Unit-V (TIME SERIES)

Components of time series – Additive and multiplicative models - Resolving components of a time series-measuring trend: Graphic, semi-averages, moving average and least squares methods. Seasonal variation- measuring seasonal variation: method of simple averages, ratio-to- trend method, ratio-to-moving average method and link relative method, Cyclical and Irregular fluctuations, Smoothing Methods.

TIME SERIES

1.1. Meaning

An arrangement of statistical data in accordance with time of occurrence or in a chronological order is called a time series. The numerical data which we get at different points of time-the set of observations is known as time series.

In time series analysis, current data in a series may be compared with past data in the same series. We may also compare the development of two or more series over time. These comparisons may afford important guide lines for the individual firm. In Economics, statistics and commerce it plays an important role.

Definition and Examples

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

For example, if we observe agricultural production, sales, National Income etc., over a period of time, say over the last 3 or 5 years, the set of observations is called time series. Thus a time series is a set of time, quantitative readings of some various recorded at equal intervals of time. The interval may be an hour, a day, a week, a month, or a calendar year. Hourly temperature reading, daily sales in a shop, weekly sales in a shop, weekly sales in a market, monthly production in an industry, yearly agricultural production, population growth in ten years, are examples of time series.

From the comparison of past data with current data, we may seek to establish what development may be expected in future. The analysis of time series is done mainly for the purpose of forecasts and for evaluating the past performances. The chronological variations will be object of our study in time series analysis.

The essential requirements of a time series are:

- The time gap, between various values must be as far as possible, equal.
- It must consist of a homogeneous set of values.
- Data must be available for a long period.

Symbolically if "t" stands for time and "y_t" represents the value at time t then the paired values (t, yt) represents a time series data.

Ex 1: Production of rice in Tamilnadu for the period from 2010-11 to 2016-17.

Table 1.1. Production of rice in Tamilnadu (in '000 metric tons)

Year	**Production**
2010-11	400
2011-12	450
2012-13	440
2013-14	420
2014-15	460
2016-17	520

Uses of Time Series

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

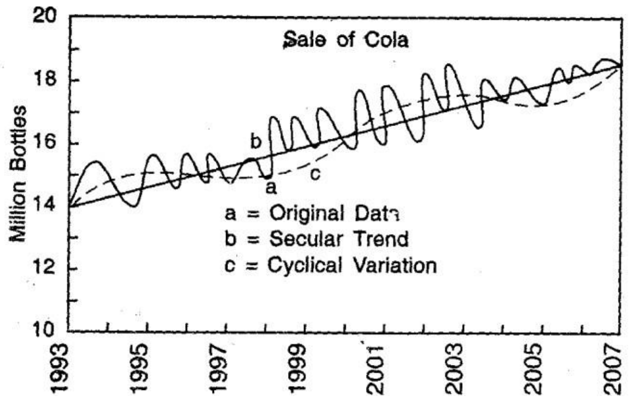
- It helps in understanding past behavior.
- It helps in planning future operations.
- It helps in evaluating current accomplishments.
- It facilitates comparison.

Components of Time Series

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

These are four types

- Secular Trend
- Seasonal Variations
- Cyclical Variations
- Irregular Variations



