RVS College of Arts and Science, Inamkulathur, Trichy 620009.
(Affiliated to Bharathidasan University)

Subject: Business Tools for Decision Making
Class: II B.Com.

MEASURES OF CENTRAL TENDENCY

Def: - Given a series of observations measured on a quantitative variable, there is a general tendency among the values to cluster around a central value. Such clustering is called central tendency and measures put forward to measure these tendency are called measures of central tendency or averages.

Average: - Average is a value which is the representative of a set of data.

Functions or objects of an average:

- It facilitates quick understanding of complex data.
- It facilitates comparison
- To know about the universe from the sample
- To get the single value that describes the characteristic of the entire group.

Requisites of a good average:

- It should be easy to understand
- It should be easy to calculate.
- It should be based on all the observations of the data.
- It should not be affected by the extreme values.
- It should be strictly defined, so that it has one and only one interpretation.
- It should be capable of further algebraic treatment.
- It's definition should be in the form of a mathematical formula.
- It should have sampling stability.

Types of averages (or) Measures of central tendency:-

The following are the important types of averages.

1) Arithmetic mean.
    - → Simple arithmetic mean.
    - → Weighted arithmetic mean.
2) Geometric mean.
    - → Simple geometric mean.
    - → Weighted geometric mean
3) Harmonic mean
4) Median
5) Quartiles, deciles and percentiles.
6) Mode

The first three are called **'mathematical averages'** where as other three are called **'measures of location'** or **'measures of position'** or **'positional averages'**

Arithmetic mean

The most popular and widely used measure of representing the entire data by one value is what a layman call an 'average' and what the statisticians called is 'arithmetic mean'. Arithmetic mean may be

1) Simple arithmetic mean
2) Weighted arithmetic mean.

1) Simple arithmetic mean:

Case-(i) Calculation of simple arithmetic mean -Individual series:

The process of computing mean in case of individual series (i.e. where frequencies are not given) is very simple. Add together the various values of the variable and divide the total by the no of items.

Direct method:

If $x_1, x_2, x_3, ...x_n$ are 'n' individual observed values of a variable X, then the A.M is denoted by $\bar{x}$ and is defined as

$$\bar{x} = \frac{x_1 + x_2 + x_3 + ... + x_n}{n}$$

$$\Rightarrow \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Short cut method:

Under this method the formula for calculating mean is $\Rightarrow \bar{x} = A + \frac{1}{n}\sum_{i=1}^{n} d_i$

Where A=assumed mean d_i = deviations of items taken from the assumed mean.

n = Number of observations

Note: Any value whether existing in the data or not can be taken as the assumed mean and the final answer would be the same. However it's better to take assumed mean nearer to the actual mean for lesser calculations.

Example: The following table gives the monthly income of 10 employees in an office. Income (in Rs): 1780, 1760, 1690, 1750, 1840, 1920, 1100, 1810, 1050, 1950.
Calculate the A.M.

Sol.

Calculation of arithmetic mean

Employee	Income(in Rs) x	d=(x-1800)
1	1780	-20
2	1760	-40
3	1690	-110
4	1750	-50
5	1840	+40
6	1920	+120
7	1100	-700
8	1810	+10
9	1050	-750
10	1950	+150
n=10	$\sum x = 16{,}650$	$\sum d = -1350$

(i)Direct method:

$$\text{Mean } \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{16{,}650}{10} = 1665.$$

Thus the average income is Rs. 1665.

(ii) Shortcut method:

$$\text{Mean } \bar{x} = A + \frac{1}{n}\sum_{i=1}^{n} d_i$$

Since A=1800

$$\bar{x} = 1800 - \frac{1350}{10}$$

$$= 1800-135$$

$$\Rightarrow \bar{x} = 1665.$$

Thus the average income is Rs. 1665.

Case-(ii) Calculation of A.M -Discrete series (or) ungrouped frequency distribution:

Direct method:

If $x_1, x_2, x_3, ... x_n$ are 'n' individual observed values of a variable X occurred with frequencies $f_1, f_2, f_3, ..., f_n$ then the arithmetic mean is defined as

$$\Rightarrow \bar{x} = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i}$$

Short cut method:

Under this method the formula for calculating mean is

$$\Rightarrow \bar{x} = A + \frac{\sum_{i=1}^{n} f_i d_i}{\sum_{i=1}^{n} f_i}$$

Where A=assumed mean d_i= deviations of items taken from the assumed mean.

n = number of observations f_i = frequency of the ith observation.

Example: From the following data of the marks obtained by 60 students of a class.

Marks	20	30	40	50	60	70
No of students	8	12	20	10	6	4

Sol :

Calculation of Arithmetic mean

Marks (x)	No. of students(f)	fx	d=(x-40)	fd
20	8	160	-20	-160
30	12	360	-10	-120

40	20	800	0	0
50	10	500	10	100
60	6	360	20	120
70	4	280	30	120
	N=60	$\sum fx = 2{,}460$		$\sum fd = 60$

(i) **Direct method**:

Here N= total frequency=60

$$\text{Mean } \bar{x} = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i}$$

$$\Rightarrow \bar{x} = \frac{2460}{60} = 41$$

Hence the average marks = 41.

(ii) **Short cut method:**

$$\text{Mean } \bar{x} = A + \frac{\sum_{i=1}^{n} f_i d_i}{\sum_{i=1}^{n} f_i}$$

Since A = 40, $\bar{x} = 40 + \frac{60}{60} = 40 + 1 = 41$.

Hence the average marks = 41.

Case-(iii) Continuous series (or) Grouped frequency distribution:

Direct method:

If $m_1, m_2, m_3, \ldots\ldots., m_n$ are 'n' mid points of the classes and $f_1, f_2, f_3, \ldots\ldots\ldots., f_n$ be the corresponding frequencies then the arithmetic mean is defined as

$$\Rightarrow \bar{x} = \frac{\sum_{i=1}^{n} f_i m_i}{\sum_{i=1}^{n} f_i}$$

Where m_i is the mid point of the ith class.

Short cut method:

Under this method the formula for calculating mean is

$$\Rightarrow \bar{x} = A + \frac{\sum_{i=1}^{n} f_i d_i}{\sum_{i=1}^{n} f_i}$$

Where A=assumed mean $d_i = m_i - A$

n = number of observations

f_i = frequency of the ith observation.

Step deviation method:

Under this method the formula for calculating mean is

$$\Rightarrow \bar{x} = A + \frac{\sum_{i=1}^{n} f_i d_i}{\sum_{i=1}^{n} f_i} \times c$$

Where

$$d_i = \frac{m_i - A}{c}$$

A=assumed mean

C= class length

n = number of observations

f_i = frequency of the ith observation.

Properties of arithmetic mean:

- The sum of the deviations of the items taken from their arithmetic mean is zero.

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

- The sum of the squares of the deviations of the items taken from arithmetic mean is minimum.

  i.e. $\sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2$ is minimum

  i.e. $\sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2 < \sum_{i=1}^{n} \left(x_i - A\right)^2$, Where A is any arbitrary value.

- If a series of n observations consists of two components having n_1 and n_2 observations ($n_1 + n_2 = n$), and means $\overline{x_1}$ and $\overline{x_2}$ respectively then the Combined mean $\bar{x}$ of n observations is given by

  Combined mean $\bar{x} = \frac{n_1 \overline{x_1} + n_2 \overline{x_2}}{n_1 + n_2}$

- Let the values in $x-$series be $x_1, x_2, x_3, \ldots\ldots\ldots\ldots x_n$ and values of $y-$series be $y_1, y_2, y_3, \ldots\ldots\ldots\ldots y_n$ and define $z_i = x_i \pm y_i$ then $\bar{z} = \bar{x} \pm \bar{y}$.

Weighted arithmetic mean:

In calculating simple arithmetic mean, it is assumed that all the items in the series carry equal importance. But in practice, there are many cases where relative importance should be given to different items. Hence the limitation of not giving equal importance in case of simple arithmetic mean can be eliminated by giving relative importance i.e. weights by computing weighted arithmetic mean.

Formula:

If $w_1, w_2, w_3, \ldots\ldots\ldots\ldots w_n$ are weights of n observations in a series $x_1, x_2, x_3, \ldots\ldots\ldots\ldots x_n$ then the weighted mean is calculated as

$$\bar{x}_w = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i}$$

Note: If the weights of all the observations are equal i.e. $w_1 = w_2 = w_3, \ldots\ldots\ldots\ldots = w_n = w$ then the weighted A.M is equal to simple A.M i.e. $\bar{x}_w = \bar{x}$.

Merits of Arithmetic Mean:

- Arithmetic mean is most popular among averages used in statistical analysis.
- It is very simple to under stand and easy to calculate.
- The calculation of A.M is based on all the observations in the series.
- The A.M is responsible for further algebraic treatment.
- It is strictly defined.
- It provides a good means of comparison.
- It has more sampling stability.

Demerits of Arithmetic Mean:

- The A.M is affected by the extreme values in a series.
- In case of a missing observation in a series it is not possible to calculate the A.M.
- In case frequency distribution with open end classes the calculation of A.M is theoretically impossible.
- The arithmetic mean is an unsuitable average for qualitative data.

Median

The median is the middle most or central value of the observations made on a variable when the values are arranged either in ascending order or descending order.

As distinct from the arithmetic mean which is calculated from each and every item in the series, the median is what is called 'positional average'. The term position refers to the place of value in a series. The place of the median in a series is such that an equal number of items lie on either side of it, i.e. it splits the observations into two halves.

Calculation of median – Individual series:

Step-1: Arrange the data in ascending or descending order of magnitude.

Step-2:

Case-i If the number of observations is odd then median is the $\frac{n+1}{2}th$ observation in the arranged order.

Case-ii If the number of observations is even then the median is the mean of $\frac{n}{2}th$ and $\left(\frac{n}{2}+1\right)th$ observations in the arranged order.

Calculation of median-Discrete series:

Step-i: Arrange the data in ascending or descending order of magnitude.

Step-ii: find out the cumulative frequency (c.f)

Step-iii: Apply the formula: Median =size of $\frac{N+1}{2}$

Step-iv: Now look at the cumulative frequency column and find that total which is either equal to $\frac{N+1}{2}$ or next higher to that and determine the value of the variable corresponding to it. That gives the value of median.

Computation of median-Continuous series:

The median of a continuous series can be calculated by the below interpolation formula.

$$\text{Median} = l_1 + \frac{(l_2 - l_1)}{f}(m - c)$$

Where l_1 =lower limit of the median class.

l_2 = upper limit of the median class

f = frequency corresponding to the median class

m = $\frac{N}{2}$, N = total frequency

c = cumulative frequency of the class preceding to the median class.

Mathematical property of median:

The sum of the deviations of the items from median, ignoring signs is the least.

i.e $\sum_{i=1}^{n}|x_i - md|$ is least.

Merits:

- The median is useful in case of frequency distribution with open-end classes.
- The median is recommended if distribution has unequal classes.
- Extreme values do not affect the median as strongly as they affect the mean.
- It is the most appropriate average in dealing with qualitative data.
- The value of median can be determined graphically where as the value of mean can not be determined graphically.
- It is easy to calculate and understand.

Demerits:

- For calculating median it is necessary to arrange the data, where as other averages do not need arrangement.
- Since it is a positional average its value is not determined by all the observations in the series.
- Median is not capable for further algebraic calculations.
- The sampling stability of the median is less as compared to mean.

Determining the median graphically:

Median can be determined graphically by applying any one of the following methods.

Method-1:

Step-1: Draw two ogives- one by less than method and other by more than method.

Step-2: From the point where these both curves intersect each other draw a Perpendicular on the X-axis.

Step -3: The point where this perpendicular touches the X-axis gives the value of median.

Method-2:

Step-1: Draw only one ogive by less than method or more than method by taking variable on the X-axis and frequency on the Y-axis.

Step-2: Determine the median value by the formula median=size of $\frac{N}{2}$ th item.

Step-3: Locate this value on the Y-axis and from it draw a perpendicular on the ogive

Step -3: The point where this perpendicular touches the X-axis gives the value of median.

Note: The other partition values like quartiles, deciles, etc can be also determined graphically by this

method no. II

OTHER POSITIONAL MEASURES

Quartiles:

Quartile is that value which divides the total distribution into four equal parts. So there are three quartiles, *i.e.* Q_1, Q_2 and Q_3. Q_1, Q_2 and Q_3 are termed as first quartile, second quartile and third quartile or lower quartile, middle quartile and upper quartile respectively. Q_1 (quartile one) covers the first 25% items of the series and it divides the first half of the series into two equal parts. Q_2 (quartile two) is the median or middle value of the series and Q_3 (quartile three) covers 75% items of the series.

Calculation of Quartiles:

The calculation of quartiles is done exactly in the same manner as it is in case of the calculation of median.

In case of Individual and Discrete Series:

$$Q_i = Size\ of\ \frac{i(N+1)}{4}\ th\ item\ of\ the\ series$$

i = 1, 2, 3

In case of continuous series:

$$Q_i = Size\ of\ \frac{iN}{4}\ th\ item\ of\ the\ series\ ,\ i = 1, 2, 3$$

Interpolation formula for continuous series:

$$Q_i = l_1 + \frac{l_2 - l_1}{f}\left(\frac{iN}{4} - c\right),\ i = 1, 2, 3$$

Where,

l_1 = Lower limit of ith quartile class

l_2 = upper limit of ith quartile class

c = Cumulative frequency preceding the ith quartile class

f = Frequency of ith quartile class.

Example:

Find the Q} and Q3 of the following:

(a) 4, 5, 6, 7, 8, 9, 12, 13, 15, *10, 20*

(b) 100, 500, 1000, 800, 600, 400, 7000 and 1200

Answer:

(a) Values of the variable are in ascending order:

i.e. 4, 5, 6, 7, 8, 9, 10, 12, 13, 15, 20, So N = 11 (No. of Values)

$$Q_1 = Size\,of\,\frac{(N+1)}{4}\,th\,item\,of\,the\,series$$

= size of 3rd item = 6

$$Q_3 = Size\,of\,\frac{3(N+1)}{4}\,th\,item\,of\,the\,series$$

= size of 9th item = 13

:. Required Q_1 and Q_3 are 6 and 13 respectively,

(b) The values of the variable in ascending order are:

100, 400, 500, 600, 700, 800, 1000, 1200, N = 8

$$Q_1 = Size\,of\,\frac{(N+1)}{4}\,th\,item\,of\,the\,series$$

$$= Size\,of\,\frac{(8+1)}{4}\,th\,item\,of\,the\,series$$

= size of 2.25th item

= size of {Second item + 0.25(Third item - Second item)}

= 400 + 0.25 (500 - 400) = 400 + 25 = 425

$$Q_3 = Size\,of\,\frac{3(N+1)}{4}\,th\,item\,of\,the\,series$$

$$= Size\,of\,\frac{3(8+1)}{4}\,th\,item\,of\,the\,series$$

= size of 6.75th item

= size of [6th item + 0.75(7th item - 6th item)]

= 800 + 0.75 (1000 - 800) = 800 + 150 = 950

Required Q_1 and Q_3 are 425 and 950 respectively

Deciles:

Deciles are those values which divide the series into ten equal parts. There are nine deciles *i.e.* D_1, D_2, D_3........ D_9 in a series and 5th decile is same as median and 2nd quartile, because those values divide the series in two equal parts.

Calculation of Deciles:

The calculation of deciles is done exactly in the same manner as it is in case of calculation of median.

In case of Individual and Discrete Series:

$$D_i = Size\,of\,\frac{i(N+1)}{10}th\,item\,of\,the\,series$$, i = 1, 2, 3, ……,9

In case of continuous series:

$$D_i = Size\,of\,\frac{iN}{10}th\,item\,of\,the\,series$$, i = 1, 2, 3,……..,9

Interpolation formula for continuous series:

$$D_i = l_1 + \frac{l_2 - l_1}{f}\left(\frac{iN}{10} - c\right)$$, i = 1, 2, 3,…….,9

Where,

l_1 = Lower limit of ith decile class

l_2 = upper limit of ith decile class

c = Cumulative frequency preceding the ith decile class

f = Frequency of ith decile class.

Percentiles:

Percentiles are the values which divides the series into hundred equal parts. There are 99 percentiles *i.e.* P_1, P_2, P_3, ……..,P_{99} in a series, The value of 50th percentile = the value of 5th decile = Value of median. The 50th percentile divides the series into two equal parts. Similarly the value of $Q_1 = P_{25}$ and value of $Q_3 = P_{75}$

Calculation of Percentiles:

The calculation of percentiles is done exactly in the same manner as it is in the case of the calculation of median. The series must have to organize either in ascending or descending order, then necessary formula be put in order to get the percentile value.

In case of Individual and Discrete Series:

$$P_i = Size\,of\,\frac{i(N+1)}{100}th\,item\,of\,the\,series$$, i = 1, 2, 3, ……,99

In case of continuous series:

$$P_i = Size\,of\,\frac{iN}{100}th\,item\,of\,the\,series$$, i = 1, 2, 3,……..,99

Interpolation formula for continuous series:

$$P_i = l_1 + \frac{l_2 - l_1}{f}\left(\frac{iN}{100} - c\right)$$, i = 1, 2, 3,…….,99

Where,

l_1 = Lower limit of ith percentile class

l_2 = upper limit of ith percentile class

c = Cumulative frequency preceding the ith percentile class

f = Frequency of ith percentile class.

Advantages of Quartiles, Deciles and Percentiles:

(i) These averages can be directly determined in case of open end class intervals without knowing the lower limit of lowest class and upper limit of the largest class.

(ii) These averages can be calculated easily in absence of some data in a series.

(iii) These averages are helpful in the calculation of measures of dispersion.

(iv) These averages are not affected very much by the extreme items.

(v) These averages can be located graphically.

Disadvantages of Quartiles, Deciles and Percentiles:

(i) These averages are not easily understood by a common man. These are not well defined and easy to calculate.

(ii) These averages are not based on all the observations of a series.

(iii) These averages cannot be computed if items are not given in ascending or descending order.

(iv) These averages are affected very much by the fluctuation of sampling.

(v) The computation of these averages is not so easy in case of continuous series as the formula of interpolation is to be used.

Mode:

The mode or the modal value is that value in a series of observations which occurs with the greatest frequency.

Ex: The mode of the series 3, 5,8,5,4,5,9,3 would be 5.

