then after arranging the values in ascending or descending order Median is calculated by the following rules:

$$\text{Median} = \text{Size of } \left(\frac{n+1}{2}\right)\text{th item} \quad (\text{if } n \text{ is odd})$$

or

$$\text{Median} = \frac{1}{2}\left[\text{Size of } \frac{n}{2}\text{th} + \text{Size of } \left(\frac{n+2}{2}\right)\text{th item}\right] (\text{if } n \text{ is even})$$

**Example 13** Find the median for the following values.

2, 4, 5, 6, 3

**Solution** Writing the given data in increasing order

2, 3, 4, 5, 6

| S.No | x (value) |
|---|---|
| 1 | 2 |
| 2 | 3 |
| 3 | 4 |
| 4 | 5 |
| 5 | 6 |

By using formula

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

$$= \text{size of } \left(\frac{5+1}{2}\right)\text{th item}$$

$$= \text{size of } 3^{rd} \text{ item}$$

$$= 4$$

So Median = 4

**What's the error?**

0, 5, 7, 9, 10, 3, 5, 8

Median = 9.5

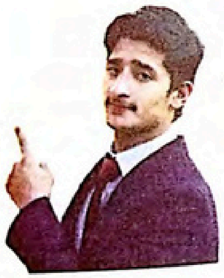
**Note**

The median for an odd number of data is the middle value when the data are arranged in ascending /descending order. The median for an even number of data is arranged is the mean of two middle values when the data are arranged in ascending/descending order.

**Example 14** The following is the daily pocket money in rupees for the children of a family 10, 20, 15, 30. Calculate the median for the data.

**Solution**

| S.No | x |
|---|---|
| 1 | 10 |
| 2 | 15 |
| 3 | 20 |
| 4 | 30 |



Since number of items is even, so

$$\text{Median} = \frac{1}{2}\left[\text{Size of } \frac{n}{2}\text{th} + \text{Size of } \left(\frac{n+2}{2}\right)\text{th}\right]\text{item}$$

$$= \frac{1}{2}\left[\text{Size of } \frac{4}{2}\text{th} + \text{Size of } \left(\frac{6}{2}\right)\text{th}\right]\text{item}$$

$$= \frac{1}{2}[\text{Size of 2nd} + \text{Size of 3rd}]\text{ item} = \frac{1}{2}(15+20) = \frac{35}{2} = 17.5$$







**(b) Median for discrete data**

In a frequency distribution if $\sum f$ is odd then the value of $x$ opposite to $\left(\frac{\sum f+1}{2}\right)$th item in the cumulative frequency column is its median. But if $\sum f$ is even, then the value of $x$ opposite to $\frac{\sum f}{2}$ th item in the cumulative frequency column is its median. This can be illustrated with the help of the following examples.

**Example 15** The following are the marks obtained by 35 students in a test.

| x | 10 | 12 | 15 | 20 | 25 | 30 |
|---|---|---|---|---|---|---|
| f | 1 | 10 | 5 | 13 | 2 | 4 |

Find median of the data.

**Solution**

| x | f | c.f |
|---|---|---|
| 10 | 1 | 1 |
| 12 | 10 | 11 |
| 15 | 5 | 16 |
| 20 | 13 | 29 |
| 25 | 2 | 31 |
| 30 | 4 | 35 |

Since $n = \sum f = 35$

So $\text{Median} = \text{Size of } \left(\frac{n+1}{2}\right)\text{th item}$

$$= \text{Size of } \left(\frac{35+1}{2}\right)\text{th item}$$

$$= \text{Size of 18th item}$$

$$= 20$$

**Example 16** Find the median marks from the following distribution:

| (Marks) $x$ | 10 | 20 | 22 | 25 |
|---|---|---|---|---|
| Number of students | 0 | 2 | 4 | 6 |

**Solution**

| Marks $x$ | Frequency | c.f |
|---|---|---|
| 10 | 0 | 0 |
| 20 | 2 | 2 |
| 22 | 4 | 6 |
| 25 | 6 | 12 |
| | $\sum f = 12$ | |

$$\text{Median} = \text{Size of} \left(\frac{n}{2}\right)\text{th item}$$

$$= \text{Size of} \left(\frac{12}{2}\right)\text{th item}$$

$$= \text{Size of 6th item}$$

$$= 22$$

So, Median = 22

### (c) Median from continuous data

For computing median in continuous data, it is important to make class boundaries and then median is calculated by the following formula:

$$\text{Median} = l + \frac{h}{f}\left(\frac{n}{2} - c\right) \text{ where,}$$

$l$ = lower limit of the median class
$h$ = width (class-interval of the median class)
$f$ = frequency of the median class
$\frac{n}{2} = \frac{\sum f}{2}$
$c$ = cumulative frequency of the class preceding the median class.