Secular Trend

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

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

Mathematically the secular trend may be classified into two types

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

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

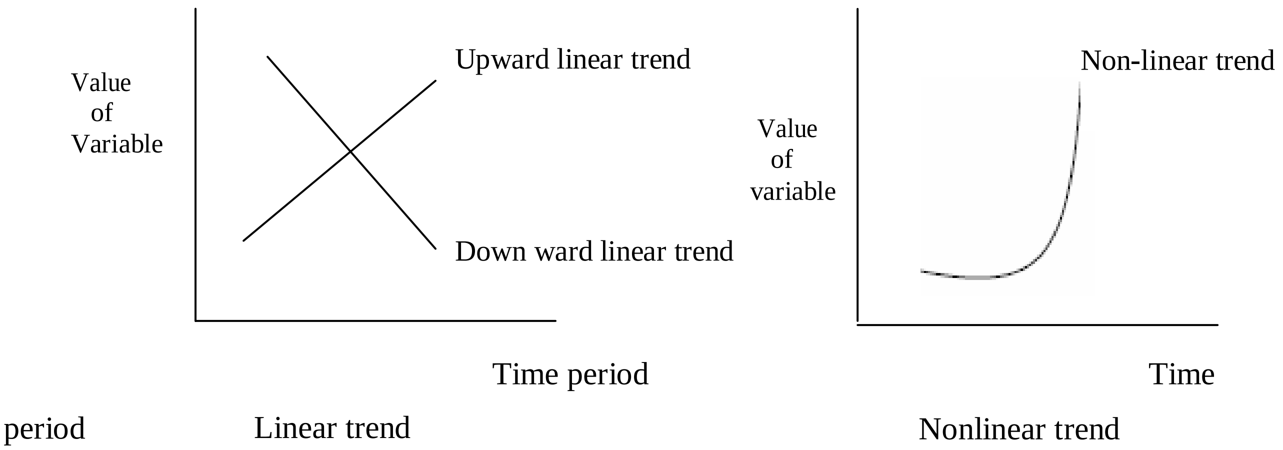


Figure 1. Linear Trend and Nonlinear Trend

Seasonal Variations

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

- Due to natural forces
- Manmade convention.

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

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

Cyclical Variations or Oscillatory Variation

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

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

- Boom
- Decline
- Depression
- Improvement

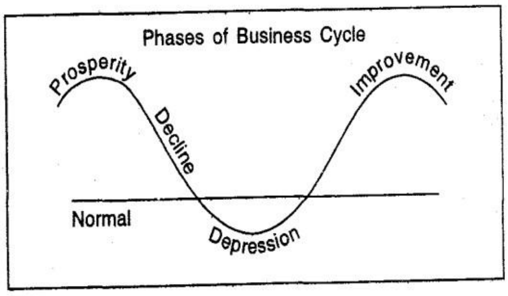


Figure 2. Phases of Business Cycle

Irregular Variation

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

Mathematical Model

In classical analysis, it is assumed that some type of relationship exists among the four components of time series. Analysis of time series requires decomposition of a series, to decompose a series we must assume that some type of relationship exists among the four components contained in it.

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

- Additive model
- Multiplicative model

Additive model

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

Where T_t = Trend value at time t

S_t = Seasonal component

C_t = Cyclical component

R_t = Irregular component

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

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

Multiplicative model:

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

The multiplicative model can be put in additive model by taking log both sides. However most business analysis uses the multiplicative model and finds it more appropriate to analyze business situations.

According to this model, the simple

One of the most important tasks before economists and businessmen these days is to make estimates for the future. For example, a businessman is interested in finding out his likely sales in the year 2016 or as a long-term planning in 2025 or the year 2030 so that he could adjust his production accordingly and avoid the possibility of either unsold stocks or inadequate production to meet the demand. Similarly, an economist is interested in estimating the likely population in the coming year so that proper planning can be carried out with regard to food supply, jobs for the people, etc. However, the first step in making estimates for the future consists of gathering information from the past. In this connection one usually deals with statistical data which are collected, observed or recorded at successive intervals of time. Such data are generally referred to as „time series". Thus when we observe numerical data at different points of time the set of observations is known as time series. For example if we observe production, sales, population, imports, exports, etc. at different points of time, say, over the last 5 or 10 years, the set of observations formed shall constitute time series. Hence, in the analysis of time series, time is the most important factor because the variable is related to time which may be either year, month, week, day and hour or even- minutes or seconds.

Measurement of Secular trend:

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

- Graphic method or eye inspection method
- Semi average method
- Method of moving average
- Method of least squares

Graphic method or eye inspection method

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

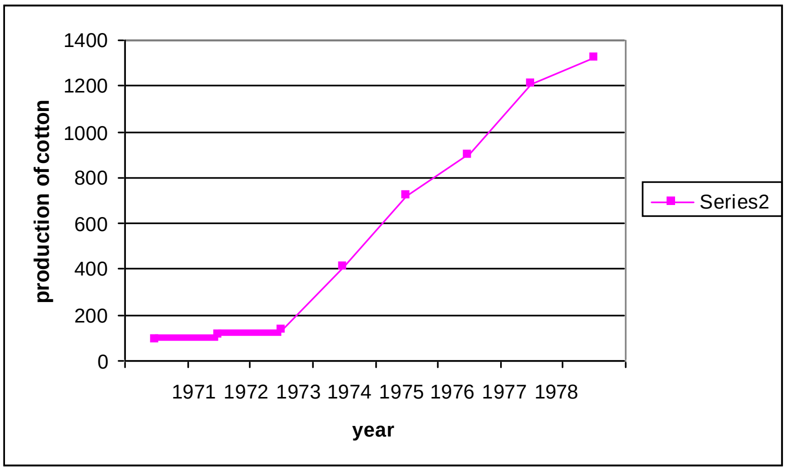


Figure 3. Graphic method for the production of cotton base on year

Advantages

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

Disadvantages:

