Unit - 12 ▶ Information Handling

After completion of this unit, the students will be able to:

- Define frequency, frequency distribution.
- Construct frequency table.
- Construct a histogram representing frequency table.
- Describe measures of central tendency.
- Calculate mean (average), weighted mean, median and mode for ungrouped data.
- Solve real life problems involving mean (average), weighted mean, median and mode.

12.1 FREQUENCY AND FREQUENCY DISTRIBUTION.

12.1.1 Definitions

- **Frequency**

The number of times a value occurs in a data is called the frequency of that value.

For example: The marks obtained out of *10* in a test by *15* students of a class are as follows:

3, 5, 7, 10, 7, 9, 3, 7, 5, 4, 6, 8, 7, 5, 2.

- The data consists of *15* values.
- Some of the values are occurring more than once e.g., *3, 5, 7.*
- The frequency of *3* marks is *2*.
- The frequency of *5* marks is *3*.
- The frequency of *7* marks is *4*.
- All other values have frequency *1*.

- **Frequency Distribution**

To write a data in the form of a table in such a way that the frequency of each class can be observed at once is called its **frequency distribution.**

10.1.2 Construction of Frequency Distribution Table

Let us consider the given weights in *kg* of *50* students selected from a school:

35, 30, 32, 36, 31, 40, 35, 42, 35, 45, 37, 41, 33, 37, 30, 28, 29,
30, 32, 33, 31, 35, 36, 30, 28, 37, 39, 28, 31, 34, 39, 45, 38, 36,
35, 28, 31, 34, 30, 41, 35, 36, 41, 28, 31, 34, 30, 29, 28, 37

We note that the weights of the selected students range from *28 kg* to *45 kg*. We arrange the data in groups in the form of a table as below:

Class Interval	Frequency
28 − 30	*14*
31 − 33	*9*
34 − 36	*13*
37 − 39	*7*
40 − 42	*5*
43 − 45	*2*
Total:	*50*

In the above table the frequency of the group of students whose weights from *28kg* to *30kg* are *14* and similarly the other class frequencies can easily be seen.

(i)   Look for the largest value and the smallest value i.e., *45* and *28* respectively.

(ii)   Number of classes to be made is *6*.

(iii)   For finding the size of class interval use the formula.

$$Size\ of\ class\ interval = \frac{largest\ value - smallest\ value}{number\ of\ classes}$$

$$= \frac{45 - 28}{6} = \frac{17}{6}$$

$$\simeq 2.8 \simeq 3$$

Example 1:

Listed below are the scores of *50* students in a *60* marks test.

25, 33, 26, 34, 28, 35, 29, 36. 30, 54, 30, 39, 36, 37, 39, 40, 37, 34,

27, 41, 37, 41, 38, 42, 48, 51, 40, 51, 43, 40, 41, 39, 48, 51, 53, 41,

37, 52, 28, 46, 44, 37, 39, 52, 51, 40, 45, 46, 43, 53

Make a frequency distribution table taking 6 classes of equal size by tally marks.

Solution:

Lowest value   =   *25*

Highest value   =   *54*

Total classes to be made   =   *6*

We take the size of class   = $\dfrac{54 - 25}{6}$

$$= \frac{29}{6} = 5\ (approx.)$$

Class Interval	Tally Mark	Frequency
25 − 29	ꟽ I	*6*
30 − 34	ꟽ	*5*
35 − 39	ꟽ ꟽ III	*13*
40 − 44	ꟽ ꟽ II	*12*
45 − 49	ꟽ	*5*
50 − 54	ꟽ IIII	*9*
	Total:	*50*

Example 2: The number of units of electricity consumed by *31* households are listed below. Construct a frequency table with *10* classes.