**Example 17** Find the median of the following distribution.

| Daily wages (in Rs.) | 60—69 | 70—79 | 80—89 | 90—99 | 100—109 |
|---|---|---|---|---|---|
| Labour | 4 | 6 | 8 | 10 | 5 |

**Solution**

| Daily wages | $f$ | (C.B) Class-Boundaries | (C.F) Cumulative Frequency |
|---|---|---|---|
| 60 — 69 | 4 | 59.5 — 69.5 | 4 |
| 70 — 79 | 6 | 69.5 — 79.5 | 10 |
| 80 — 89 | 8 | 79.5 — 89.5 | 18 |
| 90 — 99 | 10 | 89.5 — 99.5 | 28 |
| 100 — 109 | 5 | 99.5 — 109.5 | 33 |
| | $\sum f = 33$ | | |

$$\text{Median} = \frac{n}{2}\text{th item} = \frac{33}{2}\text{th item}$$

$$= 16.5\text{th item}$$

Median lies in the group 79.5 — 89.5

$$\text{Median} = l + \frac{h}{f}\left(\frac{n}{2} - c\right)$$

$$= 79.5 + \frac{10}{8}(16.5 - 10)$$

$$= 79.5 + \frac{10}{8}(6.5)$$

$$= 79.5 + 8.125$$

$$= 87.625$$

### (iii) Mode

The value that appears more times in a data, is called **mode** for the given data or the most frequent value in a data is called **mode**.

**Example 18** From the following sizes of kids trousers, find the model size.

25, 30, 31, 25, 35, 25

**Solution** In the above data, 25 is the most frequent value, so the model size is 25.

**Example 19** The following data shows the weights of the students. Find the model weight.

| Weights (in KG) | 40 | 42 | 50 | 51 | 55 |
|---|---|---|---|---|---|
| Number of students | 10 | 8 | 3 | 2 | 1 |

**Solution** Mode = 40, because 40 is the most frequent value in the data.

**Example 20** Calculate the mode from the following frequency distribution.

| Marks (C.B) | 0 — 4 | 4 — 8 | 8 — 12 | 12 — 16 | 16 — 20 |
|---|---|---|---|---|---|
| No. of students | 3 | 5 | 4 | 6 | 2 |

**Solution**

| *Marks (C.B)* | *No. of students f* |
|---|---|
| 0 — 4 | 3 |
| 4 — 8 | 5 |
| 8 — 12 | 4 |
| 12 — 16 | 6 |
| 16 — 20 | 2 |

→ Modal group

The mode lies in the group 12 — 16.
So by using the formula

$$\text{Mode} = l + \left(\frac{f_m - f_0}{2f_m - f_0 - f_1}\right) \times h$$

Where
$l$ = lower limit of the modal group
$f_m$ = Frequency of the modal group
$f_0$ = Frequency of the group preceding the modal group
$f_1$ = Frequency of the group following the modal group
$h$ = class-interval

So, $\text{Mode} = 12 + \left(\frac{6-4}{2(6)-4-2}\right) \times 4$

$= 12 + \left(\frac{2}{12-6}\right) \times 4$

$= 12 + \frac{8}{6}$

$= 12 + 1.33 = 13.33$







**(iv) Geometric Mean (G.M)**
Geometric Mean is the nth positive root of the product of $n$ values.

**(a) Geometric Mean from ungrouped data**
If $x_1, x_2, x_3, \ldots x_n$ are $n$ values in a data, then the Geometric Mean is given as

$$\text{G.M} = (x_1 \times x_2 \times x_3 \times \ldots \times x_n)^{\frac{1}{n}}$$

or $$\text{G.M} = \text{Anti-log}\left(\frac{\sum \log x}{n}\right)$$

**Example 21** Find the Geometric Mean (G.M) of the marks obtained by the 9th class students.
60, 65, 70, 80, 85, 90, 75.

**Solution**

| $x$ | Log $x$ |
|---|---|
| 60 | log 60 = 1.7781 |
| 65 | log 65 = 1.8129 |
| 70 | log 70 = 1.8450 |
| 75 | log 75 = 1.8750 |
| 80 | log 80 = 1.9030 |
| 85 | log 85 = 1.9294 |
| 90 | log 90 = 1.9542 |
| | $\sum \log x = 13.0976$ |

$\text{G.M} = \text{Anti-log} \frac{\sum \log x}{n}$

$= \text{Anti-log} \frac{13.0976}{7}$

$= \text{Anti-log } 1.8710$

$= 74.31$

**(b) Geometric Mean of Grouped data**

Let $x_1, x_2, x_3, \ldots, x_n$ be the mid-points in a frequency distribution and $f_1, f_2, f_3, \ldots f_n$ are their respective frequencies. Then the G.M is calculated by the following formula:

$$\text{G.M} = \text{Anti-log}\left(\frac{\sum f \log x}{\sum f}\right)$$

### Example 22

Calculate the Geometric Mean (G.M) of the following data.

| Marks | 0 — 20 | 20 — 40 | 40 — 60 | 60 — 80 |
|---|---|---|---|---|
| Number of students | 3 | 4 | 10 | 11 |

