

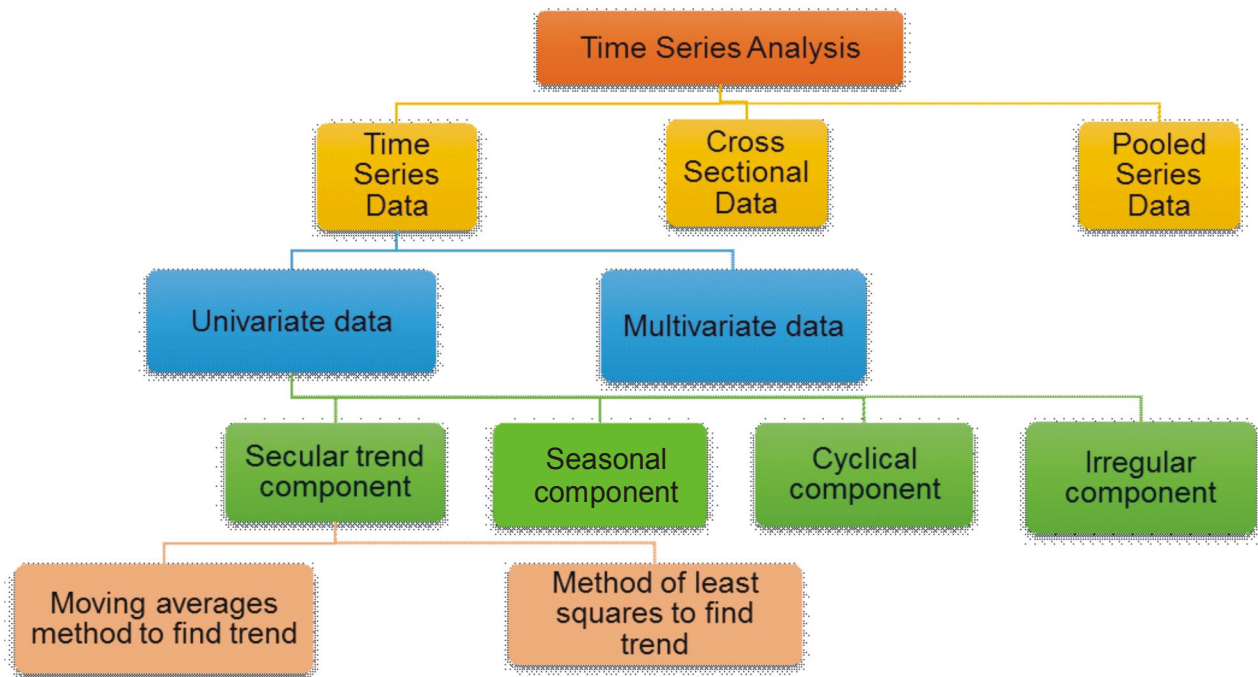
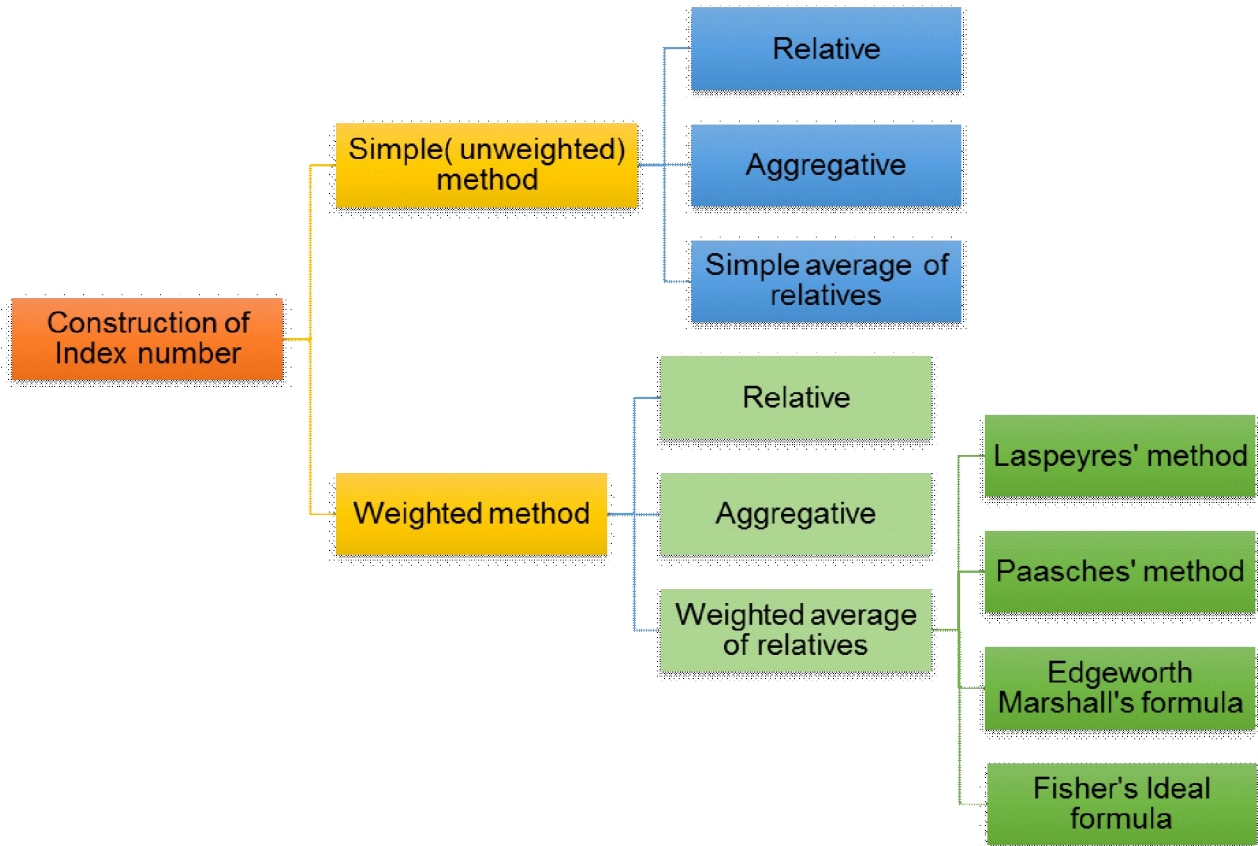
Index Numbers and Time-based Data

6.0 LEARNING OUTCOMES

At the end of this unit, the student will be able to:

- ❖ Understand the concept of an index number
- ❖ Appreciate the purpose and usefulness of index numbers in global economy
- ❖ Construct index number for simple and weighted data set
- ❖ Familiarize with limitations of index numbers (Simple and Weighted)
- ❖ Classify and validate to test adequacy of index numbers (Unit test and Time reversal test)
- ❖ Familiarize with the characteristics and components of a time series
- ❖ Acquire understanding of time series and analyse time series for univariate data sets
- ❖ Learn to compute and review trend analysis by method of moving average
- ❖ Learn to compute and review straight line trend analysis by using least squares method

6.1 Concept Map



6.2 Introduction

Recall when a teacher compares the increase in your math test scores from one term to the next, and says that performance is 25% times better or 10% lower than earlier.

We often read in the newspaper how the cost of consumer goods has increased by 30% in last decade or the economists announce that crude oil cost is up by 5.8 points.

As it turns out, these percentages or numbers are *index numbers*, numbers used in statistics and market to tell changes in various fields of economy

Index numbers is a concept applied very widely in the economic sphere, many of the prominent economic indicators are presented as index numbers, including the Consumer Prices Index (CPI), Gross Domestic Product (GDP) and the Retail Sales Index (RSI).

At NSE (National Stock exchange) and BSE (Bombay Stock Exchange) make use of Nifty and Sensex indexes.

However, the index numbers are increasingly being used in the social sphere through composite measures such as - to quantify complex concepts such as poverty and prosperity in a country.

