10

Business Forecasting and Time Series Analysis

INTRODUCTION

The growing competition, rapidity of change in business activities and the trend towards automation demand that decisions in business are not based purely on guesses and hunches rather on a careful analysis of data concerning the future course of events. More time and attention should be given to the future than to the past, and the question 'what is likely to happen ?' should take precedence over 'what has happened ?' though no attempt to answer the first can be made without the facts and figures being available to answer the second.

When estimates of future conditions are made on a systematic basis, the process is referred to as "forecasting" and the figure or statement obtained is known as a "forecast". Forecasting is a service whose purpose is to offer the best available basis for management expectations of the future and to help management understand the implications for the firm's future of the alternative courses of action to them at present. In a world where the future is not known with certainty, virtually every business and economic decision rests upon a forecast of future conditions. In fact when a person assumes the responsibility of running a business he automatically takes the responsibility for attempting to forecast the future and to a very large extent his success or failure would depend upon the ability to forecast successfully the future course of events. Forecasting aims at reducing the area of uncertainty that surrounds management decision-making with respect to costs, profit, sales, production, pricing, capital investment, and so forth. If the future were known with certainty, forecasting would be unnecessary. Decisions could be made and plans formulated on a once-for-all basis, without the need for subsequent revision. But uncertainty does exist, future outcomes are rarely assured and, therefore, organised system of forecasting is necessary.

Forecasting is concerned with two main tasks : first, the determination of the best basis available for the formation of intelligent managerial expectations; and second, the handling of uncertainty about the future, so that implication of decisions become explicit. Forecasting activity can be viewed as part of the management information system. It also impinges on the control system. Forecasts are commonly applied to capital investment decisions, strategic planning, product and market planning, production planning and stock control, budgetary control and financial planning and competitive position planning. In fact, managers are forecasters and they are forecasting for much of their time. They plan production in expectation of certain levels of sales. They set prices in expectation of certain levels of wages, raw material costs, financial constraints and sales. They build warehouses in expectation of certain levels of stocks and sales. They recruit labour, buy materials, arrange finance, or plan factories in expectation of certain levels of sales and other activity. The following are main functions of forecasting :

(1) The creation of plans of action. It is impossible to evolve a worthwhile system of business control without one acceptable system of forecasting.

(2) The second general use of forecasting is to be found in monitoring the continuing progress of plans based on forecasts. Forecasts serve the function of lighthouses to shipmasters at night, reference points for course and speed requiring action/no action decisions.

(3) The forecast provides a warning system of the critical factors to be monitored regularly because they might drastically affect the performance of the plan.

It is obvious from the above that forecasts intelligently used may serve the function of both lighthouse and compass. However, the object of business forecasting is not to determine a curve or series of figures that will tell exactly what will happen, say, a year in advance, but it is to make analysis based on definite statistical data, which will enable an executive to take advantage of future conditions to a greater extent than he could do without them. In many respects the future tends to move like the past. This is a good thing, since without some element of continuity between past, present and future, there would be little possibility of successful prediction. But history is not likely to repeat itself and we would hardly expect economic conditions next year or over the next year ten years to follow a clear-cut pattern. Yet, frequently, past patterns prevail sufficiently to justify using the past as a basis for predicting the future.

In forecasting one should note that it is impossible to forecast the future precisely—there always must be some range of error allowed for in the forecast. Statistical forecasts are those in which we can use the mathematical theory of probability to measure the risks of errors in predictions.

Steps in Forecasting

Broadly speaking, the forecasting of business fluctuations consists of the following steps :

1. *Understanding why changes in the past have occurred.* One of the basic principles of statistical forecasting—indeed of all forecasting when historical data are available—is that the forecaster should use the data on past performance to get a "speedometer reading" of the current rate (of sales, say) and of how fast that rate is increasing or decreasing. The current rate and changes in the rate—"acceleration" and "deceleration"—constitute the basis of forecasting. Once they are known, various mathematical techniques can develop projections from them. If an attempt is made to forecast business fluctuations without understanding why past changes have taken place, the forecast will be purely mechanical, based solely upon the application of mathematical formulae and subject to serious error.

2. *Determining which phases of business activity must be measured.* After it is known why business fluctuations have occurred, or if there is a reasonable supposition it is necessary to measure certain phases of business activity in order to predict what changes will probably follow the present level of activity.

3. *Selecting and compiling data to be used as measuring devices.* There is an interdependent relationship between the selection of statistical data and determination of why business fluctuations occur. Statistical data cannot be selected and compiled in an intelligent manner unless there is a sufficient understanding of business fluctuations ; likewise, it is important that reasons for business fluctuations be stated in such a manner that it is possible to secure data that are related to the reasons.

4. *Analysis of data.* In this last step, the data are analysed in the light of one's understanding of the reason why changes occur. For example, if it is reasoned that a certain combination of forces will result in a given change, the statistical part of the problem is to measure these forces and from the data available to draw conclusions on the future course of action. The methods of drawing conclusions may be called forecasting techniques, which represent any one of a large number of analytical devices for summarising data and drawing inferences from the summaries.

Requirements of a Good Forecasting System

A forecasting system to be instrumental in contributing to better management decision-making needs certain conditions :

(1) It must involve the managers whose decisions are affected.
(2) Individual forecasts and group of forecasts have to be specifically relevant to the decisions being taken.
(3) The forecasts must not claim too much validity or authority.
(4) Implications of the various probable errors in the predictions for the organisations need to be thoroughly worked through so that management can evaluate the consequences of the probable range of likely outcomes.
(5) Management must at least know how badly things could go wrong if all the guesses turned out wrong.

Methods of Forecasting

There is nothing new about business forecasting. For centuries businessmen have tried to adjust themselves in such a manner as to make the best out of the future conditions. The rule-of-thumb method has been widely practised in business. It consists of deciding about the future in terms of past experience and familiarity with the problem at hand. Even today this method is very widely used in business. However, it can lead to absurd conclusions if employed by the inexperienced.

In recent years the techniques of forecasting have improved to a marked degree and are applicable to almost every sphere of business activity. Attempts are being made to make forecasting as scientific as possible. The base of scientific forecasting is statistics, *i.e.*, numerical data on business trends which many businessmen fail to acquaint themselves with. However, forecasting business change involves more than an analysis of statistical data—it also embodies the prediction of economic change such as secular trend, seasonal variation or cyclical variation and a consideration of cause and effect. To handle the increasing variety of managerial forecasting problems, many forecasting techniques have been developed in recent years. Each has its special use, and care must be taken to select the correct technique for particular application. Also before applying a method of forecasting, the following questions should be answered :

(1) What is the purpose of the forecast—how is it to be used ?
(2) What are the dynamics and components of the system for which the forecast will be made ?
(3) How important is the past in estimating the future ?

The following are some of the important methods of forecasting :

1. Historical Analogy Method,
2. Field Surveys and Opinion Poll,
3. Business Barometers,
4. Extrapolation,
5. Regression Analysis,
6. Econometric Models,
7. Lead-Lag Analysis,
8. Exponential Smoothing,
9. Input-output or end-use Analysis,
10. Time Series Analysis,

Here only a brief description of first nine methods and detailed description of the last method, which is very popularly used in practice, is made.

1. **Historical Analogy Method.** When this method of forecasting is used, the forecast in regard to a particular phenomenon is based on some analogous conditions elsewhere in the past. For example, the

forecast for demand for steel, cement, cars, etc., in India today may be based on the same analogy of demand for these products in the U.S.A. In the year, say, 2004 if it is found that the conditions now prevailing in India are very much like those that prevailed in the United States during that period.

Analogies very much help a country in determining the various stages of growth through which a country is passing before it reaches the '*take off* stage. Through analogies one can also have an idea of the social changes such as changes in attitudes and values, norms of social life, life style, etc. Generally, we find that as a country heads towards economic advancement many of the old beliefs and values change and people start thinking in different light altogether.

This method of forecasting is considered to be a qualitative one because it is difficult to quantify most phenomena in respect to which analogies are being made. A serious limitation of this method is that a search has to be made of such a place and period in history and it may really be difficult to get exactly comparable conditions. Hence the method can only be used as a rough guide and exclusive reliance on it should not be placed.

2. **Field Surveys and Opinion Poll.** Field surveys may be conducted to obtain the necessary information which may constitute the basis for forecasting. The survey methods are used widely for forecasting demand both of the existing and new products marketed within and outside the country. Surveys can help in obtaining both qualitative and quantitative information. However, the various survey techniques are to be used with great caution so that the element of bias in responses is minimised.

The information gathered through survey methods can be discussed with various experts and other knowledgeable persons in the field and their opinion can be obtained. For example, the opinion of sales representatives, wholesalers, retailers and other intermediaries may be obtained while formulating demand projections.

It is quite likely that experts in the different fields such as production, sales, finance may have divergent views and it may be necessary for all of them to sit together, give a patient hearing to others' viewpoints, convince others and change their opinion, if necessary. A consensus view can be obtained by the use of Delphi method.

3. **Business Barometers.** Of great assistance in practical forecasting is a series that can be used as an "index" or "indicator" of the basic conditions related to the industry. The term "barometer" is also widely though loosely used in business statistics ; sometimes the term is used to mean simply an indicator of the present economic situation and sometimes it is used to designate an indicator of future conditions.

The following are some of the important business activities which aid businessmen in forecasting :

1. Gross national product,
2. Employment,
3. Wholesale prices,
4. Consumer prices,
5. Industrial production,
6. Volume of bank deposits and currency outstanding,
7. Consumer credit,
8. Disposable personal income,
9. Departmental store sales,
10. Stock prices,
11. Bond yields,

This list is by no means exhaustive, nor is the arrangement necessarily in order of importance. Several of the above series are composite average of totals—or indexes of these averages or totals. Analysis also should be made of some of the major components of these activities.

Index numbers relating to different activities in the field of production, trade, finance, etc., may also be combined into a general index of business activity. This general index refers to the general conditions of trade and industry. But the behaviour of individual industries or trades might show a different trend from that of the Composite Business Activity Index. Also, general boom or depression may be reflected in a majority of separate industries and trades, yet some industries and trades might show quite contrary tendencies. Hence, the study of general business conditions as revealed by the Composite Business Index should be supplemented by special studies of individual businesses based on separate indices. The trends indicated by barometers will guide the businessmen as to whether the stocks of goods should be increased or decreased or whether to increase investment or not, etc.

4. **Extrapolation.** Extrapolation is the simplest yet often a useful method of forecasting. In many forecasting situations the most reasonable expectation is that the variable will follow its already established path. Extrapolation relies on the relative consistency in the pattern of past movements in some time series. Strictly speaking, nothing needs to be known about causation—why the series moves as it does. But in practice the justification for extrapolation does involve the nature of the growth process being described. Extrapolation is used frequently for sales forecasts and for other estimates when "better" forecasting methods may not be justified.

Since extrapolation assumes that the variable will follow its established pattern of growth, the problem is to determine accurately the appropriate trend curve and the values of its parameters. Numerous alternative trend curves are suitable for business forecasting application. Some of the most useful ones are :

(*a*) *Arithmetic trend*. The straight-line arithmetic trend assumes that growth will be by a constant absolute amount each year.

(*b*) *Semi-log trend*. The semi-logarithmic trend assumes constant percentage increase each year. Since the annual increment is constant in logarithms, this line translates into a straight line when drawn on paper with a logarithmic vertical scale.

(*c*) *Modified exponential trend*. This curve assumes that each increment of growth will be a constant per cent (less than 100) of the previous one. The line trends generally do approach, but never quite reach a constant asymptote, which may be thought of as an upper limit.

(*d*) *Logistic curve*. The logistic curve has both an upper asymptote and a lower asymptote. It assumes a 'law of growth' involving increasing increments from an initial low value and then gradual slowing down of growth as 'maturity' is approached.

(*e*) *The Gompertz curve*. The Gompertz curve is a curve with similar properties as described above and is often used to describe growth of industrial output.

Selection of an appropriate growth curve can be guided by empirical and theoretical considerations. Empirically, it is a question of selecting the curve that best fits the past movement of the data. Theoretical matters which intervene in that logic may support a particular growth pattern. For example, population growth when there are no resources or choice restraints imply a geometric pattern of growth, as has been known since Malthus. With limited resources, however, population is sometimes thought to grow along a logistic curve. Let these theoretical notions be taken too seriously, it should be emphasised that empirical considerations may lead us quickly to a more realistic and less restrictive notion of the relevant growth curve.

5. **Regression Analysis.** The regression approach offers many valuable contributions to the solution of forecasting problems. It is the means by which we select from among the many possible or theoretically suggested relationships between variables in a complex economy. With it, one makes the jump from intuitive evaluation on the connection between two variables to precise quantified knowledge. If two variables are functionally related then a knowledge of one will make possible an estimate of the other. For example, if we know that advertising expenditure and sales are correlated then for a given advertising expenditure, we can find out the probable increase in sales or *vice versa*.

Regression analysis may involve only one predicted, or dependent, and one independent variable—simple regression, or it may involve relationships between the variable to be forecast and several independent variables—multiple regression. Statistical techniques to estimate the regression equations are often fairly complex and time-consuming, but there are many computer programs now available that estimate simple and multiple regressions quickly without much of costs involved.

There are two dangers in using regression analysis for forecasting :

(*i*) There is possibility of a mechanistic approach, accepting with little question the relationship which the calculations reveal—perhaps that with the highest r^2— and applying it to the forecast. There are many possibilities for spurious correlation among time series as many series move together over time even where there is no conceivable connection between them.

(*ii*) If the trend of observations follows a curve, the linear regression will still fit the best straight line to the data, but any projection will be nonsense. There is little which we can do with non-linearity graphically, and computers handle nonlinear forms as routinely as linear.

6. **Econometric Models.** The term econometric refers to the application of mathematical economic theory and statistical procedures to economic data in order to verify economic theorems and to establish quantitative results in economics. An econometrician is, therefore, an economist, a statistician and a mathematician, all in one. Econometric models take the form of a set of simultaneous equations. The values of the constants in such equations are supplied by a study of statistical time series, and large number of equations may be necessary to produce an adequate model. The work of computations is greatly facilitated by electronic data processing equipment like computer, etc.

At the present time, most short-term forecasting uses only statistical methods with little quantitative information. However, in the years to come when most large companies develop and refine econometric models of their major businesses, this tool of forecasting will become more popular. But, it should be remembered that the development of an econometric model requires sufficient data so that the correct relationships can be established. Hence when data are scarce—for example, when a product is first introduced into a market—this method cannot be profitably employed.

The econometric model is, in principle, the most formal, since the forecast is based on an explicit mathematical model. The model states in detail and in quantitative terms the way in which the various aspects of the economy are interrelated. Theoretically, the model makes possible a wholly mechanical forecast because once values have been estimated for the exogenous variable, the solution of the model gives specific values for the predicted variable. But, in actual practice qualitative and quantitative forecasters have tended to come together. The 'artist' forecaster has become fully aware of the fact that he needs quantitative relationships, while the 'econometric' forecaster has learned that in some instances, quantitative relationships have to be modified by qualitative factors.

The econometric model provides the forecaster with a record of the prediction with a clear statement of the assumptions concerning exogenous variables and the solution of the model—it is often possible—or at least it is made easier—to trace and reproduce the causes for success as well as failures. One can learn just where errors were made and, hopefully, where improvements can be made. Thus, discredited hypothesis may be dropped and new ones can be substituted which ultimately will lead to better understanding of the economic system and business fluctuations.

The econometric models are not very popular in practice because it is probably neither necessary nor feasible for every business forecaster to construct his own model of the economy. The effort and costs involved in a fully-developed econometric model are well beyond most forecasting operations. Thus, most forecasters will probably rely for some time on the basic aggregate models developed at research institutes or universities. These models may be used to make predictions and to test out alternative assumptions about Government policy or the other exogenous aspects of the economy. With the help of the models and, hopefully, sector analysis of his own industry, the business forecaster will be in a good position to augment other more familiar approaches. The better the understanding of the various forecasting methods and of their interrelationships, the better the forecasts will be.

7. Lead-lag Analysis

The Lead-lag approach attempts to determine the approximate lapse of time between the movement of one series and the movements of general business conditions. If one or more series can be found such that their turning points lead by a number of months with substantial regularity the turning points of general business in the past, it is quite logical to use these leading series to predict what is going to happen to general business activity.

The most important list of statistical indicators in modern times originated during the 1937-38 sharp business contraction. The list was prepared by the National Bureau of Economic Research (NBER). The list comprised 21 series that on their past performance, some dating as far back as 1854, promised to be fairly reliable indicators of business revival. The list was revised three times—the most recent list comprises 26 indicators of business expansion and contraction. Among the current 26 NBER statistical indicators, 12 are classified as leading series, 8 as coincident series and 6 as lagging series. Leading indicators, as pointed out by Chou, are mainly those series which are concerned with business decisions to expand or to curtail output. Time is required to work out their effects, and so they tend to move ahead of turns in business cycles. *Leading indicators* signal in advance a change in the basic performance of the economy as a whole. *Coincident indicators* are those whose movements coincide roughly with, and provide a measure of the current performance of aggregate economic activity. Hence they inform us whether the economy is currently experiencing a slowdown or not. Movements of *lagging indicators* usually follow rather than lead those of the coincident indicators. In general, lagging indicators move in directions opposite to those of the leading indicators throughout various phases of business cycles.

8. Exponential Smoothing*

The method of exponential smoothing for forecasting is an outgrowth of the recent attempts to maintain the smoothing function of moving averages without their corresponding drawbacks and limitations. Two major limitations to the use of moving averages are :

(*i*) to compute a moving average forecast it is necessary to store the last N observation values. This takes up considerable space which in many computer systems is costly ;

(*ii*) the method of moving averages gives equal weight to each of the last N observed and no weight at all to observations before period $(t - N)$; that is, the weight given to each of the last N observations is $1/N$ and 0 (zero) for any previous observations.

In principle, exponential smoothing operates in a manner analogous to moving averages by "smoothing" historical observations to eliminate randomness. The mathematical procedure for performing this smoothing, however, is somewhat different from that used in moving averages. The basic principle and the application of this device are quite simple. If we wish to forecast the value of a time series for the period $t + 1$ on the basis of information available just after period t, the forecast is best considered as a function of two components : the actual value of the series for period, t, and the forecasted value for the same period made in the previous period $t - 1$. The use of both realised and estimated values available now for predicting future values is better than the use of either alone.

The exponential smoothing models can be either single exponential smoothing model or double exponential smoothing model, the former being applicable in the absence of trend, and the latter being applicable when the time series is exhibiting some type of growth pattern.

---

* The term exponential smoothing is obtained from the weight attached to the preceding observations.

When we talk of single exponential smoothing model, the forecasted value of the series at time period t, $\hat{y}_t$, is equal to a fraction of the forecast error of the previous period $(y_{t-1} - \hat{y}_{t-1})$, plus the forecasted values of the previous period. Thus to predict the value of the time series at time $t + 1$, we use

$$y_{t+1} = a(y_t - \hat{y}_t) + \hat{y}_t$$

It should be noted that exponential smoothing, a special kind of weighted moving average, is found to be useful in short-term forecasting for inventories and sales. Exponential smoothing can also be employed for projections over long terms. Because of several computational advantages over the simple moving average, the exponential smoothing time series model is perhaps the most widely used time series forecasting technique today.

Some experts have been critical of exponential smoothing procedures on the ground that they are empirically inadequate when major policy decisions undergo change. They also object to the assumption that differing weights should be assigned to observations depending solely upon their relative recency. The exponential smoothing as a technique for prediction has very limited applications for sales items that have many wide variations in expected demand.

9. Input-Output Analysis

The input-output (I-O) technique was developed by Professor Leontief in the 1930s. However, it is only recently that it has caught the eye of big business. Input-output analysis gets its name from the type of data on which it is based. That is, the material requirements (input) and the product (output) of every economic activity in an I-O model are the raw data. The most widely familiar model is the pioneer work of Leontief for the United States.

When input-output method is used, an input-output table is made. An input-output table is a technique for determining the transactions taking place within and among different sectors of the economy as well as the magnitude of such transactions. An input-output table summarises "taking" and "giving" among and within all industries and between them and the final consumer. In this sense it is an accounting procedure for reporting all transactions. Its main usefulness is for planning purposes at the level of national economy. Several countries, mainly European, utilise input-output tables widely for planning purposes.

The input-output tables can also be constructed for single business, organisation or industry. The different divisions, departments or organisational units can then become the entries of the table among which transactions can take place and for which forecasts can be made.

Since the construction of an input-output table can be expensive and since the technological requirements among departments are not constant, its value in forecasting is rather doubtful. The value of input-output analysis lies more in the planning stages than in forecasting itself, but it is still an expensive planning device for the purposes of individual organisations.

10. Time Series Analysis

This is the most popular method of business forecasting and is discussed in detail.

Business Forecasting and Time Series Analysis

The first step in making estimates for the future consists of gathering information from the past data. 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

---

* "A time series is a set of observations taken at specified times, usually at 'equal intervals'. Mathematically, a time series is defined by the values Y_1, Y_2.. of a variable Y at times t_1, t_2...Thus Y is a function of t, symbolised $Y = F(t)$."

example, if we observe production, sales, imports, etc., at different points of time, say, over the last 5 or 10 years, the set of observations formed shall constitute time series. Thus, 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 minute or second.

The problem of time series analysis can best be appreciated with the help of the following example :

The following are the figures of sales (in thousand units) of a firm :

Year	*Sales of Firm A (thousand units)*	*Year*	*Sales of Firm A (thousand units)*
2003	40	2007	43
2004	32	2008	48
2005	47	2009	65
2006	41	2010	42

If we observe the above series we find that generally the sales have increased but for some years a decline is also noticed. There may be several causes responsible for increase or decrease from one period to another such as changes in tastes and habits of people, growth of population, availability of alternative products, etc. It may be very difficult to study the effect of various factors that have led either to an increase or decrease in sales. The statistician, therefore, tries to analyse the effect of the various forces under four broad heads :

(1) Changes that have occurred as a result of general tendency of the data to increase or decrease, known as 'secular variations or trend.'

(2) Changes that have taken place during a period of 12 months as a result of change in climate, weather conditions, festivals, etc. Such changes are called 'Seasonal Variations.'

(3) Changes that have taken place as a result of booms and depressions. Such changes are classified under the head 'Cyclical Variations'.

(4) Changes that have taken place as a result of such forces that could not be predicted like floods, earthquakes, famines, etc. Such changes are classified under the head 'Irregular or Erratic Variations'.