**Solution**

| Marks | $x$ | $f$ | $\log x$ | $f \log x$ |
|---|---|---|---|---|
| 0 — 20 | 10 | 3 | 1.0000 | 3.0000 |
| 20 — 40 | 30 | 4 | 1.4771 | 5.9084 |
| 40 — 60 | 50 | 10 | 1.6989 | 16.989 |
| 60 — 80 | 70 | 11 | 1.8450 | 20.295 |
| | | $\sum f = 28$ | | $\sum f \log x = 46.1924$ |

By formula,

$$G.M = \text{Anti-log} \frac{\sum f \log x}{\sum f}$$

$$= \text{Anti-log} \frac{46.1924}{28}$$

$$= \text{Anti-log } 1.6497$$

$$= 44.64$$

**Did You Know?**

When you multiply 1089 by 9, the answer you get is 9801, which is the exact reverse of the number 1089





### (v) Harmonic Mean (H.M)

It is the reciprocal of the Arithmetic Mean of the reciprocal values.

#### (a) Harmonic Mean from ungrouped data

Let there be '$n$' values i.e., $x_1, x_2, x_3, \ldots, x_n$ in a data, then the Harmonic Mean (H.M) is calculated by the following formula:

$$H.M = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} +, \ldots + \frac{1}{x_n}} = \frac{n}{\sum\left(\frac{1}{x}\right)}$$

### Example 23

Find the Harmonic Mean of the following values. 5, 6, 8, 9 and 10.

**Solution**

| $x$ | $\log \frac{1}{x}$ |
|---|---|
| 5 | $\frac{1}{5} = 0.2$ |
| 6 | $\frac{1}{6} = 0.16$ |
| 8 | $\frac{1}{8} = 0.125$ |
| 9 | $\frac{1}{9} = 0.11$ |
| 10 | $\frac{1}{10} = 0.1$ |
| | $\sum \frac{1}{x} = 0.695$ |

By using formula:

$$H.M = \frac{n}{\sum\left(\frac{1}{x}\right)} = \frac{5}{0.695} = 7.194$$

### (b) Harmonic Mean of grouped data

The Harmonic Mean of grouped data is computed by the following formula:

$$H.M = \frac{\sum f}{\sum\left(\frac{f}{x}\right)}$$

**Example 24**

Find the Harmonic Mean from the following data.

| Classes | 0 — 6 | 6 — 12 | 12 — 18 | 18 — 24 | 24 — 30 |
|---|---|---|---|---|---|
| Frequency | 1 | 2 | 5 | 4 | 6 |

**Solution**

| Classes | Frequency ($f$) | (Mid-point) | $\frac{f}{x}$ |
|---|---|---|---|
| 0 — 6 | 1 | 3 | $\frac{1}{3} = 0.33$ |
| 6 — 12 | 2 | 9 | $\frac{2}{9} = 0.22$ |
| 12 — 18 | 5 | 15 | $\frac{5}{15} = 0.33$ |
| 18 — 24 | 4 | 21 | $\frac{4}{21} = 0.19$ |
| 24 — 30 | 6 | 27 | $\frac{6}{27} = 0.22$ |
| | $\sum f = 18$ | | $\sum\left(\frac{f}{x}\right) = 1.29$ |

$$H.M = \frac{\sum f}{\sum\left(\frac{f}{x}\right)} = \frac{18}{1.29} = 13.95$$






### 6.3.2 Properties of Arithmetic Mean

- The sum of deviations of values measured from their A.M is always equal to zero.

  i.e. $\sum_{i=1}^{n}(x_i - \bar{x}) = 0$

- The sum of squared deviations of values measured from their A.M is least (minimum).

  $\sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n}(x_i - a)^2$ , where $\bar{x}$ = A.M and $a$ = any constant

- If $y_i = ax_i + b$ where $a$ and $b$ are some non-zero constants, than $\bar{y} = a\bar{x} + b$.
- If K-subgroups of a data having their respective means $\bar{x}_1, \bar{x}_2, \dots \bar{x}_k$ with respective frequencies $n_1, n_2, n_3, \dots n_k$ then the mean of the combined data is given as:

  $$\bar{x}_c = \frac{n_1\bar{x}_1 + n_2\bar{x}_2 + \dots + n_k\bar{x}_k}{n_1 + n_2 + n_3 + \dots + n_k} = \frac{\sum_{i=1}^{n} n_i\bar{x}_i}{\sum_{i=1}^{n} n_i}$$

### 6.3.3(a) Weighted mean

- The numerical values which show the relative importance of different items are called weights and the average of different items having different weights is called weighted mean. Let $x_1, x_2, x_3, \dots x_n$ are different values of items having weights $w_1, w_2, w_3 \dots w_n$ then weighted mean:

  $$\bar{x}_w = \frac{x_1 w_1 + x_2 w_2 + \dots + x_n w_n}{w_1 + w_2 + \dots + w_n} = \frac{\sum x_i w_i}{\sum w_i}$$

**Example 25** The marks obtained by a student in Maths, English, Urdu and Statistics were 70, 60, 80, 65 respectively. Find the average if weights of 2, 1, 3, 1 are assigned to the marks.