727, 773, 859, 711, 860, 747, 862, 738, 774, 852, 791, 836, 834,

752, 828, 792, 908, 839, 752, 715, 880, 838, 852, 816, 751, 834,

818, 835, 831, 778, 837

Solution: Lowest value   = *711*

Highest value = *908*

Total classes to be made   = *10*

$$\text{Size of class interval} = \frac{908 - 711}{10}$$

$$= \frac{197}{10} = 19.7 \simeq 20$$

| Class Interval	Tally Mark	Frequency
711 − 730	|||	*3*
731 − 750	||	*2*
751 − 770	|||	*3*
771 − 790	|||	*3*
791 − 810	||	*2*
811 − 830	|||	*3*
831 − 850	ℕℕ |||	*8*
851 − 870	ℕℕ	*5*
871 − 890			*1*
891 − 910			*1*
	Total:	*31*

12.1.3 Construction of Histogram

We are familiar with pie and bar graphs. Another common graphic way of presenting data is by means of a histogram. A histogram is similar to bar graph but it is constructed for a frequency table.

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.

To draw a histogram from a grouped data, the following steps are followed.

(i) Draw *X-axis* and *Y-axis*.

(ii) Mark class boundaries of the classes along *X-axis*.

(iii) Mark frequencies along *Y-axis*.

(iv) Draw a bar for each class so that the height of the bar drawn for each class is equal to the frequency of the class.

The graph is shown below:
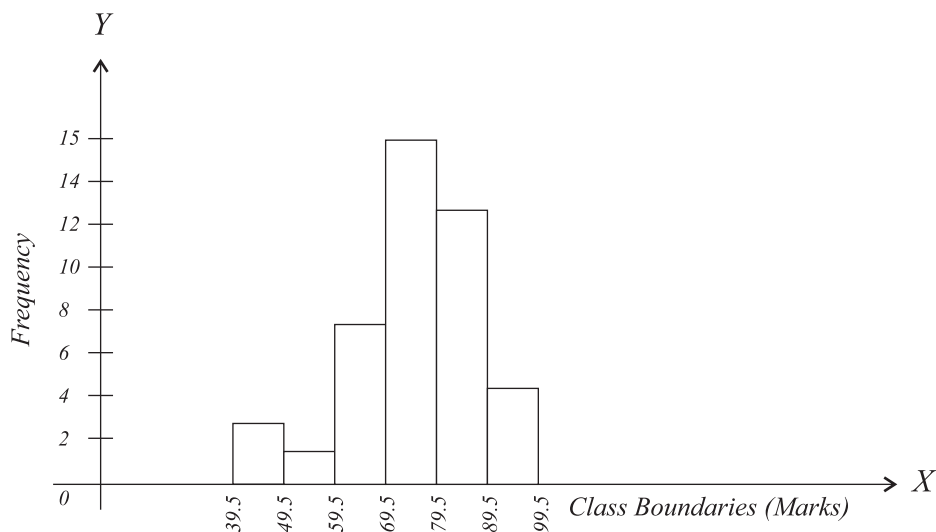
Example 3: The detail of distances travelled daily by the residents of a locality are given below. Construct a histogram for the following frequency table.

Distance travelled (in *km*)	*1 − 8*	*9 − 16*	*17 − 24*	*25 − 32*	*33 − 40*
Number of persons	*15*	*12*	*7*	*4*	*2*

Solution: Frequency distribution table is:

Distance travelled (km)	Class boundries	Frequency (No. of Persons)
1 − 8	0.5 − 8.5	15
9 − 16	8.5 − 16.5	12
17 − 24	16.5 − 24.5	7
25 − 32	24.5 − 32.5	4
33 − 40	32.5 − 40.5	2
	Total:	*40*