In certain cases there may not be a mode or there may be more than one mode.

Ex: 1) 40, 44,57,78,84 (no mode)

2) 3, 4, 5, 5, 4, 2, 1 (modes 4 and 5)

3) 8, 8, 8, 8, 8 (no mode)

A series of data which having one mode is called ‘unimodal’ and a series of data which having two modes is called ‘bimodal’. It may also have several modes and be called ‘multimodal’.

Calculation of mode – discrete series:

a) Simple inspection method:

In a discrete series the value of the variable against which the frequency is the largest, would be the modal value.

Age	5	7	10	12	15	18
No.of Boys	4	6	9	7	5	3

Ex:

From the above data we can clearly say that mode is 10 because 10 has occurred maximum number of times i.e. 9.

b) Grouping and Analysis table method:

This method is practically applied when the below problems occurs.

1) When the difference between the maximum frequency and the frequency preceding it or succeeding it is very small.
2)

Process:

In order to find mode, a grouping table and an analysis table are to be prepared in the following manner.

Grouping Table:

A grouping table consists of 6 columns.

1) Arrange the values in ascending order and write down their corresponding frequencies in the column-1.
2) In column-2 the frequencies are grouped into two's and added.
3) In column-3 the frequencies are grouped into two's, leaving the first frequency and added.
4) In column-4 the frequencies are grouped into three's, and added.
5) In column-5 the frequencies are grouped into three's, leaving the first frequency and added.
6) In column-6 the frequencies are grouped into three's, leaving the first and second frequencies and added.
7) Now in each these columns mark the highest total with a circle.

Analysis Table:

After preparing a grouping table, prepare an analysis table. While preparing this table take the column numbers as rows and the values of the variable as columns. Now for each column number see the highest total in the grouping table (Which is marked with a circle) and mark the corresponding values of the variable to which the frequencies are related by using bars in the relevant boxes. Now the value of the variable (class) which gets highest number of bars is the modal value (modal class).

Calculation of mode – continuous series:

In a continuous series, to find out the mode we need one step more than those used for discrete series. As explained in the discrete series, modal class is determined by inspection or by preparing grouping and analysis tables. Then we apply the following formula.

$$\text{Mode } (M_o) = l_1 + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times i \quad \textbf{Or} \quad l_1 + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times i$$

Where $\Delta_1 = f_1 - f_0$

$\Delta_2 = f_1 - f_2$

l_1 = lower limit of the modal class.

f_1 =frequency of the modal class.

f_0 =frequency of then class preceding to the modal class.

f_2 = frequency of the class succeeding to the modal class.

i = size of the class.

Note:

1) While applying the above formula for calculating mode, it is necessary to see that the class intervals are uniform through out. If they are unequal they should first made equal on the assumption that the frequencies are equally distributed through out.
2) In case of bimodal distribution the mode can not be found.

Finding mode in case of bimodal distribution:

In a bimodal distribution the value of mode can not be determined by the help of the above formulae. In this case the mode can be determined by using the empirical relation given below.

Mode = 3Median - 2Mean

And the mode which is obtained by using the above relation is called **'Empirical mode'**

Locating mode graphically:

In a frequency distribution the value of mode can be determined graphically.
The steps in calculation are

1) Draw a histogram of the given data.
2) Draw two lines diagonally in the inside of the modal class bar, starting from each upper corner of the bar to the upper corner of the adjacent bars.
3) Draw a perpendicular line from the intersection of the two diagonal lines to the X-axis which gives the modal value.

Note:

The graphic method of determining mode can be used only where the data is unimodal.

Merits:

1) It is easy to calculate and simple to understand.
2) It is not affected by the extreme values.
3) The value of mode can be can be determined graphically.
4) Its value can be determined in case of open-end class interval.
5) The mode is the most representative of the distribution.

Demerits:

1) It is not suitable for further mathematical treatments.
2) The value of mode can not always be determined.
3) The value of mode is not based on each and every items of the series.
4) The mode is strictly defined.
5) It is difficult to calculate when one of the observations is zero or the sum of the observations is zero.

Empirical relation between Mean, Median and Mode:

The relationship between mean, median and mode depends upon the nature of the distribution. A distribution may be symmetrical or asymmetrical.

In asymmetrical distribution the mean, median and mode are equal

i.e. **Mean(AM) = Median(M) = Mode(Mo)**

In a highly asymmetrical distribution it is not possible to find a relation ship among the averages. But in a moderately asymmetric distribution the difference between the mean and mode is three times the difference between the mean and median.

i.e**. Mean-Mode = 3(Mean-Median)**

Geometric mean:

Geometric mean is defined as the n th root of the product of n items (or) values.

Calculation of G.M- Individual series:

If $x_1, x_2, x_3, ..., x_n$ be n observations studied on a variable X, then the G.M of the observations is defined as

$$\text{G.M}=\left(x_1 x_2 x_3 ....x_n\right)^{1/n}$$

Applying log both sides

$$\log G.M = \frac{1}{n}\log\left(x_1 . x_2 ....x_n\right)$$

$$=\frac{1}{n}[\log x_1 + \log x_2 + ... + \log x_n]$$

$$= \frac{1}{n}\sum_{i=1}^{n}\log x_i$$

$$\Rightarrow G.M = anti\log\left(\frac{1}{n}\sum_{i=1}^{n}\log x_i\right)$$

Calculation of G.M- Discrete series:

If $x_1, x_2, x_3, ..., x_n$ be n observations of a variable X with frequencies $f_1, f_2, f_3, ..., f_n$ respectively then the G.M is defined as

$$\text{G.M}=\left(x_1^{f_1} x_2^{f_2} x_3^{f_3} ....x_n^{f_1}\right)^{1/N} \;\ldots\ldots\ldots\ldots\ldots\ldots(*)$$

Where N $=\sum_{i=1}^{n} f_i$ i.e. total frequency

Applying log both sides in (*) we get

$$\text{G.M= antilog}\left(\frac{1}{N}\sum_{i=1}^{n} f_i \log x_i\right)$$

Calculation of G.M-Continuous series:

In continuous series the G.M is calculated by replacing the value of x_i by the mid points of the class's i.e. m_i.

$$\text{G.M= antilog}\left(\frac{1}{N}\sum_{i=1}^{n} f_i \log m_i\right)$$

Where m_i is the mid value of the i th class interval.

Properties of G.M:

1) If G_1 and G_2 are geometric means of two components having n_1 and n_2 observations and G is the geometric mean of the combined series of n (n_1+n_2) values then

$$G = G_1^{W_1} G_2^{W_2}$$

Where $w_1 = \dfrac{n_1}{n_1 + n_2}$ & $w_2 = \dfrac{n_2}{n_1 + n_2}$

Uses of G.M:

Geometrical Mean is especially useful in the following cases.

1) The G.M is used to find the average percentage increase in sales, production, or other economic or business series.

For example, from 1992 to 1994 prices increased by 5%,10%,and 18% respectively, then the average annual income is not 11% which is calculated by A.M but it is 10.9 which is calculated by G.M.

2) G.M is theoretically considered to be best average in the construction of Index numbers.

Weighted Geometric Mean:

Like weighted Arithmetic mean, we can also find weighted geometric mean and the formula is given by

$$G.M_w = \left(x_1^{w_1} x_2^{w_2} x_3^{w_3} \ldots\ldots x_n^{w_n}\right)^{1/N}$$

Where $N = \sum_{i=1}^{n} w_i$ i.e. total weight

Applying log both sides we get

$$G.M = \text{antilog}\left(\frac{1}{N}\sum_{i=1}^{n} w_i \log x_i\right)$$

Harmonic mean:

The harmonic mean (H.M) is defined as the reciprocal of the arithmetic mean of the reciprocal of the individual observations.

Calculation of H.M -Individual series:

If $x_1, x_2, x_3, \ldots\ldots\ldots, x_n$ be 'n' observations of a variable X then harmonic mean is defined as

$$H.M = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \ldots\ldots\ldots\ldots + \frac{1}{x_n}}$$

$$\Rightarrow H.M = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$$

Calculation of H.M -Discrete series:

If $x_1, x_2, x_3, \ldots\ldots\ldots, x_n$ be 'n' observations occurs with frequencies $f_1, f_2, f_3, \ldots\ldots\ldots, f_n$ respectively then H.M is defined as

$$H.M = \frac{\sum_{i=1}^{n} f_i}{\frac{f_1}{x_1} + \frac{f_2}{x_2} + \ldots\ldots\ldots\ldots + \frac{f_n}{x_n}} \quad \ldots\ldots\ldots\ldots(*)$$

$$\Rightarrow H.M = \frac{\sum_{i=1}^{n} f_i}{\sum_{i=1}^{n} \frac{f_i}{x_i}}$$

Calculation of H.M – Continuous series:

In continuous series H.M can be calculated by replacing mid values (m_i) in place of $x_i's$ in the equation (*). Hence H.M is given by

$$\Rightarrow H.M = \frac{\sum_{i=1}^{n} f_i}{\sum_{i=1}^{n} \frac{f_i}{m_i}}$$, where m_i is the mid value of the ith class interval

Uses of harmonic mean:

1) The H.M is used for computing the average rate of increase in profits of a concern.
2) The H.M is used to calculate the average speed at which a journey has been performed.

Merits:

1) Its value is based on all the observations of the data.
2) It is less affected by the extreme values.
3) It is suitable for further mathematical treatment.
4) It is strictly defined.

Demerits:

1) It is not simple to calculate and easy to understand.
2) It can not be calculated if one of the observations is zero.
3) The H.M is always less than A.M and G.M.

Relation between A.M, G.M, and H.M:

The relation between A.M, G.M, and H.M is given by

$$\boxed{A.M \geq G.M \geq H.M}$$

Note: The equality condition holds true only if all the items are equal in the distribution.

Prove that if a and b are two positive numbers then $A.M \geq G.M \geq H.M$

Sol:

Let a and b are two positive numbers then

The Arithmetic mean of a and b $= \frac{a+b}{2}$

The Geometric mean of a and b $= \sqrt{ab}$

The harmonic men of a and b $= \frac{2ab}{a+b}$

Let us assume $A.M \geq G.M$

$$\Rightarrow \frac{a+b}{2} \geq \sqrt{ab}$$

$$\Rightarrow a+b \geq 2\sqrt{ab}$$

$$\Rightarrow (a+b)^2 \geq 4ab$$

$$\Rightarrow (a-b)^2 \geq 0$$

which is always true.

$A.M \geq G.M$ ………………………… (1)

let us assume $GM \geq HM$

$$\Rightarrow \sqrt{ab} \geq \frac{2ab}{a+b}$$

$$\Rightarrow a+b \geq 2\sqrt{ab}$$

$$\Rightarrow (a+b)^2 \geq 4ab$$

$$\Rightarrow (a-b)^2 \geq 0$$

Which is always true.

$\therefore G.M \geq HM$ .... (2)

from (1) and (2) we get $A.M \geq G.M \geq H.M$

Index Numbers

INTRODUCTION

Historically, the first index was constructed in 1764 to compare the Italian price index in 1750 with the price level in 1500. Though originally developed for measuring the effect of change in prices, index numbers have today become one of the most widely used statistical devices and there is hardly any field where they are not used. Newspapers headline the fact that prices are going up or down, that industrial production is rising or falling, that imports are increasing or decreasing, that crimes are rising in a particular period compared to the previous period as disclosed by index numbers. They are used to feel the pulse of the economy and they have come to be used as indicators of inflationary or deflationary tendencies, In fact, they are described as 'barometers of economic activity', i.e., if one wants to get an idea as to what is happening to an economy, he should look to important indices like the index number of industrial production, agricultural production, business activity, etc.

Some prominent definitions of index numbers are given below:

1. 'Index numbers are devices for measuring differences in the magnitude of a group of related variables. —Croxton & Cowdert
2. "An index number is a statistical measure designed to show changes in a variable or a group of related variables with respect to time, geographic location or other characteristics such as income, profession, etc. —Spiegel
3. "In its simplest form an index number is the ratio of two index numbers expressed as a per cent. An index number is a statistical measure—a measure designed to show changes in one variable or in a group of related variables over time, or with respect to geographic location, or in terms of some other characteristics." —Patternson

Definition:

Index numbers are statistical devices designed to measure the relative change in the level of variable or group of variables with respect to time, geographical location etc.

In other words these are the numbers which express the value of a variable at any given period called "***current period*** "as a percentage of the value of that variable at some standard period called "***base period***".

Here the variables may be

1. The price of a particular commodity like silver, iron or group of commodities like consumer goods, food, stuffs etc.
2. The volume of trade, exports, imports, agricultural and industrial production, sales in departmental store.
3. Cost of living of persons belonging to particular income group or profession etc.

Ex: suppose rice sells at Rs.9/kg at BBSR in 1995 as compare to Rs. 4.50/Kg in 1985, the index number of price in 1995 compared to 1985.
Therefore the index number of price of rice in 1995 compared to 1985 is calculated as

$$\frac{Rs.9.00}{Rs.4.50}\times 100 = 200$$

This means there is a net increase of 100% in the price of rice in 1995as compared to 1985 **[the base year's index number is always treated as 100]**

Suppose, during the same period 1995 the rice sells at Rs. 12.00/kg in Delhi. There fore, the index number of price at Bhubaneswar compared to price at Delhi is

$$\frac{Rs.9.00}{Rs.12.00}\times 100 = 75$$

This means there is a net decrease of 25% in the price of rice in 1995as compared to 1985

The above index numbers are called '***price index numbers***'.

To take another example the production of rice in 1978 in Orissa was 44, 01,780 metric c tons compare to 36, 19,500 metric tons in 1971. So the index number of the quantity produced in 1978 compared to 1971 is

$$\frac{4401780}{3619500}\times 100 = 121.61$$

That means there is a net increase of 21.61% in production of rice in 1978 as compared to 1971.

The above index number is called '***quantity index number'***

Univariate index: An index which is calculated from a single variable is called *univariate index.*

Composite index: An index which is calculated from group of variables is called *Composite index*

Characteristics of index numbers:

1. **Index numbers are specialized averages:**
As we know an average is a single figure representing a group of figures. How ever to obtain an average the items must be comparable. For example the average weight of man, woman and children of a certain locality has no meaning at all. Further more the unit of measurement must be same for all the items. How ever this is not so with index numbers. Index numbers also one type of averages which shows in a single figure the change in two or more series of different items which can be expressed in different units. For example while constructing a consumer price index number the various items which are use in construction are divided into broad heads namely food, clothing, fuel, lighting, house rent, and miscellaneous which are expressed in different units.
2. **Index numbers measures the net change in a group of related variables:**
Since index numbers are essentially averages, they describe in one single figure the increase or decrease in a group of related variables under study. The group of variables may be prices of set of commodities, the volume of production in different sectors etc.
3. **Index numbers measure the effect of changes over period of time:**

Index numbers are most widely used for measuring changes over a period of time. For example we can compare the agricultural production, industrial production, imports, exports, wages etc in two different periods.

Uses of index numbers:

Index numbers are indispensable tools of economics and business analysis. Following are the main uses of index numbers.

1) **Index numbers are used as economic barometers:**

Index number is a special type of averages which helps to measure the economic fluctuations on price level, money market, economic cycle like inflation, deflation etc. G.Simpson and F.Kafka say that index numbers are today one of the most widely used statistical devices. They are used to take the pulse of economy and they are used as indicators of inflation or deflation tendencies. So index numbers are called economic barometers.

2) **Index numbers helps in formulating suitable economic policies and planning etc.**

Many of the economic and business policies are guided by index numbers. For example while deciding the increase of DA of the employees; the employer's have to depend primarily on the cost of living index. If salaries or wages are not increased according to the cost of living it leads to strikes, lock outs etc. The index numbers provide some guide lines that one can use in making decisions.

3) **They are used in studying trends and tendencies.**

Since index numbers are most widely used for measuring changes over a period of time, the time series so formed enable us to study the general trend of the phenomenon under study. For example for last 8 to 10 years we can say that imports are showing upward tendency.

4) **They are useful in forecasting future economic activity.**

Index numbers are used not only in studying the past and present workings of our economy but also important in forecasting future economic activity.

5) **Index numbers measure the purchasing power of money.**

The cost of living index numbers determine whether the real wages are rising or falling or remain constant. The real wages can be obtained by dividing the money wages by the corresponding price index and multiplied by 100. Real wages helps us in determining the purchasing power of money.

6) **Index numbers are used in deflating.**

Index numbers are highly useful in deflating i.e. they are used to adjust the wages for cost of living changes and thus transform nominal wages into real wages, nominal income to real income, nominal sales to real sales etc. through appropriate index numbers.

Classification of index numbers:

According to purpose for which index numbers are used are classified as below.

i) Price index number
ii) Quality index number
iii) Value index number
iv) Special purpose index number

Only price and quantity index numbers are discussed in detail. The others will be mentioned but without detail.

Price index number:

Price index number measures the changes in the price level of one commodity or group of commodities between two points of time or two areas.

Ex: Wholesale price index numbers
Retail price index numbers
Consumer price index numbers.

Quantity index number:

Quantity index numbers measures the changes in the volume of production, sales, etc in different sectors of economy with respect to time period or space.

Note: Price and Quantity index numbers are called *market index numbers.*

Problems in constructing index numbers:

Before constructing index numbers the careful thought must be given into following problems

i. Purpose of index numbers.

An index number which is properly designed for a purpose can be most useful and powerful tool. Thus the first and the foremost problem are to determine the purpose of index numbers. If we know the purpose of the index numbers we can settle some related problems. For example if the purpose of index number is to measure the changes in the production of steel, the problem of selection of items is automatically settled.

ii. Selection of commodities

After defining the purpose of index numbers, select only those commodities which are related to that index. For example if the purpose of an index is to measure the cost of living of low income group we should select only those commodities or items which are consumed by persons belonging to this group and due care should be taken not to include the goods which are utilized by the middle income group or high income group i.e. the goods like cosmetics, other luxury goods like scooters, cars, refrigerators, television sets etc.

iii. Selection of base period

The period with which the comparisons of relative changes in the level of phenomenon are made is termed as ***base period.*** The index for this period is always taken as 100. The following are the basic criteria for the choice of the base period.

i) The base period must be a normal period i.e. a period frees from all sorts of abnormalities or random fluctuations such as labor strikes, wars, floods, earthquakes etc.

ii) The base period should not be too distant from the given period. Since index numbers are essential tools in business planning and economic policies the base period should not be too far from the current period. For example for deciding increase in dearness allowance at present there is no advantage in taking 1950 or 1960 as the base, the comparison should be with the preceding year after which the DA has not been increased.

iii) Fixed base or chain base .While selecting the base a decision has to be made as to whether the base shall remain fixing or not i.e. whether we have fixed base or chain base. In the fixed base method the year to which the other years are compared is constant. On the other hand, in chain base method the prices of a year are linked with those of the preceding year. The chain base method gives a better picture than what is obtained by the fixed base method.

