Unit 12 Information Handling

Students' Learning Outcomes

At the end of the unit, the students will be able to:

- Construct a grouped frequency table, histogram (with unequal class interval) and frequency polygon.
- Calculate the mean modal class and median of a grouped frequency distribution.
- Solve real life situations involving mean, weighted mean, median and mode for given data (such as allocation of funds in different projects, forecasting future demographics, marketing, forecasting government budgets).

INTRODUCTION

Before knowing about information let us think how can we answer the question like how many students were there in each class of a particular school.

How many patients visited in a hospital within a particular week.

To answer these questions, we need to have quantitative information and it can be obtained by counting.

It should be kept in mind that simple numbers 85, 96, 70, 80, 73, 70, 65, 83, 89,75 are not data but if we say that the above data indicates the marks of students of different classes, the above figures are considered as data and have precise meaning.

Hence, to know about something is known as "Information" and to represent that information in a manageable way so that useful conclusions can be drawn is called information handling. So, the collection of meaningful information in the form of facts and numerical figures is known as data.

History

In statistics, information handling is also known as data handling. "Data Handling" plays vital role to represent the information in a manageable way.

The word "Data Handling was first used by Sir Ronald Fisher.

Sir Ronald Aylmer Fisher

(17 February 1890 – 29 July 1962)

For further information scan the following QR Code:

The numerical figures are obtained from any field of study e. g., the mass of the students of your class, the number of pair of shoes sold by a shopkeeper in a month etc. Data can be obtained from existing sources i.e., office records, published papers or the same can be obtained directly from the field according to needs.

Information Handling

Information handling is the process of collecting, organizing, summarizing, analyzing and interpreting numerical data.

Data is further classified into two categories.

(i) **Discrete data:** It can take only some specific values. whole numbers are used to write discrete data. e.g., number of books sold by a shopkeeper, number of patients visited a hospital in a week etc. This data is only obtained by counting.

(ii) **Continuous data:** It can take every possible value in a given interval. Decimal numbers are used to write continuous data. The data is only obtained by measuring e.g., the mass of students in class i.e., 28.5 kg, 26.5 kg, 27.5 kg etc.

12.1 Ungrouped and Grouped Data

Data which is not arranged in any systematic order (groups or classes) is called ungrouped data. For example, the number of toys sold by a shopkeeper in a month is given below:

10, 5, 8, 12, 15, 20, 25, 30, 23, 15, 23, 21, 18, 15 17, 23, 22, 15, 20, 21, 24, 18, 16, 21, 23, 21, 17, 19, 21, 23. This data is called ungrouped data.

If we arrange the above given data in groups or classes, then it is called grouped data.

Do you know?
Ungrouped data is also known as raw data.

Classes	Tally marks	No. of toys sold
5 – 9	||	2
10 – 14	||	2
15 – 19	卌 卌	10
20 – 24	卌 卌 ||||	14
25 – 29	|	1
30 – 34	|	1

Teachers' note!
By using more examples, clear the concept of grouped data and ungrouped data to the students.

In above grouped data, 5, 10, 15, 20, 25 and 30 are lower class limits and 9, 14, 19, 24, 29 and 34 are upper class limits.

12.1.1 Frequency Distribution

A distribution or table that represents classes or groups along with their respective class frequencies is called frequency distribution. In other words, the various items of data are classified into certain groups or classes and the number of items lying in each group or class is put against that group or class. The data organised and summarized in this way is known as frequency distribution.

Think!

If the size of class limits is 6. The greatest value is 80 and the smallest value is 25. Can you find the number of class limits for the data?

Formation of Frequency distribution

In this method, the raw data or the ungrouped data is presented into a grouped data. Choice is yours to select the number of classes.

Generally, the size of class limits is determined on the basis of the greatest value, smallest value and the desired number of groups or classes.

Following are the major steps to construct frequency distribution:

(i) Find the range of the data. Range is the difference between the greatest value and the smallest value i.e., Range = $X_{max} - X_{min}$

(ii) Find the size of the class by dividing the range by the number of classes or groups you wish to make.

Keep in mind!

The number of times a value occurs in a data is called the frequency of that value. It is denoted by "f".

For example, the greatest value is 136, the smallest value is 30 and if we have to make 10 classes or groups, then the size of class limits is found by the given formula.

$$\text{Size of class} = \frac{\text{Range}}{\text{Number of classes}} = \frac{\text{Greatest Value} - \text{Smallest Value}}{\text{Number of classes}}$$

$$= \frac{136-30}{10} = \frac{106}{10} = 10.6 \approx 11$$

So, size of class limits = 11

(iii) Prepare four columns.

(a) Class limits (b) Tally marks

(c) Frequencies (d) Class Boundaries

(iv) Make classes having size of 11. Start from the smallest value.

For example, 30 – 40, 41 – 51, 52 – 62 and so on.

(v) Look for the class in which each element of ungrouped data falls. Draw a small tally mark (|) against that class and also tick the element concerned with a sign (✓). In this way you can remember that you have counted for the element. Continue this way with the next element that upto the last element of the data set. If 5 or more tallies appear in any class, mark every 5th tally diagonally as ~~||||~~.

(vi) Class boundaries usually are found by the following method:

- Chose the upper class limit of the 1st class and lower class limit of the 2nd class.
- Find the difference between these two limits.
- The difference is divided by 2 and subtract it from the lower class limit and add it to the upper class limit.

Do you know?

Class boundaries may also be obtained from the midpoints (x) as $\left[x \pm \frac{h}{2}\right]$, where h is the difference between any two consecutive values of x.

Example 1: Following are the number of telephone calls made in a week to 30 teachers of a high school.

5	8	11	25	13	16	20	17	15	16	30	21	14	18	19
6	22	26	15	19	35	29	31	23	25	20	10	9	7	26

Construct a frequency distribution with number of classes 7.

Solution: (i) Find range

Greatest value (maximum value) = 35, Smallest value (minimum value) = 5

$$\text{Range} = X_{\max} - X_{\min} = 35 - 5 = 30$$

(ii) Size of class limits $= \dfrac{\text{Range}}{\text{Number of classes}} = \dfrac{30}{7} = 4.28 \approx 5$

(iii) Make class limits having size 5. For example, 5 – 9, 10 – 14, 15 – 19 and so on. (see 1st column of table: 1).

(iv) Tally marks are used to count the values, fall in the given class limits. (See 2nd column of table: 1).

(v) Now, count the number of tally marks and write the number as frequency in the third column (see 3rd column of table: 1).

(vi) **Class boundaries**

The difference between lower class limit of the second class and upper class limit of the first class is 1. i.e., 10 – 9 = 1. Now, divide the difference of the limits by 2 i.e. $\frac{1}{2} = 0.5$.

Activity

Collect data of height of 50 students in your class, and convert the data into grouped data.

Lower class boundaries are obtained by "subtracting 0.5" from the lower class limits.