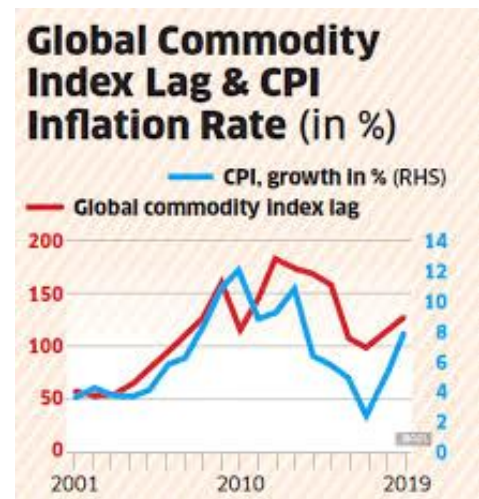
What exactly is an *index number*?

Through this unit, we shall discuss the concept of index numbers, the purpose and application of index. We shall also learn how to construct index number

6.2.1 Index Number

An index number is a measure of change in a group of related variables over two different situations with respect to time, geographical location or other characteristics. The different situations may be two different times or places. It is a measure that tracks the movement in the general level of prices of consumer goods and services.

A collection of index numbers for different years or locations etc. is called an *Index series*



6.3 Use of Index Numbers

1. Index numbers help to measure changes in the standards of living as well as prices fluctuations.
2. Government policies are framed on the basis of index numbers of prices.
3. Index Numbers not only help in the study of past and present behavior, they are also used for forecasting economic and business activities.
4. Index numbers facilitate comparative study with respect to time and place, especially where units(weights) are different.

6.4 Construction of Index Number

Let us discuss a few issues that arise in the construction of index numbers. The problem may be categorized as follows:

1. **Stating the purpose of index number:** the purpose should be clearly and unambiguously stated as an improperly constructed index may be misleading and incorrect for the analysis of data.
2. **Selection of data:** data sample of the relevant commodity should be sufficiently large sample to obtain reliable index numbers
3. **Choice of base period:** The base period should be a normal year; meaning that the prices of that period should not be subject to an erratic boom or depression or to the effect of a calamity or natural catastrophes. The base period should not be too short or too long such as the prices for a too short a period are highly unreliable
4. **Choice of commodity:** while constructing index number, including the prices of all commodities is not possible as it will be taking lot of time and efforts. Hence suitable sample of commodities is to be taken that can reflect changes in the over-all price. Selection of commodity should be by judgement sampling and not done randomly so that the results are approximate and not perfect.

All the points listed above are of varying importance and not interdependent on each other.

There are different methods of construction of index numbers that we shall learn in this module. Broadly, the calculation of price index can be divided into two subgroups, namely (a) Simple (unweighted) Aggregate method. (b) Weighted Aggregates method.

6.4.1 Relative Index Number

Example 1

Suppose we want to compare price level of crude oil for the years 2005 through 2020.

Period	Price (thousand ₹ / gallon)
December 2005	2.5
January 2010	3.5
November 2015	2.8
October 2020	2.9

Source courtesy:

<https://www.indexmundi.com/commodities/?commodity=crudeoil&months=180¤cy=INR>

In this case, year 2005 is called the *base period* and the year for which the index number is constructed, is called the *current period*.

To facilitate comparisons with other years, the actual price per gallon of the current period is converted into a *relative index number* with respect to the base period. The relative index number is calculated for prices, quantities, volume of consumption, export etc.

Choice of the base period shall be done keeping in mind that

- i) The base period shall be a normal period when the prices should not be subject to boom or a depression or to the effect of war or natural calamities
- ii) Base period shall not be too short or too long

Let p_0 and p_n denote prices of an article in two situations denoted by 0 and n respectively. A change in the price of the article can be expressed as p_n / p_0 (as a relative term)

In this case, the simplest form of price relative shows how the current price per unit is placed in comparison to a particular base period price per unit.

$$\text{Relative Price Index number in time period } n = I_n = \frac{p_n}{p_0} \times 100$$

In the same way, we can compute quantity index using the formula:

$$\text{Relative Quantity Index number in time period } n = I_n = \frac{Q_n}{Q_0} \times 100$$

For easy comparison, we shall consider the price level for the base period as 100 and the price level of a particular year in consideration is expressed relative to the base period.

Considering year 2005 as base period -

Table (i)

Period	Price (Rupees/ gallon)	Relative Index
December 2005	2,576.49	$I_{2005} = \frac{2.5}{2.5} \times 100 = 100$
January 2010	3,541.88	$I_{2005} = \frac{3.5}{2.5} \times 100 = 140$
November 2015	2,847.43	$\frac{2.8}{2.5} \times 100 = 112$
October 2020	2,931.66	$\frac{2.9}{2.5} \times 100 = 116$

NOTE: For the sake of comparison, we shall consider the price level for the base period as 100 and the price level of a particular year in consideration is expressed relative to the base period.

An index series is a list of index numbers for two or more periods of time, on the basis of same base period.

As we considered the price level on year 2005 as base period, the relative index of year 2003 will be 100 and if we compare the relative index of other years with respect to year 2005, we can say that crude oil price in year 2010 was $140 - 100 = 40\%$ above the 2005 base-year price.

In this case, the price relative index number for year 2010 is 140 (there is no need to write % symbol)

Similarly, the price relative index is 116 in year 2020 showing a 16% increase in crude oil price in year 2020 from the 2005 base period price.

The price relative index numbers, such as above are extremely useful in understanding and interpreting variations in ever so changing economic and business conditions over time.

Here, a list of index numbers have been calculated by employing same base year 2005; Therefore, the table (i) above is called an index series

Example 2

A departmental store paid annually for newspaper and television advertisements in 1990 and 2000 as shown below:

Expenditure	1990	2000
Newspaper (in ten thousand ₹)	1.3	2.9
Advertisement (in ₹)	1.8	3

Using 1990 as the base year, compute a 2000 price index for newspaper and television advertisement prices.

Also compare the relative expenditure increase between the two modes of advertisements

Solution:

I_{2000} (Newspaper)	$\frac{2.9}{1.3} \times 100 = 223$
I_{2000} (Television)	$\frac{3}{1.8} \times 100 = 167$

Clearly, Newspaper advertising cost increased at a greater rate as compared to Television advertisement cost.

NOTE: An important consideration in the construction of index numbers is the objective of the index numbers as they are constructed with specific purpose. No single index is 'all purpose' index number.

6.4.2 Simple (Unweighted) Aggregative Method

As discussed above, the relative index is useful to measure price changes over a period of time for individual items. But if we want to construct an index to be based on the price change for a group of items such as housing, food, medical care cost, stock market, transportation etc., we use an aggregated index. The purpose of aggregated index is to measure the collective change in a group of items.

This method consists of expressing aggregate of prices in any year as a percentage of their aggregate in base year

A simple aggregative index is constructed as follows:

$$\text{Price index number in time period } n = I_n = \frac{\sum p_n}{\sum p_o} \times 100$$

Where, I_n represents unweighted aggregate Index

$\sum p_0$ represents sum of unit prices for the base period 0

And, $\sum p_n$ represent sum of unit prices for the current period n

And, the quantity index:

$$\text{Quantity index number in time period } n = I_n = \frac{\sum Q_n}{\sum Q_0} \times 100$$

Example 3

A manufacturer purchases four distinct raw materials, that differ in unit price as given below:

COMMODITY	UNIT PRICE (₹) Year 2000	UNIT PRICE (₹) Year 2008
A	3.20	3.8
B	1.70	2.1
C	148.10	149.50
D	34	45

Calculate an unweighted aggregate price index for year 2008 using year 2000 as the base period.

Solution:

COMMODITY	UNIT PRICE (₹) Year 2000	UNIT PRICE (₹) Year 2008
A	3.20	3.8
B	1.70	2.1
C	148.10	149.50
D	34	45
Total	$\sum p_0 = 187$	$\sum p_n = 200.4$

$$\begin{aligned}\text{Therefore, the index number of year 2008 on the base year 2000} &= I_{2008} = \frac{\sum p_n}{\sum p_0} \times 100 \\ &= \frac{200.4}{187} \times 100 \\ &= 107.165 \\ &\approx 107.2\end{aligned}$$

From the above example, we can conclude that the price index number of year 2008 has only increased by 7.2% over the period of 2000 to 2008.

But note that the unweighted aggregate approach is heavily influenced by the commodities with large per unit pricing. Therefore, the commodities with relatively lower unit prices such as A and B are dominated by the high unit price commodities like C and D.

Because of highly sensitivity of unweighted index as shown in the above example, this form of index number is not very accurate and useful. Therefore, it is the major flaw in using absolute quantities and not relatives. Such high unit prices become the concealed weights and tend to give out biased index number.