- This method is highly subjective and curve varies from person to person who draws it.
- The work must be handled by skilled and experienced people.
- Since the method is subjective, the prediction may not be reliable.
- While drawing a trend line through this method a careful job has to be done.

Method of Semi Averages:

In this method the whole data is divided in two equal parts with respect to time. For example if we are given data from 1999 to 2016 i.e. over a period of 18 years the two equal parts will be first nine years i.e. from 1999 to 2007 and 2008 to 2016. In case of odd number of years like 9, 13, 17 etc. two equal parts can be made simply by omitting the middle year. For example if the data are given for 19 years from 1998 to 2016 the two equal parts would be from 1998 to 2006 and from 2008 to 2016, the middle year 2007 will be omitted. After the data have been divided into two parts, an average (arithmetic mean) of each part is obtained. We thus get two points. Each point is plotted against the mid year of the each part. Then these two points are joined by a straight line which gives us the trend line. The line can be extended downwards or upwards to get intermediate values or to predict future values.

Example:

Year	**Production**	**Semi averages**
2001	40	
2002	45	$\frac{40 + 45 + 40 + 42}{4} = 41.75$
2003	40	
2004	42	
2005	46	
2006	52	$\frac{46 + 52 + 56 + 61}{4} = 53.75$
2007	56	
2008	61	

Thus we get two points 41.75 and 53.75 which shall be plotted corresponding to their middle years i.e. 2002.5 and 2006.5. By joining these points we shall obtain the required trend line. This line can be extended and can be used either for prediction or for determining intermediate values.

Advantages:

- This method is simple to understand as compare to moving average method and method of least squares.
- This is an objective method of measuring trend as everyone who applies this method is bound to get the same result.

Disadvantages:

- The method assumes straight line relationship between the plotted points regardless of the fact whether that relationship exists or not.

- The main drawback of this method is if we add some more data to the original data then whole calculation is to be done again for the new data to get the trend values and the trend line also changes.
- As the Arithmetic Mean of each half is calculated, an extreme value in any half will greatly affect the points and hence trend calculated through these points may not be precise enough for forecasting the future.

Method of Moving Average:

It is a method for computing trend values in a time series which eliminates the short term and random fluctuations from the time series by means of moving average. Moving average of a period m is a series of successive arithmetic means of m terms at a time starting with 1^{st}, 2^{nd}, 3^{rd} and so on. The first average is the mean of first m terms; the second average is the mean of 2^{nd} term to $(m+1)^{th}$ term and 3^{rd} average is the mean of 3^{rd} term to $(m+2)^{th}$ term and so on.

If m is odd then the moving average is placed against the mid value of the time interval it covers. But if m is even then the moving average lies between the two middle periods which does not correspond to any time period. So further steps has to be taken to place the moving average to a particular period of time. For that we take 2-yearly moving average of the moving averages which correspond to a particular time period. The resultant moving averages are the trend values.

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

Years	Production	3-yearly moving avg (trend values)
2001-02	40	
2002-03	45	(40+45+40)/3 = 41.67
2003-04	40	(45+40+42)/3 = 42.33
2004-05	42	(40+42+46)/3 = 42.67
2005-06	46	(42+46+52)/3 = 46.67
2006-07	52	(46+52+56)/3 = 51.33
2007-08	56	(52+56+61)/3 = 56.33
2008-09	61	

Ex :2 Calculate 4-yearly moving average for the following data.