Upper class boundaries are obtained by "adding 0.5" to the upper class limits.

Lower class boundaries	**Upper class boundaries**
5 – 0.5	9 + 0.5
4.5 and so on.	9.5 and so on.

(see 4th column of the table: 1)

Table 1

Class limits	Tally marks	Frequency (f)	Class Boundaries (C.B)
5 – 9	𝍸	5	4.5 – 9.5
10 – 14	||||	4	9.5 – 14.5
15 – 19	𝍸 |||	8	14.5 – 19.5
20 – 24	𝍸	5	19.5 – 24.5
25 – 29	𝍸	5	24.5 – 29.5
30 – 34	||	2	29.5 – 34.5
35 – 39	|	1	34.5 – 39.5

12.1.2 Graph of Frequency Distribution

The following are the types of graphs which can be used to represent a frequency distribution on a graph.

(a) Histogram (b) Frequency polygon

(a) Histogram (with equal class limits)

This is a graph of adjacent rectangles constructed on xy plane. A histogram is similar to bar graph but it is constructed for a frequency distribution. In a histogram, the values of the data, (classes) are represented along the horizontal axis and the frequencies are shown by bars perpendicular to the horizontal axis. Bars of equal width are used to represent individual classes of frequency table. The procedure for making histogram is explained below:

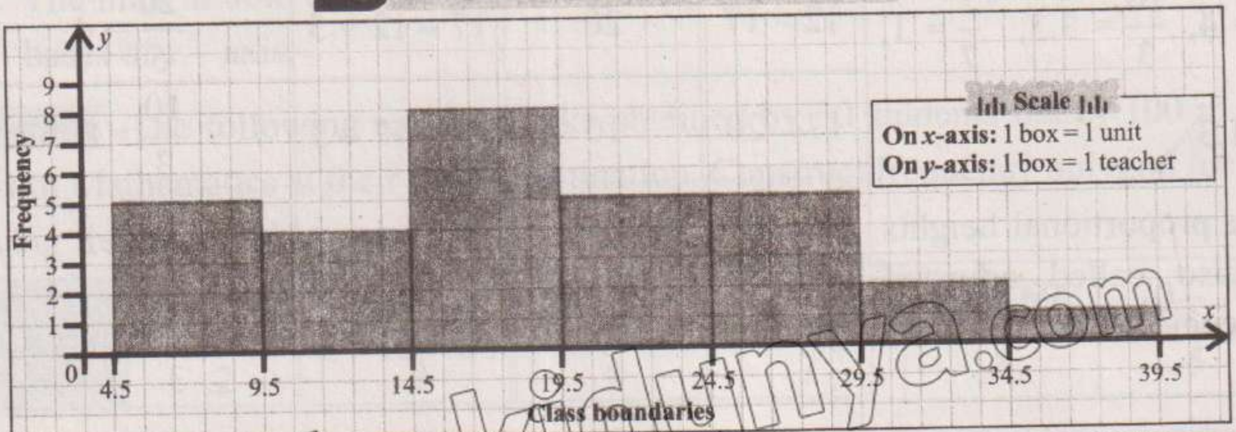
Do you know?
Continuous data is mostly represented by using histogram and frequency polygon.

(i) Draw lines as x – axis and as y – axis on a graph paper perpendicular to each other.

(ii) Class boundaries are marked on x – axis and a rectangle is made against each group with its width proportional to the size of the class limits and height proportional to the class frequencies.

(iii) Setting a scale, draw frequencies on y – axis. The resulting figure is called histogram. Histogram of table: 1 is given below:

Title Histogram of telephone calls made in a week

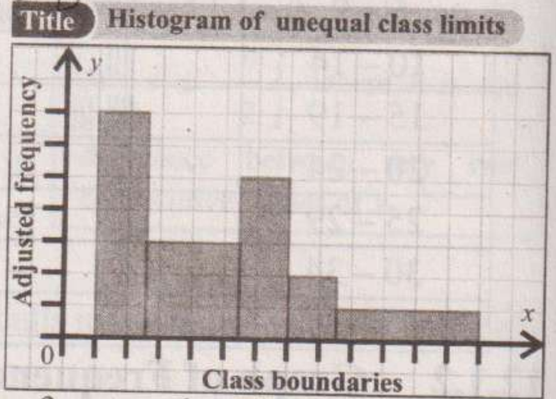


12.1.3 Histogram (with unequal class limits)

The procedure for making histogram is explained below:

i. Draw lines as x – axis and y – axis on a graph paper perpendicular to each other.

ii. Class boundaries are marked on x – axis and a rectangle is made against each group with its width proportional to the size of class limits and height proportional to the class frequencies.



iii. This can be achieved by adjusting the heights of rectangle. The height of each rectangle is obtained by dividing each class frequency on its class limit size.

Example 2: The frequency distribution of ages (in years) of 76 members of a locality is available. Draw a histogram for this data.

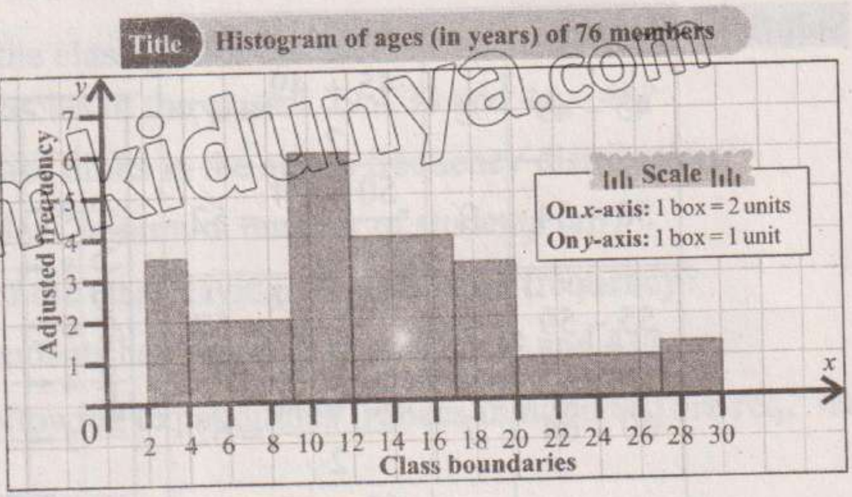
Class limits	2 – 4	4 – 9	9 – 12	12 – 17	17 – 20	20 – 27	27 – 30
Frequency (f)	7	10	18	20	10	7	4

Solution: Look at the table, indicates that the width of the class limits is not equal as first class has width 2, second has 5, the third has 3, the fourth has 5, the fifth has 3, sixth class has 7, seventh class has width 3. So, there is need to adjust the heights of the rectangles i.e., for the first class we have 2 as width of class and 7 as a frequency, so the height of the first class is $\frac{7}{2} = 3.5$, similarly for the other $\frac{10}{5} = 2$, $\frac{18}{3} = 6$, $\frac{20}{5} = 4$, $\frac{10}{3} = 3.3$, $\frac{7}{7} = 1$, $\frac{4}{3} = 1.3$.