6.4.3 Simple Average of relatives Method

To address the concerns shared for simple aggregative method, let us construct sample average of relatives. In this method, we shall convert price of each commodity in table (i) into percentage of the base period. To construct index number, we shall calculate average of all such relatives because the index number calculated from relatives will remain the same regardless of the units of each commodity.

Example 4

COMMODITY	UNIT PRICE (₹) Year 2000	UNIT PRICE (₹) Year 2008	Relative Index
A	3.20	3.8	$\frac{3.8}{3.2} \times 100 = 119$
B	1.70	2.1	$\frac{2.1}{1.7} \times 100 = 124$
C	148.10	149.50	$\frac{149.5}{148.1} \times 100 = 101$
D	34	45	$\frac{45}{34} \times 100 = 132$

Solution

$$\text{Simple average of relative index, } I_{2008} = \frac{119+124+101+132}{4} = \frac{476}{4} = 119$$

Though there is an improvement over previously calculated index number, this method is also flawed to an extent as it gives equal importance to each commodity's relative. This amounts to incorrectness in case of different weights or quantities because the individually calculated relatives disregard the absolute quantity of each commodity.

6.4.4 Weighted Aggregative Method

Due to the limitations of methods discussed above, constructing a weighted aggregate index number provides a better and more accurate comparison when the data items have variation of weights.

Since the index number does not depend on the units in which the prices are quoted, we shall weigh prices by quantities and price relatives by values

In such case, use of a weighted index number allows greater importance to be attached to some items. Moreover, the Information also includes factors such as quantity sold or quantity consumed for each item.

In this method appropriate weights are assigned to different commodities to make them comparable and thus compatible for summation. The advantage of this method of computing index number is that the allotment of weights enables the commodities of greater importance to have more impact on index number.

A weighted aggregative index is constructed as follows:

$$\text{Index number in time period } n = I_n = \frac{\sum p_{ni} Q_i}{\sum p_{0i} Q_i} \times 100$$

Where, I_n represents weighted aggregative Index

Q_i is the quantity of usage for commodity i

p_{0i} represents unit price for the base period 0

And, p_{ni} represents unit price of commodity i for the current period n

Here, quantity is fixed and assumed not to vary over time as prices change

Example 5:

With reference to example above, what will happen to the index number for year 2000 if the commodities are used in different weights (quantities)?

COMMODITY	Quantity (weights)	UNIT PRICE (₹) Year 2000	UNIT PRICE (₹) Year 2008
A	100	3.20	3.8
B	20	1.70	2.1
C	15	148.10	149.50
D	50	34	45

$$\begin{aligned} I_{2008} &= \frac{3.8 \times 100 + 2.1 \times 20 + 149.5 \times 15 + 45 \times 50}{3.2 \times 100 + 1.7 \times 20 + 148.1 \times 15 + 34 \times 50} \times 100 \\ &= \frac{4914.5}{4275.5} \times 100 \\ &= 115 \end{aligned}$$

From this calculation of weighted aggregative index, we can conclude that the cost of raw material used by the manufacturer has increased by 15% over the period from year 2000 to year 2008. In general, a weighted aggregative index along with the quantity of usage of commodities is a preferred method to establish a price index for a group of commodities.

Clearly, compared to the simple (unweighted) aggregative index, the weighted index provides more accurate indication of the price change over a period of time. Taking the quantity of usage of each commodity into account helps to find a more precise index.

But what if the quantity of usage in current period *differs* from that of base period?

6.4.5 Laspeyres' Index

In a special case of the fixed-quantity weights considered from base period usage, the weighted aggregative index is known by a new name, *Laspeyres Index*

In 1871, French economist Laspeyere recommended that quantities of commodities consumed in base year shall be taken as weights for the purpose of calculating index numbers.

According to him, the weighted aggregative index is constructed as follows:

$$\text{Index number in time period } n = I_n^{La} = \frac{\sum p_{ni} Q_{0i}}{\sum p_{0i} Q_{0i}} \times 100$$

Where, I_n^{La} represents weighted aggregative Index using Laspeyres' method

Q_{0i} is the quantity of usage for commodity i in base period

p_{0i} represents unit price for the base period 0

And, p_{ni} represents unit price of commodity i for the current period n

Hence, in the example above, if the quantities(weights) for the group of commodities are of year 2000, then the calculated Index is based on Laspeyres Index

Since the Laspeyres index uses base period weights, it has a disadvantage of overestimating the rise in the cost of living (because people may have reduced their consumption of items that have become proportionately dearer than others)

6.4.6 Paasche Index

In 1874, German statistician Paasche suggested that for determining quantity(weights) is to revise the quantity over time. When the fixed-quantity weights are considered from current period usage, the weighted aggregative index is known by another name, *Paasche Index*

In this case, the weighted aggregative index is constructed as follows:

$$\text{Index number in time period } n = I_n^{Pa} = \frac{\sum p_{ni} Q_{ni}}{\sum p_{0i} Q_{ni}} \times 100$$

Where, I_n^{Pa} represents weighted aggregative Index by Paasche's method

Q_{ni} is the quantity of usage for commodity i in current period n

p_{0i} represents unit price for the base period 0

And, p_{ni} represents unit price of commodity i for the current period n

Why do we need this weighted aggregative Index?

So in the example above, if the quantities (weights) for the group of commodities are of year 2008, then the calculated Index is based on Paasche Index.

Paasche method has the advantage of being based on current need and usage of commodities though this method may underestimate the rise in the cost of living as the calculations are based on the current period weights. Also Paasche index construction requires a new set of weights for the year in consideration, and gathering such data-information can be time-consuming and expensive.

Let us compare and analyze the application of the two stated methods of Index construction

Example 6

Following table shows the data on energy consumption and expenditure at Badarpur Thermal Power Station, in Delhi region. Construct an aggregative price index for the energy expenditures in year 2015 using

- Laspeyres' index
- Paasche index.

Sector	Quantity (weights)		Unit Price (₹ /kWh)	
	Year 1987	Year 2015	Year 1987	Year 2015
Commercial	5416	6015	1.97	10.92
Residential	15293	20262	2.32	6.16
Industrial	21287	17832	0.79	5.13
Agriculture	9473	8804	2.25	8.10

$$\begin{aligned}
 \text{(Laspeyres Index)} \ I_{2015} &= \frac{10.92 \times 5416 + 6.16 \times 15293 + 5.13 \times 21287 + 8.10 \times 9473}{1.97 \times 5416 + 2.32 \times 15293 + 0.79 \times 21287 + 2.25 \times 9473} \times 100 \\
 &= \frac{339281.21}{84280.26} \times 100 \\
 &= 403
 \end{aligned}$$

$$\begin{aligned}
 \text{(Paasche Index)} \ I_{2015} &= \frac{10.92 \times 6015 + 6.16 \times 20262 + 5.13 \times 17832 + 8.10 \times 8804}{1.97 \times 6015 + 2.32 \times 20262 + 0.79 \times 17832 + 2.25 \times 8804} \times 100 \\
 &= \frac{353288.28}{92753.67} \times 100 \\
 &= 381
 \end{aligned}$$

NOTE: Paasche value being less than the Laspeyres indicates usage has increased faster in the lower priced sectors.

6.4.7 Fisher's Ideal method

This index calculation gives the geometric mean* of Laspeyres' and Paasche's methods.

$$\text{Index number in time period } n = I_n^F = \sqrt{\frac{\sum p_1 Q_0}{\sum p_0 Q_0} \times \frac{\sum p_1 Q_1}{\sum p_0 Q_1}} \times 100$$

When a, b and c are in a geometric progression, then b is called the geometric mean as $b^2 = ac$

6.4.8 Marshall-Edgeworth's Method

The statistician duo, Marshall and Edgeworth proposed that index number is to be calculated by taking the average of the base year and the current year.

$$\text{Index number in time period } n = I_n^{ME} = \frac{\sum p_n (Q_0 + Q_n)}{\sum p_0 (Q_0 + Q_n)} \times 100$$

6.4.9 Weighted Average Of Relatives

This method makes use of price relatives. When the base and current prices of a number of commodities with varying weights or quantities are given, then this method to construct index number is recommended.

$$\text{Index number in time period } n = I_n = \frac{\sum \frac{p_n}{p_0} (p_0 Q_0)}{\sum p_0 Q_0} \times 100$$

Example 7

Calculate the price index using weighted average of relatives method for the food consumption in a student hostel in a month. Use data of year 1997 as base year for calculations.