- **How a base is selected if a normal period is not available?**
  Ans: Some times it is difficult to distinguish a year which can be taken as a normal year and hence the average of a few years may be regarded as the value corresponding to the base year.

iv. Data for index numbers

The data, usually the set of prices and of quantities consumed of the selected commodities for different periods, places etc. constitute the raw material for the construction of index numbers. The data should be collected from reliable sources such as standard trade journals, official publications etc. for example for the construction of retail price index numbers, the price quotations for the commodities should be obtained from super bazaars, departmental stores etc. and not from wholesale dealers.

v. Selection of appropriate weights

A decision as to the choice of weights is an important aspect of the construction of index numbers. The problem arises because all items included in the construction are not of equal importance. So proper weights should be attached to them to take into account their relative importance. Thus there are two type of indices.

i) Un weighted indices- in which no specific weights are attached
ii) Weighted indices- in which appropriate weights are assigned to various items.

vi. Choice of average.

Since index numbers are specialized averages, a choice of average to be used in their construction is of great importance. Usually the following averages are used.

i) A.M
ii) G.M
iii) Median

Among these averages ***G.M*** is the appropriate average to be used. But in practice G.M is not used as often as A.M because of its computational difficulties.

vii. Choice of formula.

A large variety of formulae are available to construct an index number. The problem very often is that of selecting the appropriate formula. The choice of the formula would depend not only on the purpose of the index but also on the data available.

Methods of constructing index numbers:

A large number of formulae have been derived for constructing index numbers. They can be

1) Unweighted indices
   a) Simple aggregative method
   b) Simple average of relatives.

2) Weighted indices
   a) Weighted aggregative method
      i) Lasperey's method
      ii) Paasche's method
      iii) Fisher's ideal method
      iv) Dorbey's and Bowley's method
      v) Marshal-Edgeworth method
      vi) Kelly's method
   b) Weighted average of relatives

Unweighted indices:

i) Simple aggregative method:

This is the simplest method of constructing index numbers. When this method is used to construct a price index number the total of current year prices for the various commodities in question is divided by the total of the base year prices and the quotient is multiplied by 100.

`Symbolically $P_{01} = \frac{\sum P_1}{\sum P_0} \times 100$

Where P_0 are the base year prices

P_1 are the current year prices

P_{01} is the price index number for the current year with reference to the base year.

Problem:

Calculate the index number for 1995 taking 1991 as the base for the following data

Commodity	Unit	Prices 1991 (P_0)	Prices 1995 (P_1)
A	Kilogram	2.50	4.00
B	Dozen	5.40	7.20
C	Meter	6.00	7.00
D	Quintal	150.00	200.00
E	Liter	2.50	3.00
Total		166.40	221.20

Price index number = $P_{01} = \frac{\sum P_1}{\sum P_0} \times 100 = \frac{221.20}{166.40} \times 100 = 132.93$

$\therefore$ There is a net increase of 32.93% in 1995 as compared to 1991.

Limitations:

There are two main limitations of this method

1. The units used in the prices or quantity quotations have a great influence on the value of index.
2. No considerations are given to the relative importance of the commodities.

ii) Simple average of relatives

When this method is used to construct a price index number, first of all price relatives are obtained for the various items included in the index and then the average of these relatives is obtained using any one of the averages i.e. mean or median etc.

When A.M is used for averaging the relatives the formula for computing the index is

$$P_{01} = \frac{1}{n}\sum\left(\frac{P_1}{P_0}\times 100\right)$$

When G.M is used for averaging the relatives the formula for computing the index is

$$P_{01} = Anti\log\left[\frac{1}{n}\sum\log\left(\frac{P_1}{P_0}\times 100\right)\right]$$

Where n is the number of commodities

and price relative = $\frac{P_1}{P_0}\times 100$

Problem:

Calculate the index number for 1995 taking 1991 as the base for the following data

Commodity	Unit	Prices 1991 (P_0)	Prices 1995 (P_1)	$\frac{P_1}{P_0}\times 100$
A	Kilogram	50	70	$\frac{70}{50}\times 100 =$ 140
B	Dozen	40	60	150
C	Meter	80	90	112.5
D	Quintal	110	120	109.5
E	Liter	20	20	100
Total				

Price index number = $P_{01} = \frac{1}{n}\sum\left(\frac{P_1}{P_0}\times 100\right) = \frac{1}{5}\sum 612 = 122.4$

$\therefore$ There is a net increase of 22.4% in 1995 as compared to 1991.

Merits:

1. It is not affected by the units in which prices are quoted
2. It gives equal importance to all the items and extreme items don't affect the index number.
3. The index number calculated by this method satisfies the unit test.

Demerits:

1. Since it is an unweighted average the importance of all items are assumed to be the same.
2. The index constructed by this method doesn't satisfy all the criteria of an ideal index number.

3. In this method one can face difficulties to choose the average to be used.

Weighted indices:

i) Weighted aggregative method:

These indices are same as simple aggregative method. The only difference is in this method, weights are assigned to the various items included in the index.

There are various methods of assigning weights and consequently a large number of formulae for constructing weighted index number have been designed.

Some important methods are

i. **Lasperey's method:** This method is devised by Lasperey in year 1871.It is the most important of all the types of index numbers. In this method the base year quantities are taken weights. The formula for constructing Lasperey's price index number is

$$P_{01}^{La} = \frac{\sum p_1 q_0}{\sum p_0 q_0} \times 100$$

Ernst Louis Étienne Laspeyres
(1834-1913) , Germany

ii. **Paasche's method:** In this method the current year quantities are taken as weights and the formula is given by

$$P_{01}^{Pa} = \frac{\sum p_1 q_1}{\sum p_0 q_1} \times 100$$

Hermann Paasche (1851-1925)
Germany

iii. **Fisher's ideal method:** Fishers price index number is given by the G.M of the Lasperey's and Paasche's index numbers.

Symbolically

$$P_{01}^{F} = \sqrt{P_{01}^{La} P_{01}^{Pa}}$$

$$= \sqrt{\frac{\sum p_1 q_0}{\sum p_0 q_0} \times 100 \frac{\sum p_1 q_1}{\sum p_0 q_1} \times 100}$$

$$= \sqrt{\frac{\sum p_1 q_0}{\sum p_0 q_0} \frac{\sum p_1 q_1}{\sum p_0 q_1}} \times 100$$

Sir Ronald Aylmer Fisher
(1890-1962), England

iv. **Dorbey's and Bowley's method**

Dorbey's and Bowley's price index number is given by the A.M of the Lasperey's and Paasche's index numbers.

Symbolically

$$P_{01}^{DB} = \frac{P_{01}^{La} + P_{01}^{Pa}}{2}$$

Quantity index numbers:

i. Lasperey's quantity index number: Base year prices are taken as weights

$$Q_{01}^{La} = \frac{\sum q_1 p_0}{\sum q_0 p_0} \times 100$$

ii. Paasche's quantity index number : Current year prices are taken as weights

$$Q_{01}^{Pa} = \frac{\sum q_1 p_1}{\sum q_0 p_1} \times 100$$

iii. Fisher's ideal method: $Q_{01}^{F} = \sqrt{Q_{01}^{La} Q_{01}^{Pa}} = \sqrt{\frac{\sum q_1 p_0}{\sum q_0 p_0} \times \frac{\sum q_1 p_1}{\sum q_0 p_1}} \times 100$

Fisher's index number is called ideal index number. Why?

The Fisher's index number is called ideal index number due to the following characteristics.

1) It is based on the G.M which is theoretically considered as the best average of constructing index numbers.
2) It takes into account both current and base year prices as quantities.
3) It satisfies both time reversal and factor reversal test which are suggested by Fisher.
4) The upward bias of Lasperey's index number and downward bias of Paasche's index number are balanced to a great extent.

Example: Compute price index numbers for the following data by

(*i*) *Laspeyre's method,*
(*ii*) *Paasche's method,*
(*iii*) *Fisher's ideal method,*
(*iv*) *Dorbish-Bowley's method,*
(*v*) *Marshall-Edgeworth's method.*

Year	*Commodity A*		*Commodity B*		*Commodity C*	
	Price	*Quantity*	*Price*	*Quantity*	*Price*	*Quantity*
1980	*4*	*50*	*3*	*10*	*2*	*5*
1985	*10*	*45*	*6*	*8*	*3*	*4*

Base year : 1980
Price and quantity given in arbitrary units.

Calculation of Indices

Commodities	*1980*		*1985*					
	Price p_0	*Quantity* q_0	*Price* p_1	*Quantity* q_1	p_1q_0	p_0q_0	p_1q_1	p_0q_1
A	4	50	10	45	500	200	450	180
B	3	10	6	8	60	30	48	24
C	2	5	3	4	15	10	12	8
Total	—	—	—	—	575	240	510	212

(*i*) Laspeyre's method :

$$L_{01} = \frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times 100 = \frac{575}{240} \times 100 = 239.58$$

(*ii*) Paasche's method :

$$P_{01} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1} \times 100 = \frac{510}{212} \times 100 = 240.57$$

(*iii*) Fisher's ideal method :

$$F_{01} = \sqrt{L_{01} \times P_{01}} = \sqrt{239.58 \times 240.57} = 240.07.$$

(*iv*) Dorbish-Bowley's method :

$$DB_{01} = \frac{L_{01} + P_{01}}{2}$$

$$= \frac{239.58 + 240.57}{2} \doteq 239.82.$$

Comparison of Lasperey's and Paasche's index numbers:-

In Lasperey's index number base year quantities are taken as the weights and in Paasche's index the current year quantities are taken as weights.

From the practical point of view Lasperey's index is often proffered to Paasche's for the simple reason that Lasperey's index weights are the base year quantities and do not change from the year to the next. On the other hand Paasche's index weights are the current year quantities, and in most cases these weights are difficult to obtain and expensive.

Lasperey's index number is said to be have upward bias because it tends to over estimate the price rise, where as the Paasche's index number is said to have downward bias, because it tends to under estimate the price rise.

When the prices increase, there is usually a reduction in the consumption of those items whose prices have increased. Hence using base year weights in the Lasperey's index, we will be giving too much weight to the prices that have increased the most and the numerator will be too large. Due to similar considerations, Paasche's index number using given year weights under estimates the rise in price and hence has down ward bias.

If changes in prices and quantities between the reference period and the base period are moderate, both Lasperey's and Paasche's indices give nearly the same values.

Demerit of Paasche's index number:

Paasche's index number, because of its dependence on given year's weight, has distinct disadvantage that the weights are required to be revised and computed for each period, adding extra cost towards the collection of data.

What are the desiderata of good index numbers?

Irving Fisher has considered two important properties which an index number should satisfy. These are tests of reversibility.

1. Time reversal test
2. Factor reversal test

If an index number satisfies these two tests it is said to be an ideal index number.

Weighted average of relatives:

Weighted average of relatives can be calculated by taking values of the base year (p_0q_0) as the weights. The formula is given by

$$\text{When A.M is used } P_{01} = \frac{\sum PV}{\sum V}$$

$$\text{When G.M is used } P_{01} = Anti\log \frac{\sum V \log P}{\sum V}$$

Where $P = \frac{p_1}{p_0} \times 100$ and $V = p_0 q_0$ i.e. base year value

Illustration 8. From the following data compute price index by supplying weighted average of price method using :

(*a*) arithmetic mean, and
(*b*) geometric mean.

Commodity	p_0 (Rs.)	q_0	p_1 (Rs.)
Sugar	3.0	20 kg.	4.0
Flour	1.5	40 kg.	1.6
Milk	1.0	10 lt.	1.5

Solution.

(*a*) INDEX NUMBER USING WEIGHTED ARITHMETIC MEAN OF PRICE RELATIVES

Commodity	p_0	q_0	p_1	p_0q_0 V	$\frac{p_1}{p_0} \times 100$ p	PV
Sugar	Rs. 3.0	20 kg.	Rs. 4.0	60	$\frac{4}{3} \times 100$	8,000
Flour	Rs. 1.5	40 kg.	Rs. 1.6	60	$\frac{1.6}{1.5} \times 100$	6,400
Milk	Re. 1.0	10 lt.	Rs. 1.5	10	$\frac{1.5}{1.0} \times 100$	1,500
				$\Sigma V = 130$		$\Sigma PV = 15{,}900$

$$p_{01} = \frac{\Sigma PV}{\Sigma V} = \frac{15{,}900}{130} = 122.31$$

This means that there has been a 22.3 per cent increase in prices over the base level.

(*b*) INDEX NUMBER USING GEOMETRIC MEAN OF PRICE RELATIVES

Commodity	p_0	q_0	p_1	V	p	*Log p*	*V Log p*
Sugar	Rs. 3.0	**20 kg.**	Rs. 4.0	60	133.3	2.1249	127.494
Flour	Rs. 1.5	40 kg.	Rs. 1.6	60	106.7	2.0282	121.692
Milk	Re. 1.0	10 lt.	Rs. 1.5	10	150.0	2.1761	21.761
				$\Sigma V = 130$			$\Sigma V \log p$ **= 270.947**

$$p_{01} = \text{Antilog}\left[\frac{\Sigma V \log p}{\Sigma V}\right] = \text{Antilog}\left[\frac{270.947}{130}\right] = \text{Antilog } 2.084 = 120.9$$

Test of consistency or adequacy:

Several formulae have been suggested for constructing index numbers and the problem is that of selecting most appropriate one in a given situation. The following teats are suggested for choosing an appropriate index.

The following tests are suggested for choosing an appropriate index.

1) Unit test
2) Time reversal test
3) Factor reversal test
4) Circular test

1) Unit test:

This test requires that the formula for construction of index numbers should be such, which is not affected by the unit in which the prices or quantities have been quoted.

Note: This test is satisfied by all the index numbers except simple aggregative method.

2) Time reversal test

This is suggested by R.A.Fisher. Time reversal test is a test to determine whether a given method will work both ways in time i.e. forward and backward. In other words, when the data for any two years are treated by the same method, but with the bases reversed, the two index numbers secured should be reciprocals to each other, so that their product is unity. Symbolically the following relation should be satisfied.

$$P_{01} \times P_{10} = 1$$

Where P_{01} is the index for time period 1 with reference period 0.

P_{10} is the index for time period 0 with reference period 1.

Note: This test is not satisfied by Lasperey's method and Paasche's method. It is satisfied by Fisher's method.

When Lasperey's method is used

$$P_{01}^{La} = \frac{\sum p_1 q_0}{\sum p_0 q_0} \times 100$$

$$P_{10}^{La} = \frac{\sum p_0 q_1}{\sum p_1 q_1} \times 100$$

Now,

$$P_{01}^{La} \times P_{10}^{La} = \frac{\sum p_1 q_0}{\sum p_0 q_0} \times \frac{\sum p_0 q_1}{\sum p_1 q_1} \neq 1$$

Therefore this test is not satisfied by Lasperey's method

When Paasche's method is used

$$P_{01}^{Pa} = \frac{\sum p_1 q_1}{\sum p_0 q_1} \times 100$$

$$P_{10}^{Pa} = \frac{\sum p_0 q_0}{\sum p_1 q_0} \times 100$$

Now,

$$P_{01}^{Pa} \times P_{10}^{Pa} = \frac{\sum p_1 q_1}{\sum p_0 q_1} \times \frac{\sum p_0 q_0}{\sum p_1 q_0} \neq 1$$

Therefore this test is not satisfied by Paasche's method

When Fisher's method is used

$$P_{01}^{F} = \sqrt{\frac{\sum p_1 q_0}{\sum p_0 q_0} \frac{\sum p_1 q_1}{\sum p_0 q_1}} \times 100$$

$$P_{10}^{F} = \sqrt{\frac{\sum p_0 q_1}{\sum p_1 q_1} \frac{\sum p_0 q_0}{\sum p_1 q_0}} \times 100$$

Now,

$$P_{01}^{F} \times P_{10}^{F} = \sqrt{\frac{\sum p_1 q_0}{\sum p_0 q_0} \frac{\sum p_1 q_1}{\sum p_0 q_1}} \sqrt{\frac{\sum p_0 q_1}{\sum p_1 q_1} \frac{\sum p_0 q_0}{\sum p_1 q_0}} = 1$$

Value index:

The *value* of a single commodity is the product of its price and quantity. Thus a value index 'V' is the sum of the values of the commodities of given year divided by the sum of the value of the base year multiplied by 100.

$$\text{i.e. } V = \frac{\sum p_1 q_1}{\sum p_0 q_0} \times 100$$

3) Factor reversal test:

This is also suggested by R.A.Fisher. It holds that the product of a price index number and the quantity index number should be equal to the corresponding value index. In other words the test is that the change in price multiplied by the change in quantity should be equal to change in value.

If p_1 & p_0 represents prices and q_1 & q_0 the quantities in the current year and base year respectively and if P_{01} represents the change in price in the current year 1 with reference to the year 0 and Q_{01} represents the change in quantity in the current year 1 with reference to the year 0.

$$\text{Symbolically} \quad P_{01} \times Q_{01} = V_{01} = \frac{\sum p_1 q_1}{\sum p_0 q_0}$$

Note: This test is not satisfied by Lasperey's method and Paasche's method. It is satisfied by Fisher's method.