These proportional heights are also called adjusted frequencies.

Class limits	Frequency (f)	Width of Class	Height of rectangle (Adjusted frequency)
2 – 4	7	4 – 2 = 2	$\frac{7}{2} = 3.5$
4 – 9	10	9 – 4 = 5	$\frac{10}{5} = 2$
9 – 12	18	12 – 9 = 3	$\frac{18}{3} = 6$
12 – 17	20	17 – 12 = 5	$\frac{20}{5} = 4$
17 – 20	10	20 – 17 = 3	$\frac{10}{3} = 3.3$
20 – 27	7	27 – 20 = 7	$\frac{7}{7} = 1$
27 – 30	4	30 – 27 = 3	$\frac{4}{3} = 1.3$

Taking class boundaries along x – axis and corresponding adjusted frequencies along y – axis, rectangles are drawn and the histogram is given below.



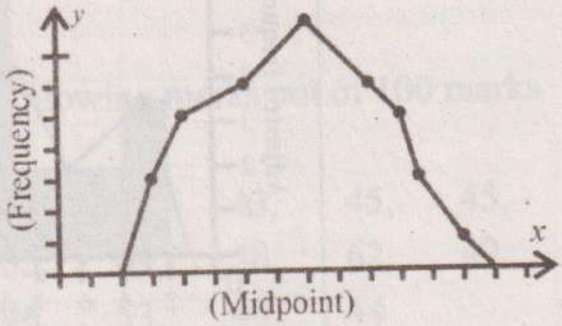
12.1.4 Frequency Polygon

A frequency polygon is a closed geometrical figure used to display a frequency distribution graphically. A line graph of a frequency distribution is known as frequency polygon in which frequencies are plotted against their midpoints.

Midpoint is the average value of the lower and upper class limits. Midpoint is also known as class mark. Midpoint is calculated by the given formula:

$$\text{Midpoint} = \frac{\text{Lower class limit} + \text{Upper class limit}}{2}$$

The following steps are followed to draw a frequency polygon for a frequency distribution:



(i) Draw lines as x – axis and y – axis perpendicular to each other.

(ii) Take midpoints on x – axis and class frequencies on y – axis.

(iii) Put a dot mark against each midpoint corresponding to its class frequency. Join all the dotted marks by straight lines to get the required frequency polygon.

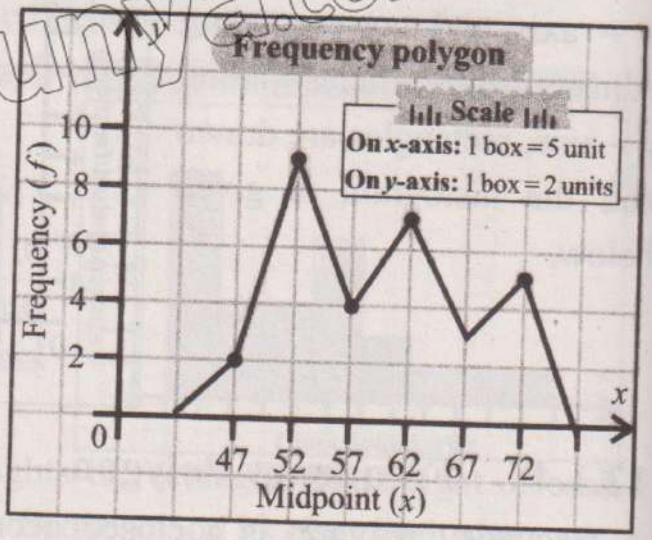
(iv) The lines at both ends are joined together with the next midpoints to touch the bases of x – axis.

Example 3: The following are the marks obtained by 30 students out of 100 in the subject of Mathematics at their final examination. Construct frequency polygon for the following frequency table.

Marks	45 – 49	50 – 54	55 – 59	60 – 64	65 – 69	70 – 74
Frequency	2	9	4	7	3	5

Solution:

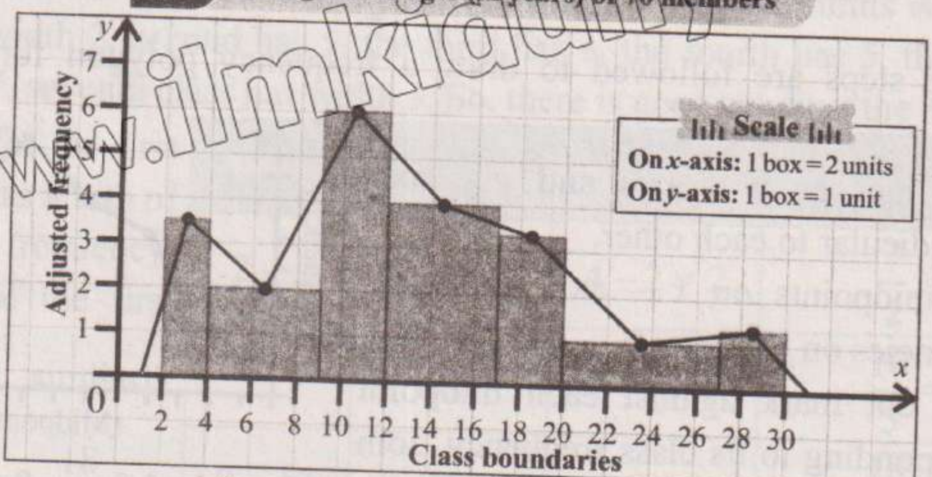
Marks	f	Midpoints
45 – 49	2	$\frac{45+49}{2} = 47$
50 – 54	9	$\frac{50+54}{2} = 52$
55 – 59	4	$\frac{55+59}{2} = 57$
60 – 64	7	$\frac{60+64}{2} = 62$
65 – 69	3	$\frac{65+69}{2} = 67$
70 – 74	5	$\frac{70+74}{2} = 72$



Remember!

Frequency polygon on histogram: In histogram, we mark the midpoints on the top of rectangles and join all the points. To touch the base of x – axis, we extend the line at both ends to the next midpoints. The resulting graph is a frequency polygon.

Title Histogram of ages (in years) of 76 members



EXERCISE 12.1

1. The following distribution represents the scores achieved by a group of chemistry students in the chemistry laboratory.

Scores	24 – 28	29 – 33	34 – 38	39 – 43	44 – 48	49 – 53	Total
No. of students	3	6	12	23	15	6	65

Answer the following questions.

(i) What is the upper limit of the last class?

(ii) What is the lower limit of the class 39 – 43?

(iii) What is the midpoint of the class (34 – 38)?

(iv) What are the class frequencies of the classes 29 – 33 and 44 – 48?

(v) What is the size of the class limits in the above frequency distribution?