Histogram:
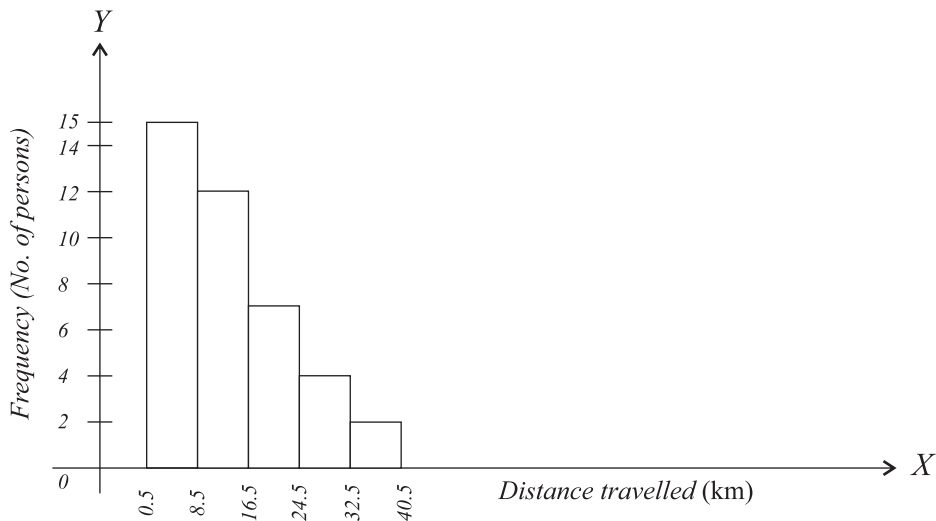
EXERCISE 12.1

1. The following data displays the number of draws of different categories of bonds.

   35, 55, 64, 70, 99, 89, 87, 65, 67, 38, 62, 60, 70, 78, 69, 86, 39, 71, 56, 75,

   51, 99, 68, 95, 86, 53, 59, 50, 47, 55, 81, 80, 98, 51, 63, 66, 79, 85, 83, 70

   Construct a frequency distribution table for the above data, with seven classes of equal size and of class interval 10.

2. Listed below are the number of electricity units consumed by *50* households in a low income group locality of Lahore.

   55, 45, 64, 130, 66, 155, 80, 102, 62, 60, 101, 58, 75, 81, 111, 90, 55, 151,

   66, 139, 77, 99, 67, 51, 50, 125, 83, 55, 136, 91, 86, 54, 78, 100, 113, 93,

   104, 111, 113, 96, 96, 87, 109, 94, 129, 99, 69, 83, 97, 97

   With *12* classes of equal width of *10*, construct a frequency table for the electricity units consumed.

3. The following list is of scores in a mathematics examination. Using the starting class *40 − 44* , set up a frequency distribution. List the class boundaries and class marks.

   63, 88, 79, 92, 86, 87, 83, 78, 40, 67, 68, 76, 46, 81, 92, 77, 84, 76, 70, 66,

   77, 75, 98, 81, 82, 81, 87, 78, 70, 60, 94, 79, 52, 82, 77, 81, 77, 70, 74, 61

4. Construct a frequency distribution for the following numbers using *1 − 10* as the starting class. List the class boundaries.

   54, 67, 63, 64, 57, 56, 55, 53, 53, 54, 44, 45, 45, 46, 47, 37, 23, 34, 44, 27,

   36, 45, 34, 36, 15, 23, 43, 16, 44, 34, 36, 35, 37, 24, 24, 14, 43, 37, 27, 36,

   33, 25, 36, 26, 5, 44, 13, 33, 33, 17

5. Following are the number of days that 36 tourists stayed in some city.

   1, 6, 16, 21, 41, 21, 5, 31, 20, 27, 17, 10, 3, 32, 2, 48, 8, 12, 21, 44, 1, 36, 5,

   12, 3, 13, 15, 10, 18, 3, 1, 11, 14, 12, 64, 10.

   Construct a frequency distribution starting with the class *1 − 7*.

6. Construct a histogram for each of the frequency tables in questions *1 − 5*.

12.2   MEASURES OF CENTRAL TENDENCY

In the previous section we have learnt to arrange data into distribution table to understand the given data easily. sometimes, the volume of data is large and it is very difficult to compare, understand and analyze. Then there is need to make that data comparable to avoid difficulty and complexity.

12.2.1 Description of Measures of Central Tendency

The **Measures of Central Tendency** are the Concept of Averages, Mean, Mode and Median.