**Solution**

| $x$ | w | $wx$ |
|---|---|---|
| 70 | 2 | 140 |
| 60 | 1 | 60 |
| 80 | 3 | 240 |
| 65 | 1 | 65 |
| | $\sum w = 7$ | 505 |

Now $\bar{x}_w = \frac{\sum wx}{\sum w} = \frac{505}{7} = 72.14$

## 6.3.3(b) Moving Averages

- It is succession of averages (Arithmetic means) derived from the successive segments of a series of values. It is continuously recomputed as new data becomes available. It progresses by dropping the earliest value and adding the latest value.

**Example 26**

During first week of May, daily temperatures were recorded as given in the table. Calculate 3-day moving average temperature.

| Days | Temperature |
|---|---|
| Saturday | 40 |
| Sunday | 37 |
| Monday | 36 |
| Tuesday | 38 |
| Wednesday | 37 |
| Thursday | 41 |
| Friday | 39 |

**Solution**

| Days | Temperature | 3 day moving average means |
|---|---|---|
| Saturday | 40 | ... |
| Sunday | 37 | $\frac{40+37+36}{3}=37.67$ |
| Monday | 36 | $\frac{37+36+38}{3}=37$ |
| Tuesday | 38 | $\frac{36+38+37}{3}=37$ |
| Wednesday | 37 | $\frac{38+37+41}{3}=38.67$ |
| Thursday | 41 | $\frac{37+41+39}{3}=39$ |
| Friday | 39 | ... |





## 6.3.4 Estimation of Median, Mode and Quartiles graphically

### (i) Median

For median the class-boundaries are plotted on the horizontal axis and their cumulative frequencies on the vertical axis. Then points are located above the boundary points against their respective cumulative frequencies. All the points are then joined by straight line to get a curve called 'Ogive'.

Horizontal and vertical lines are drawn from the point where median is located, so the value on the horizontal axis at which the vertical lines intersect the Ogive, determines the median.

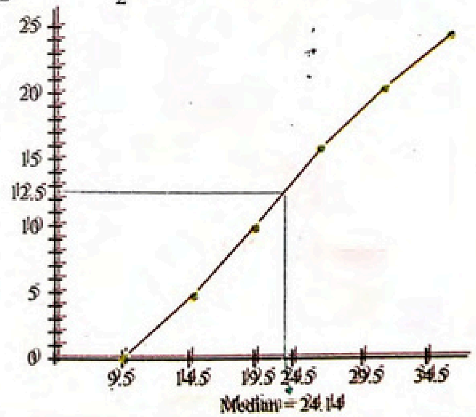
**Example 27** Find median graphically from the following frequency distribution:

| Classes | 10 — 14 | 15 — 19 | 20 — 24 | 25 — 29 | 30 — 34 | 35 — 39 |
|---|---|---|---|---|---|---|
| Frequency | 1 | 5 | 7 | 2 | 6 | 4 |

**Solution**

| Classes | $f$ | Cumulative Frequency $c.f$ | Class-Boundaries |
|---|---|---|---|
| 10 — 14 | 1 | 1 | 9.5 — 14.5 |
| 15 — 19 | 5 | 6 | 14.5 — 19.5 |
| 20 — 24 | 7 | 13 | 19.5 — 24.5 |
| 25 — 29 | 2 | 15 | 24.5 — 29.5 |
| 30 — 34 | 6 | 21 | 29.5 — 34.5 |
| 35 — 39 | 4 | 25 | 34.5 — 39.5 |

Median $= \frac{n}{2}th$ item $= \frac{25}{2}th$ item $= 12.5th$ item




**ERROR ANALYSIS** Describe and correct the error made in finding the median of the data set.

X median
10, 11, 24, 45, 41, 15, 45, 24, 50

### (ii) Mode

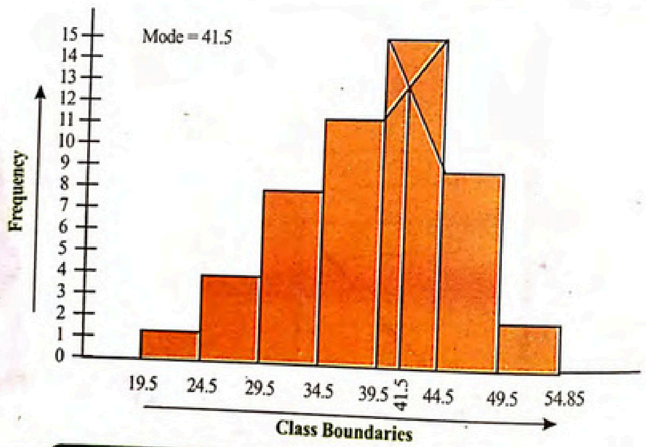
We need histogram to estimate mode graphically which is explained with the help of the following example.