COMMODITY	WEIGHT	PRICE PER UNIT	
		1997	2001
Rice	14 quintals	90	120
Wheat	20 kg	30	46
Pulses	35 kg	22	34
Milk	15 litre	50	90

Solution

Commodity	Quantity Q_0	Price per unit		Price relative (base period 1997) $\frac{p_n}{p_0}$	Value Weights $p_0 Q_0$	Weighted Price Relatives $\frac{p_n}{p_0} (p_0 Q_0) = p_n Q_0$
		Year 1997 p_0	Year 2001 p_n			
Rice	14	90	120	$\frac{120}{90}$	$14 \times 90 = 1260$	$\frac{120}{90} \times 14 \times 90 = 1680$
Wheat	20	30	46	$\frac{46}{30}$	$20 \times 30 = 600$	$\frac{46}{30} \times 20 \times 30 = 920$
Pulse	35	22	34	$\frac{34}{22}$	$35 \times 22 = 770$	$\frac{34}{22} \times 35 \times 22 = 1190$
Milk	15	50	90	$\frac{90}{50}$	$15 \times 50 = 750$	$\frac{90}{50} \times 15 \times 50 = 1350$
					$\sum p_0 Q_0 = 3380$	$\sum \frac{p_n}{p_0} (p_0 Q_0) = 5140$

$$\text{Weighted price relative for year 2001 on the base period 1997} = I_{2001} = \frac{\sum \frac{p_n}{p_0} (p_0 Q_0)}{\sum p_0 Q_0} \times 100$$

$$= \frac{5140}{3380} \times 100$$

$$= 152.07$$

6.5 Types of Index Numbers

We have learnt how to compute the index number for a single item or a group of items. Now let us consider some price indexes that are important measure of business and economy.

1. *Value Index* is the measure of the average value for a particular period with that of the average period of the base period. It is used to keep stock of inventory, sales and trading etc.

Though India improved its absolute value of the Human Development Index (0.645 in 2019 from 0.642 the previous year), it dropped a place in the overall ranking

Country	HDI rank (2019)	Change from 2018
Russia	52	-3
Sri Lanka	72	1
Brazil	84	0
China	85	2
South Africa	114	1
India	131	-2
Bangladesh	133	1
Nepal	142	1
Pakistan	154	0

Picture credit: <https://www.insightsonindia.com/2020/12/17/human-development-index-2/>

2. *Quantity Index* is the measure of change in the quantity of goods (produced/ consumed/ sold) within a stipulated period of time. An example of quantity index is the Index of Industrial Production, known as IIP
3. *Price Index* is the measure relative price change over a period of time. An example of price index is the Consumer Price Index, known as CPI



Picture credit: <https://www.investopedia.com/terms/c/consumerpriceindex.asp>

6.6 Limitations of an Index Number

- There are chances for errors, given that index numbers come as a result of samples. These samples are put together after analysis and deliberation, which creates chances for errors
- It is calculated based on items which may not be in trend which in turn will create an inaccurate analysis
- Multiple methods are used to formulate index numbers. Due to this variety of methods, outcomes may bring in different set of values which may cause confusion
- Selection of representative commodities may be skewed as they are based on samples collected or considered.

6.7 Index Series

Refer to example 1 where index numbers of two or more periods of time are constructed, on the basis of same base period.

Such a list of indexes is called Index series.

Many businesses and economies make use of various type of index series such as company sales, industry sales, and inventories, measured in dollar amounts.

The purpose of such series often is to indicate increased usage (physical, for example - volume) associated with the activities.

6.8 Test of Adequacy of Index Numbers

As discussed in 6.3, there are many methods to construct an index number. The important thing to consider is the appropriate method for constructing index number for the data analysis.

It is essential for testing the consistency of a good index number. The following tests are available for checking the adequacy of index number-

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

These tests maintain consistency by verifying their adequacy. Let us learn how to verify adequacy of indexes using the unit test and the time-reversal test

6.8.1 Unit Test

According to this test, the selection of method of construction of index number should be *independent of the units* in which the pricing or quantities of commodities are available. For example, the quantities of commodities such as wheat is in kilograms while weight of milk is in liters.

This test of adequacy can be applied to all the methods discussed above except for the simple aggregative method.

6.8.2 Time-reversal Test

The time-reversal test is used to test whether the method of constructing index number will work with *any consideration of time period*. This test says that the method used should give the same ratio between one point or another for comparison; no matter which time period is taken as base period.

Basically, if the time subscripts (p_o and p_n) of a price or quantity index number are interchangeable then the resulting price/ quantity relative should be reciprocal of the original price/quantity relative – i.e. if p_o represents price of wheat in year 2013 and p_1 represent price in year 2018;

$$\text{then } p_{o1} \times p_{1o} = 1$$

Here, p_{o1} is the index for current year '1' on the basis of base year 'o'

And, p_{1o} is the index for year '0' based on year '1'

Clearly this test of adequacy cannot be tested on Laspeyres' method and Paasche's method Because

Clearly this test cannot be tested on Laspeyres' method of index number because

$$\frac{\sum p_1 Q_o}{\sum p_o Q_o} \times \frac{\sum p_o Q_1}{\sum p_1 Q_1} \neq 1$$

Also the Paasche method of index number cannot be tested for adequacy using this test as

$$\frac{\sum p_1 Q_1}{\sum p_o Q_1} \times \frac{\sum p_o Q_o}{\sum p_1 Q_o} \neq 1$$

Whereas, the Fisher's Ideal index number satisfies the time-reversal test

Example 8

Calculate Fisher's price index number for the given data and verify that it satisfies the time-reversal test.

Commodity	Price		Quantity	
	2008 (p_o)	2012 (p_1)	2008 (Q_o)	2012 (Q_1)
Rice	10	13	4	6
Wheat	15	18	7	8
Rent	25	29	5	9
Fuel	11	14	8	10

Solution–

Commodity	Price 2008 (p_o)	Quantity			$p_o Q_o$	$p_o Q_1$	$p_1 Q_o$	$p_1 Q_1$
		2012 (p_1)	2008 (Q_o)	2012 (Q_1)				
Rice	10	13	4	6	40	60	52	78
Wheat	15	18	7	8	105	120	126	144
Rent	25	29	5	9	125	225	145	261
Fuel	11	14	8	10	88	110	112	140
			Total		359	515	435	623

$$\text{Fisher's price index number} = \sqrt{\frac{\sum p_1 Q_o}{\sum p_o Q_o} \times \frac{\sum p_1 Q_1}{\sum p_o Q_1}} \times 100 = \sqrt{\frac{435}{359} \times \frac{623}{515}} \times 100 = 121.2$$

$$\begin{aligned}
 \text{Here, : } p_{01} \times p_{10} &= \sqrt{\frac{\sum p_1 Q_0}{\sum p_0 Q_0} \times \frac{\sum p_1 Q_1}{\sum p_0 Q_1}} \times \sqrt{\frac{\sum p_0 Q_1}{\sum p_1 Q_1} \times \frac{\sum p_0 Q_0}{\sum p_1 Q_0}} \\
 &= \sqrt{\frac{\sum p_1 Q_0}{\sum p_0 Q_0} \times \frac{\sum p_1 Q_1}{\sum p_0 Q_1} \times \frac{\sum p_0 Q_1}{\sum p_1 Q_1} \times \frac{\sum p_0 Q_0}{\sum p_1 Q_0}} \quad (\text{as } \sqrt{a} \cdot \sqrt{b} = \sqrt{ab}) \\
 &= \sqrt{\frac{435}{359} \times \frac{623}{515} \times \frac{515}{623} \times \frac{359}{435}} = 1
 \end{aligned}$$

According to Time-reversal test, the adequacy of Fisher's Ideal index number is verified because $p_{01} \times p_{10} = 1$

6.9 Time Series

A time series is a sequentially recorded numerical data points for a given variable arranged in a successive order to track variation. For thorough analysis, these data points are recorded at successive times or successive periods, to provide the information being sought for analysis or forecast.