Years	Production	4-yearly moving avg	2-yealry moving avg (trend values)
2001-02	40		
2002-03	45		
		(40+45+40+42)/3 = 41.75	
2003-04	40		42.5
		(45+40+42+46)/3 = 43.15	
2004-05	42		44.12
		(40+42+46+52)/3 = 45	
2005-06	46		47
		(42+46+52+56)/3 = 49	
2006-07	52		51.38
		(46+52+56+61)/3 = 53.75	
2007-08	56		
2008-09	61		

Advantages:

- This method is simple to understand and easy to execute.
- It has the flexibility in application in the sense that if we add data for a few more time periods to the original data, the previous calculations are not affected and we get a few more trend values.
- It gives a correct picture of the long term trend if the trend is linear.
- If the period of moving average coincides with the period of oscillation (cycle), the periodic fluctuations are eliminated.
- The moving average has the advantage that it follows the general movements of the data and that its shape is determined by the data rather than the statistician's choice of mathematical function.

Disadvantages:

- For a moving average of 2m+1, one does not get trend values for first m and last m periods.

- As the trend path does not correspond to any mathematical; function, it cannot be used for forecasting or predicting values for future periods.
- If the trend is not linear, the trend values calculated through moving averages may not show the true tendency of data.
- The choice of the period is sometimes left to the human judgment and hence may carry the effect of human bias.

Method of Least Squares:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$\Rightarrow \sum t(Y_t - a - bt) = 0$$
$$\Rightarrow \sum tY_t = a\sum t + b\sum t^2 \quad \text{…………………} (4)$$

The equations (3) and (4) are called „normal equations"

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

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

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

which is the required straight line trend equation.

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

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

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

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

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

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

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

To find a and b we have two normal equations

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

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

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

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

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

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

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

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

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

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

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

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

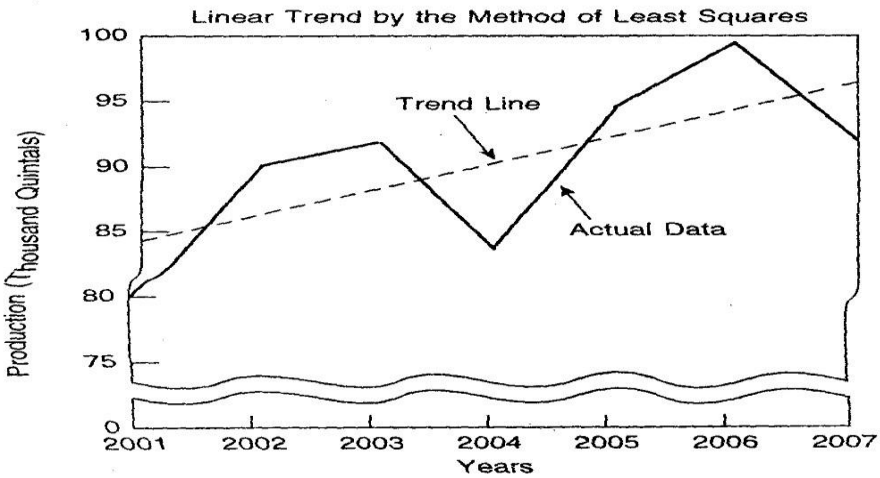


Illustration 32. Calculate trend values by the method of least-squares from the data given below :

Year	2000	2001	2002	2003	2004	2005	2006	2007
Sales	80	90	92	83	94	99	92	104

Plot the data showing also the trend line.

Solution. FITTING STRAIGHT LINE TREND BY METHOD OF LEAST SQUARES

Years	*Sales* Y	*Deviations from 2003.5*	*Deviations multiplied by 2* X	XY	X^2	Y_c
2000	80	− 3.5	− 7	− 560	49	83.0
2001	90	− 2.5	− 5	− 450	25	85.5
2002	92	− 1.5	− 3	− 276	9	88.0
2003	83	− .5	− 1	− 83	1	90.5
2004	94	+ .5	+ 1	+ 94	1	93.0
2005	99	+ 1.5	+ 3	+ 297	9	95.5
2006	92	+ 2.5	+ 5	+ 460	25	98.0
2007	104	+ 3.5	+ 7	+ 728	49	100.5
$N = 8$	$\Sigma y = 734$			$\Sigma XY = 210$	$\Sigma X^2 = 168$	$\Sigma Y_c = 734$

