# Unit No. 12

# INFORMATION HANDLING

## Basic Concepts

## Introduction (Information and Data):

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**.

## Source of Data:

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 at the same can be obtained directly from the field according to needs.

## Role of 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.

## Frequency:

Frequency, in the context of data, refers to the number of times a particular value occurs in a data set. It is often denoted by "f".

## Information Handling:

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

## Classification of 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.

## 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. Classes Tally marks No. of toys sold. Ungrouped data is also known as raw data.

**Frequency Distribution:**

A distribution or table that represents classes or groups along with their respective class frequencies is called frequency distribution.

**The difference between lower class limit of the Class boundaries:**

**Upper Class Boundaries:**

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

**Lower Class Boundaries:**

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

**Do you know?**

Continuous data is mostly represented by using histogram and frequency polygon.

**Histogram:**

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.

**Frequency Polygon:**

A frequency polygon is 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:**

Midpoint is the average value of the lower and upper class limits.

$$\text{Midpoint} = \frac{lower\ class\ limit\ +\ upper\ class\ limit}{2}$$

**Measures of Central Tendency:**

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

The following measure of central tendency will be discussed: (i) Arithmetic Mean (A.M) (ii) Median (iii) Mode (iv) Weighted mean

**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.

**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.

**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., $(\frac{n}{2})^{th}$ 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.

**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.