An essential aspect of managing any business or economy model is planning for the future. Time series analysis is useful in analyzing how a given asset, security, or economic variable changes over a period of time. These series also help to see how a business or economic variable change over a period of time. It also gives an insight on how changes associated with the chosen data points compare to the changes in other variables over the same period of time

For example, let us analyse a time series of daily opening stock prices for a particular stock over a period of one year. In such a case, you will collect a list of all opening prices of the stock for each day in chronological order as one-year, daily opening price time-series for the stock. A time series data is analyzed using technical analysis tools to see if the stock prices show any pattern or seasonality. Such information is found useful to determine when the stock goes through peaks and troughs. Analyzing the stock time series and relating it to other variables like employment rate can provide valuable information to benefit businesses and economy.

Time series forecasting tools use information based on historic data and associated patterns to predict the future activity such as trend analysis, fluctuation analysis; though the success in predicting future patterns is not guaranteed.

Such data analysis is considered in three types:

- Time series data: when data of the variable is collected at distinct time intervals, for a specified period of time.
- Cross-sectional data: when data for one or more variables is collected at the same point in time.
- Pooled data: when data in a combination of time series data and cross-sectional data is collected.

Forecasting methods can be classified as *quantitative* and *qualitative*. Quantitative method of forecasting can be used:

- When the past information about the variable is available
- When information and data of the variable can be quantified
- On the assumption that the pattern of the past will continue in the future
- The variable has a cause-and-effect relationship with one or more other variables

When forecasting is done based on historic data of past values, it is called a time series method.

Qualitative method is generally based on expert judgement and analytical opinion to develop forecasts. One of the benefits of using these methods is that they can be applied when information on the data of the variable cannot be quantified or historic information is neither available nor applicable.

6.9.1 Time series analysis

A time series in which data of only one variable is varying over time is called a *univariate* time series/data set. For example, data collected from a temperature sensor measuring the temperature of a place every second, the data will show us only one-dimensional value - temperature.

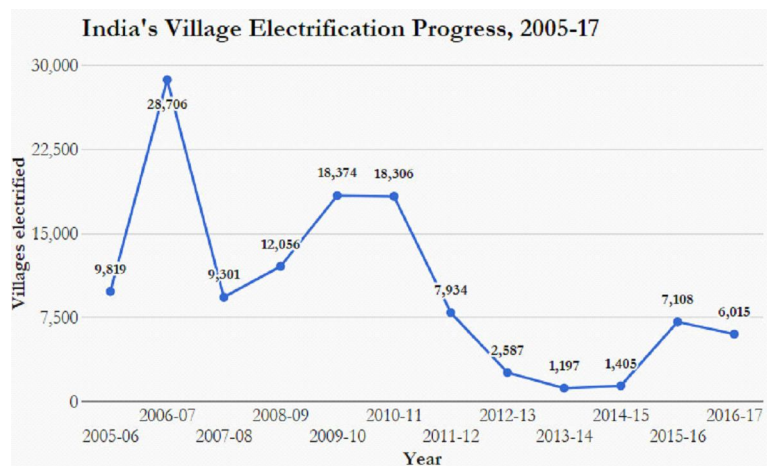


Figure 6.8.1 (i)

Source: Rural Electrification Corporation Ltd data; Power Ministry press release.

When a time series is a collection of data for multiple variables and how they are varying over time, it is called *multivariate* time series/data set.

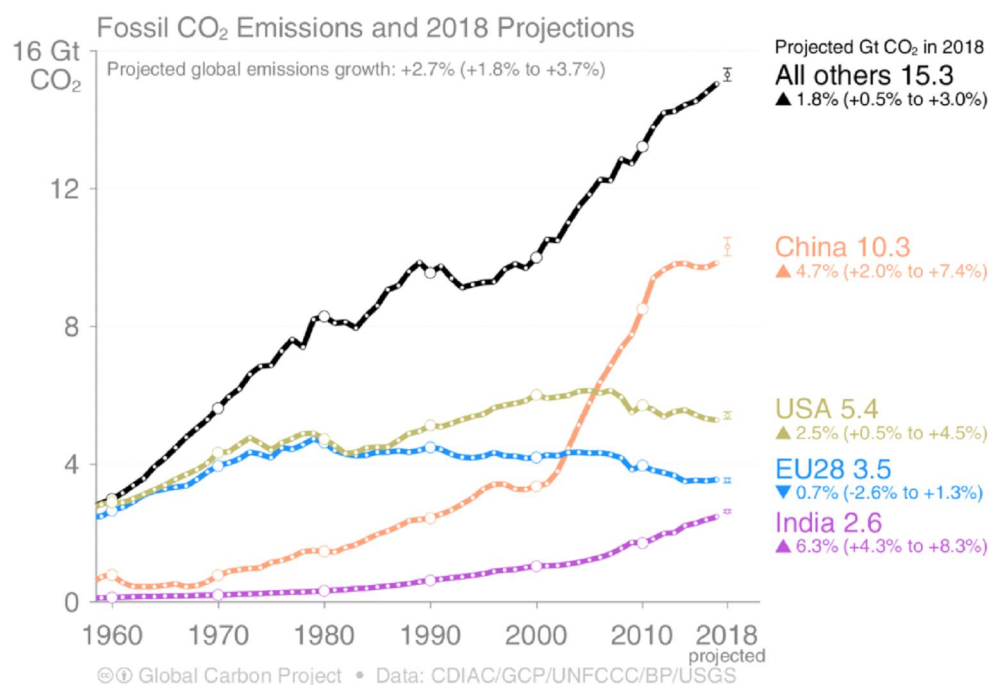


Figure 6.8.1 (ii)

Source credit: <https://www.americanexperiment.org/2019/06/global-co2-emissions-skyrocket-india-plans-build-42-shercos-coming-years/>

The patterns and behavior of the data in any time series are based on four components:

1. **Secular trend component** – also known as *trend series*, is the smooth, regular and long-term variations of the series, observed over a long period of time. Figure 6.8.1 (iii) shows an upward trend for annual electricity consumption per household in a certain residential locality from years 1990 – 2002. In general, trend variations can be either linear or non-linear.

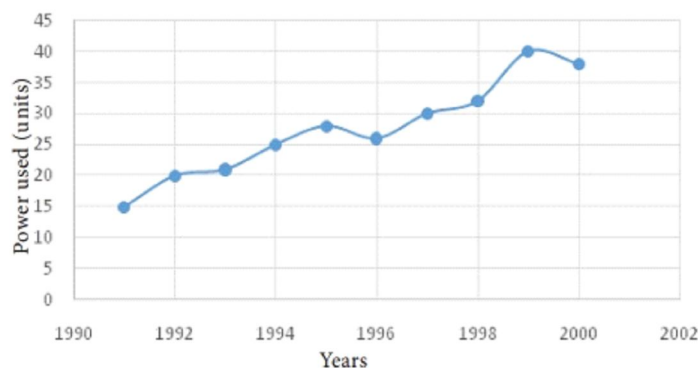


Figure 6.8.1 (iii)

2. **Seasonal component** – when a time series captures the periodic variability in the data, capturing the regular pattern of variability; within one-year periods. The main causes of such fluctuations are usually climate changes, seasons, customs and habits which people follow at different times.

Figure 6.8.1 (iv) shows seasonal electricity consumption and variations of peak demand in Nepal and India in year 2018.

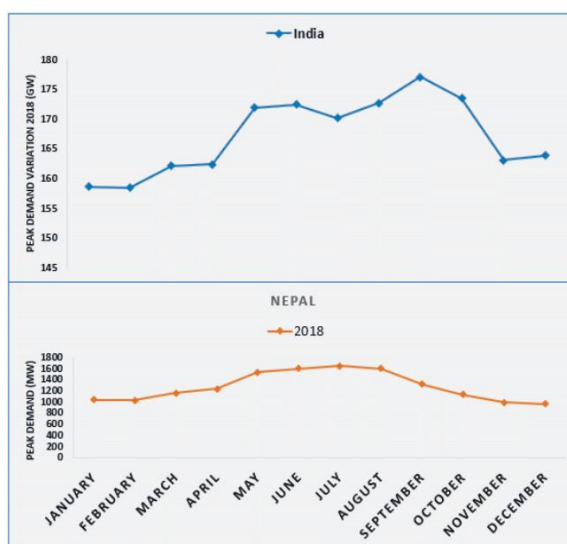


Figure 6.8.1 (iv)

3. **Cyclical component** – when a time series shows an oscillatory movement where period of oscillation is more than a year where one complete period is called a cycle.