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

To find a and b we have two normal equations

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

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

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

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

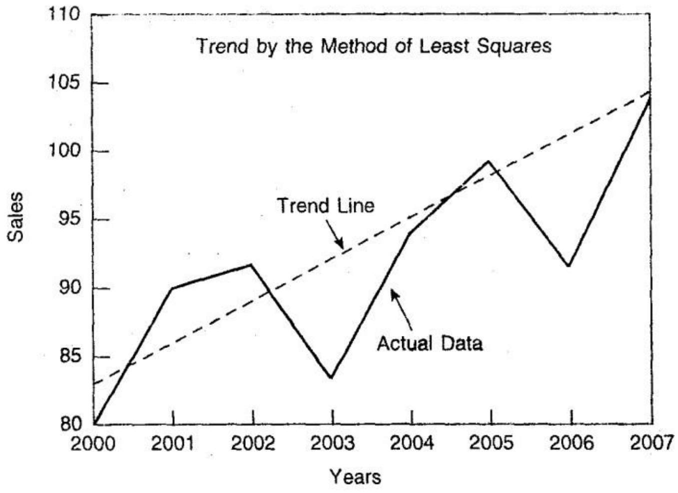
The trend values for various years are

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

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

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

and so on



Fitting of a parabolic trend by the method of least squares

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

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

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

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

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

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

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

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

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

$$\Rightarrow \sum Y_t = \sum a + b\sum t + c\sum t^2$$

$$\Rightarrow \sum Y_T = na + b\sum t + c\sum t^2 \quad \text{...} (3)$$

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

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

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

$$\Rightarrow \sum tY_t = a\sum t + b\sum t_2 + c\sum t_3 \quad \text{...................} (4)$$

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

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

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

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

The equations (3), (4) and (5) are called „normal equations"

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

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

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

Which is the required parabolic trend equation

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

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

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

Also plot the actual and trend values on the graph.

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

$$\Sigma Y = Na + b\Sigma X + c\Sigma X^2 \quad \ldots(i)$$
$$\Sigma XY = a\Sigma X + b\Sigma X^2 + c\Sigma X^3 \quad \ldots(ii)$$
$$\Sigma X^2 Y = a\Sigma X^2 + b\Sigma X^3 + c\Sigma X^4 \quad \ldots(iii)$$

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

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

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

$$848 = 6a + 3b + 19c$$
$$1{,}542 = 6a + 38b + 54c$$
$$- \quad - \quad - \quad -$$

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

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

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

$$14{,}649 = 57a + 361b + 513c$$
$$9{,}297 = 57a + 81b + 345c$$

$$5{,}352 = 280b + 168c \quad \ldots(v)$$

Multiplying equation (iv) by 8, we have

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

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

$$280b + 280c = 5{,}552$$
$$280b + 168c = 5{,}352$$
$$- \quad - \quad -$$

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

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

$$35b + (35 \times 1.786) = 694$$
$$35b = 694 - 62.5 = 631.5 \text{ or } b = 18.042$$
$$848 = 6a + 3(18.042) + 19(1.786) = 6a + 54.126 + 33.934$$
$$6a = 759.94 \quad \text{or} \quad a = 126.657$$

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

Substituting these values in the equation,

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

when $X = -2$

$$Y = 126.657 + 18.042(-2) + 1.786(-2)^2$$
$$= 126.657 - 36.084 + 7.144 = 97.717$$

when $X = -1$

$$Y = 126.657 + 18.042(-1) + 1.786(-1)^2$$
$$= 126.657 - 18.042 + 1.786 = 110.401$$

when $X = 1$,

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

when $X = 2$,

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

when $X = 3$,

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

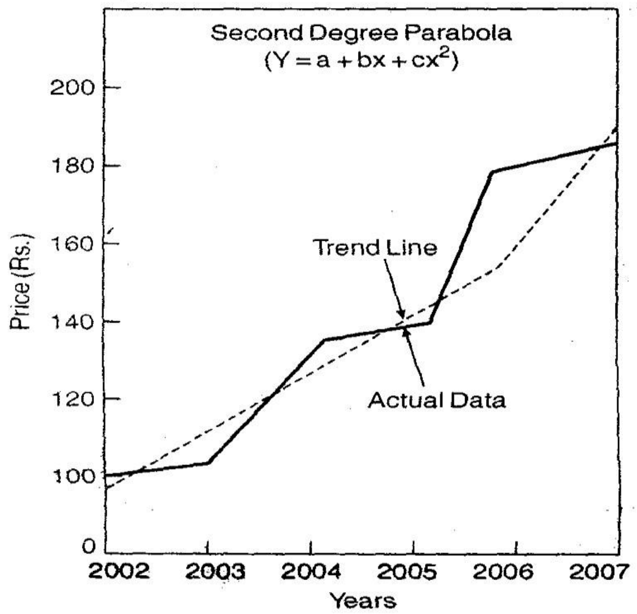
Price for the year 2008

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

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

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

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



Advantages

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

Disadvantages

- Great care should be exercised in selecting the type of trend curve to be fitted i.e. linear, parabolic or some other type. Carelessness in this respect may lead to wrong results.
- The method is more tedious and time consuming.
- Predictions are based on only long term variations i.e trend and the impact of