These variations are called components of time series and shall be discussed in detail.

Role of Time Series Analysis

Time Series Analysis is of great significance in business decision-making for the following reasons :

(1) *It helps in the understanding of past behaviour.* By observing data over a period of time, one can easily understand what changes have taken place in the past. Such analysis will be extremely helpful in predicting the future behaviour.

(2) *It helps in planning future operations.* Statistical techniques have been evolved which enable time series to be analysed in such a way that the influences which have determined the form of that series may be ascertained. If the regularity of occurrence of any feature over a sufficient long period could be clearly established then, within limits, prediction of probable future variations would become possible.

In fact, the greatest potential of a time series lies in predicting an unknown value of the series. From this information intelligent choices can be made concerning capital investment decisions, decisions concerning production and inventory, etc.

(3) *It helps in evaluating current accomplishments.* The actual performance can be compared with the expected performance and the cause of variation analysed. For example, if expected sales for 2010 were 20 lacs coloured T.V. sets and the actual sales were only 19 lacs; one can investigate the cause for the shortfall in achievement. Time Series Analysis will enable us to apply the scientific

procedure of "holding other things constant" as we examine one variable at a time. For example, if we know how much is the effect of seasonality on business we may devise ways and means of ironing out the seasonal influence or decreasing it by producing commodities with complementary seasons.

(4) *It facilitates comparison.* Different time series are often compared and important conclusions drawn therefrom.

However, one should not be led to believe that by time series analysis one can foretell with 100 per cent accuracy the course of future events. After all, statisticians are not foretellers. This could be possible only if the influences of the various forces which affect these series such as climate, customs and traditions, growth and decline factors and the complex forces which produce business cycles would have been regular in their operation. However, the facts of life reveal that this type of regularity does not exist. But this then does not mean that time series analysis is of no value. When such analysis is coupled with a careful examination of current business indicators, one can undoubtedly improve substantially upon guestimates (*i.e.,* estimates based upon pure guesswork) in forecasting future business conditions.

COMPONENTS OF TIME SERIES

It is customary to classify the fluctuations of a time series into four basic types of variations which account for the changes in the series over a period of time. These four types of patterns, variations, movements, or, as they are often called components or elements of time series, are :

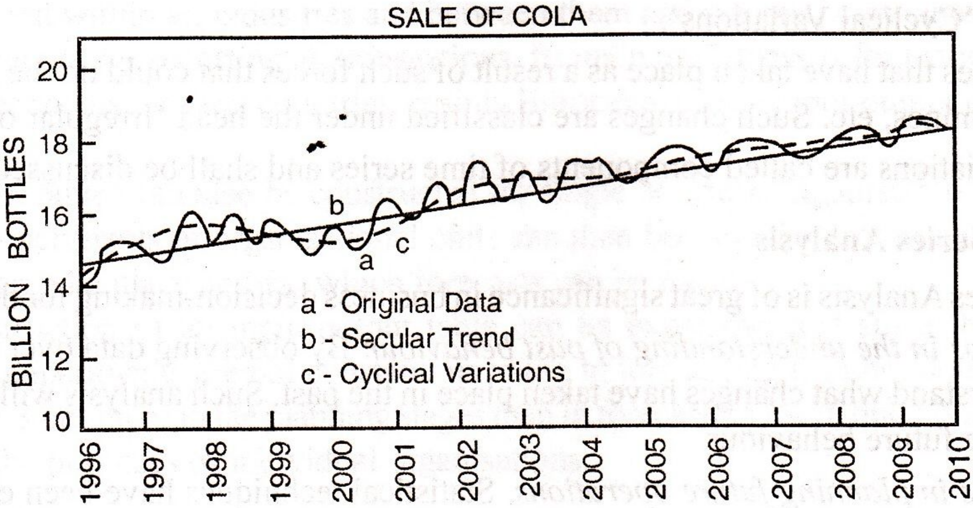
(1) Secular Trend,

(2) Seasonal Variations,

(3) Cyclical Variations,

(4) Irregular Variations.

Look at the following graph showing the sale of Cola during the years 1996 to 2010 :



The original data in this graph is represented by a curve. The general movement persisting over a long period of time represented by the diagonal line drawn through the irregular curve is called *secular trend.*

Next, if we study the irregular curve year by year, we see that in each year the curve starts with a low figure and reaches a peak about the middle of the year and then decreases again. This type of fluctuation, which completes the whole sequence of change within the span of a year and has about the same pattern year after year, is called a *seasonal variation.*

Furthermore, looking at the broken curve superimposed on the original irregular curve, we find pronounced fluctuations moving up and down every few years throughout the length of the chart. These are known as *business cycles* or *cyclical fluctuations*. They are so called because they comprise a series of repeated sequences just as wheel goes round and round.

Finally, the little saw-tooth irregularities on the original curve represent what are referred to as *irregular movements*.

In traditional or classical time series analysis, it is ordinarily assumed that there is a multiplicative relationship between these four components, that is, it is assumed that any particular value in series is the product of factors than can be attributed to the various components. Symbolically :

$$Y = T \times S \times C \times I$$

where Y is the result of the four elements :

T = Trend,
S = Seasonal Variation,
C = Cyclical Variation, and
I = Irregular Variation.

If the above model is employed, the seasonal, cyclical and random items are not viewed as absolute amounts but rather as relative influences. Thus, a seasonal index of 110 per cent would mean that the actual value is 10 per cent higher than it otherwise would be because of seasonal influences.

This particular model is appropriate for those situations in which percentage changes best represent the movement in the series.

Another approach is to treat each observation of a time series as the sum of these four components. Symbolically,

$$Y = T + S + C + I$$

When this relationship is assumed the major aim of time series analysis is to isolate those parts of the overall variation of a time series which are traceable to each of these four components and measuring each part independently. There are numerous variations of these basic models. The models

$$Y = TCS + I \text{ or } Y = TC + SI$$

are two such variations.

There is little agreement amongst experts about the validity of the different assumptions—some feel that the given classification is too crude and that there are more than four types of movements. Nothing specific is really known about how the components are related, how they combine to produce particular effects, or whether they are really separable. The effects of the various components might be additive, multiplicative or they might be combined in any one of infinitely large number of other ways. Different models (assumptions or theories) will lead to different results. Although the additive assumption is undoubtedly true in some cases, the multiplicative assumption characterises the majority of economic time series. Consequently, *the multiplicative model is not only considered the standard or traditional assumption for series analysis, it is more often employed in practice than all other possible models combined.* For this reason, we shall discuss the multiplicative model in detail in this chapter. However, it should be kept in mind that not all time series are best represented by this model and if the examination of data reveals that certain components do not react in the prescribed fashion, then a different model may be better.

The task of performing a time series analysis is to operate on the data in such a way as to bring out separately each of the components present.

1. Secular Trend

The term 'secular trend' or simply 'trend' is very popularly used in day-to-day conversation. For example, we often talk that the population, prices, production, etc., are showing an upward trend. What we really mean thereby is that if we observe such variables over a long period of time we find an increasing tendency. Similarly, we may find some variables showing downward tendency or constant tendency. For example, we find that over the last several years the death rate in our country is declining and hence we say that death rate is showing a declining trend. Similarly, with the improvement in the means of transport the number of bullock-carts on the road is declining year after year. However, in a dynamic economy such examples where either a downward or constant tendency is observed are rare—most of the variables show an upward trend. Thus, the general tendency of the data to grow or decline over a long period of time is technically called 'secular trend' or simply 'trend'. It should be noted that when we talk of trend, we mean thereby smooth, regular, long-term movement of the data—sudden and erratic movements either in upward or in downward direction have nothing to do with the trend.

There are all sorts of trends : some series increase slowly and some increase fast, others decrease at varying rates, some remain relatively constant for long periods of time, and some after a period of growth or decline reverse themselves and enter a period of decline or growth. Broadly speaking, the various types of trends are divided under two heads :

(1) Linear or Straight Line Trends, and

(2) Non-linear Trends.

Both these types will be discussed later in this chapter.

For a proper understanding of the meaning of trend, the reader's attention is drawn to the following two points :

(*i*) *The meaning of long term.* When we say that secular trend refers to the general tendency of the data to grow or decline over a long period of time, one may be interested in finding out as to what constitutes a long period of time. Does it mean several years ? The answer is 'no'. On the other hand, whether a particular period can be regarded as long or not in studying secular trend depends upon the nature of data. To take an example, if we are studying the figures of sales of a firm for 2009-10 and 2008-09 and we find that in 2009-10 sales have gone up, this increase cannot be called as secular trend because this is too short a period of time. On the other hand, if we put a strong germicide into bacterial culture, and count the number of organisms still alive after each 10 seconds for 8 minutes, these 48 observations showing a general pattern would be called secular movements. It is clear from this example that in one case year could not be regarded as a long period whereas in another case even 8 minutes constituted long period. Hence it depends on nature of data whether a particular period would be called as long or not.

Generally speaking, the longer the period covered, the more significant the trend. When the period is short, the secular movements cannot be expected to reveal themselves clearly and the general drift of the series may be unduly influenced by the cyclical fluctuations. This would make it difficult to separate the various series of variations in time series. *As a minimum safeguard, it may be said that to compute trend, the period should cover at least two or three complete cycles.*

(*ii*) Another point worth mentioning is that for concluding whether the data is showing an upward tendency or downward tendency it is *not necessary that the rise or fall must continue in the same direction throughout the period.* We have to observe the general tendency of the data. As long we can say that the period as a whole was characterised by an upward movement or by a downward movement, we say that a secular trend was present. For example, if we observe the trend of price over a period of 20 years and find that except for a year or two the prices are continuously rising, we would call it a secular rise in prices.

Factors Affecting Trend

There are several factors that affect trend in time series. The most important single factor responsible for rising trends in series like prices, production, sales, etc., has been the ever increasing population. On the other hand, declining trends in certain series are the result of the technological, institutional and cultural changes, the very things which produced much of the growing trend in most of the other series. For example, the progress in automobile industry reduced on the road and, on the other hand, increased the number of cars, buses, trucks, etc. Similarly, better medical facilities, improved sanitation, diet, etc., on the one hand reduce the death rate and on the other contribute to a rise in birth rate. Such influences as these produce gradual changes. But at times it is possible that certain innovations may take place which may cause sudden changes in the outlook for particular industries or individual concerns.

The basic objective of the study of trend is to predict the future behaviour of the data. If a trend can be determined, then the rate of change or progress can be ascertained and tentative estimates concerning the future be made accordingly. For example, by projecting the trend line one can find out expected sales of a firm, for, say 201 or the expected population for 2015 or 2030 likely, and so on. Such forecasts are of immense use in framing basic policies and planning for the future. However, such forecasts are based on the assumption that the past growth has been steady and that the conditions determining this growth may reasonably be expected to persist in the future. A change in these conditions would affect the forecasts.

2. Seasonal Variations

Seasonal variations are those periodic movements in business activity which occur regularly every year and have their origin in the nature of the year itself. Since these variations repeat during a period of 12 months they can be predicted fairly accurately. Nearly every type of business activity is susceptible to seasonal influence to a greater or lesser degree and as such these variations are regarded as normal phenomenon recurring every year. Although the word 'seasonal' seems to imply a connection with the season of the year, the term is meant to include any kind of variation which is of periodic nature and whose repeating cycles are of relatively short duration. Seasonal variation is evident when the data are recorded at weekly or monthly or quarterly intervals. Although the amplitude of seasonal variations may vary, their period is fixed being one year. As a result, seasonal variations don't appear in series of annual figures. The factors that cause seasonal variations are :

(i) Climate and weather conditions. The most important factor causing seasonal variation is the climate. Changes in the climate and weather conditions such as rainfall, humidity, heat etc., act on different product and industries differently. For example, during winter there is greater demand for woollen clothes, hot drinks, etc., whereas in summer cotton clothes, cold drinks have a greater sales. Agriculture is influenced very much by the climate. The effect of the climate is that there are generally two seasons in agriculture—the growing season and harvesting season—which directly affect the income of the farmer which in turn affects the entire business activity.

(ii) Customs, traditions and habits. Though nature is primarily responsible for seasonal variations in time series, customs, traditions, and habits also have their impact. For example, on certain occasions like Deepawali, Dussehra, Christmas, etc., there is a big demand for sweets and the bank withdrawals go up because people want money for shopping and gifts, etc. Similarly, on the first of every month there are heavy withdrawals and the banks have to keep lots of cash to meet the possible demand on the basis of past experience.

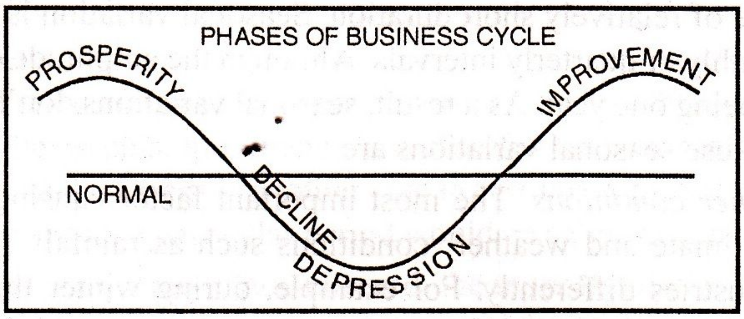
The study and measurement of seasonal patterns constitute a very important part of analysis of time series. In some cases, seasonal patterns themselves are of primary concern because little, if any, intelligent planning or scheduling (or production, inventory, personnel, advertising and the like) can be done without a knowledge based on adequate statistical measures of seasonal patterns. In other cases, the seasonal variation may not be of immediate concern, but it must be measured to facilitate the study of other types of variations based on adequate statistical measure of seasonal patterns. An accurate knowledge of seasonal behaviour is an aid in mitigating and ironing out seasonal movements through business policy. This may be done by introducing diversified products having different seasonal peaks, accumulating stock in slack seasons in order to manufacture at a more regular rate, cutting prices in slack seasons and advertising off-seasonal use for the products. Seasonal indexes are also helpful in scheduling purchases, inventory control, personnel requirement, seasonal financing and selling and advertising programmes. For example, a housewife may buy fruits for canning or preserving at the peak of the season when the prices are low and quality high. Seasonal fluctuations may also be ironed out in order that the intra-year fluctuations may be less pronounced.

3. Cyclical Variations

The term 'cycle' refers to the recurrent variations in time series that usually last longer than a year and regular, neither in amplitude nor in length.

Most of the time series relating to economics and business show some kind of cyclical variation. Cyclical fluctuations are long-term movements that represent consistently recurring rises and declines in activity. A business cycle* consists of the recurrence of the up and down movements of business activity from some sort of statistical trend or "normal". By "normal" we mean a kind of statistical average : we do not mean that there is anything very permanent or special. There are four well-defined periods or phases in the business cycle, namely : (*i*) prosperity, (*ii*) decline, (*iii*) depression, and (*iv*) improvement.

Each phase changes gradually into the phase which follows it in order given. The following diagram would illustrate a cycle. In the prosperity phase of the business cycle the public is optimistic—business is booming, prices are high and profits are easily made. There is a considerable expansion of business activity which leads to an over-development. It is then difficult to secure deliveries and there is shortage of transportation facilities, which has a tendency to cause large inventories to be



accumulated during the time of highest prices. Wages increase and labour efficiency decreases. The strong demand for money causes interest rates to rise to a high level while doubt enters the bankers mind as to the advisability of granting further loans. This situation causes businessmen to make price concessions in order to secure the necessary cash. Then follows the expectation of further reduction and the situation becomes worse instead of better. Buyers wait for lower prices and all this leads to a decline in business activity. Then follows period of pessimism in trade and industry ; factories close, businesses fail, there is widespread unemployment, while wages and prices are low. These conditions characterise the period of depression. After a period of rigid economy, liquidation and reorganisation money accumulates and

*Business cycles are a type of fluctuations found in the aggregate economic activity of nations that organize their work mainly in business enterprise, a cycle consists of expansions occurring at about the same time in many economic activities followed by similarly 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, they are not divisible into shorter cycles of similar character with amplitudes approximating their own.

—*Arthur Burns and Mitchell*

seeks a use. Then follows a period of increasing business activity with rising prices, a period of improvement or recovery. The improvement period generally develops into the prosperity period and a business cycle is completed. Then movements discussed above are constantly repeated in the order given as the cycle completes it swing every few years.

The study of cyclical variations is extremely useful in framing suitable policies for stabilising the level of business activity, *i.e.*, for avoiding periods of booms and depressions as both are bad for an economy—particularly depression which brings about a complete disaster and shatters the economy.

But despite the great importance of measuring cyclical variations, they are the most difficult type of economic fluctuations to measure. It is because of the following two reasons :

(*i*) Business cycles do not show regular periodicity—they differ widely in timing, amplitude and pattern which makes their study very tough and tedious.

(*ii*) The cyclical variations are mixed with erratic, random or irregular forces which make it impracticable to isolate separately the effect of cyclical and irregular forces.

Business cycles are distinguished from seasonal variations in the following respects :

(*i*) The cyclical variations are of a longer duration than a year. A business cycle may be of any duration but normally the period of business cycle is 2-10 years. Moreover, they do not ordinarily exhibit regular periodicity as successive cycles vary widely in timing, amplitude and pattern.

For example, the 23 cycles of general business in the United States between 1854 and 1949 averaged 40 months; in duration individual cycles differed greatly. The shortest period lasted only 20 months and the longest persisted for 29 months.

(*ii*) The fluctuations in a business cycle result from a different set of causes. The period of prosperity, decline, depression and improvement viewed as four phases of a business cycle are generated by factors other than weather, social customs, and those which create seasonal patterns.

4. Irregular Variations*

Irregular variations refer to such variations in business activity which do not repeat in a definite pattern. In fact, the category labelled irregular variation is really intended to include all types of variations other than those accounting for the trend, seasonal and cyclical movements. These latter three, if they are actually at work, act in such a way as to produce certain systematic effects. Irregular movements, on the other hand, are considered to be largely random. being the result of chance factors, which like those determining the fall of a coin, are wholly unpredictable.

Irregular variations are caused by such isolated special occurrences as flood, earthquakes, strikes and wars. Sudden changes in demand or very rapid technological progress may also be included in this category. By their very nature these movements are very irregular and unpredictable. Quantitatively, it is almost impossible to separate out the irregular movements and the cyclical movements. Therefore, while analysing time series, the trend and seasonal variations are measured separately and the cyclical and irregular variations are left altogether.

There are two reasons for recognising irregular movements:

(*i*) To suggest that on occasions it may be possible to explain certain moments in the data as due to specific causes and to simplify further analysis.

(*ii*) To emphasise the fact that prediction of economic conditions is always subject to degree of error owing to the unpredictable erratic influences which may enter.

---

*Irregular variations are also called 'erratic', 'random' or 'accidental' variations.

Problems of Classification

Although it is a simple matter to classify the factors affecting time series into four groups for analytical purposes, the actual application of the classification frequently presents serious problems. Seasonal variations are by no means always so uniform in amplitude and timing that their identification can be made with certainty. Consequently, the investigator often finds it hard to distinguish seasonal influences from cyclical or random factors. Long and severe cycles may, to some observors, appear to be changes in the direction of the regular trend. During the Great Depression of the 1930's, for example, many leading economists interpreted the existing conditions not as a cyclical depression but as "secular stagnation".

Another difficulty arises because the four components of time series data are not mutually independent of one another. An exceedingly severe seasonal influence may aggravate or even precipitate a change in the cyclical movement. Conversely, cyclical influence may seriously affect the seasonal. A very rising trend *virtually eliminates seasonal and cyclical variations.*

Finally, the fourfold breakdown of time series data when applied to general business conditions has frequently been challenged on analytical grounds. Bratt* sees not one trend, but two: a primary trend representing the long-term growth of productive capacity, and the drift away from it which he calls secondary trend. Schumpeter developed an even more detailed breakdown by identifying three cyclical components, the 3-year Kitchin cycle, the 10-year Juglar cycle and the 50-year Kondratieff cycle**. The divergence of opinion among eminent scholars indicates clearly that the fourfold breakdown is more approximation, convenient to employ but frequently subject to modification.

Preliminary Adjustments before Analysing Time Series

Before beginning the actual work of analysing a time series it is necessary to make certain adjustments in the raw data. The adjustments are:

1. Adjustment for Calendar Variation,
2. Adjustment for Population changes,
3. Adjustment for Price Changes, and
4. Adjustment for Comparability.

1. *Calendar Variation.* A vast proportion of the important time series is available in a monthly for and it is necessary to recognise that the month is a variable time unit. The actual length of the short month is about 10 per cent less than that of longest, and if we take into account holidays and weeken the variation may be even greater. Thus, the production or sales for the month of February may be l not because of any real drop in activity but because of the fact that February has fewer days. Thus purpose of adjusting for calendar variation is to eliminate certain spurious differences which are cau by differences in number of days in various months. The adjustment for calendar variation is made dividing each monthly total by the number of days in the month (sometimes by the number of wor days in the month) thus arriving at daily average for each month. Comparable (adjusted) monthly may then be obtained by multiplying each of the values by 30.4167(365/12), the average number of in a month (In a leap year this factor is 30.5).

2. *Population Changes.* Certain types of data call for adjustment for population changes. Ch in the size of population can easily distort comparisons of income, production and consumption fi For example, national income may be increasing year after year, but per capita income may be dec because of greater pressure of population. Similarly, the production of a commodity may be going the per capita consumption may be declining. In such cases where it is necessary to adjust da population changes, a very simple procedure is followed, *i.e.*, the data are expressed on a per basis by dividing the original figures by the appropriate population total.

---

*Elmer C. Bratt: *Business Cycle and Forecasting.*

**J. A. Schumpeter: *Business Cycles.*

3. *Price Changes.* An adjustment for price changes is necessary whenever we have a value series and are interested in quantity changes alone. Because of rising prices the total sale proceeds may go up even when there is a fall in the number of units sold. For example, if in 2009, 1,000 units of a commodity that is priced Rs. 10 are sold, the total sale proceeds would be $1{,}000 \times 10 =$ Rs. 10,000. Assume that in 2010 the price of the commodity increases from Rs. 10 to Rs. 11. If the sales do not decline, the total sale proceeds will be $1{,}000 \times 11 =$ Rs. 11,000. This increase in sale proceeds, *i.e.*, Rs. 1,000, is not due to increase in the demand of the commodity but purely because of the rise in price from Rs. 10 to Rs. 11. Since value is equal to price per unit multiplied by the number of units sold, the effect of price changes can be eliminated by dividing each item in a value series by an approximate price index. This in fact is the process of deflating which has been discussed in the chapter on index numbers.