(vi) In which class or group does minimum number of students fall?

(vii) What is the lower limit of the class having 15 as its class frequency?

(viii) What is the number of students having scores between 24 and 43?

2. For a school staff, the following expenditures (rupees in hundred) are required for the repair of chairs.

145, 152, 153, 156, 158, 160, 146, 152, 155, 159,
161, 163, 165, 147, 148, 151, 154, 156, 158, 160,
144, 167, 151, 150, 152, 149, 145, 153, 152, 155

Prepare a frequency distribution by tally bar method using 3 as the size of class limits and also write down what are the frequencies of the last three classes?

3. Given below are the weights in kg of 30 students of a high school.

30, 33, 24, 21, 15, 39, 37, 44, 42, 33,
33, 28, 29, 32, 31, 28, 26, 32, 34, 35,
38, 36, 41, 30, 35, 41, 23, 26, 18, 34

Taking 5 as the size of the class limit, prepare a frequency table and construct a frequency polygon.

4. A group of Grade - 10 students obtained the following marks out of 100 marks in English test.

58, 59, 58, 33, 40, 58, 45, 46, 43, 45, 45,
50, 52, 49, 50, 57, 52, 55, 49, 50, 62, 49,
48, 44, 42, 47, 46, 47, 46, 53, 40, 44

Classify the data into a frequency distribution by (direct method) taking 6 as the size of class limit. Also find the class limit with least class frequency and construct histogram for the data.

5. From the table given below. Draw a frequency polygon on histogram for the given frequency distribution.

Weight (kg)	10 – 14	15 – 19	20 – 24	25 – 29	30 – 34	35 – 39
Frequency (f)	06	17	23	30	22	13

6. The following data shows the number of heads in an experiment of 50 sets of

tossing a coin 5 times. Make a discrete frequency distribution from the information.

3, 3, 4, 0, 5, 4, 3, 3, 1, 2, 4, 5, 0, 3, 2, 4, 4, 0, 0, 0, 5, 5, 3, 2, 1
4, 3, 2, 5, 3, 2, 1, 3, 5, 4, 3, 2, 1, 3, 2, 1, 3, 1, 3, 1, 4, 3, 2, 2, 4

7. The marks obtained by the students of Grade - 10 in mathematics test were grouped into the following frequency distribution.

Marks	35 – 37	38 – 44	45 – 54	55 – 61	62 – 67	68 – 72
Frequency	2	12	16	13	9	3

Draw a histogram for the above distribution.

8. Make a frequency polygon on histogram for the following grouped data:

Class limits	5 – 8	8 – 12	12 – 20	20 – 25	25 – 27	27 – 32
Frequency (f)	2	12	25	32	14	5

12.2 Measures of Location (Central Tendency)

The measure that gives the centre of the data is called measure of central tendency. Therefore, measure of central tendency is used to find out the middle or central value of a data set.

We have seen that when the raw data has been condensed into a frequency distribution, the information was easily understood. The information given in the data can be further condensed to a single representative value for the entire distribution. It is more or less the central value around which the data appear to be crowded. For example, usually, we make statements such as:

(i) Hassan studies 6 hours daily.

(ii) The monthly expenditure of Ayesha's house is Rs.50,000.

(iii) The speed of Maham's car is 72 km per hour.

(iv) In a country, yearly income is 70,000 rupees per head.

(v) The price of onion in the market is Rs.150 per kg etc.

If we look at the first statement, we come to know that Hassan does not study exactly 6 hours daily. Sometimes, he studies more than 6 hours and sometimes less. But still why do we say that he studies 6 hours daily? As he studies near about 6 hours daily so in his study time, 6 hours becomes an important figure because of its approximated statement, which we call Average. Such an average value is known a measure of central tendency because it is a representative value of the daily study time. Similarly, other statements can also be treated as representative values. As each statement locates the centre of a distribution so it is also known as a measure of central tendency.

The following measure of central tendency will be discussed in this section:

(i) Arithmetic Mean (A.M.) (ii) Median
(iii) Mode (iv) Weighted mean

12.2.1 Arithmetic Mean (A.M.)

It is defined as a value of variable which is obtained by dividing the sum of all the values (observations) by their number of observations. Thus, the arithmetic mean of a set of values $x_1, x_2, x_3 \dots x_n$ is denoted by $\overline{X}$ (read as X-bar) and is calculated as:

$$\overline{X} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} = \frac{\Sigma x}{n} \text{ (Direct method)}$$

where, the sign Σ stands for the sum and n is the number of observations.

Example 4: The marks of a student in five examinations were 64, 75, 81, 87, 90. Find the arithmetic mean of the marks.

Try Yourself !
The mean of 10, 30, 40, x, 67 and 81 is 50. Find the value of the x.

Solution:

$$\text{A.M.} = \overline{X} = \frac{\sum x}{n}$$

$$= \frac{64 + 75 + 81 + 87 + 90}{5}$$

or $$\overline{X} = \frac{397}{5} = 79.4 \text{ marks}$$

Example 5: A government allocates funds of Rs.200,000 to five sectors of a school i.e.,

(i) School Library: Rs. 35,000
(ii) Sports facilities: Rs. 25,000
(iii) Parking area: Rs. 40,000
(iv) Room renovation: Rs. 45,000
(v) Furniture: Rs. 55,000

Try Yourself !
The mean of 15 values was 50. It was found on rechecking that the value 25 was wrongly copied as 52. Find the correct mean.

Find the average of fund allocation in each sector of a school.

Solution: To find out the average of each sector, we will find the mean of the given data.

$$\overline{X} = \frac{35{,}000 + 25{,}000 + 40{,}000 + 45{,}000 + 55{,}000}{5}$$

$$\overline{X} = \frac{200{,}000}{5}$$

$$\overline{X} = \text{Rs. } 40{,}000$$

On average, each sector takes Rs.40,000 in funding.

Method of finding Arithmetic Mean for Grouped Data

Let $x_1, x_2, x_3, \ldots, x_n$ be the midpoints of the class limits with corresponding frequencies say $f_1, f_2, f_3, \ldots, f_n$. Then the arithmetic mean is obtained by dividing sum of the products of f and x by the sum of all the frequencies.

$$\overline{X} = \frac{f_1x_1 + f_2x_2 + \cdots + f_nx_n}{f_1 + f_2 + \cdots + f_n} = \frac{\Sigma fx}{\Sigma f}$$

Example 6: Given below are the marks out of 100 obtained by 100 students in a examination. Find the average marks of the students.