Examples for Practice:

1) Define a time series. Discuss its main components.
2) Define secular trend of a time series and explain methods that are used in isolating it.
3) Explain the method of moving average for the determination of trend. What are the advantages and disadvantages of this method?
4) What are the advantages and disadvantages of the graphic method and least square method in trend analysis?
5) Explain briefly the method of moving averages for calculating the trend.
6) How does analysis of time series help business forecasting?
7) Distinguish between secular trend, seasonal variations and cyclical fluctuations. Discuss various methods of measuring each.
8) Explain briefly the additive and multiplicative models of time series. Which of these models is more popular in practice and why?
9) Explain how you would determine seasonal variation by 12-monthly moving average.
10) What are the various methods for determining trend in a time series?
11) Describe in detail the method of least squares for determining trend.
12) The production data on steel in a factory in the past 10 years are given below:

Year	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996
Production (tonnes)	75	86	98	90	96	108	124	140	150	165

Fit a staright line trend and tabulate the trend values. What is the expected production in the year 1997 on the basis of trend? ($Y = 113.2 + 5.08X$; $Y = 169.98$)

13) Fit a staright line trend of the following data by least square method. Also find an estimate for the year 1997;

Year	1989	1990	1991	1992	1993	1994	1995	1996
Production (tonnes)	12	13	13	16	19	23	21	23

($Y = 17.5 + .893X$; $Y_{1997} = 25.54$)

14) Fit a staright line trend by the method of least square to the following data. Also find an estimate for the year 2000;

Year	1990	1991	1992	1993	1994	1995	1996	1997
Production (tonnes)	38	40	65	72	69	67	95	104

($Y = 68.75 + 4.404X$;

SEASONAL VARIATIONS

Introduction

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

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

Measurement of seasonal variations:

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

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

- Method of simple averages.
- Ratio to trend method
- Ratio to moving average method
- Link relative method.

Method of simple averages

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

The method consists of the following steps.

- Arrange the data by years and months (or quarters if quarterly data is given).

- Compute the average $\overline{x}_i$ (i = 1,2,…..12 for monthly and i=1,2,3,4 for quarterly) for the i^{th} month or quarter for all the years.
- Compute the average $\overline{x}$ of the averages.
- i.e. $\overline{x} = \frac{1}{12}\sum_{i=1}^{12}\overline{x}_i$ for monthly and $\overline{x} = \frac{1}{4}\sum_{i=1}^{4}\overline{x}_i$ for quarterly
- Seasonal indices for different months (quarters) are obtained by expressing monthly (quarterly) averages as percentages of $\overline{x}$. Thus seasonal indices for i-th month (quarter) = $\frac{\overline{x}_i}{\overline{x}} \times 100$

Advantages and Disadvantages:

Method of simple average is easy and simple to execute.

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

Example: 1

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

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

What are the seasonal indices for various quarters ?

Solution. COMPUTATION OF SEASONAL INDICES

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

Notes for calculating seasonal index

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

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

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

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

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

Ratio to trend method:

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

$$Y_t = T_t S_t C_t R_t$$

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

- Obtain the trend values by the least square method by fitting a mathematical curve, either a straight line or second degree polynomial.
- Express the original data as the percentage of the trend values. Assuming the multiplicative model these percentages will contain the seasonal, cyclical and irregular components.
- The cyclical and irregular components are eliminated by averaging the percentages for different months (quarters) if the data are In monthly (quarterly), thus leaving us with indices of seasonal variations.
- Finally these indices obtained in step(3) are adjusted to a total of 1200 for monthly and 400 for quarterly data by multiplying them through out by a constant K which is

given by $K = \dfrac{1200}{\text{Total of the indices}}$ For monthly,

$K = \dfrac{400}{\text{Total of theindices}}$ for quarterly.