4. *Comparability.* For any meaningful analysis of time series, it is necessary to see that the data are strictly comparable throughout the time period under investigation. Quite often it is difficult or even impossible to get strictly comparable data. For example, if we are observing a phenomenon over the last 25 years, the comparability may be observed by differences in definition, differences in geographical average, differences in the method adopted, change in the method of reporting, etc. For example, a sale figure for January 2008 may give the average for that month, some years later the corresponding sales figure may give the total for the month or perhaps sales on the 15th or last day of the month. If such type of changes are not taken into account, the data cannot strictly be compared and its analysis would lead to fallacious conclusion.

STRAIGHT LINE TREND—METHODS OF MEASUREMENT

The following methods are used for measuring trend :

1. The Freehand or Graphic Method,
2. The Semi-average Method,
3. The Method of Least Squares, and

Each of these methods is discussed below :

1. Freehand or Graphic Method

This is the simplest method of studying trend. Under this method, the given data are plotted on a graph paper and a trend line is fitted to the data just by *inspecting the graph of the series.* There is no formal statistical criterion whereby the adequacy of such a line can be judged and it all depends on the judgment of the statistician. However, as a rough guide the line should be so drawn that it passes between the plotted points in such a manner that the fluctuations in one direction are approximately equal to those in the other direction and that it shows a general movement.

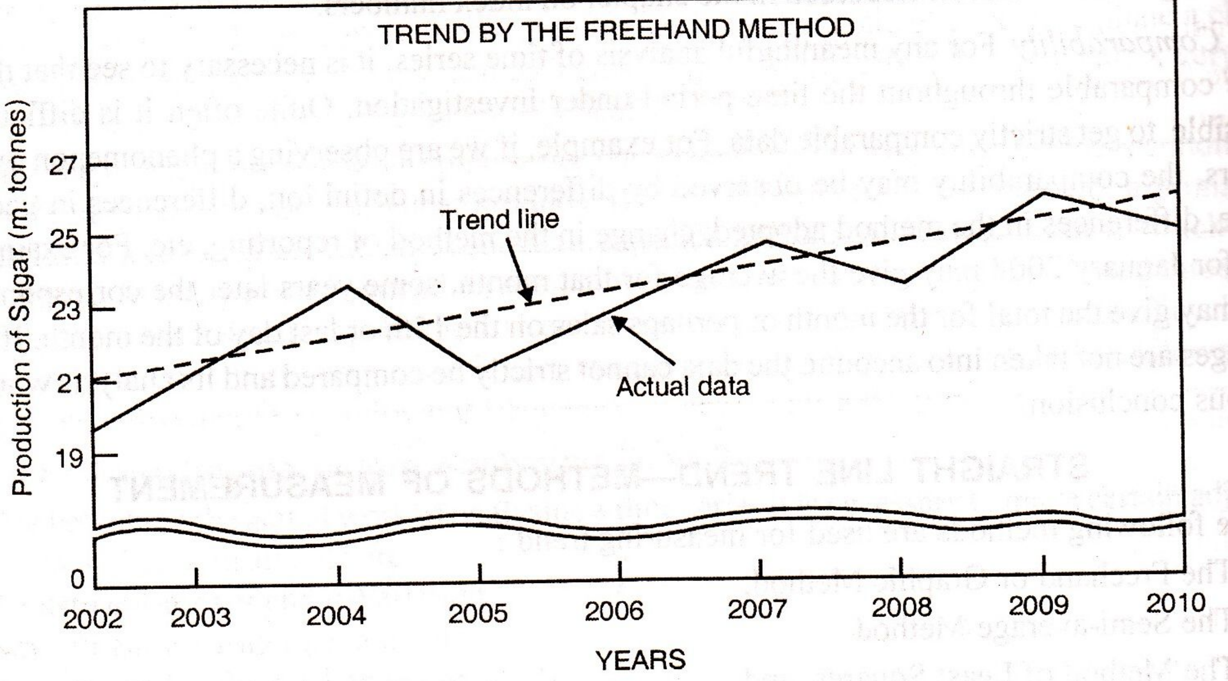
When a trend line is fitted by the freehand method, an attempt should be made to make it conform as much as possible to following conditions :

1. It should be smooth—either a straight line or a combination of long gradual curves.
2. The sum of the vertical deviations from the trend of the annual observations above the trend line should equal the sum of the vertical deviations from the trend of the observations below the trend line.
3. The sum of the squares of the vertical deviations of the observations from the trend should be as small as possible.
4. The trend should bisect the cycles so that area above the trend equals that below the trend, not only for the entire series but as much as possible for each full cycle. This last condition cannot always be met fully, but a careful attempt should be made to observe it as closely as possible.

Illustration 1. Fit a trend line to the following data by the freehand method :

Year	*Production of sugar (in million tonnes)*	*Year*	*Production of sugar (in million tonnes)*
2002	20	2007	25
2003	22	2008	23
2004	24	2009	26
2005	21	2010	25
2006	23		

The trend line drawn by the freehand method can be extended to predict future values. However, since the freehand curve fitting may be subjective, this method should not be used as a basis for prediction.



Merits and Limitations of the Freehand Method

Merits. 1. This is the simplest method of measuring trend.

2. This method is very flexible in that it can be used regardless of whether the trend is a straight line or curve.

3. The trend line drawn by a statistician experienced in computing trend and having knowledge of the economic history of the concern or the industry under analysis may be better expression of the secular movement than a trend fitted by the use of a rigid mathematical formula which, while providing a good fit to the points, may have no other logical justification. In fact a specialist of long experience who is familiar with the institutional setting, history and behaviour of the series may well be able manually to fit a trend superior to one derived by mathematical means. Although the freehand method is not recommended for beginners, it has considerable merit in the hands of experienced statisticians and is widely used in applied situations.

Limitations. 1. This method is highly subjective because the trend line depends on the personal judgment of the investigator and, therefore, different persons may draw different trend lines from the same set of data. Moreover, the work cannot be left to clerks and it must be handled by skilled and experienced people who are well conversant with the history of the particular concern.

2. Since freehand curve fitting is subjective, it cannot have much value if it is used as a basis for predictions.

3. Although this method appears simple and direct, it takes a lot of time to construct a freehand trend if a careful and conscientious job is done.

It is only after long experience in trend fitting that a statistician should attempt to fit a trend line by inspection.

2. Method of Semi-Averages

When this method is used the given data are divided into two parts, preferably, with the equal number of years. For example, if we are given data from 1993 to 2010, *i.e.*, over a period of 18 years, the two equal parts will be first nine years, *i.e.*, from 1993 to 2001 and from 2002 to 2010. In case of odd number of years like, 9, 13, 17, etc., two equal parts can be made simply by ignoring the middle year. For example, if data are given for 19 years from 1993 to 2010 the two equal parts would be from 1992 to 2000 and from 2002 to 2010—the middle year 2001 will be ignored.

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 at the mid-point of the class-interval covered by the respective part and then the two points are joined by a straight line which gives us the required trend line. The line can be extended downwards or upwards to get intermediate values or to predict future values.

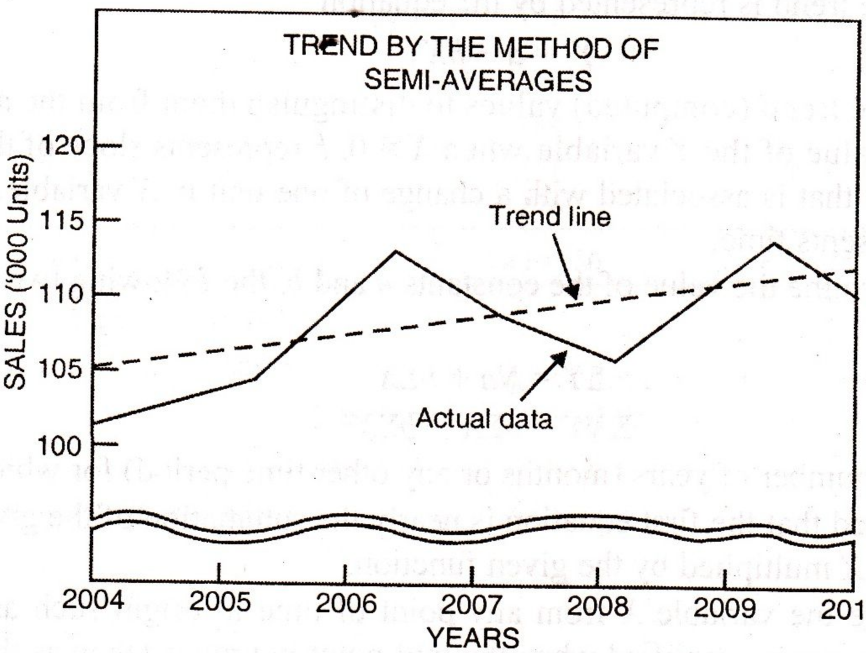
The following example shall illustrate the application of this method :

Illustration 2. Fit a trend line to the following data by the method of semi-averages :

Year	*Sales of Firm A (thousand units)*	*Year*	*Sales of Firm A (thousand units)*
2004	102	2008	108
2005	105	2009	116
2006	114	2010	112
2007	110		

Solution. Since seven years are given, the middle year shall be omitted and an average of the first three years and the last three years shall be obtained. The average of the first three years is $\frac{102+105+114}{3} = \frac{321}{3} = 107$ and the average of the last three years is $\frac{108+116+112}{3} = \frac{336}{3} = 112$. Thus, we get two points 107 and 112 which shall be plotted corresponding to their respective middle years, *i.e.*, 2005 and 2009. By joining these two points we shall obtain the required trend line. The line can be extended and can be used either for prediction or for determining intermediate values.

The actual data and the trend line are shown in the graph below.



Where there are even number of years like 6, 8, 10, etc., two equal parts can easily be formed and an average of each part obtained. However, when the average is to be centered there would be some problem in case the number of years is 8, 12, etc. For example, if the data relate to 2005-2010 which would be the middle year ? In such a case the average will be centered corresponding to 1st July, 2007, *i.e.*, middle of 2007 and 2008.

Merits and Limitations of the Semi-Average Method

Merits. 1. This method is simple to understand compared to the moving average method and method of least squares.

2. This is an objective method of measuring trend as everyone who applies the method is bound to get the same result (of course, leaving aside the arithmetical mistakes).

Limitations. 1.This method assumes straight line relationship between the plotted points regardless of the fact whether that relationship exists or not.

2. The limitations of arithmetic average shall automatically apply. If there are extremes in either half or both halves of the series, then the trend line is not a true picture of the growth factor. This danger is greatest when the time period represented by the average is small. Consequently, trend values obtained are not precise enough for the purpose either of forecasting the future trend or of eliminating trend from original data.

For the above reasons if the arithmetic average of the data is to be used in estimating the secular movement, it is sometimes better to use moving average than the semi-averages.

3. Method of Least Squares

This method is most widely used in practice. When this method is applied, a trend line is fitted to the data in such a manner that the following two conditions are satisfied :

(1) $$\Sigma(Y-Y_c) = 0$$

i.e., the sum of deviations of the actual values of Y and the computed values of Y is zero.

(2) $$\Sigma (Y - Y_c)^2 \text{ is least,}$$

i.e., the sum of the squares of the deviations of the actual and computed values is least from this line. That is why this method is called the method of least squares. The line obtained by this method is known as the line of '*best fit*'.

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

The straight line trend is represented by the equation

$$Y_c = a + bX$$

where Y_c denotes the trend (computed) values to distinguish them from the actual Y values, a is the Y intercept or the value of the Y variable when $X = 0$, b represents slope of the line or the amount of change in Y variable that is associated with a change of one unit in X variable. The X variable in time series analysis represents time.

In order to determine the value of the constants a and b, the following two normal equations are to be solved* :

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

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

where N represents number of years (months or any other time period) for which data are given.

It should be noted that the first equation is nearly the summation of the given function, the second is the summation of X multiplied by the given function.

We can measure the variable X from any point of time in origin such as the first year. But the calculations are very much simplified when the mid-point in time is taken as the origin because in that case the negative values in the first half of the series balance out the positive values in the second half so that $\Sigma X = 0$. In other words, the time variable is measured as a deviation from its mean. Since $\Sigma X = 0$, the above two normal equations would take the form :

*For details, see chapter on Regression Analysis.

$$\Sigma Y = Na$$
$$\Sigma XY = b\Sigma X^2$$

The values of a and b can now be determined easily.

Since $\Sigma Y = Na,$

$$a = \frac{\Sigma Y}{N} = \bar{Y}$$

Since $\Sigma XY = b\Sigma X^2,$

$$b = \frac{\Sigma XY}{\Sigma X^2}$$

The constant a give the arithmetic mean of Y and the constant b indicates the rate of change.

It should be noted that in case of odd number of years when the deviations are taken from the middle year, $\sum X$ would always be zero, provided there is no gap in the data given. However, in case of even number of years also $\sum X$ would always be zero if the X origin is placed midway between the two middle years. Hence both in odd as well as in even number of years we can use the simple procedure of determining the values of the constant a and b.

Illustration 3. Below are given the figures of production (in m. tonnes) of a sugar factory:

Year	2004	2005	2006	2007	2008	2009	2010
Production (in m. tonnes)	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.
(*iii*) Estimate the likely sales of the company during 2012.

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

Year	*Production (in m. tonnes)* Y	*Deviations from middle year* X	XY	X^2	*Trend Values* Y_c
2004	80	−3	−240	9	84
2005	90	−2	−180	4	86
2006	92	−1	−92	1	88
2007	83	0	0	0	90
2008	94	+1	+94	1	92
2009	99	+2	+198	4	94
2010	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: $Y_c = a + bX.$

Since $\Sigma X = 0$

$$a = \frac{\Sigma Y}{N} = \frac{630}{7} = 90;\ b = \frac{\Sigma XY}{\Sigma X^2} = \frac{56}{28} = 2$$

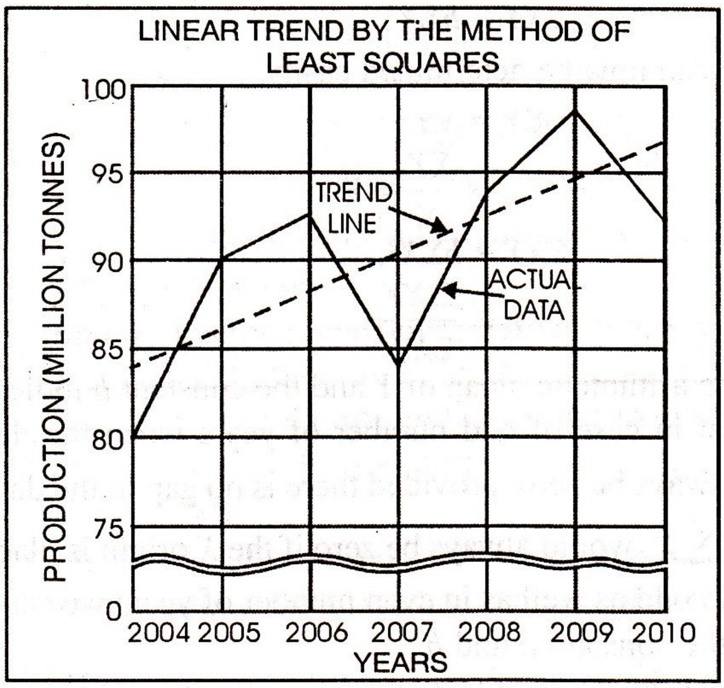
Hence the equation of the straight line trend is :

$$Y_c = 90 + 2X.$$

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, only first trend value need be obtained and then if the value of b is positive we may continue adding the value of b to every preceding value. For example, in the above case for 2004 the calculated value of Y is 84. For 2005 it will be 84 + 2 = 86; for 2006 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 :



(*iii*) For 2012, X would be +5. Putting $X = 5$ in the equation

$$Y_{2012} = 90 + 2(5) = 100$$

Hence the likely production for 2012 is 100 m. tonnes.

Illustration 4. Apply the method of least squares to obtain the trend values from the following data :

Year	*Sales (in lakh tonnes)*	*Year*	*Sales (in lakh tonnes)*
2006	100	2009	140
2007	120	2010	80
2008	110		

Also predict the sales for the year 2013.

Solution. CALCULATION OF TREND VALUES BY THE METHOD OF LEAST SQUARES

Year	*Sales* Y	*Deviations from middle year* X	XY	X^2	Y_C
2006	100	−2	−200	4	114
2007	120	−1	−120	1	112
2008	110	0	0	0	110
2009	140	+1	+140	1	108
2010	80	+2	+160	4	106
$N = 5$	$\Sigma Y = 550$	$\Sigma X = 0$	$\Sigma XY = -20$	$\Sigma X^2 = 10$	$\Sigma Y_C = 550$

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

Since $\Sigma X = 0$, $a = \dfrac{\Sigma Y}{N} = \dfrac{550}{5} = 110$ and $b = \dfrac{\Sigma XY}{\Sigma X^2} = \dfrac{-20}{10} = -2$

The required equation is : $Y_C = 110 - 2X$.

For $X = -2$, $Y_C = 110 - 2\,(-2) = 114$.

Now the other trend values will be obtained by *deducting* the value of b from the preceding value. Thus for 20 :he trend value will be 114 − 2 = 112 (since the value of b is negative). For 2013, likely sales = 100 lakh tonnes (sinc vould be 5 for 2013).

Illustration 5. Calculate the trend values by the method of least squares from the data given below and estimate sales for the year 2012-13.

Year :	2005-06	2006-07	2007-08	2008-09	2009-10
Sales of T.V. Sets (in lakh) :	12	18	20	23	27

Solution. CALCULATION OF TREND VALUES BY THE METHOD OF LEAST SQUARES

Year	*Sales* Y	*Taking 2007.5 as origin* X	XY	X^2	*Trend values* Y_c
2005-06	12	−2	−24	4	13.0
2006-07	18	−1	−18	1	16.5
2007-08	20	0	0	0	20.0
2008-09	23	+1	+23	1	23.5
2009-10	27	+2	+54	4	27.0
$N = 5$	$\Sigma Y = 100$	$\Sigma X = 0$	$\Sigma XY = 35$	$\Sigma X^2 = 10$	$\Sigma Y_c = 100$

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

Since $\Sigma X = 0$, $a = \dfrac{\Sigma Y}{N} = \dfrac{100}{5} = 20$ and $b = \dfrac{\Sigma XY}{\Sigma X^2} = \dfrac{35}{10} = 3.5$

Thus the equation of the straight line trend is : $Y = 20 + 3.5X$

$$Y_{2005\text{-}06} = 20 + 3.5\,(-2) = 13$$

$$Y_{2006\text{-}07} = 20 + 3.5\,(-1) = 16.5, \text{ etc.}$$

For 2012-13, X would be + 5.

Hence, $Y_{2012\text{-}13} = 20 + 3.5(+5) = 20 + 17.5 = 37.5$

Thus the estimated sales of television sets for the year 2012-13 is 37.5 lakh.

Illustration 6. Fit a straight line trend by the method of least squares to the following data and find the trend values :

Year	2005	2006	2007	2008	2009	2010
Sale of airconditioners (in lakh)	10	13	16	21	24	30

Solution. FITTING STRAIGHT LINE TREND BY THE METHOD OF LEAST SQUARES

Year	*Sales (in lakh)* Y	*Taking deviations from 2007* X	XY	X^2	*Trend values* Y_c
2005	10	−2	−20	4	9.143
2006	13	−1	−13	1	13.086
2007	16	0	0	0	17.029
2008	21	+1	+21	1	20.972
2009	24	+2	+48	4	24.915
2010	30	+3	+90	9	28.858
$N = 6$	$\Sigma Y = 114$	$\Sigma X = 3$	$\Sigma XY = 126$	$\Sigma X^2 = 19$	$\Sigma Y_c = 114.003$

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

Since ΣX is not zero, we have to solve the two normal equations simultaneously.

$$\Sigma Y = Na + b\Sigma X$$

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

$$114 = 6a + 3b \qquad \ldots(i)$$

$$126 = 3a + 19b \qquad \ldots(ii)$$

Multiplying Eqn. (*ii*) by 2,

$$114 = 6a + 3b$$

$$252 = 6a + 38b$$

$$-35b = -138 \qquad \text{or } b = 3.943$$

Putting the value of b in Eqn. (*i*)

$$114 = 6a + 3\,(3.943)$$

$$6a + 11.829 = 114$$

$$6a = 114 - 11.829 = 102.171 \text{ or } a = 17.029$$

The equation of the straight line trend is :

$$Y = 17.029 + 3.943\ X$$
$$Y_{2005} = 17.029 + 3.943\ (-2) = 9.143$$
$$Y_{2006} = 9.143 + 3.943 = 13.086$$

The other trend values can be obtained similarly by adding the value of b to the preceding value.

We can simplify the above calculations in case of even number of years by taking deviations from middle, *i.e.*, 2007.5 in the above price and apply the shortcut method.

Year	*Sales (in lakhs)* Y	*Taking deviation from 2007.5*	*Multiplying deviations by 2* X	XY	X^2
2005	10	−2.5	−5	−50	25
2006	13	−1.5	−3	−39	9
2007	16	−0.5	−1	−16	1
2008	21	+0.5	+1	+21	1
2009	24	+1.5	+3	+72	9
2010	30	+2.5	+5	+150	25
$N = 6$	$\Sigma Y = 114$		$\Sigma X = 0$	$\Sigma XY = 138$	$\Sigma X^2 = 70$

Since $\Sigma X = 0$, $a = \frac{\Sigma Y}{N} = \frac{114}{6} = 19$, $b = \frac{\Sigma XY}{\Sigma X^2} = \frac{138}{70} = 1.9714$

Thus the equation of the straight line trend is : $Y_c = 19 + 1.9714X$

$Y_{2005} = 19 + 1.9714\ (-5) = 19 - 9.857 = 9.143$

$Y_{2008} = 19 + 1.9714\ (-3) = 19 - 5.9142 = 13.086$, etc.

Note. Instead of calculating like this we can double the value of b (since b is giving half-yearly trend value) and add to the preceding value.

Annual trend value of $b = 1.9714 \times 2 = 3.943$

$Y_{2006} = 9.143 + 3.943 = 13.086$

$Y_{2007} = 13.086 + 3.943 = 17.029$ (as before)

Deviations have been multiplied by 2 just to simplify the calculations.