When Lasperey's method is used

$$P_{01}^{La} = \frac{\sum p_1 q_0}{\sum p_0 q_0} \times 100$$

$$Q_{01}^{La} = \frac{\sum q_1 p_0}{\sum q_0 p_0} \times 100$$

Now,

$$P_{01}^{La} \times Q_{01}^{La} = \frac{\sum p_1 q_0}{\sum p_0 q_0} \times \frac{\sum q_1 p_0}{\sum q_0 p_0} \neq \frac{\sum p_1 q_1}{\sum p_0 q_0}$$

Therefore this test is not satisfied by Lasperey's method

When Paasche's method is used

$$P_{01}^{Pa} = \frac{\sum p_1 q_1}{\sum p_0 q_1} \times 100$$

$$Q_{01}^{Pa} = \frac{\sum q_1 p_1}{\sum q_0 p_1} \times 100$$

Now,

$$P_{01}^{Pa} \times Q_{10}^{Pa} = \frac{\sum p_1 q_1}{\sum p_0 q_1} \times \frac{\sum q_1 p_1}{\sum q_0 p_1} \neq \frac{\sum p_1 q_1}{\sum p_0 q_0}$$

Therefore this test is not satisfied by Paasche's method

When Fisher's method is used

$$P_{01}^{F}=\sqrt{\frac{\sum p_1q_0}{\sum p_0q_0}\frac{\sum p_1q_1}{\sum p_0q_1}}\times 100$$

$$Q_{01}^{F}=\sqrt{\frac{\sum q_1p_0}{\sum q_0p_0}\times\frac{\sum q_1p_1}{\sum q_0p_1}}\times 100$$

$$P_{01}^{La}\times Q_{01}^{La}=\sqrt{\frac{\sum p_1q_0}{\sum p_0q_0}\frac{\sum p_1q_1}{\sum p_0q_1}}\sqrt{\frac{\sum q_1p_0}{\sum q_0p_0}\times\frac{\sum q_1p_1}{\sum q_0p_1}}$$

$$=\sqrt{\frac{\left(\sum p_1q_1\right)^2}{\left(\sum p_0q_0\right)^2}}=\frac{\sum p_1q_1}{\sum p_0q_0}$$

Therefore this test is satisfied by Fisher's method

4) **Circular test:**

This is another test of consistency of an index number. It is an extension of time reversal test. According to this test, the index should work in a circular fashion.

Symbolically

$$P_{01}\times P_{12}\times P_{20}=1$$

Note:

This test is not satisfied by Lasperey's method, Paasche's method and Fisher's method.

This test is satisfied by simple average of relatives based on G.M and Kelly's fixed base method.

Illustration 13. Construct a Fisher's Ideal Index from the following data and show that it satisfies time reversal and factor reversal tests :

	2006		*2007*	
Items	p_0	q_0	p_1	q_1
A	10	40	12	45
B	11	50	11	52
C	14	30	17	30
D	8	28	10	29
E	12	15	13	20

Solution. CONSTRUCTION OF FISHER'S IDEAL INDEX

Items	p_0	q_0	p_1	q_1	p_1q_0	p_0q_0	p_1q_1	p_0q_1
A	10	40	12	45	480	400	540	450
B	11	50	11	52	550	550	572	572
C	14	30	17	30	510	420	510	420
D	8	28	10	29	280	224	290	232
E	12	15	13	20	195	180	260	240
					$\Sigma p_1q_0=2015$	$\Sigma p_0q_0=1774$	$\Sigma p_1q_1=2172$	$\Sigma p_0q_1=1914$

Fisher's Ideal Index : $P_{01}=\sqrt{\frac{\Sigma p_1q_0}{\Sigma p_0q_0}\times\frac{\Sigma p_1q_1}{\Sigma p_0q_1}}$

$$=\sqrt{\frac{2015}{1774}\times\frac{2172}{1914}}\times 100=1{\cdot}135\times 100=113{\cdot}5$$

Time Reversal Test : Time reversal test is satisfied when :

$$P_{01} \times P_{10} = 1$$

$$P_{10} = \sqrt{\frac{\Sigma p_0 q_1}{\Sigma p_1 q_1} \times \frac{\Sigma p_0 q_0}{\Sigma p_1 q_0}} = \sqrt{\frac{1914}{2172} \times \frac{1774}{2015}}$$

$$P_{01} \times P_{10} = \sqrt{\frac{2015}{1774} \times \frac{2172}{1914} \times \frac{1914}{2172} \times \frac{1774}{2015}} = 1$$

Hence time reversal test is satisfied by the given data.

Factor Reversal Test : Factor reversal test is satisfied when :

$$P_{01} \times Q_{01} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_0}$$

$$Q_{01} = \sqrt{\frac{\Sigma q_1 p_0}{\Sigma q_0 p_0} \times \frac{\Sigma q_1 p_1}{\Sigma q_0 p_1}} = \sqrt{\frac{1914}{1774} \times \frac{2172}{2015}}$$

$$P_{01} \times Q_{01} = \sqrt{\frac{2015}{1774} \times \frac{2172}{1914} \times \frac{1914}{1774} \times \frac{2172}{2015}} = \frac{2172}{1774}$$

$\frac{\Sigma p_1 q_1}{\Sigma p_0 q_0}$ is also equal to $\frac{2172}{1774}$. Hence factor reversal test is satisfied by the given data.

- **Prove that AM of Lasperey's index numbers and Paasche's index number is greater than or equal to Fisher's index number.**

Let

Lasperey's index number = P_{01}^{La}

Paasche's index number= P_{01}^{Pa}

Fisher's index number= P_{01}^{F}

And we have $P_{01}^{F} = \sqrt{P_{01}^{La} P_{01}^{Pa}}$

Now we have to show that

$$\frac{P_{01}^{La} + P_{01}^{Pa}}{2} \geq P_{01}^{F}$$

$$\Rightarrow \frac{P_{01}^{La} + P_{01}^{Pa}}{2} \geq \sqrt{P_{01}^{La} P_{01}^{Pa}}$$

$$\Rightarrow P_{01}^{La} + P_{01}^{Pa} \geq 2\sqrt{P_{01}^{La} P_{01}^{Pa}}$$

$$\Rightarrow \left(P_{01}^{La} + P_{01}^{Pa}\right)^2 \geq 4 P_{01}^{La} P_{01}^{Pa}$$

$$\Rightarrow \left(P_{01}^{La} - P_{01}^{Pa}\right)^2 \geq 0$$

The Chain Index Numbers

In fixed base method the base remain constant through out i.e. the relatives for all the years are based on the price of that single year. On the other hand in chain base method, the relatives for each year is found from the prices of the immediately preceding year. Thus the base changes from year to year. Such index numbers are useful in comparing current year figures with the preceding year figures. The relatives which we found by this method are called link relatives.

$$\text{Thus link relative for current year} = \frac{Current\ years\ figure}{Previous\ years\ figure} \times 100$$

And by using these link relatives we can find the chain indices for each year by using the below formula

$$\text{Chain index for current year} = \frac{\text{Linkrelative of current year} \times \text{Chain index of previous year}}{100}$$

Note: The fixed base index number computed from the original data and chain index number computed from link relatives give the same value of the index provided that there is only one commodity, whose indices are being constructed.

Example: from the following data of wholesale prices of wheat for ten years construct index number taking a) 1998 as base and b) by chain base method

Year	*Price of Wheat (Rs. per 40 kg.)*	*Year*	*Price of Wheat (Rs. per 40 kg.)*
1998	50	2003	78
1999	60	2004	82
2000	62	2005	84
2001	65	2006	88
2002	70	2007	90

Solution. (*a*) CONSTRUCTION OF INDEX NUMBERS TAKING 1998 AS BASE

Year	*Price of Wheat*	*Index Number (1998 = 100)*	*Year*	*Price of Wheat*	*Index Number (1998 = 100)*
1998	50	100	2003	78	$\frac{78}{50} \times 100 = 156$
1999	60	$\frac{60}{50} \times 100 = 120$	2004	82	$\frac{82}{50} \times 100 = 164$
2000	62	$\frac{62}{50} \times 100 = 124$	2005	84	$\frac{84}{50} \times 100 = 168$
2001	65	$\frac{65}{50} \times 100 = 130$	2006	88	$\frac{88}{50} \times 100 = 176$
2002	70	$\frac{70}{50} \times 100 = 140$	2007	90	$\frac{90}{50} \times 100 = 180$

This means that from 1998 to 1999 there is a 20 per cent increase; from 1999 to 2000 there is a 24 per cent increase; from 2000 to 2001 there is a 30 per cent increase. If we are interested in finding out increase from 1998 to 1999, from 1999 to 2000, from 2000 to 2001, we shall have to compute the chain indices.

(*b*) CONSTRUCTION OF CHAIN INDICES

Year	*Price of Wheat*	*Link Relatives*	*Chain Indices (1998 = 100)*
1998	50	100.00	100
1999	60	$\frac{60}{50} \times 100 = 120.00$	$\frac{120 \times 100}{100} = 120$
2000	62	$\frac{62}{60} \times 100 = 103.33$	$\frac{103.33 \times 120}{100} = 124$
2001	65	$\frac{65}{62} \times 100 = 104.84$	$\frac{104.84 \times 124}{100} = 130$
2002	70	$\frac{70}{65} \times 100 = 107.69$	$\frac{107.69 \times 130}{100} = 140$
2003	78	$\frac{78}{70} \times 100 = 111.43$	$\frac{111.43 \times 140}{100} = 156$
2004	82	$\frac{82}{78} \times 100 = 105.13$	$\frac{105.13 \times 156}{100} = 164$
2005	84	$\frac{84}{82} \times 100 = 102.44$	$\frac{102.44 \times 164}{100} = 168$
2006	88	$\frac{88}{84} \times 100 = 104.76$	$\frac{104.76 \times 168}{100} = 176$
2007	90	$\frac{90}{88} \times 100 = 102.27$	$\frac{102.27 \times 176}{100} = 180$

Note: the chain indices obtained in (b) are the same as the fixed base indices obtained in (a). in fact chain index figures will always be equal to fixed index figure if there is only one series.

Example-2: Compute the chain index number with 2003 prices as base from the following table giving the average wholesale prices of the commodities A, B and C for the year 2003 to 2007

Average wholesale price (in Rs.)

Commodity	*2003*	*2004*	*2005*	*2006*	*2007*
A	20	16	28	35	21
B	25	30	24	36	45
C	20	25	30	24	30

Solution. COMPUTATION OF CHAIN INDICES

Commo-dity	*Relatives based on preceding year*				
	2003	*2004*	*2005*	*2006*	*2007*
A	100	$\frac{16}{20} \times 100 = 80$	$\frac{28}{16} \times 100 = 175$	$\frac{35}{28} \times 100 = 125$	$\frac{21}{35} \times 100 = 60$
B	100	$\frac{30}{25} \times 100 = 120$	$\frac{24}{30} \times 100 = 80$	$\frac{36}{24} \times 100 = 150$	$\frac{45}{36} \times 100 = 125$
C	100	$\frac{25}{20} \times 100 = 125$	$\frac{30}{25} \times 100 = 120$	$\frac{24}{30} \times 100 = 80$	$\frac{30}{24} \times 100 = 125$
Total of Link Relatives	300	325	375	355	310
Average of Link Relatives	100	108.33	125	118.33	103.33
Chain Index (2003 = 100)	100	$\frac{108.33 \times 100}{100} = 108.33$	$\frac{125 \times 108.33}{100} = 135.41$	$\frac{118.33 \times 135.41}{100} = 160.23$	$\frac{103.33 \times 160.23}{100} = 165.57$

Conversion of fixed based index to chain based index

$$\text{Current year C.B.I} = \frac{\textit{Current years F.B.I}}{\textit{Previous years C.B.I}} \times 100$$

Conversion of chain based index to fixed base index.

$$\text{Current year F.B.I} = \frac{\textit{Current years C.B.I} \times \textit{Previous years F.B.I}}{100}$$

Example: Compute the chain base index numbers

Year	*1980*	*1981*	*1982*	*1983*	*1984*
Fixed base index	*100*	*120*	*150*	*130*	*160*

Solution. **Base year 1980 = 100**

Year	*Fixed base indices*	*Chain base index* $\left(\frac{I_1}{I_0} \times 100\right)$
1980	100	100
1981	120	$\frac{120 \times 100}{100} = 120$
1982	150	$\frac{150}{120} \times 100 = 125$
1983	130	$\frac{130}{150} \times 100 = 86.67$
1984	160	$\frac{160}{130} \times 100 = 123.08$

Example: Calculate fixed base index numbers from the following chain base index numbers

Year	*1978*	*1979*	*1980*	*1981*	*1982*
Chain base index numbers	*120*	*140*	*120*	*130*	*150*

Solution. Computation of fixed base index numbers

Year	*Chain Base Index Numbers*	*Fixed Base Index Numbers*
1978	120	120
1979	140	$\frac{140 \times 120}{100} = 168$
1980	120	$\frac{120 \times 168}{100} = 201.60$
1981	130	$\frac{130 \times 201.60}{100} = 262.08$
1982	150	$\frac{150 \times 262.08}{100} = 393.12$

Note: It may be remembered that the fixed base index for the first year is same as the chain base index for that year.

Merits of chain index numbers:

1. The chain base method has a great significance in practice, because in economic and business data we are often concerned with making comparison with the previous period.
2. Chain base method doesn't require the recalculation if some more items are introduced or deleted from the old data.
3. Index numbers calculated from the chain base method are free from seasonal and cyclical variations.

Demerits of chain index numbers:

1. This method is not useful for long term comparison.
2. If there is any abnormal year in the series it will effect the subsequent years also.

Differences between fixed base and chain base methods:

Chain base	**Fixed base**
1. Here the base year changes	1. Base year does not changes

2. Here link relative method is used 3. Calculations are tedious 4. It can not be computed if any one year is missing 5. It is suitable for short period 6. Index numbers will be wrong if an error is committed in the calculation of link relatives	2. No such link relative method is used 3. Calculations are simple 4. It can be computed if any year is missing 5. It is suitable for long period 6. The error is confined to the index of that year only.

Base shifting:

One of the most frequent operations necessary in the use of index numbers is changing the base of an index from one period to another with out recompiling the entire series. Such a change is referred to as *'base shifting'*. The reasons for shifting the base are

1. If the previous base has become too old and is almost useless for purposes of comparison.
2. If the comparison is to be made with another series of index numbers having different base.

The following formula must be used in this method of base shifting is

Index number based on new base year = $\frac{\text{current years old index number}}{\text{new base years old index number}} \times 100$

Example:

The following are the index numbers of prices with 1998 as base year

2003	410
2004	400
2005	380
2006	370
2007	340

Shift the base from 1998 to 2004 and recast the index numbers.

Solution:

year	Index
1998	100
1999	110
2000	120
2001	200
2002	400

Index number based on new base year = $\frac{\text{current years old index number}}{\text{new base years old index number}} \times 100$

Index number for 1998 = $\frac{100}{400} \times 100 = 25$

…………………………………………

Index number for 2007= $\frac{340}{400} \times 100 = 85$

Year	**Index number (1998as base)**	**Index number (2004 as base)**	**Year**	**Index number (1998as base)**	**Index number (2004 as base)**
1998	100	$\frac{100}{400} \times 100 = 25$	2003	410	$\frac{410}{400} \times 100 = 102.5$
1999	110	$\frac{110}{400} \times 100 = 27.5$	2004	400	$\frac{400}{400} \times 100 = 100$
2000	120	$\frac{120}{400} \times 100 = 30$	2005	380	$\frac{380}{400} \times 100 = 95$

2001	200	$\frac{200}{400}\times 100=50$	2006	370	$\frac{370}{400}\times 100=92.5$
2002	400	$\frac{400}{400}\times 100=100$	2007	340	$\frac{340}{400}\times 100=85$

Splicing of two series of index numbers:

The problem of combining two or more overlapping series of index numbers into one continuous series is called *splicing*. In other words, if we have a series of index numbers with some base year which is discontinued at some year and we have another series of index numbers with the year of discontinuation as the base, and connecting these two series to make a continuous series is called splicing.

The following formula must be used in this method of splicing

$$\text{Index number after splicing} = \frac{\text{index number to be spliced} \times \text{old index number of existing base}}{100}$$

Example: The index A given was started in 1993 and continued up to 2003 in which year another index B was started. Splice the index B to index A so that a continuous series of index is made

Year	*Index A*	*Index B*	*Year*	*Index A*	*Index B*
1993	100		2002	138	
1994	110		2003	150	100
1995	112		2004		120
—			2005		140
—			2006		130
—			2007		150

Solution. INDEX *B* SPLICED TO INDEX *A*

Year	*Index A*	*Index B*	*Index B spliced to Index A 1993 as base*
1993	100		
1994	110		
1995	112		
—			
—			
2002	138		
2003	150	100	$\frac{150}{100}\times 100 = 150$
2004		120	$\frac{150}{100}\times 120 = 180$
2005		140	$\frac{150}{100}\times 140 = 210$
2006		130	$\frac{150}{100}\times 130 = 195$
2007		150	$\frac{150}{100}\times 150 = 225$

The spliced index now refers to 1993 as base and we can make a continuous comparison of index numbers from 1993 onwards.

In the above case it is also possible to splice the new index in such a manner that a comparison could be made with 2003 as base. This would be done by multiplying the old index by the ratio $\frac{100}{150}$. Thus the spliced index for 1993 would be $\frac{100}{150}\times 100 = 66.7$ for 1994, $\frac{100}{150}\times 110 = 73.3$, for 1995, $\frac{100}{150}\times 112 = 74.6$, etc. This process appears to be more useful because a recent year can be kept as a base. However, much would depend upon the object.

Deflating:

Deflating means correcting or adjusting a value which has inflated. It makes allowances for the effect of price changes. When prices rise, the purchasing power of money declines. If the money incomes of people remain constant between two periods and prices of commodities are doubled the purchasing power of money is reduced to half. For example if there is an increase in the price of rice from Rs10/kg in the year 1980 to Rs20/kg in the year 1982. then a person can buy only half kilo of rice with Rs10. so the purchasing power of a rupee is only 50paise in 1982 as compared to 1980.

Thus the purchasing power of money = $\frac{1}{price\ index}$

In times of rising prices the money wages should be deflated by the price index to get the figure of real wages. The real wages alone tells whether a wage earner is in better position or in worst position.