**Example 28** Find mode graphically from the following frequency distribution.

| Classes | 20 — 24 | 25 — 29 | 30 — 34 | 35 — 39 | 40 — 44 | 45 — 49 | 50—54 |
|---|---|---|---|---|---|---|---|
| Frequency | 1 | 4 | 8 | 11 | 15 | 9 | 2 |

**Solution**

| Classes | $f$ | Class-Boundaries |
|---|---|---|
| 20 — 24 | 1 | 19.5 — 24.5 |
| 25 — 29 | 4 | 24.5 — 29.5 |
| 30 — 34 | 8 | 29.5 — 34.5 |
| 35 — 39 | 11 | 34.5 — 39.5 |
| 40 — 44 | 15 | 39.5 — 44.5 |
| 45 — 49 | 9 | 44.5 — 49.5 |
| 50 - 54 | 2 | 49.5 — 54.5 |










### (iii) Quartiles

To find the values of $Q_1$ and $Q_3$ from graph the cumulative frequency polygon (ogive) is drawn.

**Example 29** Find $Q_1$ and $Q_3$ from the following distribution.

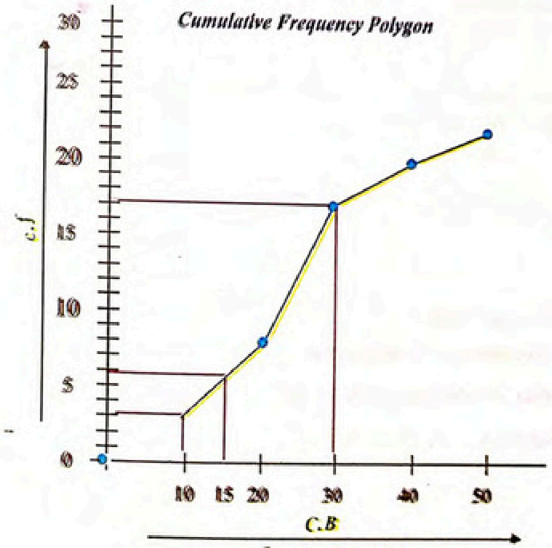
| Marks | 0 — 10 | 10 — 20 | 20 — 30 | 30 — 40 | 40 — 50 |
|---|---|---|---|---|---|
| Frequency | 3 | 5 | 9 | 3 | 2 |

**Solution**

| Marks | $f$ | $c.f$ |
|---|---|---|
| 0 — 10 | 3 | 3 |
| 10 — 20 | 5 | 8 |
| 20 — 30 | 9 | 17 |
| 30 — 40 | 3 | 20 |
| 40 — 50 | 2 | 22 |

$$\text{Location of } Q_1 = \frac{n}{4}\text{th item} = \frac{22}{4}\text{th item} = 5.5\text{th item}$$

$$\text{Location of } Q_3 = \frac{3n}{4}\text{th item} = 3\left(\frac{22}{4}\right)\text{th item} = 16.5\text{th item}$$




$Q_1 = 15,\ Q_3 = 29.4$

## Exercise 6.3

1. The following are the weights (in kg) of students of the $10^{th}$ grade.
   45, 30, 25, 36, 42, 27, 31, 43, 49 and 50.
   Calculate mean of the weights.
2. Find the mean of the weights given in Question 1 by using short-cut method.
3. Using assumed mean find the mean of the following numbers.
   1242, 1248, 1252, 1244, 1249
4. Find the mean of the following frequency distribution.
   Marks obtained by students of $10^{th}$ class in mathematics.

| Score (out of 75) | 0 – 15 | 16 – 31 | 32 – 47 | 48 – 63 | 64 – 75 |
|---|---|---|---|---|---|
| Frequency (f) | 0 | 10 | 40 | 70 | 45 |

5. Find the median of the following data.
   (i). Heights of boys in inches.
   64, 65, 65, 66, 66, 67
   (ii). Salaries of 9 workers of a factory
   7000, 6600, 8000, 4500, 7500, 11000, 9000, 7500

6. Find the Arithmetic mean, Geometric mean, Median and Mode of the following data
   58 59 60 62 64 64 65 67 67 68
   70 71 71 71 73
7. A set of data contains the values as 148, 145, 160, 157, 156 and 160. Show that Mode > Median > Mean.


8. From the following distribution

| Daily wages (in Rs) | 112—116 | 117—121 | 122—126 | 127—131 | 132—136 |
|---|---|---|---|---|---|
| No. of workers | 3 | 20 | 11 | 4 | 5 |

   (i). Construct a frequency table.
   (ii). Find the class-boundaries for each group.
   (iii). Calculate Median, Mode, Harmonic Mean, and Geometric Mean for the table.
9. Find Median, $Q_1$, $Q_3$ and Mode from the following distribution graphically.

| Classes | 10 — 14 | 15 —19 | 20 — 24 | 25 — 29 | 30 —34 |
|---|---|---|---|---|---|
| Frequency | 1 | 3 | 7 | 12 | 2 |






## 6.4 Measures of Dispersion

Dispersion is the scatterdness of values from its central value (Average).
Types of measure of dispersion are:

(i) Range
(ii) Standard deviation (S.D)
(iii) Variance

### (i) Range

The RANGE is the difference between the smallest observation and the largest observation.