Advantages:

- It is easy to compute and easy to understand.
- 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 ratio to moving average method that in this method we obtain ratio to trend values for each period for which data are available where as it is not possible in ratio to moving average method.

Disadvantages:

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

Example 1:

Calculate seasonal indices by Ratio to moving average method from the following data.

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

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

CALCULATING TREND BY METHOD OF LEAST SQUARES

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

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

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

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

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

TREND VALUES

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

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

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

GIVEN QUARTERLY VALUES AS % OF TREND VALUES

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

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

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

Ratio to moving average method:

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

- Calculate the centered 12-monthly moving average (or 4-quarterly moving average) of the given data. These moving averages values will eliminate S and I leaving us T and C components.
- Express the original data as percentages of the centered moving average values.
- The seasonal indices are now obtained by eliminating the irregular or random components by averaging these percentages using A.M or median.
- The sum of these indices will not in general be equal to 1200 (for monthly) or 400 (for quarterly). Finally the adjustment is done to make the sum of the indices to a total of 1200 for monthly and 400 for quarterly data by

multiplying them through out by a constant K which is given by

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

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

Advantages:

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

Disadvantages:

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

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

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

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

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

CALCULATION OF SEASONAL INDEX

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

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

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

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

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

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

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

Link relative method:

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

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

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

- Calculate the average of the link relatives for each period for all the years using mean or median.
- Convert the average link relatives into chain relatives on the basis of the first season. Chain relative for any period can be obtained by
- $$\frac{Avg.\ link\ relative\ for\ that\ period \times Chain\ relative\ of\ the\ previous\ period}{100}$$

  the chain relative for the first period is assumed to be 100.
- Now the adjusted chain relatives are calculated by subtracting correction factor "kd" from $(k+1)^{th}$ chain relative respectively.
- Where k = 1, 2,…….11 for monthly and k = 1,2,3 for quarterly data.

  and $d = \frac{1}{N}\left[New\ chain\ relative\ for\ first\ period - 100\right]$ where N denotes the number of periods i.e. N = 12 for monthly N = 4 for quarterly
- Finally calculate the average of the corrected chain relatives and convert the corrected chain relatives as the percentages of this average. These percentages are seasonal indices calculated by the link relative method.

Advantages:

- As compared to the method of moving average the link relative method uses data more.

Disadvantages:

- The link relative method needs extensive calculations compared to other methods and is not as simple as the method of moving average.
- The average of link relatives contains both trend and cyclical components and these components are eliminated by applying correction.

Illustration 26. Apply the method of link relatives to the following data and calculate seasonal indices :

QUARTERLY FIGURES

Quarter	*2003*	*2004*	*2005*	*2006*	*2007*
I	6.0	5.4	6.8	7.2	6.6
II	6.5	7.9	6.5	5.8	7.3
III	7.8	8.4	9.3	7.5	8.0
IV	8.7	7.3	6.4	8.5	7.1

Solution. CALCULATION OF SEASONAL INDICES BY THE METHOD OF LINK RELATIVES

Year	*Quarter* *I*	*II*	*III*	*IV*
2003	—	108.3	120.0	111.5
2004	62.1	146.3	106.3	86.9
2005	93.2	95.6	143.1	68.8
2006	112.5	80.6	129.3	113.3
2007	77.6	110.6	109.6	88.8
Arithmetic average	$\frac{345.4}{4}=86.35$	$\frac{541.4}{5}=108.28$	$\frac{608.3}{5}=121.66$	$\frac{469.3}{5}=93.86$
Chain relatives	100	$\frac{100\times 108.28}{100}$ $= 108.28$	$\frac{121.66\times 108.28}{100}$ $= 131.73$	$\frac{93.86\times 131.73}{100}$ $= 123.64$
Corrected chain relatives	100	$108.28 - 1.675$ $= 106.605$	$131.73 - 3.35$ $= 128.38$	$123.64 - 5.025$ $= 118.615$
Seasonal indices	$\frac{100\times 100}{113.4}$ $= 88.18$	$\frac{106.605}{113.4}\times 100$ $= 94.01$	$\frac{128.38}{113.4}\times 100$ $= 113.21$	$\frac{118.615}{113.4}\times 100$ $= 104.60$

The calculations in the above table are explained below :

Chain relative of the first quarter (on the basis of first quarter) = 100

Chain relative of the first quarter (on the basis of the last quarter)

$$= \frac{86.35 \times 123.64}{100} = 106.7.$$

The difference between these chain relatives = 106.7 − 100 = 6.7.