For calculating real wage, the money wages or income is divided by the corresponding price index and multiplied by 100.

i.e. Real wages = $\frac{Money\ wages}{Price\ index} \times 100$

Thus Real Wage Index= $\frac{Real\ wage\ of\ current\ year}{Real\ wage\ of\ base\ year} \times 100$

Example: The following table gives the annual income of a worker and the general Index Numbers of price during 1999-2007. Prepare Index Number to show the changes in the real income of the teacher and comment on price increase

Year	1999	2000	2001	2002	2003	2004	2005	2006	2007
income (Rs.)	3600	4200	5000	5500	6000	6400	6800	7200	7500
Price Index No.	100	120	145	160	250	320	450	530	600

Solution. INDEX NUMBER SHOWING CNANGES IN THE REAL INCOME OF THE WORKER

Year	*Income (Rs.)*	*Price Index No.*	*Real Income*	*Real income Index No.*
1999	3600	100	$\frac{3600}{100} \times 100 = 3600.00$	100.00
2000	4200	120	$\frac{4200}{120} \times 100 = 3500.00$	97.22
2001	5000	145	$\frac{5000}{145} \times 100 = 3448.27$	95.78
2002	5500	160	$\frac{5500}{160} \times 100 = 3437.50$	95.49
2003	6000	250	$\frac{6000}{250} \times 100 = 2400.00$	66.67
2004	6400	320	$\frac{6400}{320} \times 100 = 2000.00$	55.56
2005	6800	450	$\frac{6800}{450} \times 100 = 1511.11$	41.98
2006	7200	530	$\frac{7200}{530} \times 100 = 1358.49$	37.74
2007	7500	600	$\frac{7500}{600} \times 100 = 1250.00$	34.72

The method discussed above is frequently used to deflate individual values, value series or value indices. Its special use is in problems dealing with such diversified things as rupee sales, rupee inventories of manufacturer's, wholesaler's and retailer's income, wages and the like.

Cost of living index numbers (or) Consumer price index numbers:

The ***cost of living index numbers*** measures the changes in the level of prices of commodities which directly affects the cost of living of a specified group of persons at a specified place. The general index numbers fails to give an idea on cost of living of different classes of people at different places.

Different classes of people consume different types of commodities, people's consumption habit is also vary from man to man, place to place and class to class i.e. richer class, middle class and poor class. For example the cost of living of rickshaw pullers at BBSR is different from the rickshaw pullers at Kolkata. The consumer price index helps us in determining the effect of rise and fall in prices on different classes of consumers living in different areas.

Main steps or problems in construction of cost of living index numbers

The following are the main steps in constructing a cost of living index number.

1. **Decision about the class of people for whom the index is meant**

   It is absolutely essential to decide clearly the class of people for whom the index is meant i.e. whether it relates to industrial workers, teachers, officers, labors, etc. Along with the class of people it is also necessary to decide the geographical area covered by the index, such as a city, or an industrial area or a particular locality in a city.

2. **Conducting family budget enquiry**

   Once the scope of the index is clearly defined the next step is to conduct a sample family budget enquiry i.e. we select a sample of families from the class of people for whom the index is intended and scrutinize their budgets in detail. The enquiry should be conducted during a normal period i.e. a period free from economic booms or depressions. The purpose of the enquiry is to determine the amount; an average family spends on different items. The family budget enquiry gives information about the nature and quality of the commodities consumed by the people. The commodities are being classified under following heads i)Food

   ii) Clothing

   iii)Fuel and Lighting

   iv)House rent

   v) miscellaneous

3. **Collecting retail prices of different commodities**

   The collection of retail prices is a very important and at the same time very difficult task, because such prices may vary from lace to place, shop to shop and person to person. Price quotations should be obtained from the local markets, where the class of people reside or from super bazaars or departmental stores from which they usually make their purchases.

Uses of cost of living index numbers:

1. Cost of living index numbers indicate whether the real wages are rising or falling. In other words they are used for calculating the real wages and to determine the change in the purchasing power of money.

$$\text{Purchasing power of money} = \frac{1}{\text{Cost of living index number}}$$

$$Real\ Wages = \frac{\text{Money wages}}{\text{Cost of living index umbers}} \times 100$$

2. Cost of living indices are used for the regulation of D.A or the grant of bonus to the workers so as to enable them to meet the increased cost of living.
3. Cost of living index numbers are used widely in wage negotiations.
4. These index numbers also used for analyzing markets for particular kinds of goods.

Methods for construction of cost of living index numbers:

Cost of living index number can be constructed by the following formulae.

1) Aggregate expenditure method or weighted aggregative method
2) Family budget method or the method of weighted relatives

1) Aggregate expenditure method or weighted aggregative method

In this method the quantities of commodities consumed by the particular group in the base year are taken as weights. The formula is given by

$$\text{consumer price index} = \frac{\sum p_1 q_0}{\sum p_0 q_0} \times 100$$

Steps:

i) The prices of commodities for various groups for the current year is multiplied by the quantities of the base year and their aggregate expenditure of current year is obtained .i.e. $\sum p_1 q_0$

ii) Similarly obtain $\sum p_0 q_0$

iii) The aggregate expenditure of the current year is divided by the aggregate expenditure of the base year and the quotient is multiplied by 100.

Symbolically $\dfrac{\sum p_1 q_0}{\sum p_0 q_0} \times 100$

2) Family budget method or the method of weighted relatives

In this method cost of living index is obtained on taking the weighted average of price relatives, the weights are the values of quantities consumed in the base year i.e. $v = p_0 q_0$. Thus the consumer price index number is given by

$$\text{consumer price index} = \frac{\sum pv}{\sum v}$$

Where $p = \dfrac{p_1}{p_o} \times 100$ for each item

$v = p_0 q_0$, value on the base year

Note: It should be noted that the answer obtained by applying the aggregate expenditure method and family budget method shall be same.

Example: Construct the consumer price index number for 2007 on the basis of 2006 from the following data using (i) the aggregate expenditure method, and (ii) the family budget method.

Commodity	Quantity consumed in 2006	Units	Price in 2006		Price in 2007	
			Rs.	Paise	Rs.	Paise
A	6 Quintals	Quintal	5	75	6	0
B	6 Quintals	Quintal	5	0	8	0
C	1 Quintals	Quintal	6	0	9	0
D	6 Quintals	Quintal	8	0	10	0
E	4 Kg.	Kg.	2	0	1	50
F	1 Quintals	Quintal	20	0	15	0

Solution. COMPUTATION OF CONSUMER PRICE INDEX NUMBER FOR 2007 (Base 2006 = 100) BY THE AGGREGATE EXPENDITURE METHOD

Commodity	Quantities consumed	Unit	Price in 2006 p_0	Price in 2007 p_1	p_1q_0	p_0q_0
A	6 Qtl.	Qtl.	5.75	6.00	36.00	34.50
B	6 Qtl.	"	5.00	8.00	48.00	30.00
C	1 Qtl.	"	6.00	9.00	9.00	6.00
D	6 Qtl.	"	8.00	10.00	60.00	48.00
E	4 Kg.	Kg.	2.00	1.50	6.00	8.00
F	1 Qtl.	Qtl.	20.00	15.00	15.00	20.00
					$\Sigma p_1q_0 = 174$	$\Sigma p_0q_0 = 146.5$

$$\text{Consumer Price Index} = \frac{\Sigma p_1q_0}{\Sigma p_0q_0} \times 100 \; \frac{174}{146.5} \times 100 = 118.77$$

CONSTRUCTON OF CONSUMER PRICE INDEX NUMBER FOR 2007 (Base 2006 = 100) BY THE FAMILY BUDGET METHOD

Commodity	Quantities consumed q_0	Unit	Price in 2006 p_0	Price in 2007 p_1	$\frac{p_1}{p_0} \times 100$ P	p_0q_0 V	PV
A	6 Qtl:	Qtl.	5.75	6.0	104.34	34.5	3,600
B	6 Qtl.	Qtl.	5.00	8.0	160.00	30.0	4,800
C	1 Qtl.	Qtl.	6.00	9.0	150.00	6.0	900
D	6 Qtl.	Qtl.	8.00	10.0	125.00	48.0	6,000
E	4 Kg.	Kg.	2.00	1.5	75.00	8.0	600
F	1 Qtl.	Qtl.	20.00	15.0	75.00	20.0	1,500
						$\Sigma V = 146.5$	$\Sigma PV = 17,400$

$$\text{Consumer Price Index} = \frac{\Sigma PV}{\Sigma V} = \frac{17.400}{146.5} = 118.77$$

Thus, the answer is the same by both the methods. However, the reader should prefer the aggregate expenditure method because it is far more easier to apply compared to the family budget method.

Possible errors in construction of cost of living index numbers:

Cost of living index numbers or its recently popular name consumer price index numbers are not accurate due to various reasons.

1. Errors may occur in the construction because of inaccurate specification of groups for whom the index is meant.
2. Faulty selection of representative commodities resulting out of unscientific family budget enquiries.
3. Inadequate and unrepresentative nature of price quotations and use of inaccurate weights
4. Frequent changes in demand and prices of the commodity
5. The average family might not be always a representative one.

Problems or steps in construction of wholesale price index numbers (WPI):

Index numbers are the best indicators of the economic progress of a community, a nation and the world as a whole. Wholesale price index numbers can also be constructed for different economic activities such as Indices of Agricultural production, Indices of Industrial production, Indices of Foreign Trade etc. Besides some International organizations like the United Nations Organization, the F.A.O. of the U.N., the World Bank and International Labour Organization, there are a number of organizations in the country who publish index numbers on different aspects. These are (a) Ministry of Food and Agriculture, (b) Reserve Bank of India, (c) Central Statistical Organization, (d) Department of Commercial Intelligence and Statistics, (e) Labour Bureau, (f) Eastern Economist. The Central Statistical Organization of the Government of India publishes a Monthly Abstract of' Statistics which contains All India index numbers of Wholesale Prices (Revised series : Base year 1981-82) both commodity-wise and also for the aggregate.

i. Purpose or object of index numbers.

A wholesale price index number which is properly designed for a purpose can be most useful and powerful tool. Thus the first and the foremost problem are to determine the purpose of index numbers. If we know the purpose of the index numbers we can settle some related problems.

ii. Selection of commodities

Representative items should be taken into consideration. The items may be grouped into relatively homogeneous heads to make the calculation. The construction of WPI of a region or country we may group the commodities as (1) Primary Articles — (a) Food Articles (b) Non-food Articles (c) Minerals (ii) Fuel. Power, Light and Lubricants (iii) Manufactured Products (iv) Chemicals and Chemical Products (v) Machinery and Machine Equipments (vi) Other Miscellaneous Manufacturing Industries.

iii. Selection of base period

1. <u>The base period must be a normal period</u> i.e. a period frees from all sorts of abnormalities or random fluctuations such as labor strikes, wars, floods, earthquakes etc.
2. <u>The base period should not be too distant from the given period.</u> Since index numbers are essential tools in business planning and economic policies the base period should not be too far from the current period. For example for deciding increase in dearness allowance at present there is no advantage in taking 1950 or 1960 as the base, the comparison should be with the preceding year after which the DA has not been increased.
3. <u>Fixed base or chain base</u> .While selecting the base a decision has to be made as to whether the base shall remain fixing or not i.e. whether we have fixed base or chain base. In the fixed base method the year to which the other years are compared is constant. On the other hand, in chain base method the prices of a year are linked with those of the preceding year. The chain base method gives a better picture than what is obtained by the fixed base method.

iv. Data for index numbers

The data, usually the set of prices and of quantities consumed of the selected commodities for different periods, places etc. constitute the raw material for the construction of wholesale rice index numbers. The data should be collected from reliable sources such as standard trade journals, official publications etc.

v. Selection of appropriate weights

A decision as to the choice of weights is an important aspect of the construction of index numbers. The problem arises because all items included in the construction are not of equal importance. So proper weights should be attached to them to take into account their relative importance. Thus there are two type of indices.

1. Un weighted indices- in which no specific weights are attached
2. Weighted indices- in which appropriate weights are assigned to various items.

vi. Choice of average.

Since index numbers are specialized averages, a choice of average to be used in their construction is of great importance. Usually the following averages are used.

iv) A.M
v) G.M
vi) Median

Among these averages ***G.M*** is the appropriate average to be used. But in practice G.M is not used as often as A.M because of its computational difficulties.

vii. Choice of formula.

The selection of a formula along with a method of averaging depends on data at hand and purpose for which it is used. Different formulae developed for the purpose have already been discussed in earlier sections.

Wholesale price index numbers (Vs) consumer price index numbers:

1. The wholesale price index number measures the change in price level in a country as a whole. For example economic advisors index numbers of wholesale prices.
   Where as cost of living index numbers measures the change in the cost of living of a particular class of people stationed at a particular place. In this index number we take retail price of the commodities.
2. The wholesale price index number and the consumer price index numbers are generally different because there is lag between the movement of wholesale prices and the retail prices.
3. The retail prices required for the construction of consumer price index number increased much faster than the wholesale prices i.e. there might be erratic changes in the consumer price index number unlike the wholesale price index numbers.
4. The method of constructing index numbers in general the same for wholesale prices and cost of living. The wholesale price index number is based on different weighting systems and the selection of commodities is also different as compared to cost of living index number

Importance and methods of assigning weights:

The problem of selecting suitable weights is quite important and at the same time quite difficult to decide. The term weight refers to the relative importance of the different items in the construction of the index. Generally various items say wheat, rice, kerosene, clothing etc. included in the index are not of equal importance, proper weights should be attached to them to take into their relative importance. Thus there are two types of indices.

1) Unweighted indices – in which no specific weights are attached to various commodities.
2) Weighted indices – in which appropriate weights are assigned to various commodities.

The Unweighted indices can be interpreted as weighted indices by assuming the corresponding weight for each commodity being unity. But actually the commodities included in the index are all not of equal importance. Therefore it is necessary to adopt some suitable method of weighting, so that arbitrary and haphazard weights may not affect the results.

There are two methods of assigning weights.

i) Implicit weighting
ii) Explicit weighting

In *implicit weighting,* a commodity or its variety is included in the index a number of times. For example if wheat is to be given in an index twice as much times as rice then the weight of wheat is two. Where as in *explicit weighting* two types of weights can be assigned. i.e. quantity weights or value weights.

A quantity weight symbolized by q means the amount of commodity produced, distributed or consumed in some time period. A value weight in the other hand combines price with quantity produced, distributed or consumed and is denoted by v=pq.

For example quantity weights are used in the method of weighted aggregative like Lasperey's, Paasche's index numbers and value weights are used in the method of weighted average of price relatives.

Limitations or demerits of index numbers:

Although index numbers are indispensable tools in economics, business, management etc, they have their limitations and proper care should be taken while interpreting them. Some of the limitations of index numbers are

1. Since index numbers are generally based on a sample, it is not possible to take into account each and every item in the construction of index.
2. At each stage of the construction of index numbers, starting from selection of commodities to the choice of formulae there is a chance of the error being introduced.
3. Index numbers are also special type of averages, since the various averages like mean, median, G.M have their relative limitations, their use may also introduce some error.
4. None of the formulae for the construction of index numbers is exact and contains the so called *formula error*. For example Lasperey's index number has an upward bias while Paasche's index has a downward bias.
5. An index number is used to measure the change for a particular purpose only. Its misuse for other purpose would lead to unreliable conclusions.
6. In the construction of price or quantity index numbers it may not be possible to retain the uniform quality of commodities during the period of investigation.

Analysis of Time Series

➢ Introduction:

One of the most important tasks before economists and businessmen these days is to make estimates for the future. For example, a businessman is interested in finding out his likely sales in the year 2008 or as a long-term planning in 2020 or the year 2030 so that he could adjust his production accordingly and avoid the possibility of either unsold stocks or inadequate production to meet the demand. Similarly, an economist is interested in estimating the likely population in the coming year so that proper planning can be carried out with regard to food supply, jobs for the people, etc. However, the first step in making estimates for the future consists of gathering information from the past. In this connection one usually deals with statistical data which are collected, observed or recorded at successive intervals of time. Such data are generally referred to as 'time series'. Thus when we observe numerical data at different points of time the set of observations is known as time series. For example if we observe production, sales, population, imports, exports, etc. at different points of time, say, over the last 5 or 10 years, the set of observations formed shall constitute time series. Hence, in the analysis of time series, time is the most important factor because the variable is related to time which may be either year, month, week, day, hour or even- minutes or seconds.

- **Definition:** A ***time series*** is a set of observations made at specified times and arranged in a chronological order.

  The data of a time series are *bivariate data* where time is the independent variable.

For example:

1) The population of India on different census dates.
2) Year wise production of cereals.
3) Annual industrial production figures.
4) Annual sales of a departmental store etc.

Symbolically if 't' stands for time and 'y_t' represents the value at time t then the paired values (t, y_t) represents a time series data.

Ex: given below an example of production of fish in Orissa for the period from1971-72 to 1976-77.

Production of fish in Orissa (in '000 metric tons)

Year	Production
1971-72	40
1972-73	45
1973-74	40
1974-75	42
1975-76	46
1976-77	52

Uses of time series:

The analysis of time series is of great significance not only to the economists and business man but also to the scientists, astronomists, geologists, sociologists, biologists, research worker etc. for the reason below.

1) It helps in understanding past behavior.

By observing data over a period of time one can easily under stand what changes have taken place in the past. Such analysis will be extremely helpful in predicting the future behavior.

2) It helps in planning future operations.

The major use of time series analysis is in the theory of fore casting. The analysis of the past behavior enables to forecast the future. Time series forecasts are useful in planning, allocating budgets in different sectors of economy.

3) It helps in evaluating current accomplishments.

The actual performance can be compared with the expected performance and the cause of variation can be analyzed. If expected sale for 1996-97 was 10000 refrigerators and the actual sale was only 9000. One can investigate the cause for the shortfall in achievements

4) It facilitates comparison.

Different time series are often compared and important conclusions are drawn there from.

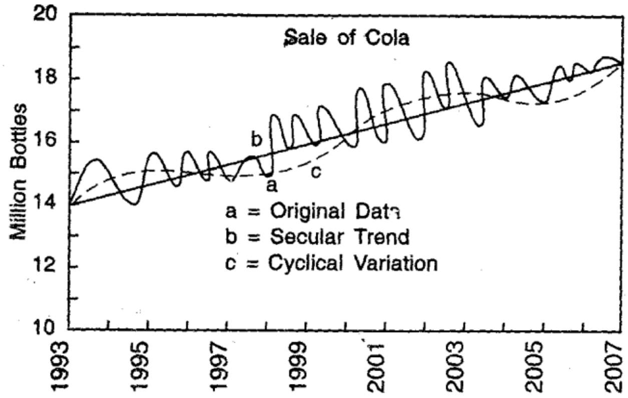
Components of time series:

The values of a time series may be affected by the number of movements or fluctuations, which are its characteristics. The types of movements characterizing a time series are called components of time series or elements of a time series.