4. **Irregular component** – these kinds of fluctuations are unaccountable, unpredictable or sometimes caused by unforeseen circumstances like – floods, natural calamities, labor strike etc. Such random variations in the time series are caused by short-term, unanticipated and nonrecurring factors that affect the time series.

Source credits: https://www.researchgate.net/figure/ndia-Nepal-Peak-Demand-seasonal-variation-in-a-year_fig2_337444939

6.9.2 Trend analysis by fitting linear trend line

Among the four components of the time series as discussed above, the secular trend analysis (also known as trend analysis) depicts the long-term direction of the series. One of the most widely used in practice mathematical techniques of finding the trend values is the method of least squares. It plays an important role in finding the trend forecasts for the future economic and business time series data

Trend can be measured using the following methods:

1. Graphical method
2. Semi averages method
3. Moving averages method
4. Method of least squares

We shall be studying two methods to compute trend line from the above list.

6.9.2 (i) Trend analysis by moving average method

This method is used to draw smooth curve for a time series data. It is mostly used for eliminating the seasonal variations for a given variable. The moving average method helps to establish a trend line by eliminating the cyclical, seasonal and random variations present in the time series. The period of the moving average depends upon the length of the time series data. As shown in the figure 6.8.2 (i), the red smooth curve is the trend-cycle, which is noticeably smoother than the original data and captures the main movement of the time series ignoring the minor fluctuations. The order of the moving average determines the smoothness of the trend-cycle estimate.

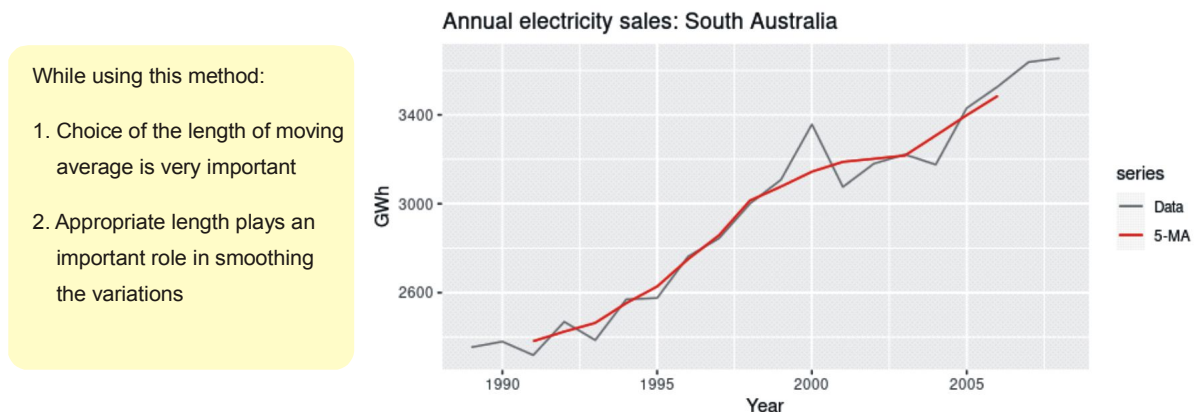


Figure 6.8.2 (i)

Picture credits: <https://otexts.com/fpp2/moving-averages.html>

Procedure for calculating Moving average for odd number of years ($n = \text{odd}$)

Let us take an example for $n = 3$ years moving averages to understand the procedure

1. Add up the values of the first 3 years and place the yearly sum against the median (middle) year. (This sum is called 3-year moving total)
2. Continue this process by leaving the first-year value, add up the next three year values and place it against its median year.
3. This process must be continued till all the values of the data are taken for calculation.
4. Calculate the n -year average by dividing each n -yearly moving total by n to get the n -year moving averages, which is our required trend values.
5. There will be no trend value for the beginning period and the ending period in this method

Example 9 :

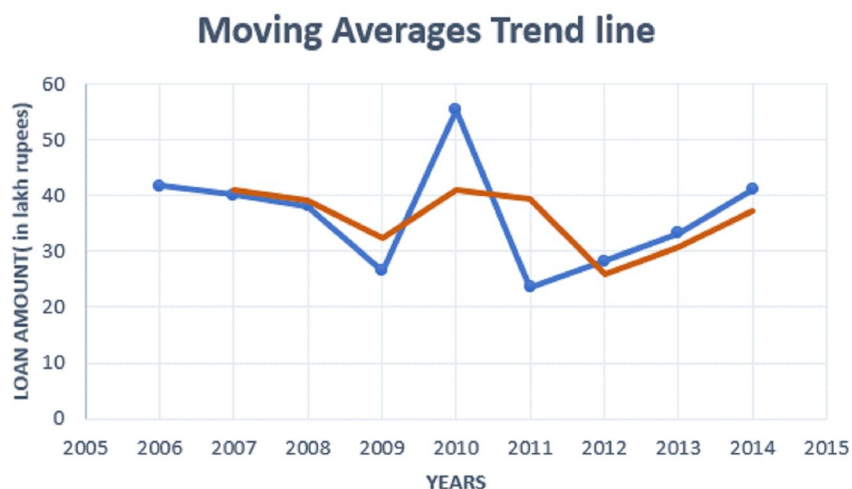
Calculate the 3-year moving averages for the loans (In lakh ₹) issued by co-operative banks for farmers in different states of India based on the values given below.

Year	2006	2007	2008	2009	2010	2011	2012	2013	2014
Loan amount (In lakh ₹)	41.85	40.2	38.12	26.5	55.5	23.6	28.36	33.31	41.1

Solution:

Year	Loan amount	3- year moving total	3- year moving average
2006	41.85	-	-
2007	40.2	120.17	$120.17/3 = 40.6$
2008	38.12	104.82	34.94
2009	26.5	120.12	40.4
2010	55.5	105.6	35.2
2011	23.6	107.46	35.82
2012	28.36	85.27	28.4
2013	33.31	102.77	34.26
2014	41.1	-	-

The graph shows the observation data in blue whereas, the red curve shows the smooth trend curve obtained by calculating moving averages of 3 years



Procedure for calculating Moving average for even number of years ($n = \text{even}$)

Let us take an example for $n = 4$ years moving averages to understand the procedure

1. Add up the values of the first 4 years and place the sum against the middle of 2nd and 3rd year. (This sum is called 4-year moving total)
2. For the next moving total, leave the first year value and add next 4 values from the 2nd till 5th year and write the sum against its middle position i.e. in the middle of 3rd year and 4th year

3. This process will continue till the value of the last observation is taken into account.
4. Now, calculate the average of each 4-year moving totals by dividing each moving total by 4
5. In the next column, calculate the sum of the first two 4-years moving averages and write the sum against 3rd year, in the middle (*known as centered total*).
6. After this, leave the first 4-year moving averages and add the next two 4-year moving total and place it against 4th year.
7. This process of finding centered totals will continue till all pairs of 4-yearly moving averages of previous column are summed up and centered.
8. Divide the newly obtained centered totals by 2 to get the moving averages which are our required trend values based on 4-year moving averages.

Example 10

Compute the trends by the method of moving averages, assuming that 4-year cycle is present in the following series.

Year	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989
Index number	400	470	450	410	432	475	461	500	480	430

Solution: The 4- year moving averages are shown in the last column as centered average

Year	Index Number	4-year Moving total	4-year Moving Average	Centered total	Centered moving average
1980	400			-	-
		-	-		
1981	470			-	-
		1730	$1730/4 = 432.5$		
1982	450			875.5	$875.5/2 = 437.75$
		1762	$1772/4 = 443$		
1983	410			884.75	$884.75/2 = 442.38$
		1767	$1767/4 = 441.75$		
1984	432			886.25	443.13
		1778	$1778/4 = 444.5$		
1985	475			911.5	455.75
		1868	$1868/4 = 467$		
1986	461			946	473
		1916	$1916/4 = 479$		

1987	500			946.75	473.38
		1871	$1871/4 = 467.75$		
1988	480			-	-
		-	-		
1989	430			-	-

6.9.2 (ii) Computation of Straight-line trend by using Method of Least squares

Method of least squares is a technique for finding the equation which best fits a given set of observations. In this technique, the sum of squares of deviations of the actual and computed values is least and eliminates personal bias.