Merits and Limitations

Merits. 1. This is a mathematical method of measuring trend and as such there is no possibility of subjectiveness.

2. The line obtained by this method is called *the line of best fit* because it is this line from where the sum of the positive and negative deviations is zero and the sum of the squares of the deviations is least, *i.e.*, $\Sigma\ (Y - Y_c) = 0$ and $\Sigma\ (Y - Y_c)^2$ least.

Limitations. Mathematical curves are useful to describe the general movement of a time series, but it is doubtful whether any analytical significance should be attached to them, except in special cases. It is seldom possible to justify on theoretical grounds any real dependence of a variable with the passage of time. Variables do change in a more or less systematic manner over time, but this can usually be attributed to the operation of other explanatory variables. Thus, many economic time series show persistent upward trends over time due to a growth of population or to a general rise in prices, *i.e.*, national income and the trend element can to a considerable extent be eliminated by expressing these series per capita or in terms of constant purchasing power. For these reasons mathematical trends are generally best regarded as tools for describing movements in time series rather than as theories of the causes of such movement. It follows that it is extremely dangerous to use trends to forecast future movements of a time series. Such forecasting, involving as it does extrapolation, can be valid only if there is theoretical justification for the particular trend as an expression of a functional relationship between the variable under consideration and the time. But if the trend is purely descriptive of past behaviour, it can give few clues about future behaviour. Often the explanation of a trend gives ridiculous results which themselves are *prima facie* evidence that the trend could not be maintained.

Hence, mathematical methods of fitting trend are not foolproof—in fact, they can be a source of some of the most serious errors that are made in statistical work. They should never be used unless rigidly controlled by a separate logical analysis. Trend fitting depends upon the judgment of the statistician, and a skilfully made freehand sketch may often be more practical than a refined mathematical formula.*

NON-LINEAR TREND

The straight line trends discussed above indicate the increase or decrease of a time series at constant amount. It is the simplest form in describing the secular trend movement and the description of the trend is frequently accurate. However, in many cases, a straight line cannot fit the data adequately. For example, a time series may have faster (or slower) increase at early stage and have a slower (or faster increase at more recent time. In such a case better description of the time series is given by a non-linear curve rather than straight line.

The following are the methods of measuring non-linear trends:

1. Freehand or Graphic Method.
2. Moving Average Method.
3. A parabolic trend by a second degree polynomial equation obtained by the method of least squares.

1. Freehand or Graphic Method

As explained earlier, this method involves an element of subjectiveness and as such is not recommended for general use.

2. Method of Moving Averages

When a trend is to be determined by the method of moving averages, the average value for a number of years (or months, or weeks) is secured, and this average is taken as the normal or trend value for the unit of time falling at the middle of the period covered in the calculation of the average. The effect of averaging is to give a smoother curve, lessening the influence of the fluctuations that pull the annual figures away from the general trend.

While applying this method, it is necessary to select a period for moving average such as 3-yearly moving average, 6-yearly moving average, 8-yearly moving average, etc. The period of moving average is to be decided in the light of the length of the cycle. Since the moving average method is most commonly applied to data which are characterised by cyclical movements, it is necessary to select a period for moving average which coincides with the length of cycle, otherwise the cycle will not be entirely removed. This danger is more severe, the shorter the time period represented by the average. When the period of moving average and the period of the cycle do not coincide the moving average will display a cycle which has the same period as the cycle in the data, but which has less amplitude than the cycle in the data. Often we find that the cycles in the data are not of uniform length. In such a case, we should take a moving average period equal to or somewhat greater than the average period of the cycle in the data. Ordinarily the necessary period will range between three and ten years for general business series but even longer periods are required for certain types of data.

*Riggleman and Frisbee: *Business Statistics.*

The formula for 3-yearly moving average will be :

$$\frac{a+b+c}{3}, \frac{b+c+d}{3}, \frac{c+d+e}{3}, \frac{d+e+f}{3}, \ldots\ldots$$

and for 5-yearly moving average

$$\frac{a+b+c+d+e}{5}, \frac{b+c+d+e+f}{5}, \frac{c+d+e+f+g}{5}, \ldots\ldots$$

Illustration 7. Calculate the 5-yearly and 7-yearly moving average for the following data of a number of commercial industrial failures in a country during 1995 to 2010.

Year	*No. of failures*	*Year*	*No. of failures*
1995	23	2003	9
1996	26	2004	13
1997	28	2005	11
1998	32	2006	14
1999	20	2007	12
2000	12	2008	9
2001	12	2009	3
2002	10	2010	1

Solution. CALCULATION OF 5-YEARLY AND 7-YEARLY MOVING AVERAGE

Year	*No. of failures*	*5-yearly moving totals*	*5-yearly moving average*	*7-yearly moving totals*	*7-yearly moving average*
1995	23	—	—	—	—
1996	26	—	—	—	—
1997	28	129	25.8 or 26	—	—
1998	32	118	23.6 = 24	153	21.9 or 22
1999	20	104	20.8 = 21	140	20.0 = 20
2000	12	86	17.2 = 17	123	17.6 = 18
2001	12	63	12.6 = 13	108	15.4 = 15
2002	10	56	11.2 = 11	87	12.4 = 12
2003	9	55	11.0 = 11	81	11.6 = 12
2004	13	57	11.4 = 11	81	11.6 = 12
2005	11	59	11.8 = 12	78	11.1 = 11
2006	14	59	11.8 = 12	71	10.1 = 10
2007	12	49	9.8 = 10	63	9.0 = 9
2008	9	39	7.8 = 8	—	—
2009	3	—	—	—	—
2010	1	—	—	—	—

If the period of moving average is even, say, four-yearly or six-yearly, the moving total and moving average which are placed at the centre of the time span from which they are computed fall between two time periods. This placement is inconvenient since the moving average so placed would not coincide with an original time period. We, therefore, synchronise moving averages and original data. This process is called centering and consists of taking a two-period moving average of the moving averages.*

* There is another method of centering the moving averages. If we are calculating 4-yearly moving average, we will first take four-yearly totals and of these totals, we will again take 2-yearly totals and divide these totals by 8.

Illustration 8. Work out the centered 4-yearly moving average for the following data :

Year	*Tonnage of cargo cleared*	*Year*	*Tonnage of cargo cleared*
1999	1102	2005	1452
2000	1250	2006	1549
2001	1180	2007	1586
2002	1340	2008	1476
2003	1212	2009	1624
2004	1317	2010	1586

Solution. CALCULATION OF THE CENTERED FOUR-YEARLY MOVING AVERAGE

Year	*Tonnage of cargo cleared*	*4-yearly moving totals*	*4-yearly moving average*	*4-yearly centered moving average*
1999	1102	—	—	—
2000	1250	—		—
	→	4872	1218.00	
2001	1180		→	1231.75
	→	4982	1245.50	
2002	1340		→	1253.87
	→	5049	1262.25	
2003	1212		→	1296.25
	→	5321	1330.25	
2004	1317		→	1356.37
	→	5530	1382.50	
2005	1452		→	1429.25
	→	5904	1476.00	
2006	1549		→	1495.87
	→	6063	1515.75	
2007	1586		→	1537.25
	→	6235	1558.75	
2008	1476		→	1563.37
	→	6272	1568.00	
2009	1624	—	—	—
2010	1586	—		

Merits and Limitations

Merits. 1. This method is simple as compared to the method of least squares.

2. It is a flexible method of measuring trend. If a few more figures are added to the data, the entire calculations are not changed—we only get some more trend values.

3. If the period of moving average happens to coincide with the period of cyclical fluctuations in the data, such fluctuations are automatically eliminated.

4. 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 a mathematical function.

Limitations. 1. Trend values cannot be computed for all the years. The longer the period of moving average, the greater the number of years for which trend values cannot be obtained. For example, in a three-yearly moving average, trend values cannot be obtained for the first year and last year, in five-yearly moving average for the first two years and the last two years, and so on. It is often these extreme years in which we are most interested.

2. Great care has to be exercised in selecting the period of moving average. No hard and fast rules are available for the choice of the period and one has to use his own judgment.

3. Since the moving average is not represented by a mathematical function, this method cannot be used in forecasting which is one of the main objectives of trend analysis.

4. Although theoretically we say that if the period of moving average happens to coincide with the period of cycle, the cyclical fluctuations are completely eliminated, but in practice since the cycles are by no means perfectly periodic, the lengths of the various cycles in any given series will usually vary considerably and, therefore, no moving average can completely remove the cycle. The best results would be obtained by a moving average whose period was equal to the average length of all the cycles in the given series. However, it is difficult to determine the average length of the cycle until the cycles are isolated from the series.

5. Finally, when the trend situation is not linear (a straight line) the moving average lies either above or below the true sweep of the data.

The moving average is appropriate for trend computation only when :

(*a*) the purpose of investigation does not call for current analysis or forecasting,

(*b*) the trend is linear, and

(*c*) the cyclical variations are regular both in period and amplitudes.

However, in practice, these conditions rarely hold true.

3. Second Degree Parabola

The simplest example of the non-linear trend is the *second degree parabola*, the equation of which is written in the form :

$$Y_c = a + bX + cX^2$$

When numerical values for constants a, b and c have been derived, the trend value for any year may be computed substituting in the equation the value of X for that year. The values of a, b and c can be determined by solving the following three normal equations simultaneously :

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

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

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

Note that the first equation is merely the summation of the given function, the second is the summation of X multiplied into the given function, and the third is the summation of X^2 multiplied into the given function.

When time origin is taken between two middle years ΣX and ΣX^3 would be zero. In that case the above equations are reduced to :

(*i*) $$\Sigma Y = Na + c\Sigma X^2$$

(*ii*) $$\Sigma XY = b\Sigma X^2$$

(*iii*) $$\Sigma X^2 Y = a\Sigma X^2 + c\Sigma X^4$$

The value of b can now directly be obtained from equation (*ii*) and that of a and c by solving (*i*) and (*iii*) simultaneously. Thus,

$$a = \frac{\Sigma Y - c\Sigma X^2}{N} \; ; \; b = \frac{\Sigma XY}{\Sigma X^2}$$

$$c = \frac{N\Sigma X^2 Y - \Sigma X^2 \Sigma Y}{N\Sigma X^4 - (\Sigma X^2)^2}$$

Illustration 9. The price (in Rs.) of a commodity during 2005-2010 is given below. Fit a parabola $Y = a+bX+cX^2$ to this data. Estimate the price of commodity for the year 2013.

Year	*Price*	*Year*	*Price*
2005	100	2008	140
2006	107	2009	181
2007	128	2010	192

Also plot the actual and trend values on the graph.

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

$$\Sigma Y = Na + b\Sigma X + c\Sigma X^2$$
$$\Sigma XY = a\Sigma X + b\Sigma X^2 + c\Sigma X^3$$
$$\Sigma X^2Y = a\Sigma X^2 + b\Sigma X^3 + c\Sigma X^4$$

Year	*Price* Y	X	X^2	X^3	X^4	XY	X^2Y	*Trend Values* Y_c
2005	100	–2	4	–8	16	–200	400	97.744
2006	107	–1	1	–1	1	–107	107	110.426
2007	128	0	0	0	0	0	0	126.680
2008	140	+1	1	+1	1	+140	140	146.506
2009	181	+2	4	+8	16	+362	724	169.904
2010	192	+3	9	+27	81	+576	1728	196.874
$N = 6$	$\Sigma Y = 848$	$\Sigma X = 3$	$\Sigma X^2 = 19$	$\Sigma X^3 = 27$	$\Sigma X^4 = 115$	$\Sigma XY = 771$	$\Sigma X^2Y = 3099$	$\Sigma Y_c = 848.134$

$$848 = 6a + 3b + 19c \quad \text{...(i)}$$
$$771 = 3a + 19b + 27c \quad \text{...(ii)}$$
$$3{,}099 = 19a + 27b + 115c \quad \text{...(iii)}$$

Solving Eqns. (*i*) and (*ii*), we get

$$35b + 35c = 694 \quad \text{...(iv)}$$

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

$$53.52 = 280b + 168c \quad \text{...(v)}$$

Solving Eqns. (*iv*) and (*v*), we get

$$c = 1.786$$

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

$$b = 18.04$$

Putting the value of *b* and *c* in Eqn. (*i*), we get

$$a = 126.68$$

Thus, $a = 126.68$, $b = 18.04$ and $c = 1.786$

Substituting the values in the equation

$$Y = 126.68 + 18.04X + 1.786X^2$$

When $X = -2$, $\quad Y = 126.68 + 18.04\,(-2) + 1.786\,(-2)^2$
$= 126.68 - 36.08 + 7.144 = 97.744$

When $X = -1$, $\quad Y = 126.68 + 18.04\,(-1) + 1.786\,(-1)^2$
$= 126.68 - 18.04 + 1.786 = 110.426$

When $X = 0$, $\quad Y = 126.68$

When $X = 1$, $\quad Y = 126.68 + 18.04 + 1.786 = 146.506$

When $X = 2$, $\quad Y = 126.68 + 18.04\,(2) + 1.786\,(2)^2$
$Y = 126.68 + 36.08 + 7.144 = 169.904$

When $X = 3$, $\quad Y = 126.68 + 18.04\,(3) + 1.786\,(3)^2$
$= 126.68 + 54.12 + 16.074 = 196.874$

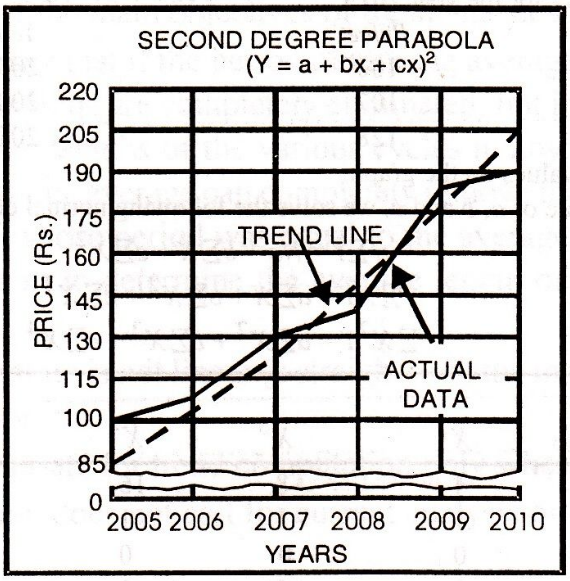
Price for 2013

For 2013 X will be 6.

When $X = 6$ $\quad Y = 126.68 + 18.04\,(6) + 1.786\,(6)^2$
$= 126.68 + 108.24 + 64.296 = 299.216$

Thus, the likely price of the commodity for the year : 2013 is Rs. 299.216.

The graph of the actual trend values is given below :



I. MEASURING TRENDS BY LOGARITHMS

The trends discussed so far were plotted on arithmetic scales. Trends may also be plotted on a semi-logarithmic (or semi-log) chart in the form of a straight line or a nonlinear curve. A straight line on the semi-log chart shows the increase of Y values of a time series at a constant rate (A straight line on an arithmetic chart indicates the increase at a constant amount). When it is a nonlinear curve on the semi-log chart an upward curve shows the increase at varying rates, depending on the shape of the slope—the steeper the slope, the higher is the rate of increase.

The types of trend usually computed by logarithms are :

1. Exponential trends, and
2. Growth curves.

Exponential Trends

The equation of the exponential curve is of the form

$$Y = ab^X$$

Putting the equation in logarithmic form, we get

$$\log Y = \log a + X \log b$$

When plotted on a semi-logarithmic graph, the curve gives a straight line. However, on an arithmetic chart the curve gives a nonlinear trend. In order to find out the values of a and b, the two normal equations to be solved are :

$$\Sigma \log Y = N \log a + \log b \, \Sigma X$$

$$\Sigma (X.\log Y) = \log a \, \Sigma X + \log b \, \Sigma X^2$$

When deviations are taken from middle year, *i.e.*, $\Sigma X = 0$, the above equations take the following form

$$\Sigma \log Y = N \log a$$

$$\Sigma (X.\log Y) = \log b \, \Sigma X^2$$

or

$$\log a = \frac{\Sigma \log Y}{N}; \text{ and } \log b = \frac{\Sigma (X.\log Y)}{\Sigma X^2}$$

Take the antilogs of these expressions to arrive at the actual trend values.

Illustration 10. The sales of a company in lakhs of rupees for the year 2004 to 2010 are given below :

Years :	2004	2005	2006	2007	2008	2009	2010
Sales :	32	47	65	92	132	190	275

(m. tonnes)

Estimate sales figures for the year 2013 using an equation of the form $Y = ab^x$, where X = years and Y = sales

Solution. FITTING EQUATION OF THE FORM $Y = ab^x$

Year	*Sales* Y	*Deviations from 2007* X	X^2	*log Y*	*X log Y*
2004	32	−3	9	1.5051	−4.5153
2005	47	−2	4	1.6721	−3.3442
2006	65	−1	1	1.8129	−1.8129
2007	92	0	0	1.9638	0
2008	132	+1	1	2.1206	+2.1206
2009	190	+2	4	2.2788	+4.5576
2010	275	+3	9	2.4393	+7.3179
$N = 7$	$\Sigma Y = 833$	$\Sigma X = 0$	$\Sigma X^2 = 28$	$\Sigma \log Y = 13.7926$	$\Sigma X \log Y = 4.3237$

We have to fit the equation $Y = ab^x$. It can be written as

$$\log Y = \log a + X \log b$$

Since deviations are taken from middle year, $\Sigma X = 0$

$$\log a = \frac{\Sigma \log Y}{N} = \frac{13.7926}{7} = 1.97$$

$$\log b = \frac{\Sigma (X \log Y)}{\Sigma X^2} = \frac{4.3237}{28} = 0.1544$$

Hence $\log Y = 1.97 + 0.1544\, X$

For 2013, X would be + 6.

$$\log Y = 1.97 + 0.1544\,(6) = 2.8964$$

$$Y = \text{Antilog } 2.8964 = 787.77$$

Thus the estimated figure of sales for the year 2013 is 787.77 m. tonnes.

Illustration 11. Fit a logarithmic straight line to the following data :

Years :	2005	2006	2007	2008	2009	2010
Production (m. tonnes of steel) :	64	70	75	82	88	95

Solution. FITTING OF LOGARITHMIC STRAIGHT LINE

Year	*Production* Y	X	*log Y*	X^2	*X.log Y*
2005	64	−3	1.8062	9	−5.4186
2006	70	−2	1.8451	4	−3.6902
2007	75	−1	1.8751	1	−1.8751
2008	82	0	1.9138	0	0
2009	88	+1	1.9445	1	+1.9445
2010	95	+2	1.9777	4	+3.9554
$N = 6$	$\Sigma Y = 474$	$\Sigma X = -3$	$\Sigma \log Y = 11.3624$	$\Sigma X^2 = 19$	$\Sigma X.\log Y = -5.084$

The logarithmic straight line trend is given by

$$\log Y = \log a + X \log b$$

The two normal equations are :

$$\Sigma \log Y = N \log a + \log b\, \Sigma X$$

$$\Sigma X \log Y = \log a\, \Sigma X + \log b\, \Sigma X^2$$

Substituting the values

$$11.3624 = 6 \log a - 3 \log b \qquad \ldots(i)$$

$$-5.084 = -3 \log a + 19 \log b \qquad \ldots(ii)$$

Multiplying eq. (*ii*) by 2 and adding to (*i*),

$$\begin{aligned} 11.3624 &= 6 \log a - 3 \log b \\ -10.168 &= -6 \log a + 38 \log b \\ \hline 35 \log b &= 1.1944 \\ \log b &= \frac{1.1944}{35} = 0.034 \end{aligned}$$

Putting the value of log b in eqn. (*i*)

$$\begin{aligned} 11.3624 &= 6 \log a - 0.102 \\ 6 \log a &= 11.4644 \\ \log a &= 1.911 \end{aligned}$$

Hence $\log Y = 1.911 - 0.034\,X$

Second Degree Curves Fitted to Logarithms

We may come across data which when plotted on semi-logarithmic graph paper may continue to show curvature, being concave either upward or downward; or in other words, the ratio of change may be either increasing or decreasing. In such cases, we may fit second degree curve to the logarithms of the Y values using

$$\log Y = \log a + X \log b + X^2 \log c$$

Taking the X origin at the middle of the period, the three normal equations are:

(*i*) $$\Sigma \log Y = N \log a + \log c \Sigma X^2$$

(*ii*) $$\Sigma (X.\ \log Y) = \log b \Sigma X^2$$

(*iii*) $$\Sigma (X^2.\ \log Y) = \log a \Sigma X^2 + \log c \Sigma X^4$$

Growth Curves

In economic data very often we come across phenomenon where at first the growth is very slow, but as the product is accepted the demand increases by a greater amount each year and finally as the market becomes more and more fully developed, the amount of growth each year becomes less. The curve continues to grow more and more slowly, approaching an upper limit but not reaching it. Such series are best represented by growth curves. The growth curves do not reach a maximum and turn down in the manner of the second degree parabola.

A number of different growth curves have been used to measure secular trend, but the curves used most widely to describe growth are the *Gompertz Curve* and the *Peart Reed* or *logistic* curve.

The equation of the Gompertz curve is

$$Y = kab^X$$

which when put to logarithmic form becomes

$$\log Y = \log k + (\log a) b^X$$

The Gompertz curve serves to describe the series which while increasing seem to approach some maximum value as a limit. Although the growth continues it does so at a decreasing rate.

The equation of the logistic (or the Peart-Reed) curve is:

$$Y = \frac{1}{k + ab^X}$$

where k, a and b are constants. The logistic curve has been applied widely to population data of various kinds, both human and non-human, and it has also provided a good fit to many economic series pertaining to industrial growth.

Both the Gompertz and the logistic curves approach a finite limit. This fact ought to be taken into account when fitting a given time series to one of the curves. Often a key resource is known to exist to some finite amount, and this can be used to establish a limit on the growth of the time series question. Increasingly, for example, new cities are being planned with an eye towards limiting growth with planned land available having an upper limit. Whenever a given time series increasing at a constant rate but is understood to be approaching a finite limit in a predictable manner, growth curve may be appropriate for assessing the secular trend component of the series.

Conversion of Annual Trend Values to Monthly Trend Values

Usually for trend computations annual figures are employed. However, it is sometimes required to obtain monthly trend ordinates. In converting straight line trends from an annual to a monthly basis, two situations must be clearly distinguished. For series such as sales, production or earnings, the annual figure is the total of monthly figures. Here it is necessary to divide both a and b by 12 to reduce them to monthly level. In other words, on the average, monthly sales or production is one-twelth of the annual total. The b values must then be divided by 12 once again in order to convert from annual to monthly increments.