12.2.2 Calculation of Measures of Central Tendency

- **Mean (Average)**

Let x_1, x_2, ... , x_n be n given quantities. Then their average is the value presenting the tendency of these quantities and is called their Mean value or Mean. It can be calculated by the formula:

$$\overline{X} = \frac{x_1 + x_2 + ... , + x_n}{n}$$

$$\overline{X} = \frac{sum\ of\ all\ values}{number\ of\ values}$$

Example 4:

The scores of a student in eight papers are *58, 72, 65, 85, 94, 78, 87, 85*. Find the mean score.

$$\overline{X} = \frac{58 + 72 + 65 + 85 + 94 + 78 + 87 + 85}{8}$$

$$\overline{X} = \frac{624}{8} = 78$$

Hence, mean score is *78*

- **Weighted Mean**

When all values of given data have same importance, then we use mean. But when different values have different importance then these values are known as weights.

If $x_1, x_2, x_3, ... , x_n$ have the weights $w_1, w_2, w_3, ... , w_n$ then:

$$Weighted\ Mean = \overline{X}w = \frac{w_1 x_1 + w_2 x_2 + w_3 x_3 , ... , + w_n x_n}{w_1 + w_2 + w_3 , ... , + w_n} = \frac{\sum xw}{\sum w}$$

Example 5:

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

Marks (x)	74	78	74	90
Weights (w)	4	3	5	6

Find its weighted mean:

Solution:

$$\text{Weighted Mean} = \overline{Xw} = \frac{4(74) + 3(78) + 5(74) + 6(90)}{4 + 3 + 5 + 6}$$

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

$$= \frac{1440}{18} = 80 \text{ marks}$$

- **Median**

  If a data is arranged in ascending or descending order then median of the data is:

  (a) The middle value of the data, if it consists of odd number of values.

  (b) The mean of the two middle values of the data is the Median of the data if the number of values in a data is even.

Example 6: The weights in kg of 9 students are as under, find the median:

29, 32, 45, 27, 30, 47, 35, 37, 33

Solution:

Arranging these values in descending order:

$\overline{47, 45,}$ $\overline{37, 35,}$ $33,$ $\overline{32, 30,}$ $\overline{29, 27}$

The middle value is *33*

So, Median $= 33\,Kg$

• **Mode**

Mode is the value that occurs most frequently in a data . In case no value is repeated in a data, then the data has no mode. If two or more values occur with the same greatest frequency, then each is a mode.

Example 7: Find the mode of the given data:

1, 2, 5, 7, 8, 2, 2, 4, 3, 5, 7

Solution: The value *2* is repeated the most, so *2* is the mode of this data.

Example 8: Find the mode of the given data:

2, 4, 6, 8, 10, 12, 14, 16, 20

Solution: This data has no mode because no value is repeated in the given data

Example 9: Find the mode of the data given below:

1, 2, 2, 2, 3, 4, 5, 5, 5, 6, 7

Solution: Since *2* is repeated *3* times and *5* is also repeated *3* times so this data has two modes i.e., *2* and *5*.

Remember that:

(i) A data can have more than one Mode.
(ii) A data may or may not have a Mode.

12.2.3 Real life problems involving Mean, Weighted Mean, Median and Mode

Example 10:

The heights of *12* students (in centimeters) of 8th class are given below:
148, 144, 145, 146, 148, 150, 145, 155, 151, 152, 145, 149

(i) Find the average height of a student.

(ii) Find the most common height.

(iii) Find the Median height.