These are four types

1. Secular Trend
2. Seasonal Variations
3. Cyclical Variations

4. Irregular Variations



Secular Trend:

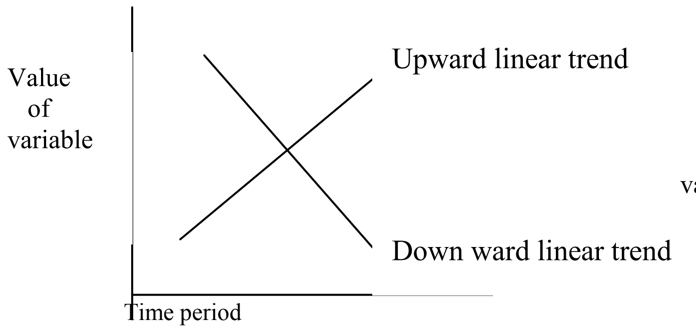
Secular Trend is also called long term trend or simply trend. The *secular trend* refers to the general tendency of data to grow or decline over a long period of time. For example the population of India over years shows a definite rising tendency. The death rate in the country after independence shows a falling tendency because of advancement of literacy and medical facilities. Here long period of time does not mean as several years. Whether a particular period can be regarded as long period or not in the study of secular trend depends upon the nature of data. For example if we are studying the figures of sales of cloth store for 1996-1997 and we find that in 1997 the sales have gone up, this increase can not be called as secular trend because it is too short period of time to conclude that the sales are showing the increasing tendency.

On the other hand, if we put strong germicide into a bacterial culture, and count the number of organisms still alive after each 10 seconds for 5 minutes, those 30 observations showing a general pattern would be called secular movement.

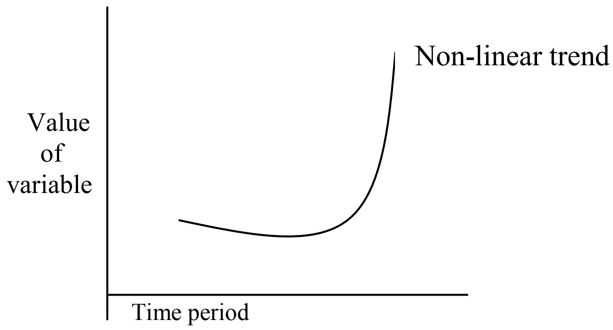
Mathematically the secular trend may be classified into two types

1. Linear Trend
2. Curvi-Linear Trend or Non-Linear Trend.

If one plots the trend values for the time series on a graph paper and if it gives a straight line then it is called a linear trend i.e. in linear trend the rate of change is constant where as in non-linear trend there is varying rate of change.



Linear trend



Non linear trend

Seasonal Variations:

Seasonal variations occur in the time series due to the rhythmic forces which occurs in a regular and a periodic manner with in a period of less than one year. Seasonal variations occur during a period of one year and have the same pattern year after year. Here the period of time may be monthly, weekly or hourly. But if the figure is given in yearly terms then seasonal fluctuations does not exist. There occur seasonal fluctuations in a time series due to two factors.

1) Due to natural forces
2) Man made convention.

The most important factor causing seasonal variations is the climate changes in the climate and weather conditions such as rain fall, humidity, heat etc. act on different products and industries differently. For example during winter there is greater demand for woolen clothes, hot drinks etc. Where as in summer cotton clothes, cold drinks have a greater sale and in rainy season umbrellas and rain coats have greater demand.

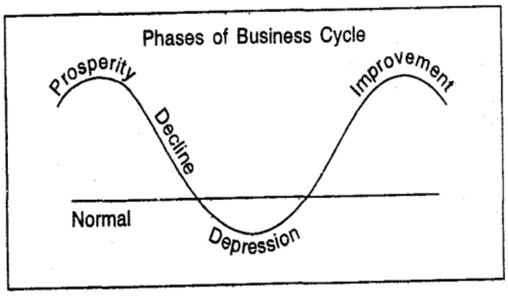
Though nature is primarily responsible for seasonal variation in time series, customs, traditions and habits also have their impact. For example on occasions like dipawali, dusserah, Christmas etc. there is a big demand for sweets and clothes etc., there is a large demand for books and stationary in the first few months of the opening of schools and colleges.

Cyclical Variations or Oscillatory Variation:

This is a short term variation occurs for a period of more than one year. The rhythmic movements in a time series with a period of oscillation(repeated again and again in same manner) more than one year is called a cyclical variation and the period is called a cycle. The time series related to business and economics show some kind of cyclical variations.

One of the best examples for cyclical variations is '*Business Cycle*'. In this cycle there are four well defined periods or phases.

1) Boom
2) Decline
3) Depression
4) Improvement



Irregular Variation:

It is also called Erratic, Accidental or Random Variations. The three variations trend, seasonal and cyclical variations are called as regular variations, but almost all the time series including the regular variation contain another variation called as random variation. This type of fluctuations occurs in random way or irregular ways which are unforeseen, unpredictable and due to some irregular circumstances which are beyond the control of human being such as earth quakes, wars, floods, famines, lockouts, etc. These factors affect the time series in the irregular ways. These irregular variations are not so significant like other fluctuations.

Combination of the various components:

The value Y_t of a time series at any time t can be expressed as the combinations of factors that can be attributed to the various components. These combinations are called as models and these are two types.

1) additive model
2) multiplicative model

Additive model:

In additive model $Y_t = T_t + S_t + C_t + R_t$

Where T_t = Trend value at time t

S_t = Seasonal component

C_t = Cyclical component

R_t = Irregular component

But if the data is in the yearly form then seasonal variation does not exist, so in that situation $Y_t = T_t + C_t + R_t$

Generally the cyclical fluctuations have positive or negative value according to whether it is in above or below the normal phase of cycle.

Multiplicative model:

In multiplicative model $Y_t = T_t.S_t.C_t.R_t$

The multiplicative model can be put in additive model by taking log both sides.

However most business analysis uses the multiplicative model and finds it more appropriate to analyze business situations.

➢ Measurement of Secular trend:

Secular trend is a long term movement in a time series. This component represents basic tendency of the series. The following methods are generally used to determine trend in any given time series. The following methods are generally used to determine trend in any given time series.

1) Free hand curve method or eye inspection method
2) Semi average method
3) Method of moving average
4) Method of least squares

year	production of cotton
1971	91
1972	111
1973	136
1974	412
1975	720
1976	900
1977	1206
1978	1322

1) Free hand curve method or eye inspection method

Free hand curve method is the simplest of all methods and easy to under stand.

The method is as follows.

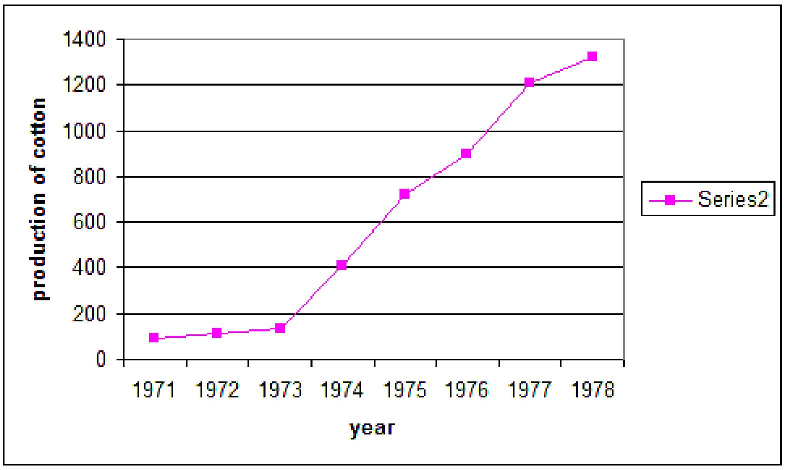
First plot the given time series data on a graph. Then a smooth free hand curve is drawn through the plotted points in such a way that it represents general tendency of the series. As the curve is drawn through eye inspection, this is also called as eye-inspection method. The free hand curve method removes the short term variations to show the basic tendency of the data. The trend line drawn through the free hand curve method can be extended further to predict or estimate values for the future time periods. As the method is subjective the prediction may not be reliable.

There is another method which is adopted while drawing a free hand curve called as method of selected points. By this method we select points on the graph of the original data and draw a smooth curve through these points. For example if we want to draw a straight line trend, two characteristic points are selected on the graph and a line is drawn through these points.

Example:

Original line

Trend line



Merits :

1) It is very simplest method for study trend values and easy to draw trend**.**
2) Some times the trend line drawn by the statistician experienced in computing trend may be considered better than a trend line fitted by the use of a mathematical formula.
3) Although the free hand curves method is not recommended for beginners, it has considerable merits in the hands of experienced statisticians and widely used in applied situations.

Demerits:

1) This method is highly subjective and curve varies from person to person who draws it.

2) The work must be handled by skilled and experienced people.
3) Since the method is subjective, the prediction may not be reliable.
4) While drawing a trend line through this method a careful job has to be done.

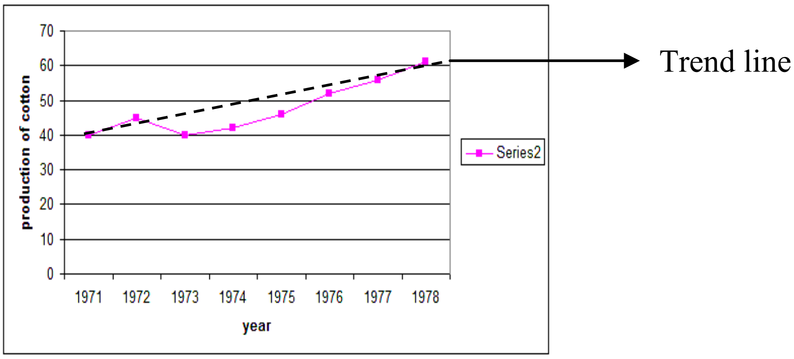
2) Method of Semi Averages:

In this method the whole data is divided in two equal parts with respect to time. For example if we are given data from 1979 to 1996 i.e. over a period of 18 years the two equal parts will be first nine years i.e. from 1979 to 1987 and 1988 to 1996. In case of odd number of years like 9, 13, 17 etc. two equal parts can be made simply by omitting the middle year. For example if the data are given for 19 years from 1978 to 1996 the two equal parts would be from 1978 to 1986 and from 1988 to 1996, the middle year 1987 will be omitted. After the data have been divided into two parts, an average (arithmetic mean) of each part is obtained. We thus get two points. Each point is plotted against the mid year of the each part. Then these two points are joined by a straight line which gives us the trend line. The line can be extended downwards or upwards to get intermediate values or to predict future values.

Example:

year	production	Semi averages
1971	40	
1972	45	$\frac{40+45+40+42}{4}=41.75$
1973	40	
1974	42	
1975	46	
1976	52	$\frac{46+52+56+61}{4}=53.75$
1977	56	
1978	61	

Thus we get two points 41.75 and 53.75 which shall be plotted corresponding to their middle years i.e. 1972.5 and 1976.5. by joining these points we shall obtain the required trend line. This line can be extended and can be used either for prediction or for determining intermediate values.



Merits:

1) This method is simple to understand as compare to moving average method and method of least squares.
2) This is an objective method of measuring trend as every one who applies this method is bound to get the same result.

Demerits:

1) The method assumes straight line relationship between the plotted points regardless of the fact whether that relationship exists or not.
2) The main drawback of this method is if we add some more data to the original data then whole calculation is to be done again for the new data to get the trend values and the trend line also changes.
3) As the A.M of each half is calculated, an extreme value in any half will greatly affect the points and hence trend calculated through these points may not be precise enough for forecasting the future.

3) Method of Moving Average:

It is a method for computing trend values in a time series which eliminates the short term and random fluctuations from the time series by means of moving average. Moving average of a period m is a series of successive arithmetic means of m terms at a time starting with 1^{st}, 2^{nd}, 3^{rd} so on. The first average is the mean of first m terms; the second average is the mean of 2^{nd} term to (m+1)th term and 3^{rd} average is the mean of 3^{rd} term to (m+2)th term and so on. If m is odd then the moving average is placed against the mid value of the time interval it covers. But if m is even then the moving average lies between the two middle periods which does not correspond to any time period. So further steps has to be taken to place the moving average to a particular period of time. For that we take 2-yearly moving average of the moving averages which correspond to a particular time period. The resultant moving averages are the trend values.

Ex:1) Calculate 3-yearly moving average for the following data.

Years	Production	3-yearly moving avg (trend values)
1971-72	40	
1972-73	45	(40+45+40)/3 = 41.67
1973-74	40	(45+40+42)/3 = 42.33
1974-75	42	(40+42+46)/3 = 42.67
1975-76	46	(42+46+52)/3 = 46.67
1976-77	52	(46+52+56)/3 = 51.33
1977-78	56	(52+56+61)/3 = 56.33
1978-79	61	

Ex:1) Calculate 4-yearly moving average for the following data.

Years	Production	4-yearly moving avg	2-yealry moving avg (trend values)
1971-72	40		
1972-73	45		
		(40+45+40+42)/3 = 41.75	
1973-74	40		42.5
		(45+40+42+46)/3 = 43.15	
1974-75	42		44.12
		(40+42+46+52)/3 = 45	
1975-76	46		47
		(42+46+52+56)/3 = 49	
1976-77	52		51.38
		(46+52+56+61)/3 = 53.75	
1977-78	56		
1978-79	61		

Merits:

1) This method is simple to under stand and easy to execute.
2) It has the flexibility in application in the sense that if we add data for a few more time periods to the original data, the previous calculations are not affected and we get a few more trend values.
3) It gives a correct picture of the long term trend if the trend is linear.
4) If the period of moving average coincides with the period of oscillation (cycle), the periodic fluctuations are eliminated.

5) The moving average has the advantage that it follows the general movements of the data and that its shape is determined by the data rather than the statistician's choice of mathematical function.

Demerits:

1) For a moving average of 2m+1, one does not get trend values for first m and last m periods.
2) As the trend path does not correspond to any mathematical; function, it can not be used for forecasting or predicting values for future periods.
3) If the trend is not linear, the trend values calculated through moving averages may not show the true tendency of data.
4) The choice of the period is some times left to the human judgment and hence may carry the affect of human bias.

4) Method of Least Squares:

This method is most widely used in practice. It is mathematical method and with its help a trend line is fitted to the data in such a manner that the following two conditions are satisfied.

1. $\sum (Y - Y_c) = 0$ i.e. the sum of the deviations of the actual values of Y and the computed values of Y is zero.
2. $\sum (Y - Y_c)^2$ is least, i.e. the sum of the squares of the deviations of the actual values and the computed values is least.

The line obtained by this method is called as the "line of best fit".

This method of least squares may be used either to fit a straight line trend or a parabolic trend.

- **Fitting of a straight line trend by the method of least squares:**

Let Y_t be the value of the time series at time t. Thus Y_t is the independent variable depending on t.

Assume a straight line trend to be of the form $Y_{tc} = a + bt$ **………… (1)**

Where Y_{tc} is used to designate the trend values to distinguish from the actual Y_t values, a is the Y-intercept and b is the slope of the trend line.

Now the values of a and b to be estimated from the given time series data by the method of least squares.

In this method we have to find out a and b values such that the sum of the squares of the deviations of the actual values Y_t and the computed values Y_{tc} is least.

i.e. $S = \sum (Y_t - Y_{tc})^2$ should be least

i.e. $S = \sum (Y_t - a - bt)^2$ **………… (2)** Should be least

Now differentiating partially (2) w.r.to a and equating to zero we get

$$\frac{\partial S}{\partial a} = 2\sum (Y_t - a - bt)(-1) = 0$$

$$\Rightarrow \sum (Y_t - a - bt) = 0$$

$$\Rightarrow \sum Y_t = \sum a + b\sum t$$

$$\Rightarrow \sum Y_t = na + b\sum t \quad \textbf{.................... (3)}$$

Now differentiating partially (2) w.r.to b and equating to zero we get

$$\frac{\partial S}{\partial b} = 2\sum (Y_t - a - bt)(-t) = 0$$

$$\Rightarrow \sum t(Y_t - a - bt) = 0$$

$$\Rightarrow \sum tY_t = a\sum t + b\sum t^2 \quad \textbf{.................... (4)}$$

The equations (3) and (4) are called '***normal equations'***

Solving these two equations we get the values of a and b say $\hat{a}$ and $\hat{b}$.

Now putting these two values in the equation (1) we get

$$Y_{tc} = \hat{a} + \hat{b}t$$

which is the required straight line trend equation.

Note: The method for assessing the appropriateness of the straight line modal is the ***method of first differences.*** If the differences between successive observations of a series are constant (nearly constant) the straight line should be taken to be an appropriate representation of the trend component.

Illustration 10. Below are given the figures of production (in thousand quintals) of a sugar factory :

Year	2001	2002	2003	2004	2005	2006	2007
Production (in '000 qtls.)	80	90	92	83	94	99	92

(*i*) Fit a straight line trend to these figures.
(*ii*) Plot these figures on a graph and show the trend line.

(M. Com., Jiwaji Univ.; M. Com., Ajmer Univ.; B. Com., HPU; B.Com., Bangalore Univ.)

Solution. (*i*) FITTING THE STRAIGHT LINE TREND

Year	*Production ('000 qtls.) Y*	*X*	*XY*	X^2	*Trend values* Y_c
2001	80	−3	−240	9	84
2002	90	−2	−180	4	86
2003	92	−1	−92	1	88
2004	83	0	0	0	90
2005	94	+1	+94	1	92
2006	99	+2	+198	4	94
2007	92	+3	+276	9	96
$N = 7$	$\Sigma Y = 630$	$\Sigma X = 0$	$\Sigma XY = 56$	$\Sigma X^2 = 28$	$\Sigma Y_c = 630$

The equation of the straight line is $Y_c = a + bX$.

To find a and b we have two normal equations

$$\sum Y = na + b\sum X$$

$$\sum XY = a\sum X + b\sum X^2$$

Since $\Sigma X = 0 ; a = \frac{\Sigma Y}{N}, \ b = \frac{\Sigma XY}{\Sigma X^2}$

$\Sigma Y = 630, \ N = 7, \ \Sigma XY = 56, \ \Sigma X^2 = 28,$

$\therefore \quad a = \frac{630}{7} = 90 \text{ ; and } b = \frac{56}{28} = 2$

Hence the equation of the straight line trend is $Y_c = 90 + 2X$.