The necessity of dividing b twice by 12, that is by 144 altogether, must be clearly understood. The division of annual change by 12 gives us only the change from same month in a given year to the corresponding month in the following year, or the annual change in monthly magnitudes. However, we are here seeking for an expression of the change in each and every month, that is, monthly change in monthly magnitude. Thus, b must be divided again by 12.

In conclusion, to convert an annual trend equation to a monthly basis when the original data are given as totals, a is divided by 12 and b is divided by 144.

If X in the trend line equation represents only 6 months, it is divided by 72 instead of 144.

Illustration 12. Convert the following annual trend equation for tea production in India to a monthly trend equation :

$$Y = 108 + 1.58\,X$$

(Origin 2010, time unit one year, Y = tea production in million kg.)

Solution. Monthly trend equation will be obtained by dividing a by 12 and b by 144. Thus the monthly trend equation will be

$$Y = \frac{108}{12} + \frac{1.58}{144}$$

(Origin July 1, 2010, time unit 1 month, Y monthly production in million kg.)

Where data are given as monthly average per year, the value of the constant 'a' in the annual trend equation is the arithmetic mean of the twelve month total. In other words, it is already at the monthly level. The value of 'b' now represents the annual change in month magnitude. As a result, to convert an annual trend equation when annual data are expressed as monthly averages, a would remain unchanged and b is divided by 12 only.

Shifting the Trend Origin

In computing trends, the middle of the time series is often used as the origin in order to cut-short the computations. But very often we need to change the origin of the trend equation to some other point in the series. This is either to facilitate comparison of trend values among neighbouring years or to convert a trend equation from any annual to a monthly basis. Shifting the origin is a very simple matter. For example, consider the trend equation

$$Y = 110 - 2X$$

(Origin 2005, time unit 1 year)

If we wish to shift the trend equation to 2010, we note that this year precedes the stated origin of 2005 by 7 time units. Consequently, we must deduct 7 times annual increment that is $b(-7)$, from the trend value of 2005 as below :

$$Y = 110 - 2\,(-7) = 110 + 14 = 124.$$

The value 124 becomes the trend value at the new origin 2010 and the trend equation may now be written as

$$Y = 124 - 2X.$$

Selecting Type of Trend

We have discussed different ways of fitting trends. However, it is not all—some other equations might also be reasonably used. Even though each series presents its own individual problem, most series can be handled by the methods which we have described. Of course, what we try to do in any particular case is to select that equation or that method of measuring trend which best describes the gradual and consistent pattern of growth.

The choice of a particular type of equation that best describes the data is often difficult and needs considerable amount of judgment and experience.

While deciding the type of trend, the first step consists of plotting the data on arithmetic paper. If the trend is not linear but either:

(*a*) concave upwards,or

(*b*) concave downwards

the data should be plotted on a semi-logarithmic paper. Examination of the plotted data often provides an adequate basis for deciding upon the type of trend to use. When further guidance is needed an approximate trend may be drawn by inspection and the following tests applied to the smoothed curve*.

1. If the first differences tend to be constant, use a straight line.
2. If the second differences tend to be constant, use a second degree curve.
3. If the approximate trend when plotted on arithmetic paper is a straight line, use a straight line.
4. If the first differences of the logarithms are constant, use an exponential curve.
5. If the second differences of the logarithms are constant, fit a second degree curve to the logarithms.
6. If the first differences tend to decrease by a constant percentage, use a modified exponential.
7. If the first differences resemble a skewed frequency curve, use a Gompertz curve or a more complex logistic curve.

Choice of the Trend Period

In order to simplify the discussion of trend computation, the illustrations in this chapter are based on 7 or 8 years data only. However, wherever possible, the period should be longer. The longer the period, the less the trend values will be distorted by cyclical or random influence. The period employed should encompass a number of business cycles and should begin and end in such a way that distortion is avoided. The purpose can be accomplished by using a period that starts and finishes either in prosperity or depression, or by beginning during recovery and ending during recession.

Trend Extrapolation

Trend analysis is often employed to forecast future levels of time series data. By substituting in the trend equation values of X for the dates for which forecast are desired, the ordinates for those dates can be obtained. This process is called extrapolation. Utmost care must be exercised while interpreting such forecasts. The following points are worth considering :

(1) The forecast obtained through trend analysis is a forecast of trend only. Actual values may be expected to diverge from it because of cyclical and random factors.

*Croxton and Cowden: *Applied General Statistics.*

(2) The forecast has meaning only if the same basic influence that shaped the trend in the past can be expected to be controlling in the future. Thus extrapolation implicitly assumes that the institutional pattern will be quite stable. However, in practice, the extrapolation of trends goes far wide off the mark because significant changes in conditions governing the values of the data cannot be or are not properly anticipated.

(3) The type of curve employed must not only properly describe the past movement of the data but also be capable of sensible extrapolation. The extrapolation of some trends especially the more complex curves may yield results that are essentially meaningless.

II. MEASUREMENT OF SEASONAL VARIATIONS

Most of the phenomena in economics and business show seasonal patterns. When data are expressed annually, there is no seasonal variation. However, monthly or quarterly data frequently exhibit strong seasonal movements and considerable interest attaches to devising a pattern of *average seasonal variation.* For example, if we observe the sales of a bookseller, we find that of the quarter July-September (when most of the students purchase books), sales are maximum. If we know by how much the sales of this quarter are usually above or below the previous quarter for seasonal reasons, we shall be able to answer a very basic question, namely, was this due to an underlying upward tendency or simply because this quarter is usually seasonally higher than the previous quarter ?

In order to analyse seasonal variation, it is necessary to assume that the seasonal pattern is superimposed on a series of values and independent of these in the sense that the same pattern is superimposed irrespective of the level of the series, *i.e.,* the June quarter always contributes so much more or so much less of the series.

Before attempting to measure seasonal variation certain preliminary decisions must be made. For example, it is necessary to decide whether weekly, quarterly or monthly indexes are required. This will be decided in the light of the nature of the problem and the type of data available.

To obtain a statistical description of a pattern of seasonal variation it will be desirable to first free the data from the effect of trend, cycles and irregular variation. Once these other components have been eliminated we can calculate, in index form, a measure of seasonal variations which is usually referred to as seasonal index. Thus the measures of seasonal variation are called *seasonal indexes* (per cent).

For monthly data, a seasonal index consists of 12 numbers, one for each month of a year, or number of years, that has taken place typically in each month. Thus a second index may be specific or typical. A *specific seasonal index* refers to the seasonal changes during a particular year. A typical seasonal index is obtained by averaging a number of specific seasonals. It is thus a generalised expression of seasonal variations for a series. Seasonal indexes are given as percentages of their average, *i.e.,* each month is represented by a figure expressing it as a percentage of the average month. For example, if a seasonal index for January is 75, this means that for the month of January, sales, orders, purchases or whatever our data happen to be are 75 per cent of those of the average month.

There are several methods of measuring seasonal variation. However, the following methods that are more popularly used in practice are discussed below :

1. Method of Simple Averages (Weekly, Monthly or Quarterly),
2. Ratio-to-Trend Method,
3. Ratio-to-moving Average Method,
4. Link Relatives Method.

1. Method of Simple Averages

This is the simplest method of obtaining a seasonal index. The following steps are necessary for calculating the index :

(i) Average the unadjusted data by years and months (or quarters if quarterly data are given).

(ii) Find totals of January, February, etc.

(iii) Divide each total by the number of years for which data are given. For example, if we are given monthly data for five years then we shall first obtain total for each month for five years and divide each total by 5 to obtain an average.

(iv) Obtain an average of monthly averages by dividing the total of monthly averages by 12.

(v) Taking the average of monthly averages as 100, compute the percentage of various monthly averages as follows :

Seasonal Index for January

$$= \frac{\text{Monthly average for January}}{\text{Average of monthly averages}} \times 100$$

If, instead of the average of each month, the totals of each month are obtained, we will get the same result.

The following example shall illustrate the method.

Illustration 13. Consumption of monthly electric power in million of Kw hours for street lighting in one of the states in India during 2006-2010 is given below :

Year	*Jan.*	*Feb.*	*Mar.*	*Apr.*	*May*	*June*	*July*	*Aug.*	*Sept.*	*Oct.*	*Nov.*	*Dec.*
2006	318	281	278	250	231	216	223	245	269	302	325	347
2007	342	309	299	268	249	236	242	262	288	321	342	364
2008	367	328	320	287	269	251	259	284	309	345	367	394
2009	392	349	342	311	290	273	282	305	328	364	389	417
2010	420	378	370	334	314	296	305	330	356	396	422	452

Find out seasonal variation by the method of monthly averages.

Solution. COMPUTATION OF SEASONAL INDICES BY THE METHOD OF MONTHLY AVERAGES

Month (1)	*Consumption of monthly electric power* 2006 (2)	2007 (3)	2008 (4)	2009 (5)	2010 (6)	*Monthly total for 5 years* (7)	*Five-yearly average* (8)	*Percentage* (9)
Jan.	318	342	367	392	420	1,839	367.8	116.1
Feb.	281	309	328	349	378	1,645	329.0	103.9
March	278	299	320	342	370	1,609	321.8	101.6
April	250	268	287	311	334	1,450	290.0	91.6
May	231	249	269	290	314	1,353	270.6	85.4
June	216	236	251	273	296	1,272	254.4	80.3
July	223	242	259	282	305	1,311	262.2	82.8
Aug.	245	262	284	305	330	1,426	285.2	90.1
Sept.	269	288	309	328	356	1,550	310.0	97.9
Oct.	302	321	345	364	396	1,728	345.6	109.1
Nov.	325	342	367	389	422	1,845	369.0	116.5
Dec.	347	364	394	417	452	1,974	394.8	124.7
					Total	19,002	3,800.4	1,200
					Average	1.5835	316.7	100

The above calculations are explained below :

1. Column No. 7 gives the total for each month for five years.
2. In column No. 8 each total of column No. 7 has been divided by 5 to obtain an average for each month.
3. The average of monthly averages is obtained by dividing the total of monthly averages by 12.

4. In column No. 9 each monthly average has been expressed as a percentage of the average of monthly averages. Thus, the percentage for January

$$= \frac{367.8}{316.7} \times 100 = 116.1$$

$$\text{Percentage for February} = \frac{329.0}{316.7} \times 100 = 103.9$$

If instead of monthly data, we are given weekly or quarterly data, we shall compute weekly or quarterly averages by following the same procedure as explained above.

Merits and Limitations of the Method of Monthly Averages

This method is the simplest of all methods of measuring seasonality. However, it is not a very good method. It assumes that there is no trend component in the series, *i.e.*, $O = CSI$. But this is not a justified assumption. Most economic series have trends and, therefore, the seasonal index computed by this method is actually an index of trends and seasonals. Furthermore, the effects of cycles on the original values may or may not be eliminated by the averaging process. This depends on the duration of the cycle and the term of the average, that is, on the number of months included in the average. Thus, this method is seldom of any value. In its simplest form, the method only serves the purpose where no definite trend exists.

2. Ratio-to-Trend Method

This method of calculating a seasonal index (also known as the percentage-to-trend method) is relatively simple and yet an improvement over the method of simple average explained in the preceding section. This method assumes that seasonal variation for a given month is constant fraction of trend. The ratio-to-trend method presumably isolates the seasonal factors in the following manner. Trend is eliminated when the ratios are computed. In effect :

$$\frac{T \times S \times C \times I}{T} = S \times C \times I$$

Random elements are supposed to disappear when the ratios are averaged. A careful selection of the period of years used in the computation is expected to cause the influences of prosperity or depression to offset each other and thus remove the cycle. For series that are not subject to pronounced cyclical or random influences and for which trend can be computed accurately, this method may suffice. The steps in the computation of seasonal index by this method are :

1. Trend values are obtained by applying the method of least squares.
2. The next step is to divide the original data month by month by the corresponding trend values and to multiply these ratios by 100. The values so obtained are now free from trend and the problem that remains is to free them also of irregular and cyclical movements.
3. In order to free the values from irregular and cyclical movements, the figures given for the various years for January, February, etc., are averaged with any one of the usual measures of central value, for instance, the *median* or the *mean.* If the data are examined month by month, it is sometimes possible to ascribe a definite cause to usually high or low values. When such causes are found to be associated with irregular variations (extremely bad weather, an earthquake, famine and the like) they may be cast out and the mean of the remaining items is referred to as a *modified mean.* Since such scrutiny of the data requires considerable knowledge of prevailing condition and is to a large extent subjective, it is often described to use the *median* which is generally not affected by very high or very low values.

4. The seasonal index for each month is expressed as a percentage of the average month. The sum of 12 values must equal 1,200 or 100%. If it does not, an adjustment is made by multiplying each index by a suitable factor $\left(\dfrac{1200}{\text{the sum of 12 values}}\right)$. This gives the final seasonal index.

Illustration 14. Find seasonal variations by ratio-to-trend method from the data given below :

Year	*1st Quarter*	*2nd Quarter*	*3rd Quarter*	*4th Quarter*
2006	30	40	36	34
2007	34	52	50	44
2008	40	58	54	48
2009	54	76	68	62
2010	80	92	86	82

Solution. For determining seasonal variation by ratio-to-trend method, first we will determine the trend of yearly data 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 val* Y_c
2006	140	35	–2	–70	4	32
2007	180	45	–1	–45	1	44
2008	200	50	0	0	0	56
2009	260	65	+1	+65	1	68
2010	340	85	+2	+170	4	80
		$\Sigma Y = 280$	$\Sigma X = 0$	$\Sigma XY = 120$	$\Sigma X^2 = 10$	

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

Since $\Sigma X = 0$, $\quad a = \dfrac{\Sigma Y}{N} = \dfrac{280}{5} = 56; \quad \dfrac{\Sigma XY}{\Sigma X^2} = \dfrac{120}{10} = 12$

Quarterly increment $= \dfrac{12}{4} = 3$

Calculation of Quarterly Trend Values. Consider 2006. Trend value for the middle *i.e.*, half of 2nd and half of 3rd is 32. Quarterly increment is 3. So the trend value of 2nd c 32–3/2, *i.e.*, 30.5 and for 3rd quarter is 32 + 3/2, *i.e.*, 33.5. Trend value for the 1st quart – 3, *i.e.*, 27.5 and of 4th quarter is 33.5 + 3, *i.e.*, 36.5. We thus get quarterly trend va given values are to be expressed as the percentages of the corresponding trend values.

TREND VALUES

Year	*1st Quarter*	*2nd Quarter*	*3rd Quarter*	*4th Qua*
2006	27.5	30.5	33.5	36.5
2007	39.5	42.5	45.5	48.
2008	51.5	54.5	57.5	60.
2009	63.5	66.5	69.5	72.
2010	75.5	78.5	81.5	84

QUARTERLY VALUES AS % OF TREND VALUES

Year	*Ist Quarter*	*2nd Quarter*	*3rd Quarter*	*4th Q*
2006	109.1	131.1	107.5	9
2007	86.1	122.4	109.9	9
2008	77.7	106.4	93.9	[illegible]
2009	85.0	114.3	97.8	[illegible]
2010	106.0	117.1	105.5	
Total	463.9	591.3	514.6	4
Average	92.78	118.26	102.92	
Seasonal Index adjusted	92.0	117.4	102.1	

Total of averages = 92.78+118.26 + 102.92 + 89.12 = 403.08

Seasonal Index adjusted for Ist quarter = $\frac{92.78}{100.77}$ (avg. of total averages)

Since the total is more than 400 an adjustment is made by multiplying each average by $\frac{400}{403.08}$ and final indices are obtained.

Merits and Limitations of the Ratio-to-Trend Method

Merits. (1) Compared with the method of monthly averages this method is certainly a more logical procedure for measuring seasonal variations. It has an advantage over the moving average procedure too, for it has a ratio-to-trend value for each month for which data are available. Thus there is no loss of data as occurs in the case of moving averages. This is a distinct advantage, especially when the period covered by time series is very short.

(2) It is simple to compute and easy to understand.

Limitations. The main defect of the ratio-to-trend method is that if there are pronounced 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-month moving average does. In consequence 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.

3. Ratio-to-Moving Average Method*

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

1. Eliminate seasonality from the data by ironing it out from the original data. Since seasonal variations recur every year—that is, since the fluctuations have a time span of 12 months—a centered 12-month moving average tends to eliminate these fluctuations. (In case of quarterly data, a centered 4 quarter moving average must be used.) The centered 12-month moving average which aims to eliminate seasonal and irregular fluctuations (*S* and *I*) represents the remaining elements of the original data, namely, trend and cycles. Thus, the centered 12-month moving average approximates *T.C*.

2. Express the original data for each month as percentage of the centered 12-month moving average corresponding to it.

3. Divide each monthly item of the original data by the corresponding 12-month moving average, and list the quotients as 'Percent of Moving Average'. We have now succeeded in eliminating from the original data to a considerable extent the disturbing influence of trend and cycles. It remains to rid the data of irregular variations. By averaging these percentages, for a given month (step 4) the irregular factors tend to cancel out and the average itself reflects the seasonal influence alone.

4. The purpose of this step is to average, and—in process of averaging—to eliminate the irregular factors. We assume that the relatively high or extremely low values of seasonal relatives for any month are caused by irregular factors. The elimination of extremes may be achieved while we are averaging all Januarys, Februarys and the like. We do this by using an appropriate type to average. The median is appropriate since it is not affected by extremes. Thus, by using the median as an average we can obtain the typical seasonal relative for each month which will not be affected by irregular factors.

Sometimes a so-called modified mean is used as an average for each month. Here, extreme values are omitted before the arithmetic mean is taken. In an array of seasonal relatives for each month, a value

*The computation by this method is identical with computations of the ratio-to-trend seasonal index just described, except that a moving average trend is substituted for the least square trend used in the previous calculation.

or several values on one end or both ends may be dropped and then the arithmetic mean of the remaining seasonal relatives is taken. A separate table is prepared in which the calculations involved in this step are shown. These means are preliminary seasonal indexes. They should average 100 per cent or total 1,200 for 12 months by definition.

5. If the total is not equal to 1,200 or 100 per cent, an adjustment is made to eliminate the discrepancy. The adjustment consists of multiplying average of each month obtained in step 4 by

$$\frac{1{,}200}{\text{The total of the modified mean for 12 months}}$$

This adjustment is made not only to achieve accuracy, but also because when we come to eliminate seasonality from the original data we do not wish to raise or lower the level of the data unduly. Thus, if a seasonal index aggregates more than 1,200 (or averages more than 100) then the original data adjusted in terms of it will total less than the unadjusted original data. If it totals less than 1,200, the opposite would be true.

The logical reasoning behind this method follows from the fact that 12-month moving average can be considered to represent the influence of cycle and trend $C \times T$. If the actual value for any month is divided by the 12-month moving average centred to that month, presumably cycle and trend are removed. This may be represented by the following expression :

$$\frac{T \times S \times C \times I}{T \times C} = S \times I$$

Thus the ratio to the moving average, from which this method gets its name, represents irregular and seasonal influences. If the ratios for each worked over a period of years are then averaged most random influences will usually be eliminated. Hence, in effect,

$$\frac{S \times I}{I} = S$$

Illustration 15. Apply ratio to moving average method to ascertain seasonal indices from the following data :

Year and month	*Sales (in thousand units)*	*Year and month*	*Sales (in thousand units)*
2007		2009	
Jan.	10	Jan.	10
Feb.	12	Feb.	12
March	13	March	11
April	15	April	12
May	16	May	13
June	16	June	15
July	17	July	15
Aug.	18	Aug.	17
Sept.	18	Sept.	18
Oct.	19	Oct.	20
Nov.	22	Nov.	22
Dec.	22	Dec.	24
2008		2010	
Jan.	11	Jan.	12
Feb.	11	Feb.	13
March	12	March	13
April	13	April	15
May	14	May	16
June	14	June	18
July	15	July	20
Aug.	15	Aug.	20
Sept.	15	Sept.	21
Oct.	16	Oct.	22
Nov.	18	Nov.	24
Dec.	20	Dec.	25

Solution. COMPUTATION OF 12-MONTH MOVING AVERAGE

Year and month	*Sales (Thousand units)*	*12-Month moving total*	*12-Month moving average*	*2-Month moving total of col.4*	*Centered 12-month moving average (col. 5 ÷ 2)*	*Percentage of centered 12-month moving average (col. 2 ÷ col. 6)*
(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*
2007						
Jan.	10					
Feb.	12					
March	13					
April	15					
May	16					
June	16					
		198	16.50			
July	17			33.03	16.54	102.8
		199	16.58			
Aug.	18			33.03	16.54	108.8
		198	16.50			
Sept.	18			32.92	16.46	109.4
		197	16.42			
Oct.	19			32.67	16.33	116.3
		195	16.25			
Nov.	22			32.33	16.16	136.1
		193	16.08			
Dec.	22			32.00	16.00	137.5
		191	15.92			
2008						
Jan.	11			31.67	15.83	69.5
		189	15.75			
Feb.	11			31.25	15.62	70.4
		186	15.50			
March	12			30.75	15.37	78.1
		183	15.25			
April	13			30.25	15.12	86.0
		180	15.00			
May	14			29.67	14.83	94.4
		176	14.67			
June	14			29.17	14.59	96.0
		174	14.50			
July	15			28.82	14.46	103.7
		173	14.42			
Aug.	15			28.92	14.46	103.7
		174	14.50			
Sept.	15			28.92	14.46	103.7
		173	14.42			
Oct.	16			28.75	14.37	111.3
		172	14.33			
Nov.	18			28.58	14.29	126.0
		171	14.25			
Dec.	20			28.58	14.29	140.00

Year and month	Sales (Thousand units)	12-Month moving total	12-Month moving average	2-Month moving total of col.4	Centered 12-month moving average (col. 5 ÷ 2)	Percentage of centered 12-month moving average (col. 2 ÷ col. 6)
(1)	(2)	(3)	(4)	(5)	(6)	(7)
2009		172	14.33			
Jan.	10			28.66	14.33	70.0
		172	14.33			
Feb.	12			28.83	14.41	83.3
		174	14.50			
March	11			29.25	14.62	75.2
		177	14.75			
April	12			29.83	14.91	80.5
		181	15.08			
May	13			30.50	15.25	85.2
		185	15.42			
June	15			31.17	15.58	96.3
		189	15.75			
July	15			31.67	15.83	94.7
		191	15.92			
Aug.	17			31.92	15.96	106.5
		192	16.00			
Sept.	18			32.17	16.08	111.9
		194	16.17			
Oct.	20			32.59	16.29	122.8
		197	16.42			
Nov.	22			33.09	16.54	133.0
		200	16.67			
Dec.	24			33.59	16.79	142.9
		203	16.92			
2010						
Jan.	12			34.25	17.12	70.1
		208	17.33			
Feb.	13			34.91	17.45	74.5
		211	17.58			
March	13			35.41	17.70	73.4
		214	17.83			
April	15			35.83	17.91	83.7
		216	18.00			
May	16			36.17	18.08	88.5
		218	18.17			
June	18			36.42	18.21	98.8
		219	18.25			
July	20					
Aug.	20					
Sept.	21					
Oct.	22					
Nov.	24					
Dec.	25					

COMPUTATION OF SEASONAL INDICES

	2007	2008	2009	2010	Median	Seasonal index
Jan.		69.5	70.0	70.1	70.0	70.28
Feb.		70.4	83.3	74.5	74.5	74.80
March		78.1	75.2	73.4	75.2	75.50
April		86.0	80.5	83.7	83.7	84.03
May		94.4	85.2	88.5	88.5	88.85
June		96.0	96.3	98.8	96.3	96.38
July	102.8	103.7	94.7		102.8	103.21
Aug.	108.8	103.7	106.5		106.5	106.93
Sept.	109.4	103.7	111.9		109.4	109.84
Oct.	116.3	111.3	122.8		116.3	116.77
Nov.	136.1	126.0	133.0		133.0	133.53
Dec.	137.5	140.0	142.9		140.0	140.56
					1,196.2	1,200.68*

It should be noted that there are only three values for each month since the moving average failed to provide averages for the first half of 2007 and the iast half of 2010. Median has been used to average the figures given for the individual months. The sum of 12 values obtained is 1,196.2. It is necessary, therefore, to make an adjustment so that the total is 1,200. The adjustment is done by multiplying the average (median) values by $\frac{1200}{1196.2} = 1.003$. The final result thus obtained gives us the seasonal indices. The interpretation of this index is very simple. Typical April sales are 84.03 per cent of those of the average month, typical November sales are 133.53 per cent of those of the average month, and so on.