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

Adjusted chain relatives are obtained by subtracting 1×1.675, 2×1.675, 3×1.675 from the chain relatives of the 2nd, 3rd and 4th quarters respectively.

Average of corrected chain relatives

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

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

Deseasonalisation

When the seasonal component is removed from the original data, the reduced data are free from seasonal variations and is called deseasonalised data. That is, under a multiplicative model

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

Deseasonalised data being free from the seasonal impact manifest only average value of data.

Seasonal adjustment can be made by dividing the original data by the seasonal index.

$$\text{That is, Deseasonalised data} = \frac{\text{Original data}}{\text{Seasonal index}} \times 100$$

where an adjustment-multiplier 100 is necessary because the seasonal indices are usually given in percentages.

In case of additive model

$$Y_t = T + S + C + I,$$

$$\text{Deseasonalised data} = \text{Original data} - \frac{\text{Seasonal index}}{100}$$

$$= Y_t - \frac{\text{Seasonal index}}{100}.$$

Uses and limitations of seasonal indices

Seasonal indices are indices of seasonal variation and provide a quantitative measure of typical seasonal behavior in the form of seasonal fluctuations.

Measurement of cyclical variations:

The various methods used for measuring cyclical variations are

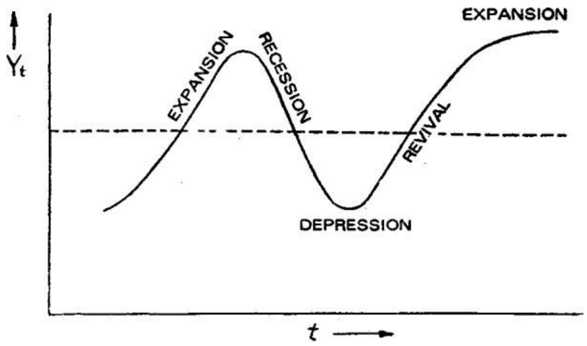
- Residual method
- Reference cycle analysis method
- Direct method
- Harmonic analysis method

Business Cycle

According to Mitchell, "Business cycle are a type of fluctuation found in the aggregate economic activity of nations that organize their work mainly in business enterprises : a cycle consists of expansions occurring at about the same time in many activities, followed by general recessions, contractions and revivals which merge into the expansion phase of the next cycle; this sequence of changes is recurrent but not periodic; in duration business cycles vary from more than one year to ten or twelve years.

There are four phases of a business cycle, such as

(a) Expansion (prosperity)
(b) Recession
(c) Depression (contraction)
(d) Revival (recovery).



A cycle is measured either from trough-to-trough or from peak-to-peak. Recession and contraction are the result of cumulative downswing of a cycle whereas revival and expansion are the result of cumulative upswing of a cycle.

Question bank:

1. Distinguish between seasonal variations, and cyclical fluctuations. How would you measure secular trend in any given data?
2. Describe the method of link relatives for calculating the seasonal variation indices.
3. How would you determine seasonal variation in the absence of trend?
4. Briefly describe the relative merits and demerits of ratio to trend and ratio to moving average method.
5. What do you understand by cyclical fluctuations in time series?

6. What do you understand by random fluctuation in time series?
7. Explain the term „Business cycle" and point out the necessity of its study in time series analysis.
8. Calculate seasonal variation for the following data of sale in thousands Rs. of a firm by the Ratio to trend method.

Year	**1st Quarter**	**2nd Quarter**	**3rd Quarter**	**4th Quarter**
1979	30	40	36	34
1980	34	52	50	44
1981	40	58	54	48
1982	52	76	68	62

9. Calculate seasonal indices by Ratio to moving average method from the following data.

Year	**1st Quarter**	**2nd Quarter**	**3rd Quarter**	**4th Quarter**
1980	75	60	54	59
1981	86	65	63	80
1982	90	72	66	85
1983	100	78	72	93

10. The data below gives the average quarterly prices of a commodity for five years. Calculate seasonal indices by method of link relatives.

Year	**1st Quarter**	**2nd Quarter**	**3rd Quarter**	**4th Quarter**
1979	30	26	22	31
1980	35	28	22	36
1981	31	29	28	32
1982	31	31	25	35
1983	34	36	26	33

****