Origin, 2004 : X units, one year; Y units, production in thousand quintals.

For $X = -3$, $Y_c = 90 + 2(-3) = 84$

For $X = -2$, $y_c = 90 + 2(-2) = 86$

For $X = -1$, $y_c = 90 + 2(-1) = 88$.

Similarly, by putting $X = 0$, 1, 2, 3, we can obtain other trend values. However, since the value of b is constant, first trend value need be obtained and then if the value b is positive we may continue adding the value of b to every preceding value. For 2002 it will be 84 + 2 = 86, for 2003 it will be 86 + 2 = 88, and so on. If b is negative, then instead of adding we will deduct.

(*ii*) The graph of the above data is given below :

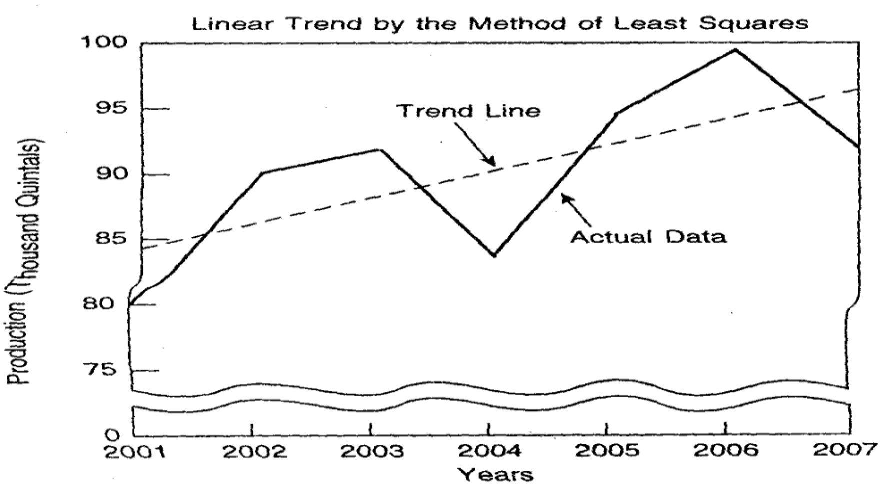


Illustration
below :
Year
Sales
Plot the (

Solution.

Years
2000
2001
2002
2003
2004
2005
2006
2007
$N = 8$

The equation of the straight line is $Y_c = a + bX$.

To find a and b we have two normal equations

$$\sum Y = na + b\sum X$$
$$\sum XY = a\sum X + b\sum X^2$$

Since $\Sigma X = 0$, $a = \frac{\Sigma Y}{N} = \frac{734}{8} = 91.75$, $b = \frac{\Sigma XY}{\Sigma X^2} = \frac{210}{168} = 1.25$

The required line equation is $Y = 91.75 + 1.25X$

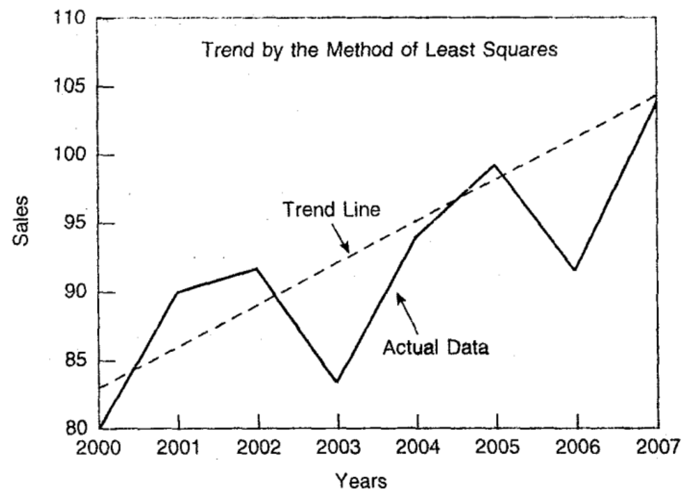
The trend values for various years are

$Y_{2000} = 91.75 + 1.25(-7) = 91.75 - 8.75 = 83$

For finding these trend values, double the value of b, *i.e.*, $1.25 \times 2 = 2.5$ and add to the preceding value :

$Y_{2001} = 83 + 2.5 = 85.5$

and so on



- **Fitting of a parabolic trend by the method of least squares**

Let Y_t be the value of the time series at time t. Thus Y_t is the independent variable depending on t.

Assume a parabolic trend to be of the form $Y_{tc} = a + bt + ct^2$ **............ (1)**

Now the values of a, b and c to be estimated from the given time series data by the method of least squares.

In this method we have to find out a, b and c values such that the sum of the squares of the deviations of the actual values Y_t and the computed values Y_{tc} is least.

i.e. $S = \sum (Y_t - Y_{tc})^2$ should be least

i.e. $S = \sum (Y_t - a - bt - ct^2)^2$ **........... (2)** Should be least

Now differentiating partially (2) w.r.to a and equating to zero we get

$$\frac{\partial S}{\partial a} = 2\sum (Y_t - a - bt - ct^2)(-1) = 0$$
$$\Rightarrow \sum (Y_t - a - bt - ct^2) = 0$$
$$\Rightarrow \sum Y_t = \sum a + b\sum t + c\sum t^2$$
$$\Rightarrow \sum Y_t = na + b\sum t + c\sum t^2 \text{ .... (3)}$$

Now differentiating partially (2) w.r.to b and equating to zero we get

$$\frac{\partial S}{\partial b} = 2\sum (Y_t - a - bt - ct^2)(-t) = 0$$

$$\Rightarrow \sum t(Y_t - a - bt - ct^2) = 0$$

$$\Rightarrow \sum tY_t = a\sum t + b\sum t^2 + c\sum t^3 \qquad \text{..................... (4)}$$

Now differentiating partially (2) w.r.to c and equating to zero we get

$$\frac{\partial S}{\partial c} = 2\sum (Y_t - a - bt - ct^2)(-t^2) = 0$$

$$\Rightarrow \sum t^2(Y_t - a - bt - ct^2) = 0$$

$$\Rightarrow \sum t^2 Y_t = a\sum t^2 + b\sum t^3 + c\sum t^4 \qquad \text{..................... (5)}$$

The equations (3), (4) and (5) are called '*normal equations'*

Solving these three equations we get the values of a, b and c say $\hat{a}, \hat{b}$ and $\hat{c}$.

Now putting these three values in the equation (1) we get

$$Y_{tc} = \hat{a} + \hat{b}t + \hat{c}t^2$$

Which is the required parabolic trend equation

Note: The method for assessing the appropriateness of the second degree equation is the ***method of second differences.*** If the differences are taken of the first differences and the results are constant (nearly constant) the second degree equation be taken to be an appropriate representation of the trend component.

Illustration 14. The prices of a commodity during 2002-2007 are given below. Fit a parabola $Y = a + bX + cX^2$ to these data. Estimate the price of the commodity for the year 2008 :

Year	*Prices*	*Year*	*Prices*
2002	100	2005	140
2003	107	2006	181
2004	128	2007	192

Also plot the actual and trend values on the graph. *(B.Com. (H), DU; M. Com., M.D. Univ.)*

Solution : To determine the values of a, b and c, we solve the following normal equations :

$$\Sigma Y = Na + b\Sigma X + c\Sigma X^2 \quad ...(i)$$

$$\Sigma XY = a\Sigma X + b\Sigma X^2 + c\Sigma X^3 \quad ...(ii)$$

$$\Sigma X^2 Y = a\Sigma X^2 + b\Sigma X^3 + c\Sigma X^4 \quad ...(iii)$$

Year	Prices (Rs.) Y	X	X^2	X^3	X^4	XY	X^2Y	Trend Values (Y_c)
2002	100	− 2	4	− 8	16	− 200	400	97.717
2003	107	− 1	1	− 1	1	− 107	107	110.401
2004	128	0	0	0	0	0	0	126.657
2005	140	+ 1	1	+ 1	1	+ 140	140	146.485
2006	181	+ 2	4	+ 8	16	+ 362	724	169.885
2007	192	+ 3	9	+ 27	81	+ 576	1728	196.857
$N = 6$	$\Sigma Y = 848$	$\Sigma X = 3$	$\Sigma X^2 = 19$	$\Sigma X^3 = 27$	$\Sigma X^4 = 115$	$\Sigma XY = 771$	$\Sigma X^2Y = 3,099$	$\Sigma Y_c =$ 848.002

$$848 = 6a + 3b + 19c \quad ...(i)$$

$$771 = 3a + 19b + 27c \quad ...(ii)$$

$$3,099 = 19a + 27b + 115c \quad ...(iii)$$

Multiplying the second equation by 2 and keeping the first as it is, we get .

$$848 = 6a + 3b + 19c$$

$$1,542 = 6a + 38b + 54c$$

$$- \quad - \quad - \quad -$$

$$-694 = -35b - 35c \quad ...(iv)$$

or $\quad 35b + 35c = 694$

Multiplying Eqn. (*ii*) by 19 and Eqn. (*iii*) by 3, we get

$$14,649 = 57a + 361b + 513c$$

$$9,297 = 57a + 81b + 345c$$

$$5,352 = 280b + 168c \quad ...(v)$$

Multiplying equation (iv) by 8, we have

$$280b + 280c = 5,552$$

Solving equations (*iv*) and (*v*)

$$280b + 280c = 5,552$$

$$280b + 168c = 5,352$$

$$- \quad - \quad -$$

$$112c = 200 \quad \text{or} \quad c = 1.786$$

Substituting the value of c in Eqn. (*iv*),

$$35b + (35 \times 1.786) = 694$$

$$35b = 694 - 62.5 = 631.5 \text{ or } b = 18.042$$

$$848 = 6a + 3(18.042) + 19(1.786) = 6a + 54.126 + 33.934$$

$$6a = 759.94 \quad \text{or} \quad a = 126.657$$

Thus $\quad a = 126.657,\ b = 18.042$ and $c = 1.786$

Substituting these values in the equation,

$$Y = 126.657 + 18.042X + 1.786X^2$$

when $X = -2$

$$Y = 126.657 + 18.042(-2) + 1.786(-2)^2$$

$$= 126.657 - 36.084 + 7.144 = 97.717$$

when $X = -1$

$$Y = 126.657 + 18.042(-1) + 1.786(-1)^2$$

$$= 126.657 - 18.042 + 1.786 = 110.401$$

when $X = 1$,

$$Y = 126.657 + 18.042 + 1.786 = 146.485$$

when $X = 2$,

$$Y = 126.657 + 18.042\ (2) + 1.786\ (2)^2 = 169.885$$

when $X = 3$,

$$Y = 126.657 + 18.042\ (3) + 1.786\ (3)^2 = 196.857$$

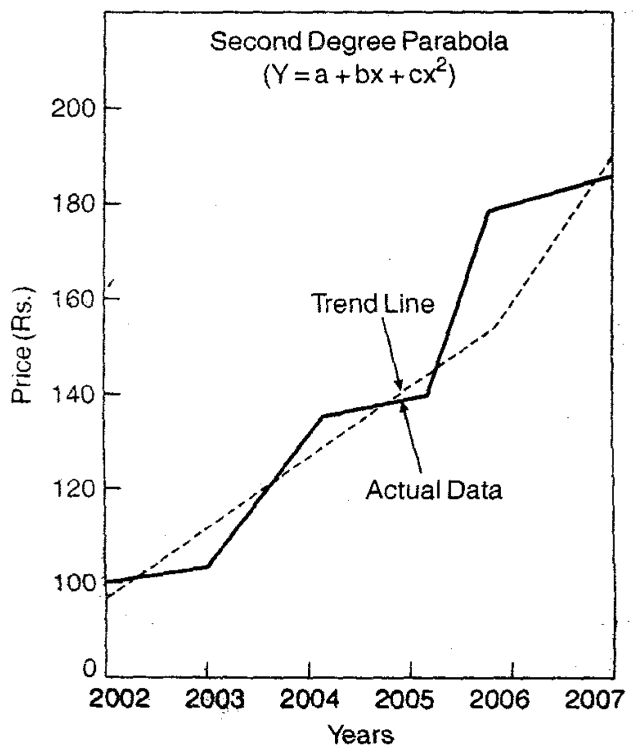
Price for the year 2008

For 2008 X would be equal to 4. Putting $X = 4$ in the equation,

$$Y = 126.657 + 18.042\ (4) + 1.786\ (4)^2$$
$$= 126.657 + 72.168 + 28.576 = 227.401.$$

Thus the likely price of the commodity for the year 2008 is Rs. 227.41 approx.

The graph of the actual and trend values is given below:



Merits:

1. This is a mathematical method of measuring trend and as such there is no possibility of subjectiveness i.e. every one who uses this method will get same trend line.
2. The line obtained by this method is called *the line of best fit.*
3. Trend values can be obtained for all the given time periods in the series.

Demerits:

1. Great care should be exercised in selecting the type of trend curve to be fitted i.e. linear, parabolic or some other type. Carelessness in this respect may lead to wrong results.
2. The method is more tedious and time consuming.
3. Predictions are based on only long term variations i.e trend and the impact of cyclical, seasonal and irregular variations is ignored.
4. This method can not be used to fit the growth curves like Gompertz curve $\left(Y = K\,a^{b^X}\right)$, logistic curve $\left(Y = \dfrac{1}{K + ab^X}\right)$ etc.

SHIFTING THE TREND ORIGIN

For simplicity and ease of computation, trends are usually fitted to annual data with the middle of the series as origin. At times it may be necessary to change the origin of the trend equation to some other point in the series. For example, annual trend values must be changed to monthly or quarterly values if we wish to study seasonal or cyclical patterns.

The shifting of the trend origin is a simple process. For an arithmetic straight line we have to find out new Y intercept, i.e., the value of a. The value of b' remains unchanged, since the slope of the trend line is the same irrespective of the origin. The procedure of shifting the origin may be generalized by the expression.

$$Yt = a + b\,(X + k)$$

where k is the number of time units shifted. If the origin is shifted forward in time, k is positive, if shifted backward in time, k is negative.

Illustration 16. You are given the trend equation

$$Y = 110 = 2\,X \qquad \text{(origin 2001, time unit = 1 year)}$$

Shift the origin to 2005.

Solution. We are required to shift the origin to 2005, *i.e.*, 4 years forward. Here $k = 4$. The required equation can be obtained as follows :

$$Y_1 = a + b\,(X + k)$$
$$= 110 + 2\,(X + 4) = 110 + 2\,X + 8 = 118 + 2\,X \qquad \text{(origin 2005, unit = 1 year)}$$

(*b*) You are given the trend equation

$$Y = 210 - 1.5\,X \qquad \text{(origin 2005, time unit = 1 year)}$$

Shift the origin to 2000.

Solution. Changing origin from 2005 to 2000 means going back by 5 years. Using the formula

$$Y_1 = a + b\,(X + k)$$
$$= 210 - 1.5\,(X - 5) = 210 - 1.5\,X + 7.5 = 217.5 - 1.5\,X$$

(origin 2000, time period one year)

The formula explained above can be expanded to cover parabolic trend equations.

Illustration. You are given the following equation :

$$Y = 126.55 + 18.04\,X + 1.786\,X^2 \qquad \text{(origin 2004-05)}$$

If we wish to shift the origin for this equation to 2005, we may follow the procedure suggested above

$$Y_1 = 126.55 + 18.04\,(X + 0.5) + 1.786\,(X + 0.5)^2$$
$$= 126.55 + 18.04\,X + 9.02 + 1.786\,(X^2 + x + 0.25)$$
$$= 126.55 + 18.04\,X + 9.02 + 1.786\,x^2 + 1.786\,X + 0.4465$$
$$= 136.0165 + 19.826\,X + 1.786\,X^2$$

It should be noted that in the parabolic trend equation C is constant and hence its value would remain unchanged.

CONVERSION OF ANNUAL TREND VALUES TO MONTHLY VALUES

From annual trend equations we can obtain monthly trend equations without any loss in accuracy. When the Y units are annual totals then an annual trend equation can be converted into an equation for monthly totals by dividing the computed constant 'a' by 12 and the value of 'b' by 144. Justification of dividing 'a and 'b' by 12 and 144 is that the data are sums of 12 months hence 'a' and 'b' must be divided by 12 and 'b' is again divided by 12 so that the time units (X's) will be in months as well, i.e., 'b' would give monthly increments. Thus the monthly trend equation becomes

$$Y = \frac{a}{12} + \frac{b}{144}\,X.$$

The annual trend equation can also be reduced to quarterly trend equation which will be given by:

$$Y = \frac{a}{4} + \frac{b}{4 \times 12}\,X \quad \text{or} \quad \frac{a}{4} + \frac{b}{48}\,X.$$

Illustration 17. The trend of the annual sales of ABC Co. Ltd. is described by the following equation :

$Y_C = 30 + 3.6\,X$ (origin 2006, X unit = 1 year, Y unit = annual sales)

Convert the equation to monthly basis.

Solution. To convert an annual trend equation to monthly basis, the value of '*a*' is divided by 12 and the value of '*b*' by 144. The equation on monthly basis is :

$$Y_C = \frac{30}{12} + \frac{3.6}{144} X$$

$$Y_C = 2.5 + 0.025\,X.$$

If the annual trend equation is of second degree, the corresponding monthly trend equation is obtained by dividing '*a*' by 12, '*b*' by 144 and '*c*' by 1,728 (the last being identical to dividing '*c*' by 12 three times).

Illustration 18. Convert the following annual trend equation on a monthly basis :

$Y = 10.6 + 0.8X + 0.64\,X^2$

Solution. To convert annual trend equation of the second degree on monthly basis divide '*a*' by 12, '*b*' by 144 and '*c*' by 1,728. Thus, the required equation is :

$$Y = \frac{10.6}{12} + \frac{0.8}{144} X + \frac{0.64}{1728} X^2$$

$$= 0.883 + 0.0056\,X + 0.00037\,X^2$$

Where data are given as monthly averages per year, the value of '*a*' remains unchanged and '*b*' is divided by 12 only once. The reason is that '*a*' is already at the monthly level and '*b*' now represents the annual change in monthly magnitudes. In case of a second-degree trend equation, the value of '*c*' is divided by 144.