Merits and Limitations of the Ratio-to-Moving Average Method

Merits. This method of measuring seasonal variation is considered to be the most satisfactory and as such is more widely used in practice than other methods. The index obtained by the ratio-to-moving average method ordinarily does not fluctuate so much as the index based on straight-line trends. Mathematical methods of avoiding the effects of the business cycle are not usually needed, for the 12-month moving average follows the cyclical course of the actual data quite closely. Therefore, the index ratios are often more representative of the data from which they are obtained than in the ratio-to-trend method. Also ratio-to-moving average method allows for greater flexibility.

Limitations. However, one drawback of this method is that seasonal indices cannot be obtained for each month for which data are available. When a 12-month moving average is taken, six months in the beginning and six months in the end are left out for which we cannot calculate seasonal indices.

4. Link Relatives Method

Among all the methods of measuring seasonal variation, link relatives method is the most difficult one. When this method is adopted, the following steps are taken to calculate the seasonal variation indices:

1. Calculate the link relatives of the seasonal figures. Link relatives are calculated by dividing the figure of each season** by the figure of immediately preceding season and multiplying it by 100.

*The difference is due to approximation.

**The word season refers to the time period. In case of monthly data, season would refer to a month and in case of quarterly data to a quarter.

$$\frac{\text{Current season's figure}}{\text{Previous season's figure}} \times 100$$

These percentages are called link relatives since they link each month (or quarter or other time period) to the preceding one.

2. Calculating the average of the link relatives for each season. While calculating average we might take arithmetic average but the median is probably better. The arithmetic average would give undue weight to extreme cases which were not due primarily to seasonal influences.

3. Convert these averages into chain relatives on the base of the first season.

4. Calculate the chain relatives of the first season on the base of the last season. There will be some difference between the chain relative of the first season and the chain relative calculated by the previous method, This difference will be due to the effect of long-term changes. It is, therefore, necessary to correct these chain relatives.

5. For correction, the chain relative of the first season calculated by first method is deducted from the chain relative (of the first season) calculated by the second method. The difference is divided by the number of seasons. The resulting figure multiplied by 1, 2, 3 (and so on) is deducted respectively from the chain relatives of the 2nd, 3rd, 4th (and so on) seasons. These are correct chain relatives.

6. Express the corrected chain relatives as percentage of their averages. These provide the required seasonal indices by the method of link relatives.

The following example will illustrate the process :

Illustration 16. Apply method of link relatives to the following data and calculate seasonal indices.

QUARTERLY FIGURES

Quarter	*2006*	*2007*	*2008*	*2009*	*2010*
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 METHOD OF LINK RELATIVES

Year	*Quarter* I	*II*	*III*	*IV*
2006	—	108.3	120.0	111.5
2007	62.1	146.3	106.3	86.9
2008	93.2	95.6	143.1	68.8
2009	112.5	80.6	129.3	113.3
2010	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 relative	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 relative	100	$108.28 - 1.675 = 106.605$	$131.73 - 3.35 = 128.38$	$123.64 - 5.025 = 118.615$
Seasonal Indices	100	$\frac{106.605}{113.4} \times 100 = 94.00$	$\frac{128.38}{113.4} \times 100 = 113.21$	$\frac{118.615}{113.4} \times 100 = 104.60$

In the above table the correction factor has been calculated as follows :
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)

$$\frac{86.35 \times 123.64}{100} = 106.76$$

The difference between these chain relatives = 106.76 – 100 = 6.76

$$\text{Difference per quarter} = \frac{6.7}{4} = 1.675$$

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.

Seasonal variation indices have been calculated as follows:

$$\frac{100 + 106.605 + 128.38 + 118.615}{4} = \frac{453.6}{4} = 113.4$$

$$\text{Seasonal variation index} = \frac{\text{Correct chain relatives} \times 100}{113.4}$$

Which Method to use. For different methods of measuring seasonal variations have been discussed above. The question now arises which method to adopt in a particular case. The choice will very much depend upon the nature of data and the object of investigation. Amongst all the methods, method of monthly averages is the simplest. But it is a crude method as it assumes that there is no trend component in time series. This method can be used only if seasonal rhythm dominates the data, and trend and cycle are negligible. The method of link relatives was widely used at one time, but its disadvantages seem to outweigh its advantages and it has currently fallen in some disfavour*. On weighing the merits and demerits of ratio-to-trend method and ratio-to-moving average method, one finds that ratio-to-moving average method has several advantages over the ratio-to-trend method. Hence, in general it may be said that because of theoretical and practical advantages, ratio-to-moving average method should be preferred to other methods.

Selecting the Period to Compute Seasonal Indices

In order to simplify the example a period of only 4-5 years was employed to compute the seasonal indexes. In actual fact it is suggested that many more years be included. It is because of the fact that a seasonal index based on a short period is often unduly affected by conditions prevailing during one phase of the business cycle or by powerful random influences. The period should encompass at least one and, if possible, several business cycles. The long span of years offers greater likelihood that irregular and cyclical forces will cancel out or at least have their influence minimized. Ten years is often viewed as a practical minimum. In selecting the period care should be taken to have the period begin and end at the same phase of the business cycle in order to avoid distortions that could result if more years of prosperity than of depression were included.

Average in Computing Seasonals

In each of the methods described for computing seasonal variations, the individual monthly averages were averaged in order to eliminate random influences and any remaining cyclical elements. In one of these example, the average selected was the arithmetic mean and in another the median. This poses the question of the relative merits of these or other averages for the purpose at hand. Since the mean is affected by every item in the series, it should be used when the number of years is large. However, when the period is shorter, the use of the mean is not recommended because extreme items, occasioned by the very random or cyclical factors that the calculation is designed to eliminate,

*Freund and Williams: ***Modern Business Statistics.***

distort its value. The median, on the other hand, is a positional average. As such it is not affected in any way by extreme values, but it may be unduly influenced by the inclusion or exclusion of a year or two in the calculation. A positional mean as suggested by Wessel and Willett avoids the disadvantages of both the mean and the median. It is computed by taking the arithmetic mean of the central items in the series. Suppose the following are the arranged ratios to the moving average for the month of May :

80 85 90 95 96 98 100 102 105 107 112 120

In this case of arithmetic of the mean the middle six items would be employed as the seasonal index. It is obvious that extreme items cannot influence this value and detail of position alone is not significant.

Eliminating Seasonal Influences

The seasonal influences may be removed from time-series data by dividing the actual values for each month by the seasonal index. This adjustment may symbolically be expressed as follows :

$$\frac{T \times S \times C \times I}{S} = T \times C \times I$$

Such adjustments are frequently made for series that manifest significant seasonals when these are being studied for other characteristics.

Illustration 17. From the following data of the production of *XYZ* Co. Ltd. for the year **2010**, remove the seasonal influence :

Month	Jan.	Feb.	March	April	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
Production	90	100	110	112	118	150	125	118	110	107	102	99
Seasonal Index	87.0	95.2	102.4	104	106	115	110	103.6	99	105	108	85

Month	*Production*	*Seasonal index*	*Adjusted production*
Jan.	90	87.0	103.45
Feb.	100	95.2	105.04
March	110	102.4	107.42
April	112	104.0	107.69
May	118	106.0	111.32
June	150	115.0	130.43
July	125	110.0	113.64
Aug.	118	103.6	113.90
Sept.	110	99.0	111.11
Oct.	107	105.0	101.90
Nov.	102	108.0	94.44
Dec.	99	85.0	116.47

Uses and Limitations of Seasonal Index

Uses. A seasonal index may be used either analytically or synthetically. Analytically a seasonal index is employed to adjust original data in order to yield deseasonalised data that permit the study of short-run fluctuations of a series not associated with seasonal variations. The procedure of adjusting data for seasonal variations is a simple one. It involves merely the division of each of the original observations by the appropriate seasonal index for that month, *i.e.*,

$$TCI = \frac{TSCI}{S}$$

Synthetically, seasonal index is extremely useful in planning sales of production for specific periods. For example, if a firm expects to sell Rs. 36,00,000 worth of goods during the forthcoming year, average monthly sales of Rs. 3,00,000 is anticipated. If, however, the volume of sales is subjected to seasonal fluctuations, the actual monthly values will deviate significantly from this average. Should the seasonal index for September be 120, the firm can expect sales of Rs. 3,60,000 during that month, in comparison of an index of 90 for July would lead them to anticipate sales only Rs. 2,70,000.

Forecasts for future periods are frequently made by combining what is known about trend and seasonal elements. First, the trend ordinate for a given month is computed. Then this ordinate is multiplied by the seasonal index for that month. For example, if the equation for the trend of a company's sales is $y = 30,000 + 250x$, where x represents 1 month and has a value of 0 in December 2009, and the seasonal index for May is +10 the sales for October 2010 may be estimated as follows. In October 2010 the value of x will be +10 and the trend ordinate will, therefore, be 32,500. When this figure is multiplied by 1.1, the estimated sales will be Rs. 35,750. Though this type of forecast ignores cyclical and random influences it is found to be highly useful in practice. By special price and advertising policies a producer confronted with a strong seasonal demand for his product may try to stabilise sales by encouraging off-season consumption.

The most promising solution for seasonality is diversification. It benefits not only the firm but also society at large. Whenever diversification is possible, real costs of seasonal variations can be reduced or even eliminated. Diversification involves the development of production lines having complementary seasonal movements. While some expand seasonally, others contract. Consequently, labour and facilities can be transferred from one line to another as seasonal changes take place. However, diversification is possible only in those line of production that have approximately the same labour and equipment requirements.

Limitations. While making use of seasonal indexes in business and economic problems, the following precautions should be kept in mind:

1. No technique can measure seasonal variations precisely. The various methods of measuring seasonal variations are based on rather unrealistic assumption that the seasonals are changing in some regular and systematic pattern.

2. In developing seasonal index we obtain a series of measures—measure for January, measure for February, and so forth—each of which generally differs from 100. However, we must remember that these measures are only rough estimates. Hence, if we obtain a seasonal index in which the values are all close to 100—for example, if the index values for the consecutive months are 102, 99, 103, 98, etc., it may well be that no real monthly seasonal variation exists in the series and that the small differences from 100 are only due to random influences or imperfect measurement.

3. Even if the computer index of seasonal variation indicates a pronounced pattern, it may have no significance for a particular year. It must be remembered that any seasonal index of the type we have described represents an average pattern during a number of years. If the pattern of seasonal variation in the series is not a stable one, any average pattern may be a poor representation of the actual seasonal variation taking place during a given year.

III. MEASUREMENT OF CYCLICAL VARIATIONS

Business cycles are perhaps the most important type of fluctuation in economic data. Certainly they have received a lot of attention in economic literature. Despite the importance of business cycles, they are most difficult type of economic fluctuation to measure. This is because successive cycles vary widely in timing, amplitude and pattern, and because the cyclical rhythm is inextricably mixed with irregular factors. Because of these reasons it is impossible to construct meaningful typical cycle indexes of curves similar to those that have been developed for trends and seasonals. The various methods* used for measuring cyclical variations are :

1. Residual method,
2. Reference cycle analysis method,
3. Direct method, and
4. Harmonic analysis method.

Only the first two methods which are in popular use are discussed below.

Residual Method

Among all the methods of arriving at estimates of the cyclical movements of time series, the residual method is most commonly used. This method consists of eliminating seasonal variation and trend, thus obtaining the cyclical irregular movements. Symbolically,

$$\frac{T \times S \times C \times I}{S} = T \times C \times I$$

and $$\frac{T \times C \times I}{T} = C \times I$$

The data are usually smoothed in order to obtain the cyclical movements, which are sometimes termed the *cyclical relatives,* since they are always percentages. It is because cyclical, irregular or the cyclical movements remain as residuals that this procedure is referred to as the *residual* method.

Limitations of the Residual Method. If the trend ordinates perfectly depicted the pattern of secular change and if the seasonal index exactly reflected seasonal influence, the residual method would leave values reflecting only cyclical and irregular influences. Because such perfection is rarely encountered, the computed values almost always contain some trend and seasonal elements. This condition will be more or less serious depending on how well or poorly the trend line and the seasonal index represent secular and seasonal forces. If a straight line trend is employed to describe an essentially curvilinear secular movement, figures presumably adjusted for trend will be grossly distorted. The distortion would also occur if the seasonal index were not descriptive of the seasonal pattern at the time in question. Thus the residual method is based on the assumption that trend and seasonals can be accurately measured and therefore be removed at least in large part.*

Reference Cycle Analysis or the National Bureau Method

The National Bureau of Economic Research has developed a different method of analysing cyclical variations which it has used in the study of more than 1,000 specific time series. The method is of value in analysing past cycles only. The National Bureau procedure aims to answer two sets of questions:

(1) Is there in a given series a pattern of change that repeats itself (with more or less variation) in successive cycles in business at large? If so, what are its characteristics?

(2) Is there in a given series a wave movement peculiar to that series? If so, what are its characteristics?

The questions under (1) are concerned with the behaviour of individual series during successive waves of expansion and contraction in the *general economic,* those under (2) relate to periodic or semi-periodic fluctuations in *individual series.* A procedure involving 'reference dates' has been designed by the National Bureau of Economic Research as a device which allows one not only to compare each series with a standard set of dates and to observe the behaviour of individual series during expansion and contraction of general business but also to compare the results for the various individual series.

*For detail refer to Croxton and Cowden: ***Applied General Statistics.***

*Wessel and Willett : ***Statistics as Applied to Economics and Business.***

The first step is the selection of the reference dates which are the dates of the peaks and troughs of business cycles. This reference dates which cover a duration of over one year and not over ten or twelve years were chosen after examination of large number of economic time series and after study of the "contemporary" reports of observers of the business scene.

The next step consists of processing the data of the individual series in order to obtain a cyclical pattern for each series for the period between each two successive reference troughs. Each period is the same for all series, enabling one to compare the results for the various series. The processing of each series proceeds as follows:

(1) The data are adjusted for seasonal variation.*

(2) The seasonally adjusted data are divided into reference cycle segments, these segments corresponding to the intervals between adjacent reference troughs.

(3) For each segment, the monthly values are expressed as percentages of the average of the values in the segment. These are "reference cycle relatives". As a result of this step, all series, no matter what the original unit, are in percentage form. This step eliminates inter-cycle trend, since the average of the relatives for each cycle is 100, but it does not eliminate intra-cycle trend. The inclusion of intra-cycle trend is regarded as desirable, since it "helps to reveal and to explain what happens during business cycles".

(4) Each reference cycle segment is broken into nine stages, to correspond to the same nine stages in the business cycle, and the reference cycles relatives are averaged for each of nine stages. The nine stages are identified as follows :

(*i*) The 3 months centered on the initial trough.

(*ii*) The first third of the expansion period.

(*iii*) The second third of the expansion period.

(*iv*) The last third of the expansion period.

(*v*) The 3 months centered on the peak.

(*vi*) The first third of the contraction period.

(*vii*) The second third of the contraction period.

(*viii*) The last third of the contraction period.

(*ix*) The 3 months centered on the terminal trough.

The nine-stage average for each reference cycle segment serves to reduce the erratic movement in a series and gives a reference cycle pattern for the particular series under consideration.

Although the National Bureau method of cycle analysis may seem more complicated and cumbersome than the residual technique, it has proved to be the simplest and most accurate way of comparing the cyclical variations of individual series with those of general business. In addition, it is free of errors that might be introduced were secular trend improperly estimated. The latter advantage is indeed significant when series whose trend patterns are not clear are under analysis. Its principal shortcoming is found in the fact that, because no cycle can be studies in this way until it is completed, the method cannot be applied to current data.

MEASUREMENT OF IRREGULAR VARIATIONS

The irregular component in a time series represents the residue of fluctuations after trend cyclical and seasonal movements have been accounted for. Thus, if the original data is divided by *T, S* and *C;* we

*Trend influences are not removed under this method.

get I, *i.e.*, $\left(\frac{TSCI}{TSC} = I\right)$. In practice, the cycle itself is so erratic and is so interwoven with irregular movements that it is impossible to separate them. In the analysis of a time series into its component fluctuations, therefore, trend and seasonal movements are usually measured directly, while cyclical and irregular fluctuations are left altogether after the other elements have been removed.

Selecting the Appropriate Forecasting Technique

Numerous forecasting techniques with varying degrees of complexity have been devised during the last few decades. The problem very often is that of selecting the best one in a prticular situation. The following are some of the important factors that affect the decision about the appropriate forecasting technique :

(*i*) **The Time Horizon.** The period of time over which a decision will have an impact and for which the manager must plan clearly affects the choice of forecasting techniques. Time horizons are generally divided into four heads—immediate term (less than one month) ; short term (one to three months), medium term (three months to two years) and long term (more than two years). Though the exact length of time that may describe each of these categories may vary from company to company, some set of guidelines is necessary so that the forecast will be appropriate for the planning horizon used by the decision maker. Some techniques are appropriate for forecasting only one or two periods in advance whereas others can be used for several periods. Generally speaking, quantitative methods of forecasting are more appropriate for intermediate and shorter term whereas qualitative methods of forecasting are used much more for longer-term forecasts.

The time allowed for preparing the forecast must also be considered. Some methods are much time-consuming and may have to be ignored on the ground only even though they may ensure better results.

(*ii*) **Number of items.** In a situation in which only a single item is being forecast the rules used in preparing that forecast can be much more detailed and complex. But if forecasts are to be made for hundreds or thousands of products, it would be better to develop simple decision rules that can be applied mechanically to each of the items.

(*iii*) **Details required.** In selecting a forecasting technique for a specific situation, one must be aware of the level of detail that will be required for that forecast to be useful in making decisions.

(*iv*) **The pattern of data.** The forecasting methods generally involve an assumption of the type of pattern found in the data. For example, some series depict a seasonal as well as a trend pattern whereas others consist of an average value with random fluctuation surrounding it. Since different forecasting methods vary in their ability to identify different patterns, it is important to match the presumed pattern in the data with the appropriate technique.

(*v*) **Type of model.** In addition to assuming some basic underlying pattern in the data, more forecasting methods also assume some model of the situation being forecast. The model may be a casual model that represents the forecast as being dependent on the occurrence of a number of different events and one may have to use regression or correlation analysis or it may be a series in which time is viewed as the important element in determining changes in the pattern or it may be a mixed model combining in itself a number of different models. The assumptions underlying different models are different and the capabilities of different models in various decision-making situations also vary.

(*vi*) **Cost.** Cost is a very important factor which has to be taken into account while selecting appropriate forecasting technique. The variation in cost affects the attractiveness of different methods for different situations. Generally, four elements of cost that should be considered in the application of forecasting procedure are development, storage, actual operation and opportunity in terms of other techniques.

(*vii*) **Accuracy desired.** The choice of appropriate method of forecasting would also depend upon the accuracy desired. In some situations, variation anywhere between plus and minus 10% may be sufficient for the purpose whereas in other cases a variation as low as 2% to 3% may spell disaster for the company.

(*viii*) **Easc of Application.** The forecasting technique may vary from simple to highly complicated one. Since the manager is held responsible for his decisions, he should not base them on forecasts that he does not understand or in which he has no confidence. Hence, in addition to meeting the requirements of the situation, the forecasting technique must fit with the particular manager who will use the forecast. The manager will have to use his own judgment in order to be able to evaluate and mark selections in his own situation. If a manager can use a more straightforward and less expensive forecasting method rather than the most sophisticated technique available and still achieve the required level of accuracy, he should do so.

It should be borne in mind that a continuous review process must be established in order that the forecast may be compared with the actual results and improvements, and perhaps changes in the technique itself can be made. If no evaluation is made, the manager may feel disappointed with the technique used and may begin to discount it. However, at this stage there may be no record of actual *versus* forecast and other evaluation measures and thus it may be difficult, if not impossible, to determine where improvements are required.

Cautions while using Forecasting Techniques

Forecasting business conditions is a complex task which cannot be accomplished with exactness. The economic, social and political forces which shape the future are many and varied; their relative importance change almost constantly. It is obvious, therefore, that statistical methods cannot claim to be able to make the uncertain future certain—after all, forecasters are not prophets. It does not follow from this disclaimer that statistical methods have nothing to contribute to business forecasting. The choice is not between forecasting and not forecasting, because the lack of a forecast implies a dangerous type of forecast, the mere warning of a possibility of a change is better than no warning at all, as is wisely said "*forewarned is forearmed*".