Range = Largest value – Smallest value

**Example 30** What is the range of the data 209, 260, 270, 311, 311?

**Tidbit**
The range tells you how different the observations are. A large range means that the observations are very different. A smaller range means that they are more similar.

**Solution**

The largest value = 311
The smallest value = 209 m
Range = largest value–smallest value
Range = 311 – 209 = 102.

So the range is 102.

**Example 31** Following are the names and heights of mountains in Karakoram. Find the range of heights.

| | |
|---|---|
| K– 2 | 8611m |
| Ghasherbrum I | 8068m |
| Broad | 8047m |
| Ghasherbrum II | 8035m |
| Ghasherbrum III | 7952m |
| Ghasherbrum IV | 7925m |
| Rakaposhi | 7788m |


**Solution**

The maximum height = 8611 m
The minimum height = 7788 m
Range = maximum – minimum
= 8611 – 7788 = 823

So the range is 823 m.

**Note**
Range is very rarely used as it does not tell us about the observations in-between the largest and smallest values.

**Example 32** Calculate the range from the given data.

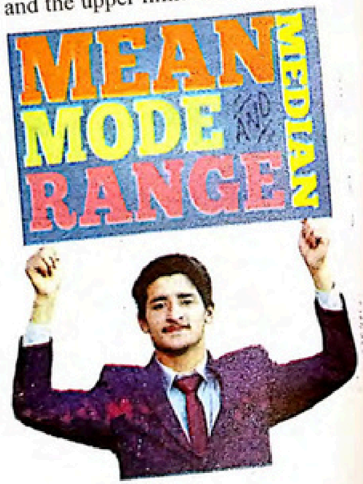
| Classes | 5 — 9 | 10 — 14 | 15 — 19 | 20 — 24 | 25 — 29 |
|---|---|---|---|---|---|
| Frequency | 10 | 15 | 12 | 21 | 3 |

**Solution**

| Classes | $f$ | Class-Boundaries |
|---|---|---|
| 5 — 9 | 10 | 4.5 — 9.5 |
| 10 — 14 | 15 | 9.5 — 14.5 |
| 15 — 19 | 12 | 14.5 — 19.5 |
| 20 — 24 | 21 | 19.5 — 24.5 |
| 25 — 29 | 3 | 24.5 — 29.5 |

In the given data the lower limit of the first group is 4.5 and the upper limit of the last group is 29.5. So,

Range = 29.5 − 4.5 = 25




**Example 33**

The numbers of grams in various candy bars are listed below.

Find the mean, median, mode, and range. Round to the nearest tenth if necessary. Then select the appropriate measure of central tendency or range to describe the data. Justify your answer.

9, 8, 9, 8, 9, 13, 24

**Solution**

mean: $\dfrac{8+8+9+9+9+13+24}{7} = 11.6$ g

median: 8, 8, 9, 9, 9, 13, 24

mode: 9 g occurs most frequently.

range: 24 − 8 = 16 g

The appropriate measure of central tendency or range to describe the data is the median or the mode. The mean is affected by the highest value, 24 grams.

**Did You Know?**

$\pi$ = 3.14159 26535 89793 23846 26433 83279 50288 41971 69399 37510 58209 74944 59230 78164 06286 20899 86280 34825 34211 70679 82148 08651 32823 ...





**(ii) Standard Deviation (S.D)**

It is the positive square root of the average of squared deviations measured from A.M (arithmetic mean).

i.e. $\text{S.D} = \sqrt{\dfrac{\sum(x-\bar{x})^2}{n}}$ for ungrouped data

and $\text{S.D} = \sqrt{\dfrac{\sum f(x-\bar{x})}{\sum f}}$ for frequency distribution

**(iii) Variance**

Variance is the square of standard deviation. Variance is usually denoted by the symbol 'S'. Mathematically it is expressed as:

$$S^2 = \frac{\sum(x-\bar{x})^2}{n} \text{ or } \frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2 \text{ for ungrouped data}$$

$$\text{and } S^2 = \frac{\sum f(x-\bar{x})^2}{\sum f} \text{ or } \frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2 \text{ for discrete and continuous data}$$

**Example 34**

Find the variance and standard deviation for the following data:

6, 8, 10, 12, 14

**Solution**

$$\bar{x} = \frac{6+8+10+12+14}{5} = \frac{50}{5} = 10$$

| $x$ | $x-\bar{x}$ | $(x-\bar{x})^2$ |
|---|---|---|
| 6 | 6 − 10 = −4 | 16 |
| 8 | 8 − 10 = −2 | 4 |
| 10 | 10 − 10 = 0 | 0 |
| 12 | 12 − 10 = 2 | 4 |
| 14 | 14 − 10 = 4 | 16 |
| | | 40 |

$$\text{S.D} = \sqrt{\frac{\sum(x-\bar{x})^2}{n}} = \sqrt{\frac{40}{5}} = \sqrt{8}$$

Variance $(S^2) = (\sqrt{8})^2 = 8$

**Example 35** The following is the distribution for the number of rotten eggs found in 60 crates. Find the standard deviation and variance of the rotten eggs.