Marks	30 – 35	35 – 40	40 – 45	45 – 50	50 – 55	55 – 60
No. of students	14	16	18	23	18	11

Solution:

Marks	Midpoint (x)	Frequency (f)	fx
30 – 35	32.5	14	455.0
35 – 40	37.5	16	600.0
40 – 45	42.5	18	765.0
45 – 50	47.5	23	1092.5
50 – 55	52.5	18	945.0
55 – 60	57.5	11	632.5
Total	—	$\Sigma f = 100$	$\Sigma fx = 4490$

$$\overline{X} = \frac{\Sigma fx}{\Sigma f} = \frac{4490}{100}$$

or $\overline{X} = 44.9$ marks

Hence, the average marks is 44.9 of the surdents.

Short Formula for Computing Arithmetic Mean

The computation of arithmetic mean using direct method for ungrouped data as well as for grouped data is no doubt easy for small values. If x and f become very large, it becomes difficult to deal with the problems so to minimize our time and calculations we take deviations from an assumed or provisional mean. Let A be considered as assumed or provisional mean (may be any value from the values of x or any number) and D denotes the deviations of x from A i.e., $D = x - A$. For $x = D + A$, the formula of arithmetic mean becomes;

$$\overline{X} = A + \frac{\Sigma D}{n} \quad \text{(for ungrouped data)} \quad ...(i)$$

$$\overline{X} = A + \frac{\Sigma fD}{\Sigma f} \quad \text{(for grouped data)} \quad ...(ii)$$

Example 7: Find the arithmetic mean using short formula for the runs made by a batsman.

Runs: 40, 45, 50, 52, 50, 60, 56, 70.

Try Yourself !

If $\overline{X} = 120$; $A = 85$ and $n = 25$, then can you find the value of ΣD?

Solution: Taking deviations from $A = 52$ (assumed mean)

x	40	45	50	52	50	60	56	70
$D = x - A$	−12	−7	−2	0	−2	8	4	18

Now: $\Sigma D = -23 + 30 = 7$

$\because \quad \overline{X} = A + \frac{\Sigma D}{n}$

So, $\quad \overline{X} = 52 + \frac{7}{8}$

$= 52 + 0.875 = 52.88$ or 53 runs.

Example 8: Deviations from 12.5 of ten different values are 6, −2, 3.5, 9, 8.7, −5.5, 14, 11.3, −6.8, −4.2, find the arithmetic mean.

Solution: Deviations from 12.5 are:

6, −2, 3.5, 9, 8.7, −5.5, 14, 11.3, −6.8, −4.2

Now, $\Sigma D = 34$. Also, $A = 12.5$, using the formula we have.

$\overline{X} = A + \frac{\Sigma D}{n}$

$= 12.5 + \frac{34}{10}$

or $\quad \overline{X} = 12.5 + 3.4 = 15.9$

Example 9: The heights (in inches) of 200 students are recorded in the following frequency distribution. Find the mean height of the student by short formula.

Height (x) (in inches)	51	52	53	54	55	56	57	58	59	60
Frequency (f)	2	5	8	24	55	45	38	16	6	1

Solution:

Heights (x) (in inches)	Frequency (f)	$A = 55$ $D = x - A$	fD
51	2	–4	–8
52	5	–3	–15
53	8	–2	–16
54	24	–1	–24
$A \leftarrow$ 55	55	0	0
56	45	1	45
57	38	2	76
58	16	3	48
59	6	4	24
60	1	5	5
Total	$\Sigma f = 200$		$\Sigma fD = 135$

Now, using the formula (ii), we get

$$\overline{X} = A + \frac{\Sigma fD}{\Sigma f}$$

$$\overline{X} = 55 + \frac{135}{200}$$

or $\overline{X} = 55 + 0.675$

$\therefore$ $\overline{X} = 55.68$ inches approx.

Hence, the mean height of the students is 55.68 inches.

Example 10: Ten students each from Grade-V section A and B of a well reputed school were taken randomly. Their weights were measured in kg. and recorded as given below:

Weights (kg) Section A	30	28	32	29.5	35	34	31	33	40	37.5
Weights (kg) Section B	35	31.5	34.5	35	32.8	38	29.5	36	36.5	34

(i) Compute the mean weight for section A and B.

(ii) Conclude which section is better on Average?

Solution: (i) We find arithmetic mean for both the sections by direct method. (Any method can be applied).

As number of observations $n = 10$

and $\overline{X}_{(A)} = \dfrac{\Sigma X_{(A)}}{n}$

$\therefore \overline{X}_{(A)} = \dfrac{330}{10} = 33$ kg

and $\overline{X}_{(B)} = \dfrac{\Sigma X_{(B)}}{n}$

$\therefore \overline{X}_{(B)} = \dfrac{342.8}{10} = 34.28$ kg

$X_{(A)}$	$X_{(B)}$
30	35
28	31.5
32	34.5
29.5	35
35	32.8
34	38
31	29.5
33	36
40	36.5
37.5	34
$\Sigma X_{(A)} = 330$	$\Sigma X_{(B)} = 342.8$

(ii) We have seen from the results that $\overline{X}_{(B)}$ is greater than $\overline{X}_{(A)}$. Therefore, we conclude that section B is better on the average.

12.2.2 Median

Median is the middle most value in an arranged (ascending or descending order) data set. Median is the value which divides the data into two equal parts i.e., 50% data is before the median and 50% data after it. Median is denoted by $\widetilde{X}$.

Median for ungrouped data

The median of n observations $x_1, x_2, \ldots, x_n$ is obtained as:

$$\text{Median } \left(\widetilde{X}\right) = \left(\frac{n+1}{2}\right)^{\text{th}} \text{ observation} \qquad \begin{pmatrix}\text{when } n \text{ is}\\ \text{odd number}\end{pmatrix}$$

$$\text{Median } \left(\widetilde{X}\right) = \frac{1}{2}\left(\left(\frac{n}{2}\right)^{\text{th}} \text{ observation} + \left(\frac{n+2}{2}\right)^{\text{th}} \text{ observation}\right) \qquad \begin{pmatrix}\text{when } n \text{ is}\\ \text{even number}\end{pmatrix}$$

Example 11: The following are the scores made by a batsman. Find the median of the data. 8, 12, 18, 13, 16, 5, 20.

Solution: Writing the scores in an ascending order, we have

<u>5, 8, 12, 13, 16, 18, 20</u>

Since, number of observations is odd i.e., $n = 7$

$$\text{Median } (\widetilde{X}) = \left(\frac{n+1}{2}\right)^{\text{th}} \text{ observation}$$

$$= \left(\frac{7+1}{2}\right)^{\text{th}} \text{ observation} = 4^{\text{th}} \text{ observation} = 13$$

Hence, 13 is the median of the given data.

Example 12: Following are the marks out of 100 obtained by 10 students in English. 23, 15, 35, 48, 41, 5, 8, 9, 11, 51. Find the median of the data.

Solution: Arranging the data in an ascending order.