No matter what method of forecasting is used it is essential that the forecasts be checked by the judgment of individual who is familiar with the business. While it is true that the use of statistical data is an attempt to substitute facts for subjective judgment it does not mean that knowledge gained through experience in a given situation should be ignored in favour of quantitative data. It is particularly important to take into consideration any specific plans of the business that might affect the pattern of sales in relation to indicators used for forecasting. More successful forecasting will result by combining with statistical forecasting the judgment and knowledge of current business trends.

Also it is important to emphasise that any forecast should be reviewed frequently and revised in the light of the most recent information. Forecasting is not a one-shot operation. To be effective it requires continuous attention. Unanticipated developments will often change our picture of the future, or at least clarify it. In terms of any original decisions and actions that have been taken, this rule implies continuous modification where possible. The technique of flexible budgets has been developed to permit the revision of the budget estimates, and everyone dealing with forecasts should be alert to the need for constantly checking to see if anything has happened to change the outlook. Keeping accurately informed about the current level of business is probably the simplest insurance that can be secured against making wrong decisions regarding the future.

Last but not the least it should be kept in mind that as is the case with any method employed to forecast the future, the prediction is no better than the data used no matter how elaborate or complicated the mathematical procedure. As one expert has stated, "It is far better to be approximately correct than precisely wrong." Too often the mathematically-oriented person forgets this point in his zeal to apply his newly discovered tools.

MISCELLANEOUS ILLUSTRATIONS

Illustration 18. Fit a straight line trend to the following time series data :

Year	:	2006	2007	2008	2009	2010
Sale of sugar (in m tonnes)	:	80	90	92	83	94

Eliminate trend from the series. What components are left over ?

Solution. FITTING STRAIGHT LINE TREND

Year	*Sales Y*	*X*	*XY*	X^2	Y_c	$(Y-Y_c)$
2006	80	−2	−160	4	83.6	− 3.6
2007	90	−1	−90	1	85.7	+ 4.3
2008	92	0	0	0	87.8	+ 4.2
2009	83	+1	+83	1	89.9	− 6.9
2010	94	+2	+188	4	92.0	+ 2.0
$N = 5$	$\Sigma Y = 439$	$\Sigma X = 0$	$\Sigma XY = 21$	$\Sigma X^2 = 10$		

$$Y_c = a + bX$$

$$a = \frac{\Sigma Y}{N} = \frac{439}{5} = 87.8; \quad b = \frac{\Sigma XY}{\Sigma X^2} = \frac{21}{10} = 2.1$$

Hence $Y = 87.8 + 2.1X$

$Y_{2006} = 87.8 + 2.1\,(-2) = 87.8 - 4.2 = 83.6$; and similarly other trend values are computed.

After eliminating trend what is left is the effect of seasonal, cyclical and irregular variations.

Illustration 19. Below are given the figures of production (in million tonnes) of a cement factory :

Year	:	2001	2003	2004	2005	2006	2007	2010
Production (in m. tonnes)	:	77	88	94	85	91	98	90

(*i*) Fit a straight line trend by the 'least squares method' and tabulate the trend values.
(*ii*) Eliminate the trend. What components of the Time Series are thus left over ?
(*iii*) What is monthly increase in the production of cement ?

Solution. FITTING STRAIGHT LINE TREND BY THE METHOD OF LEAST SQUARES

Year	*Production (in m. tonnes) Y*	*Deviations from 2005 X*	*XY*	X^2	*Trend values* Y_c
2001	77	−4	−308	16	83.299
2003	88	−2	−176	4	86.051
2004	94	−1	−94	1	87.427
2005	85	0	0	0	88.803
2006	91	+1	+91	1	90.179
2007	98	+2	+196	4	91.555
2010	90	+5	+450	25	95.683
$N = 7$	$\Sigma Y = 623$	$\Sigma X = 1$	$\Sigma XY = 159$	$\Sigma X^2 = 51$	

(*i*) The equation of the straight line trend is $Y_c = a + bX$. Since ΣX is not zero, we will solve the two normal equations

$$\Sigma Y = Na + b\Sigma X$$
$$\Sigma XY = a\Sigma X + b\Sigma X^2$$
$$623 = 7a + b$$
$$159 = a + 51b$$

Multiplying eq. (*ii*) by 7 and subtracting from eq. (*i*), we get $b = 1.376$

Putting the value of b in eq. (*i*)

$7a + 1.376 = 623$ or $7a = 621.624$ or $a = 88.803$

Hence $Y = 88.803 + 1.376X$ is the equation of the straight line

$Y_{2001} = 88.803 + 1.376\,(-4) = 88.803 - 5.504 = 83.299$

$Y_{2003} = 88.803 + 1.376\,(-2) = 88.803 - 2.752 = 86.051$

$Y_{2004} = 88.803 + 1.376\,(-1) = 88.803 - 1.376 = 87.427$

$Y_{2005} = 88.803$

$Y_{2006} = 88.803 + 1.376\,(+1) = 88.803 + 1.376 = 90.179$

$Y_{2007} = 88.803 + 1.376\,(+2) = 88.803 + 2.752 = 91.555$

$Y_{2010} = 88.803 + 1.376\,(+5) = 88.803 + 6.88 = 95.683$

(*ii*) After eliminating the trend we are left with seasonal, cyclical and irregular variations.

(*iii*) Monthly increase in the production of cement shall be given by

$$\frac{b}{12} = \frac{1.376}{12} = 0.115$$

Illustration 20. The sale of a commodity (in tonnes) varied from January 2010 to December, 2010 in the following manner :

280	300	280	280	270	240
230	230	220	200	210	200

Fit a trend line by the method of semi-averages.

Solution.

FITTING TREND LINE BY METHOD OF SEMI-AVERAGES

Month	*Sales (in tonnes)*		*Month*	*Sales (in tonnes)*	
January	280		July	230	
February	300	1,650 (Total)	August	230	1,290 (Total)
March	280	of first six	September	220	of last six
April	280	months	October	200	months
May	270		November	210	
June	240		December	200	

$$\text{Average of the first half} = \frac{1650}{6} = 275 \text{ tonnes.}$$

$$\text{Average of the second half} = \frac{1290}{6} = 215 \text{ tonnes.}$$

These two figures, namely, 275 and 215, shall be plotted at the middle of their respective periods, *i.e.*, middle of March-April 2010 and that of September-October 2010. By joining these two points we get a trend line which describes the given data.

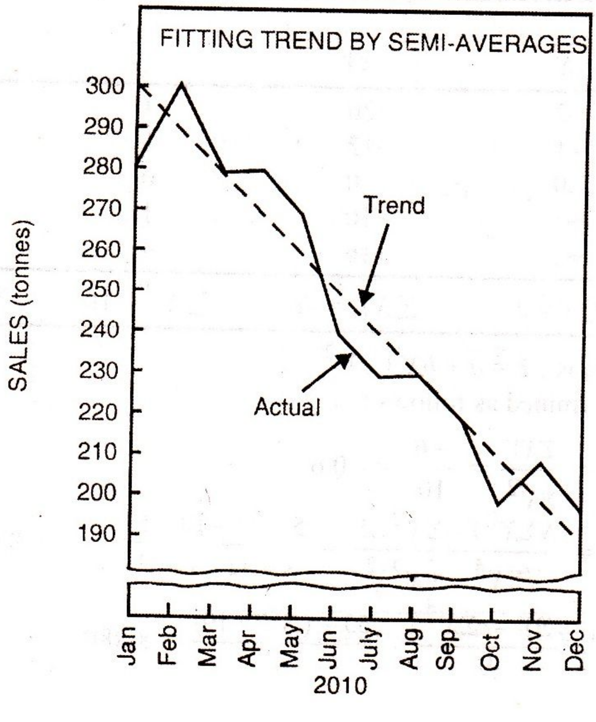


Illustration 21. (*i*) Given the trend equation :

$$Y_C = 35 + 5X + 3X^2$$

(Origin : 2004, X unit = 1 year) change the origin of the equation to 2010.

(*ii*) Given the equation $Y_C = 10(1.5)^X$

(Origin : 2004, X unit = 1 year)

Shift the origin forward by two years.

(*iii*) The trend of the annual sales of an Aluminium Company is described by the following equation :

$$Y_C = 12 + 0.7X$$

(Origin : 2004, X unit = 1year and Y unit = annual production)

Shift the origin to January, 2010 and write the equation on monthly basis.

Solution. (*i*) To shift the origin from 2004 to 2010, *i.e.*, 6 years forward, put $X = 6$ in the equation.

$$Y_C = 35 + 5(6) + 3(6)^2 = 35 + 30 + 108 = 173$$

The new trend equation is

$$Y_C = 173 + 5X + 3X^2$$

(Origin : 2004, X unit = 1 year)

(*ii*) For shifting the origin forward by two years, we put $X + 2$ instead of X in the given equation. The new trend equation is

$$Y_C = 10(1.5)^{X+2}$$
$$= 10\,(1.5)^2\,(1.5)^X = 22.5\,(1.5)^X$$

(*iii*) To shift the origin to January, 2010, we should subtract $1/2b$ from trend value of July, 2010.

$$1/2b = 0.7/2 = 0.35; \text{ therefore, } a = 12 - 0.35 = 11.65$$

The new trend equation is

$$Y_C = 11.65 + 0.7X$$

(Origin : January, 2010, X unit = 1 year)

To obtain the equation on monthly basis, we divide the value of a by 12 and value of b by 144. The new trend equation on monthly basis can be written as :

$$Y_C = \frac{11.65}{12} + \frac{0.7}{144}\,X = 0.971 + .0049\,X$$

(Origin : January 2010, X unit = 1 month).

Illustration 22. Fit a parabolic curve of the second degree to the data given below and estimate the value for 2012 and comment on it.

Year	:	2006	2007	2008	2009	2010
Sales (in '000 Rs.)	:	10	12	13	10	8

Solution. COMPUTATION OF SECOND DEGREE PARABOLA

Year	*Sales ('000 Rs.)* Y	X	XY	X^2	X^2Y	X^4
2006	10	–2	–20	4	40	16
2007	12	–1	–12	1	12	1
2008	13	0	0	0	0	0
2009	10	+1	+10	1	10	1
2010	8	+2	+16	4	32	16
$N = 5$	$\Sigma Y = 53$	$\Sigma X = 0$	$\Sigma XY = -6$	$\Sigma X^2 = 10$	$\Sigma X^2Y = 94$	$\Sigma X^4 = 34$

The equation of the second parabola is : $Y = a + bX + cX^2$

The values of a, b and c can be determined as follows :

$$b = \frac{\Sigma XY}{\Sigma X^2} = \frac{-6}{10} = -0.6$$

$$c = \frac{N\Sigma X^2Y - \Sigma X^2 \Sigma Y}{N\Sigma X^4 - (X^2)^2} = \frac{5 \times 94 - 10 \times 53}{5 \times 34 - (10)^2} = -0.857$$

$$a = \frac{\Sigma Y - c\Sigma X^2}{N} = \frac{53 - (.857 \times 10)}{5} = 8.886$$

Thus, $Y = 8.886 - 0.6X - 0.857X^2$

For 2012, X would be 4 $Y_{2012} = 8.886 - 0.6(4) - 0.857(4)^2$

$= 8.886 - 2.4 - 13.712 = -7.226.$

The expected sales for 2012 comes to be negative. The second degree parabola does not seem to describe the data well.

Illustration 23. Assuming that trend is absent, determine if there is any seasonality in the data given below :

Year	*1st Quarter*	*2nd Quarter*	*3rd Quarter*	*4th Quarter*
2007	3.7	4.1	3.3	3.5
2008	3.7	3.9	3.6	3.6
2009	4.0	4.1	3.3	3.1
2010	3.3	4.4	4.0	4.0

What are the seasonal indices for various quarters ?

Solution. COMPUTATION OF SEASONAL INDICES

Year	*1st Quarter*	*2nd Quarter*	*3rd Quarter*	*4th Quarter*
2007	3.7	4.1	3.3	3.5
2008	3.7	3.9	3.6	3.6
2009	4.0	4.1	3.3	3.1
2010	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	98.7	110.8	95.3	95.3

Notes for calculating seasonal index

The average of averages $= \dfrac{3.675+4.125+3.55+3.55}{4} = \dfrac{14.900}{4} = 3.725$

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

$$\text{Seasonal Index for first quarter} : \frac{3.675}{3.725} \times 100 = 98.7$$

$$\text{Seasonal Index for second quarter} : \frac{4.125}{3.725} \times 100 = 110.7$$

$$\text{Seasonal Index for the third quarter} : \frac{3.55}{3.725} \times 100 = 95.3$$

$$\text{Seasonal Index for the fourth quarter} : \frac{3.55}{3.725} \times 100 = 95.3$$

Illustration 24. Given below are the figures of production of a sugar factory :

Year :	2004	2005	2006	2007	2008	2009	2010
Production (m. tonnes) :	40	45	46	42	47	49	46

Fit a straight line trend by the method of least squares and estimate the value for 2014

Solution. FITTING STRAIGHT LINE BY METHOD OF LEAST SQUARES

Year	*Production (m. tonnes)* Y	*Deviations from 2007* X	XY	X^2
2004	40	–3	– 120	9
2005	45	–2	– 90	4
2006	46	–1	– 46	1
2007	42	0	0	0
2008	47	+1	+ 47	1
2009	49	+2	+ 98	4
2010	46	+3	+ 138	9
$N = 7$	$\Sigma Y = 315$	$\Sigma X = 0$	$\Sigma XY = 27$	$\Sigma X^2 = 28$

$$Y_c = a + bX$$

$$a = \frac{\Sigma Y}{N} = \frac{315}{7} = 45; \; b = \frac{\Sigma XY}{\Sigma X^2} = \frac{27}{28} = 0.964$$

$$Y = 45 + 0.964X$$

$$Y_{2014} = 45 + .964\,(7) = 45 + 6.748 = 51.748$$

Thus, the estimated production for 2014 is 51.748 m. tonnes.

Illustration 25. Fit a straight line trend by the method of least squares to the data given below :

Year :	2004	2005	2006	2007	2008	2009	2010
Sales (m. tonnes) :	9	11	13	12	14	15	17

Estimate the likely sales for the year 2013.

Solution. CALCULATION OF STRAIGHT LINE TREND BY METHOD OF LEAST SQUARES

Year	*Sales (m. tonnes)* Y	X	XY	X^2
2004	9	–3	–27	9
2005	11	–2	–22	4
2006	13	–1	–13	1
2007	12	0	0	0
2008	14	+1	+14	1
2009	15	+2	+30	4
2010	17	+3	+51	9
$N = 7$	$\Sigma Y = 91$	$\Sigma X = 0$	$\Sigma XY = 33$	$\Sigma X^2 = 28$

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

Since $\Sigma X = 0, \; a = \frac{\Sigma Y}{N} = \frac{91}{7} = 13 \; ; \; b = \frac{\Sigma XY}{\Sigma X^2} = \frac{33}{28} = 1.179$

Hence $Y = 13 + 1.179X$; For 2011, X would be + 6.

$$Y_{2013} = 13 + 1.179(6) = 13 + 7.074 = 20.074$$

Therefore, the estimated sales for the year 2013 is 20.074 m. tonnes.

Illustration 26. The seasonal indices of a commodity manufactured by a company for four quarters of a year are respectively 100, 90, 80, and 130. If the total sale in the first quarter is worth Rs. 25,000, how much worth of sale is expected during the whole year ?

Solution. EXPECTED SALES IN VARIOUS QUARTERS

Quarter	*Seasonal index*	*Estimated sales (Rs.)*
I	100	25,000
II	90	22,500
III	80	20,000
IV	130	32,500

$$\text{Estimated sales for 2nd Qtr.} = \frac{\text{Figure for 1st Qtr.}}{\text{S.I. for Ist Qtr.}} \times \text{S.I. for 2nd Qtr.}$$

$$= \frac{25{,}000 \times 90}{100} = 22{,}500$$

$$\text{Estimated sales for 3rd Qtr.} = \frac{\text{Figure for 1st Qtr.}}{\text{S.I. for Ist Qtr.}} \times \text{S.I. for 3rd Qtr.}$$

$$= \frac{25{,}000 \times 80}{100} = 20{,}000$$

$$\text{Estimated sales for 4th Qtr.} = \frac{25{,}000 \times 130}{100} = 32{,}500.$$

Illustration 27. Fit a straight line trend by the method of least squares to the following data on sales (Rs. in lakh) for the period 2003-2010.

Year :	2003	2004	2005	2006	2007	2008	2009	2010
Sales (Rs. lakh) :	76	80	130	144	138	120	174	190

Also :

(*a*) Calculate the trend values from 2003 to 2010.

(*b*) What will be predicted sales for 2013, assuming that the same rate of change continues.

Solution. FITTING STRAIGHT LINE TREND BY THE METHOD OF LEAST SQUARES

Year	*Sales (Rs. Lakh) Y*	*Deviations from 2006.5 × 2*	*X*	*XY*	*X*^{2}	Y_c
2003	76	–3.5	–7	– 532	49	80.169
2004	80	–2.5	–5	– 400	25	94.835
2005	130	–1.5	–3	–390	9	109.501
2006	144	–0.5	–1	–144	1	124.167
2007	138	+0.5	+1	+138	1	138.833
2008	120	+1.5	+3	+360	9	153.499
2009	174	+2.5	+5	+870	25	168.165
2010	190	+3.5	+7	+1330	49	182.831
$N = 8$	$\Sigma Y = 1052$		$\Sigma X = 0$	$\Sigma XY = 1232$	$\Sigma X^2 = 168$	$\Sigma Y_c = 1052$

Since $\Sigma X = 0$, $a = \frac{\Sigma Y}{N} = \frac{1052}{8} = 131.5$; and $b = \frac{\Sigma XY}{\Sigma X^2} = \frac{1232}{168} = 7.333$*

Hence $Y = 131.5 + 7.333X$

For the year 2013, *X* would be 13, therefore $Y_{2013} = 131.5 + 7.333\ (13) = 131.5 + 95.329 = 226.829 =$ Rs. 2,26,829.

Illustration 28. Fit a straight line trend for the following data and find the trend values. Estimate the sales for 2016.

Year :	2004	2005	2006	2007	2008	2009	2010
Sales (Rs. lakh) :	33	35	60	67	68	82	90

Solution. FITTING STRAIGHT LINE TREND

Year	*Sales Y*	*X*	*XY*	*X*^{2}	Y_c
2004	33	–3	–99	9	32.893
2005	35	–2	–70	4	42.643
2006	60	–1	–60	1	52.393
2007	67	0	0	0	62.143
2008	68	+1	+68	1	71.893
2009	82	+2	+164	4	81.643
2010	90	+3	+270	9	91.393
$N = 7$	$\Sigma Y = 435$	$\Sigma X = 0$	$\Sigma XY = 273$	$\Sigma X^2 = 28$	$\Sigma Y_c = 435$

The equation of the straight line trend is :

$$Y = a + bX$$

$$a = \frac{\Sigma Y}{N} = \frac{435}{7} = 62.143\ ; \quad b = \frac{\Sigma XY}{X^2} = \frac{273}{28} = 9.75$$

Hence,

$$Y = 62.143 + 9.75X$$

$$Y_{2004} = 62.143 + 9.75(-3) = 62.143 - 29.25 = 32.893$$

$$Y_{2005} = 62.143 + 9.75(-2) = 42.643, \text{ etc.}$$

For 2016, *X* shall be + 9; $Y_{2016} = 62.143 + 9.75(9) = 62.143 + 87.75 = 149.893$

The estimated sales for the year 2016 is Rs. 149.893 lakh.

*The calculated value of *b*, *i.e.*, 7.333 is multiplied by 2, *i.e.* (7.333 × 2 = 14.666) to obtain yearly change.

Illustration 29. Fit a straight line trend by the method of least squares to the following data :

Year :	2004	2005	2006	2007	2008	2009	2010
Production of steel (m. tonnes) :	12	10	14	11	13	15	16

Calculate the trend values and estimate the likely production for the year 2017. Interpret the values of a and b.

Solution.

CALCULATION OF TREND VALUES

Year	*Production (m. tonnes)* Y	*Deviations from 2007* X	XY	X^2	*Trend values* Y_c
2004	12	–3	–36	9	10.75
2005	10	–2	–20	4	11.50
2006	14	–1	–14	1	12.25
2007	11	0	0	0	13.00
2008	13	+1	+13	1	13.75
2009	15	+2	+30	4	14.50
2010	16	+3	+48	9	15.25
$N = 7$	$\Sigma Y = 91$	$\Sigma X = 0$	$\Sigma XY = 21$	$\Sigma X^2 = 28$	$\Sigma Y_c = 91$

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

Since $\Sigma X = 0,\ a = \dfrac{\Sigma Y}{N} = \dfrac{91}{7} = 13\ ;\ b = \dfrac{\Sigma XY}{X^2} = \dfrac{21}{28} = 0.75$

Hence, $Y = 13 + 0.75X$; For 2004 : $X = -3$

Estimated production for the year 2004 :

$$Y = 13 + 0.75(-3) = 13 - 2.25 = 10.75$$

Estimated value for 2005 $= 13 + 0.75(-2) = 11.5$, etc.

For the year 2017, X would be +10.

$$Y_{2017} = 13 + 0.75(10) = 13 + 7.5 = 20.5$$

Hence, the estimated production of steel for the year 2017 is 20.5 m. tonnes.

Illustration 30. Fit a second degree parabola, $Y = a + bX + cX^2$, to the following population data of a city :

Year :	2002	2003	2004	2005	2006	2007	2008	2009	2010
Population (in lakh) :	5	6	6	7	7	8	9	10	10

(Take the year 2006 as the working origin)

Solution.