| No. of rotten eggs | 0—4 | 4—8 | 8—12 | 12—16 | 16—20 | 20—24 |
|---|---|---|---|---|---|---|
| No. of crates | 5 | 10 | 15 | 20 | 6 | 4 |

**Solution**

| C.B | $f$ | $x$ | $fx$ | $x^2$ | $fx^2$ |
|---|---|---|---|---|---|
| 0—4 | 5 | 2 | 10 | 4 | 20 |
| 4—8 | 10 | 6 | 60 | 36 | 360 |
| 8—12 | 15 | 10 | 150 | 100 | 1500 |
| 12—16 | 20 | 14 | 280 | 196 | 3920 |
| 16—20 | 6 | 18 | 108 | 324 | 1944 |
| 20—24 | 4 | 22 | 88 | 484 | 1936 |
| | | | 696 | | 9680 |

$$\text{Variance} = \frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2$$

$$= \frac{9680}{60} - \left(\frac{696}{60}\right)^2$$

$$= 161.33 - 134.56 = 26.77$$

$$\text{S. D} = \sqrt{Variance} = \sqrt{26.77} = 5.18 \text{ (Approximately)}$$

**Activity**

Substances with pH values less than 7 are acids, those with pH values greater than 7 are bases, and substances with pH values equal to 7 are neutral. Test several solutions and list the pH values. Make the *frequency* table of pH values by completing the table at the right.

| pH number | Tally | Frequency |
|---|---|---|
| Less than 7 | | |
| Equal to 7 | | |
| Greater than 7 | | |





## Exercise 6.4

1. Find the range for the following items:
11, 13, 15, 21, 19, 23.

2. A bank branch manager was interested in the waiting times of customers. For this he carried out a survey. A random sample of 12 customers was selected and yielded the following data.
5.90 9.66 5.79 8.02 8.73 8.01 10.49 8.35 6.68 5.64 5.47 9.91
Calculate the following
(i). Average (ii). Median (iii). Standard Deviation

3. Calculate the Range, Variance and Standard Deviation for the following discrete data:

| $x$ | 5 | 10 | 11 | 13 | 15 |
|---|---|---|---|---|---|
| $f$ | 2 | 3 | 4 | 1 | 5 |

4. The Following table shows the marks obtained by 10 students of two sections of 10th class.

| Section A | 7 | 9 | 6 | 9 | 4 | 7 | 5 | 8 | 8 | 7 |
|---|---|---|---|---|---|---|---|---|---|---|
| Section B | 6 | 10 | 6 | 4 | 2 | 8 | 10 | 6 | 9 | 9 |

Find their
(a) Arithmetic Mean (b) Variance

5. Following are the marks (out of 75) of the eight students in two subjects.

| Student | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| Marks in Mathematics | 54 | 63 | 59 | 45 | 52 | 35 | 61 | 68 |
| Physics | 52 | 55 | 57 | 51 | 56 | 58 | 50 | 59 |

Compare the standard deviation of the marks and tell that in which subject students are more consistent.

6. The following is the distribution for the number of defective bulbs in 30 cartons (Packs). Find Variance and Standard Deviation of defective bulbs.

| No. of defective bulbs | 0—2 | 2—4 | 4—6 | 6—8 | 8—10 |
|---|---|---|---|---|---|
| No. of packs (f) | 1 | 3 | 15 | 10 | 2 |

# Review Exercise 6

At the end of each question, four circles are given. Fill in the correct circle only.

(i). The difference between upper limit of two consecutive classes in a frequency table is called.
○ Class limit ○ Class interval ○ Class mark ○ Range

(ii). A cumulative frequency histogram is also called
○ Histogram ○ Pie chart
○ Ogive ○ frequency polygon

(iii). The number of times a value appears on a set of data is called
○ frequency ○ average ○ mode ○ median

(iv). The data below represents the number of televisions that 11 students have in their homes. Find the mode of the data.

3, 2, 1, 1, 1, 5, 3, 1, 2, 1, 2
○ 1 ○ 2 ○ 3 ○ 4

(v). In which data set are the mean, median, mode, and range all the same number?
○ 1, 2, 3, 3, 2, 1, 2 ○ 1, 2, 3, 1, 2, 3, 1
○ 1, 3, 3, 3, 2, 3, 1 ○ 2, 2, 1, 2, 3, 2, 3

(vi). The $n^{th}$ root of product of 'n' number of values is called.
○ arithmetic ○ geometric mean
○ harmonic ○ standard deviation

(vii). In a set of data.