Suppose we are given n number of observations and it is required to fit a straight line to these data.

Note that n , the number of observations can be odd or even.

Recall that the general linear equation to represent a straight line is:

$$y = a + bx, \text{ ----- (i)}$$

where y is the actual value, x is time; a and b are real numbers

In order to fit the best fitted trend line with the help of general equation $y = a + bx$ for the given time series, we will try to find the estimated values of y_i say $\hat{y}_i$ close to the observed values y_i for $i = 1, 2, \dots, n$.

According to the principle of least squares, the best fitting equation is obtained by minimizing the sum of squares of differences which leads us to two conditions:

1. The sum of the deviations of the actual values of y and $\hat{y}$ (estimated value of y) is zero
 $\Rightarrow \Sigma(y - \hat{y}_i) = 0$
2. The sum of squares of the deviations of the actual values of y and $\hat{y}$ (estimated value of y) is least
 $\Rightarrow \Sigma(y - \hat{y}_i)^2$ is least

For the purpose of plotting the best fitted line for trend analysis, the real values of constants ' a ' and ' b ' are estimated by solving the following two equations:

$$\Sigma Y = n a + b \Sigma X \text{ ----- (ii)}$$

$$\Sigma XY = a \Sigma X + b \Sigma X^2 \text{ ----- (iii)}$$

Where ' n ' = number of years given in the data.

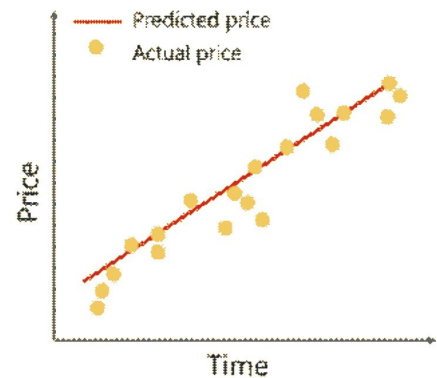
1. Remember that the time unit is usually of successive uniform duration. Therefore, when the middle time period is taken as the point of origin, it reduces the sum of the time variable x to zero

Which means that by taking the mid-point of the time as the origin,

$$\text{we get } \Sigma X = 0$$

2. When $\Sigma X = 0$, the equations (ii) and (iii) reduce to:

$$\Sigma Y = na + b (0)$$



$$\Rightarrow a = \frac{\sum Y}{n}$$

And,

$$\sum XY = a (0) + b \sum X^2$$

$$\Rightarrow b = \frac{\sum XY}{\sum X^2}$$

5. By substituting the obtained values of 'a' and 'b' in equation (i), we get the trend line of best fit.

Example 11

Given below are the consumer price index numbers (CPI) of the industrial workers.

Year	2014	2015	2016	2017	2018	2019	2020
Index Number	145	140	150	190	200	220	230

Find the best fitted trend line by the method of least squares and tabulate the trend values.

Solution

Note that the number of years is Odd

$$\Rightarrow n = \text{odd}$$

Procedure:

1. Take middle year value as A i.e. $A = 2017$
2. Find $X = x_i - A$
3. Find X^2 and XY

Year (x_i)	Index number (Y)	$X = x_i - A$ $= x_i - 2017$	X^2	XY	Trend value $Y_t = a + bX$
2014	145	-3	9	-435	$152.1 + (-3) \times 16.6 = 102.3$
2015	140	-2	4	-280	$152.1 + (-2) \times 16.6 = 118.9$
2016	150	-1	1	-150	$152.1 + (-1) \times 16.6 = 135.5$
2017	190	0	0	0	$152.1 + (0) \times 16.6 = 152.1$
2018	200	1	1	200	$152.1 + (1) \times 16.6 = 168.7$
2019	220	2	4	440	$152.1 + (2) \times 16.6 = 185.3$
2020	230	3	9	690	$152.1 + (3) \times 16.6 = 201.9$
n = 7	$\sum Y = 1065$	$\sum X = 0$	$\sum X^2 = 28$	$\sum XY = 465$	$\sum Y_t = 1064.7$

$$a = \frac{\sum Y}{n} = \frac{1065}{7} = 152.14$$

and
$$b = \frac{\sum XY}{\sum X^2} = \frac{465}{28} = 16.6$$

Therefore, the required equation of the straight-line trend is given by

$$y = a + bx \Rightarrow y = 152.1 + 16.6x$$

Example 12

Based on the data available for the sales of an item in a district, by the method of least squares

- tabulate the trend values
- find the best fit for a straight-line trend
- compute expected sale trend for year 2002

Year	1996	1997	1998	1999	2000	2001
Sales (In lakh ?)	6.5	5.3	4.3	6.1	5.6	7.8

Note that the number of years is even

$$\Rightarrow n = \text{even}$$

Procedure:

- Take middle year value as i.e. $= \frac{\text{sum of two middle years}}{2} = \frac{1998 + 1999}{2} = 1998.5$
- Find $X = \frac{x_i - A}{0.5}$ (we divide by 0.5 to avoid cumbersome calculations)
- Find X^2 and XY

Solution:

Year (x_i)	Index number (Y)	$X = \frac{x_i - A}{0.5}$ $= \frac{x_i - 1998.5}{0.5}$	X^2	XY	Trend value $Y_t = a + bX$
1996	6.5	-5	25	-32.5	$5.9 + (-5) \times 0.13 = 5.25$
1997	5.3	-3	9	-15.9	$5.9 + (-3) \times 0.13 = 5.51$
1998	4.3	-1	1	-4.3	$5.9 + (-1) \times 0.13 = 5.77$
1999	6.1	1	1	6.1	$5.9 + (1) \times 0.13 = 6.03$
2000	5.6	3	9	16.8	$5.9 + (3) \times 0.13 = 6.29$
2001	7.8	5	25	39	$5.9 + (5) \times 0.13 = 6.55$
n = 6	$\Sigma Y = 35.6$	$\Sigma X = 0$	$\Sigma X^2 = 70$	$\Sigma XY = 9.2$	$\Sigma Y_t = 35.4$

$$a = \frac{\Sigma Y}{n} = \frac{35.6}{6} = 5.9$$

and
$$b = \frac{\Sigma XY}{\Sigma X^2} = \frac{9.2}{70} = 0.13$$

Therefore, the required equation of the straight-line trend is given by

$$y = a + bx \Rightarrow y = 5.9 + 0.13x$$

According to the line trend, the predicted sales for year 2002:

$$y = 5.9 + 0.13\left(\frac{x_i - 4}{0.5}\right) = 5.9 + 0.13\left(\frac{2002 - 1998.5}{0.5}\right) = ₹ 6.81 \text{ lakh}$$

Note: 1. Future trend forecast made by using this method are based only on the trend values

2. The predicted trend values by using this method are more reliable than any other method

6.10 CHECK YOUR UNDERSTANDING REFLECTIVE QUESTIONS

Q1. Judge the correctness or otherwise of the following statements:

- i) An index number is a pure number
- ii) Index numbers are independent of choice of unit
- iii) An index number can be a negative quantity
- iv) The purchase power of money decreases as the wholesale index increases

Q2. A price index which is based on the prices of the items in the composite, weighted by their relative index is called:

- i) price relatives
- ii) Consumer price index
- iii) Weighted aggregative price index
- iv) Simple aggregative index

Answer: iii)

Q3. A weighted aggregate price index in which the weight for each variable is considered its current-period quantity is:

- i) Aggregative index
- ii) Consumer Price index
- iii) Laspeyres Index
- iv) Paasche's index

Answer: iv)

Q4. An index constructed to measure changes in quantities over a period of time is:

- i) Quantity index
- ii) Time series index
- iii) Quality index
- iv) Value index

Answer: i)

Q5. For calculating the weighted index number, which of the following uses quantities consumed in the base period as weights:

- i) Fisher's method
- ii) Paasche's method
- iii) Laspeyres method
- iv) Aggregative method

Answer : i)

Q6. What is the index number of the base period?