Illustration 19. You are given the following trend equation :

$Y = 280 - 1.8\,X$ (origin : June 30, 2003 ; Y unit = annual monthly averages)

Convert this equation into monthly terms and shift the origin half a month forward.

Solution. The desired equation can be obtained as follows :

$$Y = 280 - \frac{1.8}{12}(X + 0.5)$$

$$= 280 - 0.15\,X - 0.075 = 279.925 - 0.15\,X$$

(origin : July 15, 1993, X = 1 month)

Measurement of seasonal variations:

Seasonal variations are regular and periodic variations having a period of one year duration. Some of the examples which show seasonal variations are production of cold drinks, which are high during summer months and low during winter season. Sales of sarees in a cloth store which are high during festival season and low during other periods.

The reason for determining seasonal variations in a time series is to isolate it and to study its effect on the size of the variable in the index form which is usually referred as seasonal index.

The study of seasonal variation has great importance for business enterprises to plan the production schedule in an efficient way so as to enable them to supply to the public demands according to seasons.

There are different devices to measure the seasonal variations. These are

1. Method of simple averages.
2. Ratio to trend method
3. Ratio to moving average method
4. Link relative method.

1. Method of simple averages.

This is the simplest of all the methods of measuring seasonality. This method is based on the additive modal of the time series. That is the observed values of the series is expressed by $Y_t = T_t + S_t + C_t + R_t$ and in this method we assume that the trend component and the cyclical component are absent.

The method consists of the following steps.

1. Arrange the data by years and months (or quarters if quarterly data is given).
2. Compute the average $\bar{x}_i$ (i = 1,2,…..12 for monthly and i=1,2,3,4 for quarterly) for the i th month or quarter for all the years.

3. Compute the average $\bar{x}$ of the averages.

i.e. $\bar{x} = \frac{1}{12}\sum_{i=1}^{12}\bar{x}_i$ for monthly and $\bar{x} = \frac{1}{4}\sum_{i=1}^{4}\bar{x}_i$ for quarterly

4. Seasonal indices for different months (quarters) are obtained by expressing monthly (quarterly) averages as percentages of $\bar{x}$. Thus seasonal indices for i th month (quarter) = $\frac{\bar{x}_i}{\bar{x}} \times 100$

Merits and Demerits:

Method of simple average is easy and simple to execute.

This method is based on the basic assumption that the data do not contain any trend and cyclic components. Since most of the economic and business time series have trends and as such this method though simple is not of much practical utility.

Example:

Assuming that the trend is absent, determine if there is any seasonality in the data given below.

Year	Ist Quarter	2nd Quarter	3rd Quarter	4th Quarter
2004	3.7	4.1	3.3	3.5
2005	3.7	3.9	3.6	3.6
2006	4.0	4.1	3.3	3.1
2007	3.3	4.4	4.0	4.0

What are the seasonal indices for various quarters ? *(M. Com.. M.K. Univ.)*

Solution. COMPUTATION OF SEASONAL INDICES

Year	Ist Quarter	2nd Quarter	3rd Quarter	4th Quarter
2004	3.7	4.1	3.3	3.5
2005	3.7	3.9	3.6	3.6
2006	4.0	4.1	3.3	3.1
2007	3.3	4.4	4.0	4.0
Total	14.7	16.5	14.2	14.2
Average	3.675	4.125	3.55	3.55
Seasonal Index	98.66	110.74	95.30	95.30

Notes for calculating seasonal index

The average of averages $= \frac{3.675 + 4.125 + 3.55 + 3.55}{4} = \frac{14.9}{4} = 3.725$

$$\text{Seasonal Index} = \frac{\text{Quarterly average}}{\text{General average}} \times 100$$

Seasonal Index for the first quarter $= \frac{3.675}{3.725} \times 100 = 98.66$

Seasonal Index for the second quarter $= \frac{4.125}{3.725} \times 100 = 110.74$

Seasonal Index for the third and fourth quarters $= \frac{3.55}{3.725} \times 100 = 95.30$

2. Ratio to trend method:

This method is an improvement over the simple averages method and this method assumes a multiplicative model i.e

$$Y_t = T_t S_t C_t R_t$$

The measurement of seasonal indices by this method consists of the following steps.

1. Obtain the trend values by the least square method by fitting a mathematical curve, either a straight line or second degree polynomial.

2. Express the original data as the percentage of the trend values. Assuming the multiplicative model these percentages will contain the seasonal, cyclical and irregular components.

3. The cyclical and irregular components are eliminated by averaging the percentages for different months (quarters) if the data are In monthly (quarterly), thus leaving us with indices of seasonal variations.

4. Finally these indices obtained in step(3) are adjusted to a total of 1200 for monthly and 400 for quarterly data by multiplying them through out by a constant K which is given by

$$K = \frac{1200}{Total\ of\ the\ indices} \text{ for monthly}$$

$$K = \frac{400}{Total\ of\ the\ indices} \text{ for quarterly}$$

Merits:

1. It is easy to compute and easy to understand.
2. Compared with the method of monthly averages this method is certainly a more logical procedure for measuring seasonal variations.
3. It has an advantage over the ratio to moving average method that in this method we obtain ratio to trend values for each period for which data are available where as it is not possible in ratio to moving average method.

Demerits:

1. The main defect of the ratio to trend method is that if there are cyclical swings in the series, the trend whether a straight line or a curve can never follow the actual data as closely as a 12- monthly moving average does. So a seasonal index computed by the ratio to moving average method may be less biased than the one calculated by the ratio to trend method.

Year	*1st Quarter*	*2nd Quarter*	*3rd Quarter*	*4th Quarter*
2003	30	40	36	34
2004	34	52	50	44
2005	40	58	54	48
2006	54	76	68	62
2007	80	92	86	82

Solution. For determining seasonal variation by ratio-to-trend method, first we will determine the trend for yearly data and then convert it to quarterly data.

CALCULATING TREND BY METHOD OF LEAST SQUARES

Year	*Yearly totals*	*Yearly average Y*	*Deviations from mid-year X*	*XY*	X^2	*Trend values*
2003	140	35	− 2	− 70	4	32
2004	180	45	− 1	− 45	1	44
2005	200	50	0	0	0	56
2006	260	65	+ 1	+ 65	1	68
2007	340	85	+ 2	+ 170	4	80
$N = 5$		$\Sigma Y = 280$		$\Sigma XY = 120$	$\Sigma X^2 = 10$	

The equation of the straight line trend is $Y = a + bX$.

$$a = \frac{\Sigma Y}{N} = \frac{280}{5} = 56 \qquad b = \frac{\Sigma XY}{\Sigma X^2} = \frac{120}{10} = 12$$

Quarterly increment = $\frac{12}{4} = 3$.

Calculation of Quarterly Trend Values. Consider 2003, trend value for the middle quarter, *i.e.*, half of 2nd and half of 3rd is 32. Quarterly increment is 3. So the trend value of 2nd quarter is $32 - \frac{3}{2}$, *i.e.*, 30.5 and for 3rd quarter is $32 + \frac{3}{2}$, *i.e.*, 33.5. Trend value for the 1st quarter is 30.5 − 3, *i.e.*, 27.5 and of 4th quarter is 33.5 + 3, *i.e.*, 36.5. We thus get quarterly trend values as shown below :

TREND VALUES

Year	*1st Quarter*	*2nd Quarter*	*3rd Quarter*	*4th Quarter*
2003	27.5	30.5	33.5	36.5
2004	39.5	42.5	45.5	48.5
2005	51.5	54.5	57.5	60.5
2006	63.5	66.5	69.5	72.5
2007	75.5	78.5	81.5	84.5

The given values are expressed as percentage of the corresponding trend values.

Thus for 1st Qtr. of 2003, the percentage shall be $(30/27.5) \times 100 = 109.09$, for 2nd Qtr. $(40/30.5) \times 100 = 131.15$, etc.

GIVEN QUARTERLY VALUES AS % OF TREND VALUES

Year	*1st Quarter*	*2nd Quarter*	*3rd Quarter*	*4th Quarter*
2003	109.09	131.15	107.46	93.15
2004	86.08	122.35	109.89	90.72
2005	77.67	106.42	93.91	79.34
2006	85.04	114.29	97.84	85.52
2007	105.96	117.20	105.52	97.04
Total	463.84	591.41	514.62	445.77
Average	92.77	118.28	102.92	89.15
S.I. Adjusted	92.05	117.36	102.12	88.46

Total of averages = 92.77 + 118.28 + 102.92 + 89.15 = 403.12.

Since the total is more than 400 an adjustment is made by multiplying each average by $\frac{400}{403.12}$ and final indices are obtained.

3. Ratio to moving average method:

The ratio to moving average method is also known as percentage of moving average method and is the most widely used method of measuring seasonal variations. The steps necessary for determining seasonal variations by this method are

1. Calculate the centered 12-monthly moving average (or 4-quarterly moving average) of the given data. These moving averages values will eliminate S and I leaving us T and C components.
2. Express the original data as percentages of the centered moving average values.
3. The seasonal indices are now obtained by eliminating the irregular or random components by averaging these percentages using A.M or median.
4. The sum of these indices will not in general be equal to 1200 (for monthly) or 400 (for quarterly). Finally the adjustment is done to make the sum of the indices to a total of 1200 for monthly and 400 for quarterly data by multiplying them through out by a constant K which is given by $K = \frac{1200}{Total\ of\ the\ indices}$ for monthly

$$K = \frac{400}{Total\ of\ the\ indices} \text{ for quarterly}$$

Merits:

1. Of all the methods of measuring seasonal variations, the ratio to moving average method is the most satisfactory, flexible and widely used method.
2. The fluctuations of indices based on ratio to moving average method is less than based on other methods.

Demerits:

1. This method does not completely utilize the data. For example in case of 12-monthly moving average seasonal indices cannot be obtained for the first 6 months and last 6 months.

Illustration 24. Calculate seasonal indices by the ratio to moving average method, from the following data :

Year	*1st Quarter*	*2nd Quarter*	*3rd Quarter*	*4th Quarter*
2005	68	62	61	63
2006	65	58	66	61
2007	68	63	63	67

Solution. CALCULATION OF SEASONAL INDICES BY 'RATIO TO MOVING AVERAGE' METHOD

Year	*Quarter*	*Given figures*	*4-figure moving totals*	*2-figure moving totals*	*4-figure moving average*	*Given figure as % of moving average*
2005	I	68				
	II	62				
			254			
	III	61		505	63.186	96.54
			251			
	IV	63		498	62.260	101.19
			247			
2006	I	65		499	62.375	104.21
			252			
	II	58		502	62.750	92.43
			250			
	III	66		503	62.875	104.97
			253			
	IV	61		511	63.875	95.50
			258			
2007	I	68		513	64.125	106.04
			255			
	II	63		516	64.500	97.67
			261			
	III	63				
	IV	67				

CALCULATION OF SEASONAL INDEX

Year	*Percentage to Moving Average*			
	1st Quarter	*2nd Quarter*	*3rd Quarter*	*4th Quarter*
2005	—	—	96.63	101.20
2006	104.21	92.43	104.97	95.50
2007	106.04	97.67	—	—
Total	210.25	190.10	201.60	196.70
Average	105.125	95.05	100.80	98.35
Seasonal Index	105.30	95.21	100.97	98.52

Arithmetic average of averages $= \frac{399.32}{4} = 99.83$

By expressing each quarterly average as percentage of 99.83, we will obtain seasonal indices.

Seasonal index of 1st Quarter $= \frac{105.125}{99.83} \times 100 = 105.30$

Seasonal index of 2nd Quarter $= \frac{95.05}{99.83} \times 100 = 95.21$

Seasonal index of 3rd Quarter $= \frac{100.80}{99.83} \times 100 = 100.97$

Seasonal index of 4th Quarter $= \frac{98.35}{99.83} \times 100 = 98.52$

4. Link relative method:

This method is slightly more complicated than other methods. This method is also known as Pearson's method. This method consists in the following steps.

1. The link relatives for each period are calculated by using the below formula

$$Link\ relative\ for\ any\ period = \frac{Current\ periods\ figure}{\Pr evious\ periods\ figure} \times 100$$

2. Calculate the average of the link relatives for each period for all the years using mean or median.
3. Convert the average link relatives into chain relatives on the basis of the first season. Chain relative for any period can be obtained by

$$\frac{Avg\ link\ relative\ for\ that\ period \times Chain\ relative\ of\ the\ previous\ period}{100}$$

   the chain relative for the first period is assumed to be 100.
4. Now the adjusted chain relatives are calculated by subtracting correction factor 'kd' from (k+1)th chain relative respectively.

   Where k = 1,2,…….11 for monthly and k = 1,2,3 for quarterly data.

   and $d = \frac{1}{N}\left[New\ chain\ relative\ for\ first\ period - 100\right]$

   where N denotes the number of periods

   i.e. N = 12 for monthly

   N = 4 for quarterly
5. Finally calculate the average of the corrected chain relatives and convert the corrected chain relatives as the percentages of this average. These percentages are seasonal indices calculated by the link relative method.

Merits:

1. As compared to the method of moving average the link relative method uses data more. completely.

Demerits:

1. The link relative method needs extensive calculations compared to other methods and is not as simple as the method of moving average.
2. The average of link relatives contains both trend and cyclical components and these components are eliminated by applying correction.

Illustration 26. Apply the method of link relatives to the following data and calculate seasonal indices :

QUARTERLY FIGURES

Quarter	*2003*	*2004*	*2005*	*2006*	*2007*
I	6.0	5.4	6.8	7.2	6.6
II	6.5	7.9	6.5	5.8	7.3
III	7.8	8.4	9.3	7.5	8.0
IV	8.7	7.3	6.4	8.5	7.1

Solution. CALCULATION OF SEASONAL INDICES BY THE METHOD OF LINK RELATIVES

	Quarter			
Year	*I*	*II*	*III*	*IV*
2003	—	108.3	120.0	111.5
2004	62.1	146.3	106.3	86.9
2005	93.2	95.6	143.1	68.8
2006	112.5	80.6	129.3	113.3
2007	77.6	110.6	109.6	88.8
Arithmetic average	$\frac{345.4}{4}=86.35$	$\frac{541.4}{5}=108.28$	$\frac{608.3}{5}=121.66$	$\frac{469.3}{5}=93.86$
Chain relatives	100	$\frac{100\times 108.28}{100}$ $= 108.28$	$\frac{121.66\times 108.28}{100}$ $= 131.73$	$\frac{93.86\times 131.73}{100}$ $= 123.64$
Corrected chain relatives	100	$108.28 - 1.675$ $= 106.605$	$131.73 - 3.35$ $= 128.38$	$123.64 - 5.025$ $= 118.615$
Seasonal indices	$\frac{100\times 100}{113.4}$ $= 88.18$	$\frac{106.605}{113.4}\times 100$ $= 94.01$	$\frac{128.38}{113.4}\times 100$ $= 113.21$	$\frac{118.615}{113.4}\times 100$ $= 104.60$

The calculations in the above table are explained below :

Chain relative of the first quarter (on the basis of first quarter) = 100

Chain relative of the first quarter (on the basis of the last quarter)

$$= \frac{86.35 \times 123.64}{100} = 106.7.$$

The difference between these chain relatives = 106.7 − 100 = 6.7.

Difference per quarter $= \frac{6.7}{4} = 1.675.$

Adjusted chain relatives are obtained by subtracting 1 × 1.675, 2 × 1.675, 3 × 1.675 from the chain relatives of the 2nd, 3rd and 4th quarters respectively.

Average of corrected chain relatives

$$= \frac{100 + 106.605 + 128.38 + 118.615}{4} = \frac{453.6}{4} = 113.4$$

$$\text{Seasonal variation index} = \frac{\text{Correct chain relatives}}{113.4} \times 100$$

Deseasonalisation:

When the seasonal component is removed from the original data, the reduced data are free from seasonal variations and is called ***deseasonalised data***. That is, under a multiplicative model

$$\frac{T \times S \times C \times I}{S} = T \times C \times I.$$

Deseasonalised data being free from the seasonal impact manifest only average value of data. Seasonal adjustment can be made by dividing the original data by the seasonal index.

$$\text{That is, Deseasonalised data} = \frac{\text{Original data}}{\text{Seasonal index}} \times 100$$

where an adjustment-multiplier 100 is necessary because the seasonal indices are usually given in percentages.

In case of additive model

$$Y_t = T + S + C + I,$$

$$\text{Deseasonalised data} = \text{Original data} - \frac{\text{Seasonal index}}{100}$$

$$= Y_t - \frac{\text{Seasonal index}}{100}.$$

Uses and limitations of seasonal indices

Seasonal indices are indices of seasonal variation and provide a quantitative measure of typical seasonal behavior in the form of seasonal fluctuations.

Measurement of cyclical variations:

The various methods used for measuring cyclical variations are

1. Residual method
2. Reference cycle analysis method
3. Direct method
4. Harmonic analysis method

BUSINESS CYCLE

According to Mitchell, "Business cycle are a type of fluctuation found in the aggregate economic activity of nations that organize their work mainly in business enterprises : a cycle consists of expansions occurring at about the same time in many activities, followed by general recessions, contractions and revivals which merge into the expansion phase of the next cycle; this sequence of changes is recurrent but not periodic; in duration business cycles vary from more than one year to ten or twelve years.

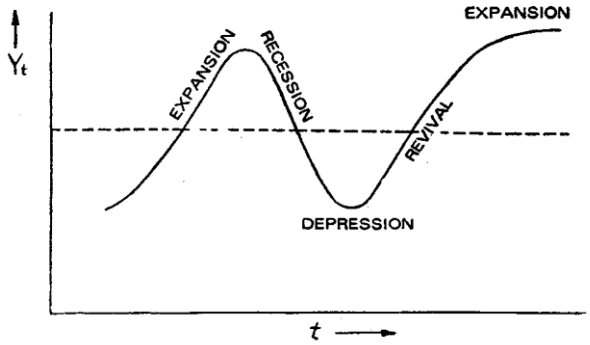
There are four phases of a business cycle, such as

(a) Expansion (prosperity)

(h) Recession

(c) Depression (contraction)

(d) Revival (recovery).



A cycle is measured either from trough-to-trough or from peak-to-peak. Recession and contraction are the result of cumulative downswing of a cycle whereas revival and expansion are the result of cumulative upswing of a cycle.