5, 8, 9, 11, 15, 23, 35, 41, 48, 51

Since, number of observation is even, i.e., $n = 10$

$$\therefore \quad \text{Median } (\widetilde{X}) = \frac{1}{2}\left(\left(\frac{n}{2}\right)^{th} \text{observation} + \left(\frac{n+2}{2}\right)^{th} \text{observation}\right)$$

$$\text{As, } \frac{n}{2} = \frac{10}{2} = 5 \text{ and } \frac{n+2}{2} = \frac{12}{2} = 6$$

$$\therefore \quad \text{Median} = \frac{1}{2}\left[5^{th} \text{ observation} + 6^{th} \text{ observation}\right]$$

$$\text{or} \quad \text{Median} = \frac{1}{2}[15+23] = \frac{38}{2} = 19$$

Hence, 19 is the median of the data.

Median for Grouped Data

The median for grouped data is obtained by the following formula:

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

Where, ℓ = Lower class boundary of median class,

h = The size of class limits of median class,

f = Frequency of the median class,

n = Total frequency i.e., Σf,

and c = Cumulative frequency preceding the median class.

Remember the following points:

(i) The groups of classes must be in a continuous form i.e., we need class boundaries.

(ii) Make the column of cumulative frequencies ($c.f.$) from the column of frequencies.

(iii) Locate median class i.e., $\left(\frac{n}{2}\right)^{th}$ see value in c.f. column wherever it lies.

(iv) Underline the median class, then take the values of f and h of the median class thus obtained

Example 13: The heights of 100 athletes, measured to the nearest (inches) are given in the following table. Find the median.

Heights (in inches)	62.5–63.5	63.5–64.5	64.5–65.5	65.5–66.5	66.5–67.5	67.5–68.5	68.5–69.5	69.5–70.5	70.5–71.5
No. of Students	4	6	10	20	30	13	12	3	2

Solution: In the above data, class boundaries have already been given:

Heights (inches)	Frequency (f)	c.f.
62.5 – 63.5	4	4
63.5 – 64.5	6	6 + 4 = 10
64.5 – 65.5	10	10 + 10 = 20
65.5 – 66.5	20	20 + 20 = 40 → c
66.5 – 67.5	30	30 + 40 = 70 → Median class
67.5 – 68.5	13	13 + 70 = 83
68.5 – 69.5	12	12 + 83 = 95
69.5 – 70.5	3	3 + 95 = 98
70.5 – 71.5	2	2 + 98 = 100 → n
Total	$\Sigma f = 100$	---

Here, $n = 100$

so, $\dfrac{n}{2} = \dfrac{100}{2} = 50$

50^{th} item lies in the class boundaries 66.5 – 67.5.

$\ell = 66.5,\ h = 1,\ f = 30,\ c = 40$

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

$$= 66.5 + \frac{1}{30}(50 - 40) \qquad \text{(Putting the values)}$$

$$= 66.5 + \frac{10}{30}$$

$$= 66.5 + 0.33$$

$$\therefore \quad \text{Median} = 66.83 \text{ inches}$$

Example 14: Following are the weights (in kg) of 50 men. Find the median weight.

Weights (kg)	110 – 114	115 – 119	120 – 124	125 – 129	130 – 134
No. of men	5	12	23	6	4

Solution: As class boundaries are not given so, first of all we make class boundaries by the usual procedure.

Weight (kg)	Frequency(f)	Class Boundaries	c.f.	
110 – 114	5	109.5 – 114.5	5	
115 – 119	12	114.5 – 119.5	17 → c	
120 – 124	23	119.5 – 124.5	40 →	Median class
125 – 129	6	124.5 – 129.5	46	
130 – 134	4	129.5 – 134.5	50 → n	
Total	$\Sigma f = 50$	---	---	

Here $n = 50$ so, $\frac{n}{2} = \frac{50}{2} = 25$. 25^{th} item lies in 119.5 – 124.5.

$\ell = 119.5,\ h = 5,\ f = 23,\ c = 17$

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

$$= 119.5 + \frac{5}{23}(25 - 17) \qquad \text{(Putting the values)}$$

$$= 119.5 + \frac{40}{23} = 119.5 + 1.74$$

$\therefore$ Median = 121.24 kg

12.2.3 Mode

In a data the values (observation) which appears or occurs most often is called mode of the data. It is the most common value. Mode is denoted by $\hat{X}$.

Mode for Ungrouped Data

Example 15: The marks in mathematics of Jamal in eight monthly tests were 75, 76, 80, 80, 82, 82, 82, 85. Find the mode of the marks.

Solution: As 82 is repeated more than any other number so, clearly mode is 82.

Example 16: Ten students were asked about the number of questions they have solved out of 20 questions last week. Records were 13, 14, 15, 11, 16, 10, 19, 20, 18, 17. Find the mode of the data.

Solution: It is obvious that the given data contains no mode. It is ill-defined. Sometimes data contains several modes. If the data is: 10, 15, 15, 15, 20, 20, 20, 25, 32, then data contains two modes i.e., 15 and 20.

Example 17: A survey was conducted from the 15 students of a school and asked the students about their favourite colour.

The responses are: purple, yellow, purple, yellow, yellow, red, blue, green, yellow, yellow, red, blue, yellow, purple, green. Find mode of the data.

Solution: Mode is the most frequent colour.

Mode = yellow

So, the colour "yellow" is the mode of the given data.

> **Remember!**
> A data can has more than one mode. A data may or may not have a mode.

Mode for Grouped Data

Mode can be calculated by the following formula:

$$\text{Mode} = \ell + \frac{(f_m - f_1)}{(f_m - f_1)(f_m - f_2)} \times h$$

Where, ℓ = Lower class boundary of the modal class.

f = Frequency of the modal class.

f_1 = Frequency preceding the modal class.

f_2 = Frequency following the modal class and

h = Size of the modal class.

> **Note:**
> Mode cannot be easily calculated from the data presented in a frequency distribution. As it has no individual values, so we do not know which value appears most frequently. We only assume the class with the highest frequency as a modal class.

Example 18: Following are the heights in (inches) of 40 students in Grade - 8.

Heights (inches)	48 – 50	50 – 52	52 – 54	54 – 56	56 – 58	58 – 60
No. of students	5	7	10	9	6	3

Find mode of the above data.

Solution:

Heights (inches)	Frequency (f)
48 – 50	5
50 – 52	$7 \rightarrow f_1$
52 – 54	$10 \rightarrow f_m$
54 – 56	$9 \rightarrow f_2$
56 – 58	6
58 – 60	3
Total	$\Sigma f = 40$

> **Activity**
> Collect data of weights of 50 students. Make a frequency distribution and find mean, median and mode of the data.

In the above data, class boundaries have already been given. Using the formula for grouped data we find mode as:

Skill practice!

Find the mean, median and mode of the first twenty whole numbers.