FITTING OF SECOND DEGREE PARABOLA

Year	*Population* Y	*Deviations from 2006* X	XY	X^2	X^2Y	X^3	X^4
2002	5	–4	–20	16	80	–64	256
2003	6	–3	–18	9	54	–27	81
2004	6	–2	–12	4	24	–8	16
2005	7	–1	–7	1	7	–1	1
2006	7	0	0	0	0	0	0
2007	8	+1	+8	1	8	+1	1
2008	9	+2	+18	4	36	+8	16
2009	10	+3	+30	9	90	+27	81
2010	10	+4	+40	16	160	+64	256
$N = 9$	$\Sigma Y = 68$	$\Sigma X = 0$	$\Sigma XY = 39$	$\Sigma X^2 = 60$	$\Sigma X^2Y = 459$	$\Sigma X^3 = 0$	$\Sigma X^4 = 708$

$$Y = a + bX + cX^2$$

Since $\Sigma X = 0$, the three normal equations would be

$$\Sigma Y = Na + c\Sigma X^2$$
$$\Sigma XY = b\Sigma X^2$$
$$\Sigma X^2Y = a\Sigma X^2 + c\Sigma X^4$$

Substituting the values from the table

$$68 = 9a + 60c \quad ...(i)$$
$$39 = 60b \quad ...(ii)$$
$$459 = 60a + 708c \quad ...(iii)$$

From eq. (*ii*) $60b = 39$ or $b = 0.65$

Multiplying eq. (*i*) by 20 and eq. (*iii*) by 3

$$1360 = 180a + 1200c$$
$$1377 = 180a + 2124c$$
$$- \quad - \quad -$$
$$17 = 924c \text{ or } c = 0.0184$$

Substituting the value of c in eq. (*i*)

$$68 = 9a + 60\ (0.0184) \text{ or } 9a + 1.104 = 68$$
$$9a = 66.896 \text{ or } a = 7.433$$

Hence $Y = 7.433 + 0.65X + 0.0184X^2$ is the required equation.

Illustration 31. The time series given below shows the number of T.V. sold by a company since 2001.

Years	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
T.V. sold ('000)	42	50	61	75	92	111	120	127	140	138

Find the linear equation that describes the trend in the number of T.V. sold. Also estimate the sale of T.V. in 2012.

Solution. FITTING LINEAR EQUATION

Year	*T.V. Sold ('000)*	*Taking deviations from 2005.5*	*Multiplying deviations by 2* X	XY	X
2001	42	–4.5	–9	–378	81
2002	50	–3.5	–7	–350	49
2003	61	–2.5	–5	–305	25
2004	75	–1.5	–3	–225	9
2005	92	–0.5	–1	–92	1
2006	111	+0.5	+1	+111	1
2007	120	+1.5	+3	+360	9
2008	127	+2.5	+5	+635	25
2009	140	+3.5	+7	+980	49
2010	138	+4.5	+9	+1242	81
$N = 10$	$\Sigma Y = 956$		$\Sigma X = 0$	$\Sigma XY = 1978$	$\Sigma X^2 = 330$

The linear equation would be of the form :

$$Y = a + bX$$

The two normal equations shall be

$$\Sigma Y = Na + b\Sigma X$$
$$\Sigma XY = a\Sigma X + b\Sigma X^2$$

Since $\Sigma X = 0$

$$a = \frac{\Sigma Y}{N} = \frac{956}{10} = 95.6; \text{ and } b = \frac{\Sigma XY}{\Sigma X^2} = \frac{1978}{330} = 5.994$$

Thus the linear equation is

$$Y_c = 95.6 + 5.994$$

For 2012, X would be +13.

$$Y_{2012} = 95.6 + 5.994\ (13) = 95.6 + 77.92 = 173.52 \text{ thousand}$$

Hence the expected number of T.V. that would be sold in 2012 shall be 1,73,520.

PROBLEMS

1-A : Answer the following questions, each question carries **one** marks:

(*i*) What do you mean by Time series ? Explain the objectives of the analysis of a time series. *(MBA, UPTech. Univ., 2007)*

(*ii*) Write a short note on "Secular trend".

(*iii*) What is seasonal variation ? *(MBA, Madurai-Kamaraj Univ., 2006)*

(*iv*) Write down the most important factors causing seasonal variations. *(MBA, Madurai-Kamaraj Univ., 2003)*

(*v*) What are the normal equations for the straight line $Y = a + bx$?

(*vi*) What are cyclical fluctuations ?

(*vii*) What is Business forecasting ?

(*viii*) Name a few methods of Business forecasting.

(*ix*) How irregular variations are caused ?

(*x*) "Forewarned is Forearmed." Comment.

1-B : Answer the following questions, each question carries **four** marks:

(*i*) What are seasonal variations ? Bring out the factors that cause seasonal variations. *(M.A. Econ., Madras Univ., 2003)*

(*ii*) Narrate the merits and limitations in the use of moving average method. *(M.A. Econ., Madras Univ., 2003)*

(*iii*) Distinguish between the additive and multiplicative models of time series analysis.

(*iv*) Suggest the important adjustments to the mode before analysing time series.

(*v*) Explain the freehand (graphic method) of measuring trend.

2. (*a*) What is Business Forecasting ? Explain clearly its role and limitations.

(*b*) How does analysis of time series helps in making business forecast ?

(*c*) What is forecasting ? Discuss in brief the various theories and methods of business forecasting. *(MBA, Delhi Univ., 1998)*

3. (*a*) Explain clearly the different components into which a time series may be analysed. Explain any method in isolating trend values in a time series. *(MBA, Delhi Univ., MBA, Vikram Univ., 2005)*

(*b*) Explain clearly the meaning of Time Series Analysis. Mention its important components. Explain these components with examples, indicating the importance of each component in business. *(B.Com., Andhra Univ., 2003)*

(*c*) Describe the seasonal variation and cyclical fluctuations in a time series. *(MBA, Anna Univ., 2003)*

4. Explain what do you understand by Time Series. Why is Time Series considered to be an effective tool of forecasting ? *(MBA, BHU, 2002)*

5. (*a*) What is business forecasting ? What are the assumptions on which business forecasts are made? Describe the techniques of forecasting that are commonly employed by big business houses.

(*b*) Explain briefly the additive and multiplicative models of time series. Which of these models is more popular in practice and why?

6. (*a*) Critically examine the various methods that are used for measuring trend. Which method do you think is the best and why ?

(*b*) Explain briefly the different methods of measuring trend. *(MBA, Madras Univ., 2002)*

7. (*a*) How seasonal variations are accounted for in the analysis of Time Series ?

(*b*) What are the common methods in use for eliminating seasonality from a time series data ? Explain any one method taking imaginary figures.

8. Critically examine the various methods that are used for business forecasting. Why is time series considered to be an effective tool for forecasting analysis? Explain.

9. Explain the following terms in the study of time series :

(*i*) Secular trend, (*ii*) Seasonal variation, (*iii*) Cyclical fluctuations.

10. (*a*) What do you understand by 'seasonal variation' in time series data ? Explain their uses.

(*b*) Why do we measure seasonal variations in a time series ?

(*c*) How would you eliminate seasonal influences ? Illustrate with the help of an example.

(*d*) Explain clearly with the help of an illustration how seasonal index is useful in planning sales or production for specific periods. Are there any limitations of seasonal index ?

11. (*a*) Explain the method of Moving Averages in estimating the trend of a time series. What are the disadvantages in using this method ?

(*b*) Explain the concept of 'auto correlation' and its use for time series analysis.Give an example of a single variable with two different time lags. *(MBA, IGNOU 2001)*

12. (a) Why do we deseasonalize data ? Explain the ratio-to-moving average method to compile the seasonal index.
(b) Explain the following statements :
(i) "... the business analyst who uses moving averages to smoothen his data while in the process of trying to discover business cycles, is likely to come up with some non-existent cycles."
(ii) "There is nothing sacred in computing seasonal indices by the method of moving average using exclusively monthly data."
(iii) "Despite great limitations of statistical forecasting, the forecasting techniques are invaluable to the economist, the businessman and the Government."
13. Suppose you are provided with a given time series data and asked to analyse its general pattern and fluctuations. Describe in detail the steps you would follow in determining the pattern of trend and whether a seasonal and/or a cyclical component contributed to movements in the series.
14. (a) (i) "A key assumption in the classical method of time series analysis is that each of the component movements in the time series can be isolated individually from a series". Do you agree with this statement ? Does this assumption create any serious limitation to such analysis?
(ii) "A 12-month moving average of time series data removes trend and cycle." Do you agree ? Why or why not ?
(b) Examine critically the time-lag and the action and the reaction theory of business forecasting. Which of these, in your opinion, is better and why ?
15. Answer the following by a brief statement on each :
(i) Why must short-term forecasts be more precise than long-term ones ?
(ii) What is the major objective of seasonal analysis ?
(iii) What purpose does a seasonal index solve ?
16. (a) What is the difference between seasonal fluctuations and cyclical variations in a time series data.
(b) Illustrate the historical analogy theory of business forecasting.
(c) What is a time series ? What are its components ? Which components of the series is mainly applicable in the following cases ? *(M.B.A., UPTech. Univ., 2005)*
(i) A fire in a factory delaying production for one month.
(ii) Formation of rocks.
(iii) Decrease in the employment in sugar factory during the off-season.
(iv) Sale of New Year greeting cards.
(v) Fall in death rate due to a advances in science.
(vi) An after Deepawali sales in a departmental store.
(vii) A need for increased rice production due to a constant increase in population.
17. Indicate three categories of forecasting models and list out five techniques from each category. Describe Delphi technique in detail.
18. (a) Critically examine the time-lag and the action and reaction theory of business forecasting. Which of these two is better and why ?
(b) While fitting a straight line trend of the type $Y = a + bX$, what is signified by Y, X, a and b ?
19. (a) Discuss the role of forecasting as a business tool.
(b) How do we manage long-range forecasting and technical change for any organisation ?
(c) Write short notes on Delphi method and Historical analogy method for business forecasting.
(d) Explain how can we use market surveys as a method of forecasting. Illustrate.
(e) Write a lucid note on Box and Jenkin's method of forecasting. *(MBA, Kurukshetra Univ., 2005)*
(f) Explain with appropriate example different methods of estimating seasonal variations.
(g) What do you understand by Naive (Time series) Quantitative Models of forecasting? *(MBA, Jamia Millia, 2003)*
20. Business today generate a large amount of data continuously. This data may be used to gain information about the system. For one such system, it is known that the relation between variables is non-linear, *i.e.*, in the form $y = axb$, where a and b are constants. Use a transformation to make it linear and discuss how would you use the method of least squares to fit a straight line to the transformed linear model.
21. Apply the method of semi-averages for determining trend to the following data and estimate the value for 2015 :

Year	*Sales (Thousand units)*	*Year*	*Sales (Thousand units)*
2005	20	2008	30
2006	24	2009	28
2007	22	2010	32

If the actual figure of sales for 2011 is 35,000 units, how do you account for difference between the figure you obtain and the actual figure given to you ?

22. Plot the following data on graph paper and ascertain trend by the method of semi- averages :

Year	*Sales (million tonnes)*	*Year*	*Sales (million tonnes)*
2004	100	2008	108
2005	120	2009	102
2006	95	2010	112
2007	105		

23. Apply the method of semi-average to depict the long-term tendency of following data and estimate the value for 2013:

Year	*Production (million tonnes)*	*Year*	*Production (million tonnes)*
2003	40	2007	51
2004	44	2008	50
2005	42	2009	54
2006	48	2010	56

24. The following series relate to the profits of a commercial concern for 8 years :

Year	*Profits Rs.*	*Year*	*Profits Rs.*
2003	15,420	2007	26,120
2004	14,420	2008	31,950
2005	15,520	2009	35,360
2006	21,020	2010	35,670

Find the trend of profits. (Assume a three-year cycle and ignore decimals.)

25. Find out the trend values for the following time series of steel production by the method of moving average using 5-point time period for your purpose. State briefly the procedure that would have been adopted if you were to choose a 4-point time period. How does one choose the proper 'period of the moving average' ?

Year	*Production (m. tonnes)*	*Year*	*Production (m. tonnes)*	*Year*	*Production (m. tonnes)*
1993	351	1999	410	2005	502
1994	366	2000	420	2006	540
1995	361	2001	450	2007	557
1996	362	2002	500	2008	571
1997	400	2003	518	2009	586
1998	419	2004	455	2010	612

26. Below are given the figures of production of a sugar factory :

Year	*Production (thousand tonnes)*	*Year*	*Production (thousand tonnes)*
2005	92	2008	92
2006	83	2009	92
2007	94	2010	110

Apply the method of least squares to determine the trend values. Also find out the short-term fluctuations.
[$Y = 95 + 1473X$]

27. Fit a straight line trend by the method of least squares :

Year	*Milk consumption (million litres)*	*Year*	*Milk consumption (million litres)*
2002	102.3	2007	118.7
2003	101.9	2008	124.5
2004	105.8	2009	129.9
2005	112.0	2010	134.8
2006	114.8		

[$Y = 116.1 + 4.3X$]

28. The following are annual profits (in thousands of rupees) of a business firm :

Year :	2004	2005	2006	2007	2008	2009	2010
Profits (in '000 Rs.) :	60	72	75	65	80	85	95

(*a*) Use the method of least squares to fit a straight line to the above data.
(*b*) Plot the above figures and draw the line.
(*c*) Also make an estimate of the profits for the year 2011.
[$Y = 76 + 4.86X$; $Y_{2011} = 100.3$]

29. Fit a straight line trend by the method of least squares to the following data :

Year (X) :	2003-04	2004-05	2005-06	2006-07	2007-08	2008-09	2009-10
Sales (in lakh Rs.) (Y) :	278	309	335	378	424	481	521

(Clearly specify the origin and the units of the variables in the trend equation obtained.)
[$Y = 389.43 + 41.5X$]

30. Fit an equation of the type $Y = a + bX + cX^2$ to the following data :

Year	*Production (in '000 tonnes)*	*Year*	*Production (in '000 tonnes)*
2006	70	2009	80
2007	72	2010	90
2008	88		

31. The following table shows the number of letters posted in a particular area during a typical period of four weeks. Assuming that the trend value during the period remains the same, calculate 'seasonal indices' (here daily indices) as percentage of the grand average :

Week	*Sun.*	*Mon.*	*Tue.*	*Wed.*	*Thurs.*	*Fri.*	*Sat.*	*Total*
1	18	161	170	164	153	181	76	923
2	18	165	169	147	148	190	80	917
3	21	162	169	153	155	190	82	932
4	20	165	170	155	150	180	85	925
Total	77	653	678	619	606	741	323	3697

32. The working capital requirements of the *XYZ* Ltd. have been subject to seasonal fluctuations. At the same time, a steady secular advance can be noted. In order to evaluate comprehensively future working capital needs, the treasurer calculated a straight line trend and the seasonal indices. The trend equation is $Y_c = 10{,}000 + 500X$, where X represents a period of 1 month and has a value of 0 in 2010. The seasonal indices are as follows :

Jan.	80	July	125
Feb.	95	Aug.	99
Mar.	90	Sep.	90
Apr.	100	Oct.	102
May.	116	Nov.	105
June	120	Dec.	87

(*a*) Prepare a schedule of estimated working capital requirements for 2010.
(*b*) What factors could cause these estimates to be incorrect?
(*c*) What might be done to compensate for inaccuracies as they become apparent?
(*d*) Would you as a banker have any interest in estimates of this type?

33. In order to find quarterly seasonal indices, first of all the quarter wholesale price for five years (2006-2010) were reduced as percentage of their centred moving averages of four quarters. These percentages are set out in the following table. You are required to calculate the quarterly seasonal indices.

Year	*I*	*II*	*III*	*IV*
2006	—	—	127	134
2007	130	122	122	132
2008	120	120	118	128
2009	126	116	121	130
2010	127	118	—	—

I	*II*	*III*	*IV*
101.0	95.6	98.04	105

34. Consult a copy of Business Statistics in your library and select a series of your own. The series should be for a period of minimum eight years. Then do the following :

(a) Compute an appropriate trend line for the series, with first month of series as the origin. Plot the original series and its trend in one diagram.

(b) Compute a typical seasonal index for the series by the ratio-to-moving average method. Plot the actual data and moving average figures.

(c) What comments can you make on T, S, C and I?

35. a) What do you understand by seasonal fluctuation in time series? Give an example.

(b) What are the major uses of seasonal indices in time series analysis? Name four methods by which one can compute a seasonal index from time series data.

(c) The sales of a company rose from Rs. 60,000 in the month of August to Rs. 69,000 in the month of September. The seasonal indices for these two months are 105 and 140 respectively. The owner of the company was not at all satisfied with the rise of sale in the month of September by Rs. 9,000. He expected much more because of the seasonal index for the month. What were his estimates of sales for the month of September?

[The expectation was $\dfrac{60{,}000 \times 140}{105}$ = Rs. 80,000]

36. (a) Given the following trend information :

$Y_C = 60 + 2.48X$ Origin : July 1, 2008

Y in million of rupees X in terms of years

Convert this equation in monthly terms. Be sure your response is in the most practical form.

(b) The annual trend equation for the *XYZ* Co. Ltd. is represented by the following :

$Y_C = 468 + 0.20X$ X = years

Y = thousands of rupees Origin : July, 2008

(i) Based on the past several years, monthly sales during January have been around Rs. 50,000. What is the typical seasonal relative for January?

(ii) If your seasonal relative for January was greater than 100, does it necessarily indicate that at least one of the 11 months has a seasonal relative, that is less than 100? Explain.

37. Fit a straight line trend to the following data and show the original observations and trend values on the graph paper :

Year	2004	2005	2006	2007	2008	2009	2010
Gross ex-factory value of output	672	824	967	1204	1464	1758	2057

[$Y = 1278 + 23.86X$]

38. The number of units of a product exported during 2003-2010 is given below. Fit a straight line trend to the data. Plot the given data showing also the trend.

Year	2003	2004	2005	2006	2007	2008	2009	2010
No. of units (in thousands)	12	13	13	16	19	23	21	23

[$Y = 17.5 + 0.893X$]

39. Calculate seasonal indices by the 'ratio-to-moving average method' from the following data :

Year	*I Quarter*	*II Quarter*	*III Quarter*	*IV Quarter*
2008	68	62	61	63
2009	65	58	66	61
2010	68	63	63	67

[105.3; 95.21; 100.97; 98.52]

40. The sales of a company increased from Rs. 4,00,000 in March to Rs. 4,80,000 in July 2010. The company's seasonal indices for these two months are 105 and 140 respectively. The owner of the company expressed dissatisfaction with the July sales but the Sales Manager said that he was quite pleased with the Rs. 80,000 increase. What argument should the owner of the company have used to reply the Sales Manager?

The Sales Manager also predicted on the basis of the July sales that the total 2012 sales were going to be Rs. 6,76,000. Criticise the Sales Manager's estimate.

41. The following table shows the number of salesmen working in a certain concern :

Year	2006	2007	2008	2009	2010
No. of salesmen	28	38	46	40	56

Use the method of least squares to fit a straight line trend and estimate the number of salesmen in 2015.

42. The materials manager of a company has projected 10, 15 and 18 truckloads of a product for three consecutive months. The seasonal indices for these are 141.5, 125.8 and 82.6 respectively. Work out the seasonalised forecast for each month of three months.

43. The seasonal indices of the sale of readymade garments of a particular type in a departmental store are given below :

	Quarter	*Seasonal index*
I	Jan.—March	95
II	April—June	80
III	July—Sept.	90
IV	Oct.—Dec.	125

If the total sales in the first quarter of the year be worth Rs. 50,000, determine how much worth of garments of this type should be kept in store to meet the demand in each of the remaining quarters.

[50,000; 42,145.26; 47,368.42; 65,789.47]

44. A company estimates its sales for a particular year to be Rs. 24,00,000. The seasonal indices for sales are as follows :

Months	*Seasonal index*	*Months*	*Seasonal index*
January	75	July	102
February	80	August	104
March	98	September	100
April	128	October	102
May	137	November	82
June	119	December	73

Using the given information, calculate estimates of monthly sales of the company. Assume that there is no trend.

(*MBA, Osmania Univ., 2002*)

45. The following figures are the production data of a cement factory :

Year	*Production ('000 tonnes)*	*Year*	*Production ('000 tonnes)*
2000	17	2006	35
2001	20	2007	35
2002	19	2008	51
2003	26	2009	74
2004	24	2010	79
2005	40		

Fit the trend of the type $Y = a + bX + cX^2$ to the above data. Select the year 2005 as the working origin.

46. Using the data given below, explain how would you determine seasonal fluctuations in a time series :

Year	*Summer*	*Monsoon*	*Autumn*	*Winter*
2006	30	81	62	199
2007	33	104	86	171
2008	42	153	99	221
2009	56	172	129	235
2010	67	201	136	302

47. The number of units produced during 2003-2010 are given below :

Year :	2003	2004	2005	2006	2007	2008	2009	2010
Units produced :	56	55	51	47	42	38	35	32

(*i*) Fit a straight line trend and obtain the trend values.
(*ii*) Eliminate the trend. What components of the time series are thus left over?
(*iii*) What is the monthly increase in the number of units produced?

48. Compute a nonlinear trend of the form $Y = a + bX + cX^2$ for the data showing the production of wheat (in thousand tonnes) during the years 2002 to 2010.

Year	:	2002	2003	2004	2005	2006	2007	2008	2009	2010
Production of wheat ('000 tonnes)	:	9	10	12	15	13	10	8	16	15

(Take the year 2006 as working origin)

49. Find the trend values by the method of least squares for the following time series :

Year	:	2003	2004	2005	2006	2007	2008	2009	2010
Production ('000 tonnes)	:	351	366	362	400	419	420	450	518

Estimate the likely production for the year 2013.

50. Use method of least squares to determine sales for the year 2012.

Year	:	2006	2007	2008	2009	2010
Sales	:	100	110	130	125	160

51. Fit a straight line trend by the method of least squares to the following data :

Year	:	2000	2001	2002	2003	2004	2005	2006	2007
Earnings (Rs. Lakh)	:	38	40	65	72	69	60	87	95

(*M.Com. Madurai-Kamaraj Univ., 2007*)

52. The projected number of women of child bearing age (15–49) for India from 2000 to 2007 are as follows :

Year	:	2000	2001	2002	2003	2004	2005	2006	2007
No. of Women (in millions)	:	152.6	156.4	160.3	164.4	168.5	172.7	176.9	181.2

Fit a trend line.

(*MBA, Anna Univ., 2007*)

53. What is meant by moving average ? Find the trend for the following series by three year weighted average with weights 1, 2, 1 :

Year (coded value)	:	–3	–2	–1	0	1	2	3
Sales (in thousand units)	:	2	4	5	7	8	10	13

(*MBA, M.D. Univ., 2006*)

54. The following are the annual profits in lakh of rupees, in a certain business :

Year	*Profits (Rs. lakh)*	*Year*	*Profits (Rs. lakh)*
2004	60	2008	80
2005	72	2009	85
2006	75	2010	95
2007	65		

(*i*) Use the method of least square to fit a straight line trend to the above data.
(*ii*) Also make an estimate of the profits for the year 2014.

* * * *