Solution:

Arrange the given data in ascending order:

144, 145, 145, 145, 146, 148, 148, 149, 150, 151, 152, 155

(i) $Mean\ (average) = \dfrac{144 + 145 + 145 + 145 + 146 + 148 + 148 + 149 + 150 + 151 + 152 + 155}{12}$

$= \dfrac{1778}{12} = 148.16$

Therefore, average height of a student is *148.16cm*

(ii) The most occurred value is *145* (*3* times)

(iii) The total number of values is *12*. So, 6th and 7th terms are the middle values of data.

$$\therefore \quad Median = \left(\dfrac{6^{th}\ term + 7^{th}\ term}{2} \right)$$

$$= \dfrac{148 + 148}{2} = \dfrac{296}{2} = 148$$

Therefore, the median is *148cm*

EXERCISE 12.2

1. Compute the mean, median and mode of the following data:
    (i)   *10, 8, 6, 0, 8, 3, 2, 5, 8, 4*
    (ii) *1, 3, 5, 3, 5, 3, 7, 5, 7, 5, 7*
    (iii) *5 , 4 , 1 , 4 , 0 , 3 , 4 , 119*
    (iv) *62, 90, 71, 83, 75*
    (v)   *45, 65, 80, 92, 80, 75, 56, 96, 62, 78*
    (vi) Number of letters in first *20* words in a book.

    3, 2, 5, 3, 3, 2, 3, 3, 2, 4, 2, 2, 3, 2, 3, 5, 3, 4, 4, 5

    (vii) The number of calories in nine different beverages of *250mm* bottles:

    99, 106, 101, 103, 108, 107, 107, 106, 108

(viii) Number of rooms in *15* houses of a locality city

5, 9, 8, 6, 8, 7, 6, 7, 9, 8, 7, 9, 7, 8, 5

(ix) Number of books in *10* school libraries, (in hundreds)

78, 215, 35, 267, 39, 17, 418, 286, 335, 50.

(x) Cost per day on a patient in *10* private hospitals (in rupees)

4125, 2500, 3115, 6580, 7150, 3750, 5920, 4575, 3225, 2500

2. A person purchased the following food items:

Food items	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 average cost of food items per *kg*?

3. The following distances (in *km*) were travelled by *40* students to reach their school.

2, 8, 1, 5, 9, 5, 14, 10, 31, 20, 15, 4, 10, 6, 5, 10, 5, 18, 12, 25, 30, 27, 20, 3, 9, 15, 15, 18, 10, 1, 1, 6, 25, 16, 7, 12, 1, 8, 21, 12.

Compute the mean, median and mode of the distances travelled.

4. Following table lists the size of *127* families:

Size of Family	*2*	*3*	*4*	*5*	*6*	*7*	*8*
Frequency	*51*	*31*	*27*	*12*	*4*	*1*	*1*

Compute the mean, median and mode.

5. Find the class mark and mean of the following frequency table:

Class Interval	*0 - 39*	*40 - 79*	*80 - 119*	*120 - 159*	*160 - 199*
Frequency	*17*	*41*	*80*	*99*	*4*

6. Find the mean of the following frequency table:

Class interval	1 - 5	6 - 10	11 - 15	16 - 20	21 - 25	26 - 30	31 - 35	36 - 40
Frequency	19	24	18	21	23	20	16	15

7. The diagram illustrates the number of children per family of a sample of *100* families in a certain housing estate:

(a) State the modal number of children per family.

(b) Calculate the mean number of children per family.

(c) Find the median number of children per family.
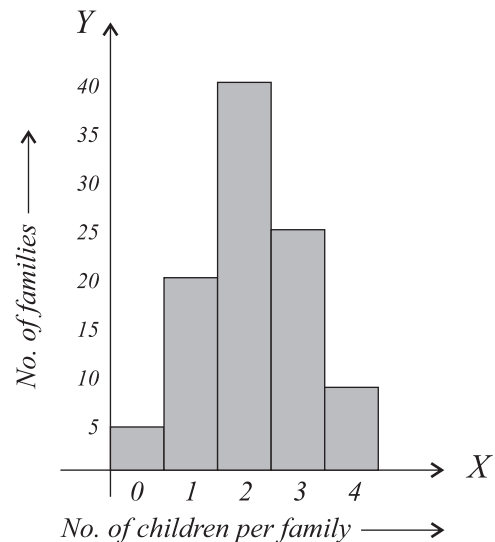
REVIEW EXERCISE 12

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

i. What is the missing number of data *14, 24, _____, 18, 30,* when mean is *23*.