$\ell = 52,\ h = 2,\ f_m = 10,\ f_1 = 7,\ f_2 = 9$

$$\text{Mode} = \ell + \frac{(f_m - f_1) \times h}{(f_m - f_1) + (f_m - f_2)}$$

or $$\text{Mode} = 52 + \frac{(10-7) \times 2}{(10-7) + (10-9)}$$

or $$\text{Mode} = 52 + \frac{3 \times 2}{3+1} = 52 + \frac{6}{4}$$

or $$\text{Mode} = 52 + 1.5 = 53.5 \text{ (inches)}$$

12.2.4 Weighted Mean

Arithmetic Mean is used when all the observations are given equal importance / weight but there are certain situations in which the different observations get different weights. In this situation, weighted mean denoted by $\overline{X}_w$ is preferred. The weighted mean of $X_1, X_2, X_3, \dots, X_n$ with corresponding weights $W_1, W_2, W_3, \dots, W_n$ is calculated as:

$$\overline{X}_w = \frac{W_1X_1 + W_2X_2 + W_3X_3 + \dots + W_nX_n}{W_1 + W_2 + W_3 + \cdots W_n} = \frac{\sum_{i=1}^{n} W_iX_i}{\sum_{i=1}^{n} W_i} = \frac{\sum WX}{\sum W}$$

Example 19: The following data describes the marks of a student in different subjects and weights assigned to these subjects are also given:

Mark(X)	74	78	74	90
Weights(W)	4	3	5	6

Find its weighted mean.

Solution: Weighted mean$\left(\overline{X}_w\right) = \dfrac{\Sigma WX}{\Sigma W}$

$$\overline{X}_w = \frac{4(74) + 3(78) + 5(74) + 6(90)}{4+3+5+6}$$

$$= \frac{296 + 234 + 370 + 540}{18} = \frac{1440}{18}$$

$$\overline{X}_w = 80$$

Example 20: A medicine company started marketing of a sample of medicine in seven different areas of a city. The company distributed the packets of medicine in each area of the city and the weight of each area based on the demand of the medicine. Find the mean and weighted mean of the given data.

Areas of a city	Number of packets (X)	Weights (W)
A	15	5
B	25	4
C	18	3
D	23	4
E	15	2
F	10	1
G	8	2

Solution: $\text{Mean} = \dfrac{\Sigma X}{n}$

$$= \frac{15+25+18+23+15+10+8}{7}$$

$$= \frac{114}{7} = 16.29 \approx 16 \text{ packets}$$

So, the average number of packets of the medicine distributed by the company per area is 16.

$$\text{Weighted mean} = \frac{\Sigma WX}{\Sigma W}$$

$$= \frac{15(5)+25(4)+18(3)+23(4)+15(2)+10(1)+8(2)}{5+4+3+4+2+1+2}$$

$$= \frac{377}{21} = 17.95 \approx 18$$

12.2.5 Real Life Situations Involving Mean, Weighted Mean, Median and Mode

Sales and Marketing

Example 21: A toy factory sold toys in a month. Consider the following data:

Class limits	10 – 20	20 – 30	30 – 40	40 – 50	50 – 60
f	15	28	45	29	20

(i) Calculate mean, median and mode of the number of toys sold by the factory.

(ii) Also tell the modal class of the distribution.

Solution: (i) **For mean**

Class limits	f	X	fX	c.f.	
10 – 20	15	15	225	15	
20 – 30	28	25	700	28 + 15 = 43	
30 – 40	45	35	1575	45 + 43 = 88	Modal class
40 – 50	29	45	1305	29 + 88 = 117	Median class
50 – 60	20	55	1100	20 + 117 = 137	
Total	**$\Sigma f = 137$**		**4905**		

$$\text{Mean}\left(\overline{X}\right) = \frac{\Sigma fx}{\Sigma f} = \frac{4905}{137} = 35.8 \approx 36$$

Average sale of the toys is 36.

For median: Here, $n = 137$, so, $\frac{137}{2} = 68.5$; 68.5 lies in 40 – 50.

$\ell = 40,\ h = 10,\ f = 29,\ n = 137,\ c = 48$

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

$$= 40 + \frac{10}{29}\left(\frac{137}{2} - 48\right)$$

$$= 40 + \frac{10}{29}(68.5 - 48)$$

$$= 40 + \frac{10}{29}(20.5)$$

$$= 40 + 7.07$$

$$\text{Median} = 47.07 \approx 47$$

Thus, median of the sold toys by the factory is 47.07.

For mode: $\ell = 30,\ h = 10,\ f_m = 45,\ f_1 = 28,\ f_2 = 29$

$$\text{Mode}\ (\hat{X}) = \ell + \frac{(f_m - f_1)}{(f_m - f_1) + (f_m - f_2)} \times h$$

$$= 30 + \frac{(45 - 28)}{(45 - 28) + (45 - 29)} \times 10$$

$= 30 + \frac{17}{17+16} \times 10$

$= 30 + \frac{17}{33} \times 10$

$= 30 + 5.15$

Mode $(\hat{X}) = 35.15 \approx 35$

Thus, mode of the sold toys by the factory is 35.

(ii) The modal class of sold toys by the factory is (30 – 40).

EXERCISE 12.2

1. Find the arithmetic mean in each of the following:
   (i) 4, 6, 10, 12, 15, 20, 25, 28, 30.
   (ii) 12, 18 19, 0, –19, –18, –12
   (iii) 6.5, 11, 12.3, 9, 8.1, 16, 18, 20.5, 25
   (iv) 8, 10, 12, 14, 16, 20, 22
2. Following are the heights in (inches) of 12 students. Find the median height.
   55, 53, 54, 58, 60, 61, 62, 56, 57, 52, 51, 63.
3. Following are the earnings (in Rs.) of ten workers:
   88, 70, 72, 125, 115, 95, 81, 90, 95, 90. Calculate
   (i) Arithmetic Mean (ii) Median (iii) Mode
4. The Marks obtained by the students in the subject of English are given below.

Marks obtained	15 – 19	20 – 24	25 – 29	30 – 34	35 – 39
Frequency	9	18	35	17	5

Find: (i) Arithmetic mean of their marks by direct and short formula.
(ii) Median of their marks.

5. Given below is a frequency distribution.

Class Interval	5 – 9	10 – 14	15 – 19	20 – 24	25 – 29
Frequency	1	8	18	11	2

Find the mode of the frequency distribution.

6. Ten boys work on a petrol pump station. They get weekly wages as follows:
   Wages (in Rs.) 4250, 4350, 4400, 4250, 4350, 4410,4500, 4300, 4500, 4390.
   Find the arithmetic mean by short formula, median and mode of their wages.

7. The arithmetic mean of 45 numbers is 80. Find their sum.

8. Five numbers are 1, 4, 0, 7, 9. Find their mean, median and mode.

9. A set of data contains the values as 148, 145, 160. 157, 156, 160.

Show that Mode > Median > Mean.

10. The monthly attendance of 10 students for their lunch in the hostel is recorded as: 21, 15, 16, 18, 14, 17, 15, 12, 13, 11.