63, 65, 66, 67, 69, median is
○ 63 ○ 66 ○ 67 ○ 69

(viii). In a set of data

41, 43, 47, 51, 57, 52, 59 median is
○ 51 ○ 47 ○ 47 ○ none of these

(ix). In the given set of data 5, 7, 7, 5, 3, 7, 2, 8, 2 mode is
○ 9 ○ 5 ○ 2 ○ 7

(x). In the given set of data 5, 5, 5, 5, 5, 5, 5 the standard deviation is
○ 5 ○ 0 ○ 7 ○ none of these

(xi). The average pocket money of 30 student is Rs. 20/-. The total amount in the class is
○ Rs. 20/- ○ Rs. 30/- ○ Rs. 300/- ○ Rs. 600/-






(xii). The sum of 30 observations is 1500. Its average will be
○ 1500 ○ 150 ○ 15 ○ none of these

(xiii). The difference of the largest and smallest value in the data is called
○ Mean ○ Mode ○ Range ○ Standard deviation

(xiv). The formula $\frac{\sum x}{n}$ determines
○ Arithmetic Mean ○ Median
○ Mode ○ G.M

(xv). What is the difference between the mean of Set B and the median of Set A?

Set A: {2, −1, 7, −4, 11, 3}
Set B: {12, 5, −3, 4, 7, −7}
○ −0.5 ○ 0 ○ 0.5 ○ 1

(xvi). $\frac{\sum f(x-\bar{x})^2}{\sum f}$ is called ______
○ Range ○ Median ○ S.D ○ Variance

(xvii). The most frequent value in the data is called its ______
○ Mean ○ Median ○ Mode ○ G.M

2. The ages (in years) of 27 students of $10^{th}$ class are given. Prepare a frequency distribution of suitable class interval.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 17, | 17, | 16, | 16, | 17, | 16, | 16, | 17, | 18, | 18, |
| 15, | 17, | 19, | 18, | 18, | 17, | 16, | 15, | 16, | 17, |
| 15, | 19, | 19, | 15, | 15, | 16, | 18 | | | |

3. Prepare a histogram of the following table;

| Brand of car | A | B | C | D | E |
|---|---|---|---|---|---|
| Sale in 1 month | 100 | 120 | 110 | 72 | 169 |

4. Prepare a frequency polygon of the following frequency distribution.

| Score in test | Frequency |
|---|---|
| 0 – 10 | 2 |
| 11 – 21 | 7 |
| 22 – 32 | 25 |
| 33 – 43 | 11 |
| 44 – 50 | 5 |

... his table summaries the weights in kilograms of 250 boys. Represent these data by means of a cumulative frequency polygon.

l. At

| Weight (kg) | Number of boys |
|---|---|
| 44.0 – 47.9 | 3 |
| 48.0 – 51.9 | 17 |
| 52.0 – 55.9 | 50 |
| 56.0 – 59.9 | 81 |
| 60.0 – 63.9 | 57 |
| 64.0 – 67.9 | 23 |
| 68.0 – 71.9 | 9 |

PROJECT

The following scores were made on a 60-item test:

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 25 | 30 | 34 | 37 | 41 | 42 | 46 | 49 | 53 |
| 26 | 31 | 34 | 37 | 41 | 42 | 46 | 50 | 53 |
| 28 | 31 | 35 | 37 | 41 | 43 | 47 | 51 | 54 |
| 29 | 33 | 36 | 38 | 41 | 44 | 48 | 52 | 54 |
| 30 | 33 | 36 | 39 | 41 | 44 | 48 | 52 | 55 |
| 30 | 33 | 37 | 40 | 42 | 45 | 48 | 52 | |

(a) Group the data into class intervals of size 2, beginning with the interval 24.5–25.5.
(b) Set up a frequency table for the above data.
(c) Build the histogram for the data.
(d) Draw the frequency polygon for this histogram.
(e) Find the cumulative frequencies.
(f) Draw the ogive.
(g) Find the range of the data.
(h) Find the mean of the data.
(i) Find the standard deviation of the data.
(j) Find the variance of the data.



Mathematics X NOT FOR SALE



# Summary

- A set of raw data can be organized and then arranged in an orderly way in the form of a frequency table.
- A frequency table can be represented graphically by a Histogram.
- A histogram is a vertical bar graph with no space in between the bars.
- The area of each bar is proportional to the frequency it represents.
- If the midpoints of the tops of the consecutive bars in a histogram are joined by line segments and mid points at both ends are joined to the horizontal axis, a frequency polygon is obtained.
- The cumulative frequency table provide another way of representing a set of data. A cumulative frequency table can be represented by a cumulative frequency curve known as an Ogive.
- The median is the middle value when a set of data is arranged in order of increasing magnitude. the median divides a set of data into two equal halves. The middle value of the lower half is the lower quartile and the middle value of the upper half is the upper quartile. The lower quartile, the median and the upper quartile which divides the data into four equal parts are called quartiles.
- The inter quartile range of a set of numbers is the difference between the upper quartile and the lower quartile.
- A set of data can be described by numerical quantities called averages.
- The three common averages are the mean, the median and the mode.
- The mode is the number that occurs most frequently.
- The mean is the sum of values divided by the number of values in a set of data.
- The mean of a set of grouped data is $\overline{x} = \frac{\sum fx}{\sum f}$, where $x$ is the middle value of the class interval and f is the frequency of the class interval.
- The median for an odd number of data is the middle value when the data are arranged in ascending/descending order.
- The median for an even number of data is the mean of the two middle values when the data are arranged in ascending/descending order.