    (a) *28*   **(b)** *29*   **(c)** *30*   **(d)** *31*

ii. What is the missing number of data *40, 28, 16, 18, 37, 20, _____,*
35, when median is *26*.

    (a) *20*   **(b)** *22*   **(c)** *24*   **(d)** *28*

iii. A number which indicates how often a value occurs is called:

    (a) Frequency **(b)** Mode **(c)** Median   **(d)** Average

iv. An arrangement of the values that one or more variables taken in data is called:

    (a) Frequency   **(b)** Frequency distribution
    (c) Median   **(d)** Mode

v. A representation of a frequency distribution by means of rectangles whose widths represent class intervals and whose areas are proportional to the corresponding frequencies is called:

    (a) Histogram **(b)** Bar chart **(c)** Pie chart **(d)** Line graph

vi. A measure of central tendency that attempts to describe a data by identifying the central position within that data:

(a) is a single value   **(b)** are multiple values

(c) are duplicate values   **(d)** are repeating values

vii. The statistical measure that identifies a single value as representative of an entire distribution is called:

(a) frequence distribution   **(b)** histogram

(c) mean   **(d)** central tendency

viii. The value which occupies the middle position when all the observations are arranged in an ascending/descending order is called:

(a) Frequency distribution   **(b)** Median

(c) Mode   **(d)** Mean

ix. The value that occurs most frequently in the data is called:

(a) Frequency distribution   **(b)** Median

(c) Mode   **(d)** Mean

2. Calculate the mean, median and mode for each set of data given below:

(a) *3, 6, 3, 7, 4, 3, 9*   **(b)** *11, 10, 12, 12, 9, 10, 14, 12, 9*

(c) *2, 9, 7, 3, 5, 5, 6, 5, 4, 9*   **(d)** *6, 8, 11, 5, 2, 9, 7, 8*

(e) *153.8, 154.7, 156.9, 154.3, 152.3, 156.1, 152.3*

3. Test scores of a class of *20* students are as follows:

93, 84, 97, 98, 100, 78, 86, 100, 85, 92, 72, 55, 91, 90, 75, 94, 83, 60, 81, 95

Draw a frequency distribution table and histogram for grouped data.

4. The price of *10* litre of drinking water was recorded at several stores, and the results are displayed in the table below:

Price (*Rs.*)	Frequency
74	*1*
75	*2*
76	*8*
77	*10*
78	*2*
79	*1*
80	*1*

Find the mean, median and mode of the price.

SUMMARY

- Frequency is a number which indicates how often value occurs.

- A frequency distribution is a summary of how often different scores occur within a sample of scores.

- A frequency distribution table is one way we can organize data so that it makes more sense. We could draw a frequency distribution table, which will give a better picture of our data than a simple list.

- A histogram is a representation of a frequency distribution by means of rectangles whose widths represent class intervals and whose areas are proportional to the corresponding frequencies.

- A measure of central tendency is a single value that attempts to describe data by identifying the central position within that data.

- Central tendency is defined as "the statistical measure that identifies a single value as representative of an entire distribution".

- Arithmetic mean (or, simply, "mean or average") is the most popular and well known measure of central tendency.

- The mean is equal to the sum of all the values in the data divided by the number of values in the data:

$$Mean = \frac{sum\ of\ data}{number\ of\ observations} \quad or \quad \bar{x} = \frac{x_1 + x_2 + x_3 + ... + x_n}{n}$$

- Median is the value which occupies the middle position when all the observations are arranged in an ascending / descending order.

    (a)   The middle value of the data, if it consists of odd number of values.

    (b)   The mean of the two middle values of the data is the Median of the data if the number of values in a data is even.

- Mode is defined as the value that occurs most frequently in the data. Some data do not have a mode because each value occurs only once.