Find the median and mode of the attendance. Also find the mean if $D = A - 20$.

11. On a prize distribution day, 50 students brought pocket money as under:

Rupees	5 – 10	10 – 15	15 – 20	20 – 25	25 – 30
Frequency (f)	12	9	18	7	4

(i) Find the median and mode of the above data.

(ii) Find the arithmetic mean of the data given above using coding method.

12. The arithmetic mean of the ages of 20 boys is 13 years, 4 months and 5 days. Find the sum of their ages. If one of the boys is of age exactly 15 years. What is the average age of the remaining boys?

13. Calculate the arithmetic mean from the following information:

(i) If $D = X - 140$, $\Sigma D = 500$ and $n = 10$

(ii) If $U = \dfrac{x-130}{6}$, $\Sigma U = -150$ and $n = 15$

(iii) If $D = x - 25$, $\Sigma fD = 300$ and $\Sigma f = 20$

(vi) If $U = \dfrac{x-120}{5}$, $\Sigma fU = 60$ and $\Sigma f = 100$

14. The three children Haris, Maham and Minal made the following scores in a game conducted by a group of teachers in the school.

Haris scores	50	55	70	85	90
Maham scores	75	60	60	45	53
Minal scores	80	77	66	42	48

It is decided that the candidate who gets the highest average score will be awarded rupees 1000. Who will get the awarded amount?

15. Given below is a frequency distribution derived by making a substitution as $D = X - 20$. Calculate the arithmetic mean.

D	–6	–4	–2	0	2	4	6
f	1	3	6	20	26	12	2

16. Being partners Hafsa and Fatima took part in a quiz programme. They made the following number of points 45, 51, 58, 61, 74. 48, 46 and 50. Compute the average number of points using deviation $D = x - 58$.

17. A person purchased the following food items:

Food item	Quantity (in Kg)	Cost per Kg (in Rs.)
Rice	10	96
Flour	12	48
Ghee	4	190
Sugar	3	49
Mutton	2	650

What is the weighted mean of cost of food items per kg?

18. For the following data, find the weighted mean.

Item	Quantity	Cost of item (in thousands)
Washing Machine	5	35
Heater	3	5
Stove	2	13
Dispenser	6	18

19. A company is planning its next year marketing budget across five years: yearly budgets (in million) are: 5, 7, 8, 6, 7. Find the average budget for the next year.

20. Ahmad obtained the following marks in a certain examination. Find the weighted mean if weights 5, 4, 2, 3, 2, 4 respectively are allotted to the subjects.

Urdu	English	Science	Math	Islamiyat	Computer
78	65	80	90	85	72

REVIEW EXERCISE 12

1. Four options are given against each statement. Encircle the correct option.

(i) Which data takes only some specific values?
(a) continuous data (b) discrete data
(c) grouped data (d) ungrouped data

(ii) The number of times a value occurs in a data is called:
(a) frequency (b) relative frequency
(b) class limit (d) class boundaries.

(iii) Midpoint is also known as:
(a) mean (b) median
(c) class limit (d) class mark

(iv) Frequency polygon is also drawn /constructed by using:
(a) histogram (b) bar graph
(c) class boundaries (d) class limit

(v) The difference between the greatest value and the smallest value is called:
(a) class limits (b) midpoint
(c) relative frequency (d) range

(vi) Measure of central tendency is used to find out the ________ of a data set.
(a) class boundaries (b) cumulative frequency
(c) middle or centre value (d) frequency

(vii) If the mean of 5, 7, 8, 9 and x is 7.5, what will be the value of x?
(a) 10 (b) 8 (c) 8.5 (d) 5.8

(viii) Find the mode of the given data: 2, 5, 8, 9, 0, 1, 3, 7 and 10
(a) 5 (b) 7 (c) 0 (d) no mode

(ix) In a data the values (observations) which appears or occurs most often is called:
(a) mean (b) mode
(c) median (d) weighted mean

(x) Find the median of the given data: 110, 125, 122, 130, 124, 127 and 120
(a) 124 (b) 120 (c) 125 (d) 127

2. Define the following:
(i) frequency distribution (ii) histogram (unequal class limits)
(iii) mean (iv) median

3. Following are the weights of 40 students recorded to the nearest (lbs).
138, 164, 150, 132, 144, 125, 149, 157, 146, 158, 140, 147, 136, 148, 152, 144, 168, 126, 138, 176, 163, 119, 154, 165, 146, 173, 142, 147, 135, 153, 140, 135 161, 145, 135, 142, 150, 156, 145, 128. (a) Make a frequency table taking size of class limits as 10 (b) Draw histogram. (c) Draw a frequency polygon of the given data.

4. From the table given below. Draw a frequency polygon on histogram for the given frequency distribution.

Weight (kg)	50 – 56	57 – 59	60 – 64	65 – 72	73 – 75	76 – 80
Frequency (f)	25	32	40	30	15	8

5. Given below are marks obtained by 45 students in the monthly test of Biology:

Marks	20 – 24	25 – 29	30 – 34	35 –39	40 –44	45 – 49
No. of students	05	08	12	15	03	02

With reference to the above table find the following:
(i) upper class boundary of the 5th class.
(ii) lower class boundaries of all the classes.
(iii) midpoint of all the classes.
(iv) the class interval with the least frequency.

6. Given below is frequency distribution.
Draw frequency polygon and histogram for the distribution.

Class limits	5 – 9	10 – 14	15 – 19	20 – 24	25 – 29	30 – 34
Frequency	1	8	18	11	2	5

7. For the following data, find the weighted mean.

Item	Quantity	Cost of item (Rs.)
Chair	20	500
Table	20	400
Black board	10	750
Tube light	25	230
Cupboard	09	950

8. A principal of a school allocates funds of Rs.50, 000 to five different sectors:
(i) chairs: Rs. 15000 (ii) tables: Rs. 12,000
(iii) black boards: Rs.6,000 (iv) room renovation: Rs. 10,000
(v) gardening: Rs. 7,000
Find the average of funds allocation in each sector of the school.

9. The marks of a student Saad in six tests were 84, 91, 72, 68, 87, 78. Find the arithmetic mean of his marks.

10. Adjoining distribution showed maximum load (in kg) supported by certain ropes. Find the mean load using short method.

Max-Load kg	93 – 97	98 – 102	103 – 107	108 – 112	113 – 117	118 – 122
No. of ropes	2	5	8	12	6	2

11. Usman rolled a fair dice eight times. Each time their sum was recorded as 8, 5, 6, 6, 9, 4, 3, 11. Find the median and mode of the sum.

12. Two partners Mr. Aslam and Mrs. Kalsoom run a company. In the following data the weekly wages (in Rs.) of employees who work in the company are given:

Wages (Rs.)	600 – 700	700 – 800	800 – 900	900 – 1000	1000 – 1100
Employees	3	5	7	21	11