- i) 200
- ii) 300
- iii) 10
- iv) 100

Answer: iv)

Q7. Index number is a special type of :

- i) Average
- ii) Dispersion
- iii) Correlation
- iv) None of the above

Answer: i)

Q8. Index number is always expressed in

- i) Percentage
- ii) Ratio
- iii) Proportion
- iv) None of the above

Answer : i)

Q.3 Taking 1995 as base year calculate relative index number for the years 1997-2005

Year	1995	1997	1999	2001	2003	2005
Price (in ₹)	12	14	13	20	25	21

Q.4 Compute the weighted aggregative index number for the following data:

Variable	Price		Weights
	Current year	Base year	
X	5	4	60
Y	3	2	50
Z	2	1	30

Q.5 Calculate price index number for 2004 taking 1994 as the base year from the following data by simple aggregative method:

Item	Rice	Wheat	Pulses	Millets	Oil
Price in year 1990 (in ₹)	60	40	100	60	90
Price in year 2010 (in ₹)	140	60	205	70	100

Q.6 Based on the data on the expenses of middle-class families in a certain city, calculate the cost-of-living index during the year 2003 as compared with 1990:

Expenses	Year	Food	Fuel	Clothing	Rent	Miscellaneous
Price (in ₹)	2003	1500	250	750	300	425
Price (in ₹)	1990	1400	200	400	200	250

Q7. From the data given below, obtain the index of retail sales in India for years 1982, 1983, 1984 with the year 1981 as base period.

Year	Index of sales volume	Index of sales value
1995	101	105
1996	113	108
1997	106	124

Q8. Calculate the price index number for the following data using weighted aggregative method:

Commodity	Unit	Weight	Price	
			Base year	Current year
P	Quintal	14	90	120
Q	Kg	20	10	17
R	Dozen	35	40	60
S	Litre	15	50	93

Q9. Based on the given data, check whether i) Paasche's formula and, ii) Fisher's formula will satisfy the time reversal test:

Commodity	Base Year		Current Year	
	Price	Quantity	Price	Quantity
P	4	10	6	15
Q	6	15	4	20
R	8	5	10	4

Q10. The annual rainfall (in mm) was recorded for Cherrapunji, Meghalaya:

Year	Rainfall (in cm)
2001	1.2
2002	1.9
2003	2
2004	1.4
2005	2.1
2006	1.3
2007	1.8
2008	1.1
2009	1.3

Determine the trend of rainfall by 3-year moving averages

Q11. Compute the seasonal indices by 4-year moving averages from the given data of production of paper (in thousand tons)

Year	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989
Index number	2450	1470	2150	1800	1210	1950	2300	2500	2480	2680

Q12. Given below is the data of workers welfare expenses (in lakh ₹) in steel industries during 2001 - 2005. Use method of least squares to:

- tabulate the trend values
- find the best fit for a straight-line trend
- compute expected sale trend for year 2006

Year	2001	2002	2003	2004	2005
Welfare expenses (in lakh ₹)	160	185	220	300	510

Q13. Fit a straight-line tend by method of least squares for the following data and also find the trend value for year 1998:

Year	1992	1993	1994	1995	1996	1997
Production (in tons)	210	225	275	220	240	235

6.11 UNIT SUMMARY

1. An index number is a measure of change in a group of related variables over two different situations with respect to time, geographical location or other characteristics.
2. Factors influencing construction of index numbers:
 - Selection of data
 - Selection of weights
 - Base period
 - Choice of variables
3. Index number for time period n is represented as I_n
4. A list of indexes is called a Index series
5. Methods to construct index number:
 - i. Relative Index number = $\frac{P_n}{P_0} \times 100$
 - ii. Simple (Unweighted) Aggregative Method = $\frac{\sum P_n}{\sum P_0} \times 100$
 - iii. Simple Average of relatives Method
 - iv. Weighted Aggregative Method = $\frac{\sum P_{ni} Q_i}{\sum P_{oi} Q_i} \times 100$
 - v. Laspeyers' method = $\frac{\sum P_{ni} Q_{oi}}{\sum P_{oi} Q_{oi}} \times 100$
 - vi. Paasche's' method = $\frac{\sum P_{ni} Q_{ni}}{\sum P_{oi} Q_{ni}} \times 100$
 - vii. Fisher's Ideal method = $\sqrt{\frac{\sum P_1 Q_0}{\sum P_0 Q_0} \times \frac{\sum P_1 Q_1}{\sum P_0 Q_1}} \times 100$
 - viii. Marshall-Edgeworth's method = $\frac{\sum P_n (Q_0 + Q_n)}{\sum P_0 (Q_0 + Q_n)} \times 100$
 - ix. Weighted averages of relatives = $\frac{\sum \frac{P_n}{P_0} (P_0 Q_0)}{\sum P_0 Q_0} \times 100$
6. Types of index numbers:
 - i. Value index
 - ii. Quantity Index
 - iii. Price Index
7. There are tests to check consistency to verify adequacy of an index number:
 - i. Unit test
 - ii. Time reversal test
 - iii. Factor reversal test
 - iv. Circular test
8. Time reversal test: $p_{01} \times p_{10} = 1$
 Here, p_{01} is the index for current year '1' on the basis of base year '0'
 And, p_{10} is the index for year '0' based on year '1'
9. A time series is a sequentially recorded numerical data points for a given variable arranged in a successive order to track variation
10. The purpose of time series is to show an increasing growth pattern over time for a variable

11. Time reversal test of adequacy cannot be tested on Laspeyers' method and Paasche's method
12. When data of the variable is collected at distinct time intervals, for a specified period of time, it is called time series data.
13. When data for one or more variables is collected at the same point in time, it is called cross-sectional data.
14. When data is collected in a combination of time series data and cross-sectional data, it is called pooled data.
15. A time series in which data of only one variable is varying over time is called a *univariate* time series.
16. When a time series is a collection of data for multiple variables and how they are varying over time, it is called *multivariate* time series.
17. Secular trend component or simply called trend series, is the smooth, regular and long-term variations of the series, observed over a long period of time.
18. Seasonal component is a time series captures the periodic variability in the data, capturing the regular pattern of variability; within one-year periods.
19. Cyclical component is a time series shows an oscillatory movement where period of oscillation is more than a year where one complete period is called a cycle. The real Gross Domestic Product (GDP) provides good examples of a time series that displays cyclical behavior.
20. Irregular component is a time series in which fluctuations are unaccountable, unpredictable or sometimes caused by unforeseen circumstances like – floods, natural calamities, labor strike etc.
21. Trend can be measured using by the following methods:
 - i. Graphical method
 - ii. Semi averages method
 - iii. Moving averages method
 - iv. Method of least squares

6.12 Check your Understanding Answer Key

ANSWERS TO REFLECTIVE QUESTIONS

2. iii) 3. iv) 4. i) 5. i) 6. iv) 7. i) 8. i) 9. iii) 10. i) 11. ii) 12. ii) 13. iii) 14. iii) 15. ii)

ANSWERS TO PRACTICE QUESTIONS

1. 120
2. Laspeyre's = 146; Paasche's = 149; Fisher's = 147

3.	Year	1995	1997	1999	2001	2003	2005
	Price (in ₹)	12	14	13	20	25	21
		100	117	108	167	208	175

4. 137
5. 164
6. 132

7.	Year	
	1995	104
	1996	96
	1997	117

8. 153
9. Paasche's formula – no, Fisher's formula - yes

10

Year	3-year moving average
2001	-
2002	2.55
2003	2.65
2004	2.75
2005	2.4
2006	2.6
2007	2.1
2008	2.25
2009	-

11

Year	4-year moving averages
1980	-
1981	-
1982	1812.5
1983	1712.5
1984	1791.25
1985	1897.5
1986	2138.75
1987	2393.75
1988	-
1989	-

12

Year (x_i)	Trend value $Y_t = 275 + 81.5 X$
2001	112
2002	193.5
2003	275
2004	381.5
2005	438

Trend line : $Y_t = 275 + 81.5 X$

Predicted trend for year 2006 = 519.5 lakh rupees

13

Year (x_i)	Trend value $Y_t = 234 + 1.6 X$
1992	226
1993	229.2
1994	232.4
1995	235.6
1996	238.8
1997	242

Trend line: $Y_t = 234 + 1.6 X$

Predicted trend for 1998 = 245.2 tons