METHODS OF TIME SERIES

TIME SERIES

Time series is set of data collected and arranged in accordance of time. According to **Croxton and Cowdon**,"A Time series consists of data arranged chronologically." Such data may be series of temperature of patients, series showing number of suicides in different months of year etc. The analysis of time series means separating out different components which influences values of series. The variations in the time series can be divided into two parts: long term variations and short term variations.Long term variations can be divided into two parts: Trend or Secular Trend and Cyclical variations. Short term variations can be divided into two parts: Seasonal variations and Irregular Variations.

METHODS FOR TIME SERIES ANALYSIS

In business forecasting, it is important to analyze the characteristic movements of variations in the given time series. The following methods serve as a tool for this analysis:

1. **Methods for Measurement of Secular Trend**
    i. Freehand curve Method (Graphical Method)

ii. Method of selected points
iii. Method of semi-averages
iv. Method of moving averages
v. Method of Least Squares

2. **Methods for Measurement of Seasonal Variations**
   i. Method of Simple Average
   ii. Ratio to Trend Method
   iii. Ratio to Moving Average Method
   iv. Method of Link Relatives

3. **Methods for Measurement for Cyclical Variations**

4. **Methods for Measurement for Irregular Variations**

METHODS FOR MEASUREMENT OF SECULAR TREND

The following are the principal methods of measuring trend from given time series:

1. GRAPHICAL OR FREE HAND CURVE METHOD

This is the simple method of studying trend. In this method the given time series data are plotted on graph paper by taking time on X-axis and the other variable on Y-axis. The graph obtained will be irregular as it would include short-run oscillations. We may observe the up and down movement of the curve and if a smooth freehand curve is drawn passing approximately to all points of a curve previously drawn, it would eliminate the short-run oscillations (seasonal, cyclical and irregular variations) and show the long-period general tendency of the data. This is exactly what is meant by **Trend**. However, It is very difficult to draw a freehand smooth curve and different persons are likely to draw different curves from the same data. The following points must be kept in mind in drawing a freehand smooth curve:

1. That the curve is smooth.
2. That the numbers of points above the line or curve are equal to the points below it.
3. That the sum of vertical deviations of the points above the smoothed line is equal to the sum of the vertical deviations of the points below the line. In this way the positive deviations will cancel the negative deviations. These deviations are the effects of seasonal cyclical and irregular variations and by this process they are eliminated.
4. The sum of the squares of the vertical deviations from the trend line curve is minimum. (This is one of the characteristics of the trend line fitted by the method of lest squares)

The trend values can be read for various time periods by locating them on the trend line against each time period.

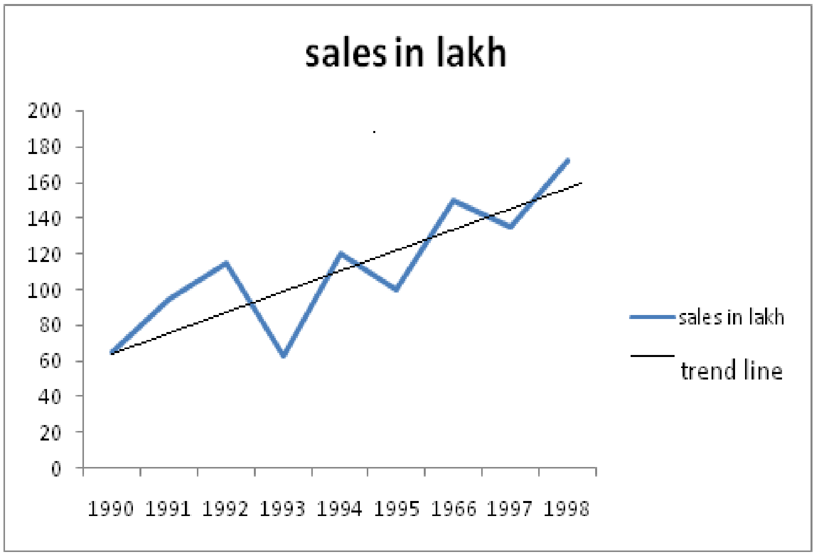
The following example will illustrate the fitting of a freehand curve to set of time series values:

Example:

The table below shows the data of sale of nine years:-

Year	1 990	1 991	1 992	1 993	1 994	1 995	1 996	1 997	1 998
Sales in (lakh units)	65	95	115	63	120	100	150	135	172

If we draw a graph taking year on x-axis and sales on y-axis, it will be irregular as shown below. Now drawing a freehand curve passing approximately through all this points will represent trend line (shown below by black line).



MERITS:

1. It is simple method of estimating trend which requires no mathematical calculations.
2. It is a flexible method as compared to rigid mathematical trends and, therefore, a better representative of the trend of the data.
3. This method can be used even if trend is not linear.
4. If the observations are relatively stable, the trend can easily be approximated by this method.
5. Being a non mathematical method, it can be applied even by a common man.

DEMERITS:

1. It is subjective method. The values of trend, obtained by different statisticians would be different and hence, not reliable.

2. Predictions made on the basis of this method are of little value.

2. METHOD OF SELECTED POINTS

In this method, two points considered to be the most representative or normal, are joined by straight line to get secular trend. This, again, is a subjective method since different persons may have different opinions regarding the representative points. Further, only linear trend can be determined by this method.

3. METHOD OF SEMI-AVERAGES

Under this method, as the name itself suggests semi-averages are calculated to find out the trend values. By semi-averages is meant the averages of the two halves of a series. In this method, thus, the given series is divided into two equal parts (halves) and the arithmetic mean of the values of each part (half) is calculated. The computed means are termed as semi-averages. Each semi-average is paired with the centre of time period of its part. The two pairs are then plotted on a graph paper and the points are joined by a straight line to get the trend. It should be noted that if the data is for even number of years, it can be easily divided into two halves. But if it is for odd number of years, we leave the middle year of the time series and two halves constitute the periods on each side of the middle year.

MERITS:

1. It is simple method of measuring trend.
2. It is an objective method because anyone applying this to a given data would get identical trend value.

DEMERITS:

1. This method can give only linear trend of the data irrespective of whether it exists or not.
2. This is only a crude method of measuring trend, since we do not know whether the effects of other components are completely eliminated or not.

4. METHOD OF MOVING AVERAGE

This method is based on the principle that the total effect of periodic variations at different points of time in its cycle gets completely neutralized, i.e. $\sum S_t = 0$ in one year and $\sum C_t = 0$ in the periods of cyclical variations.

In the method of moving average, successive arithmetic averages are computed from overlapping groups of successive values of a time series. Each group includes all the observations in a given time interval, termed as the **period of moving average.** The next group is obtained by replacing the oldest value by the next value in the series. The averages of such groups are known as the moving averages. The moving averages of a group are always shown at the centre of its period.

The process of computing moving averages smoothens out the fluctuations in the time series data. It can be shown that if the trend is linear and the oscillatory variations are regular, the moving average with the period equal to the period of oscillatory variations would get minimized because the average of a number of observations must lie between the smallest and the largest observation. It should be noted that the larger is the period of moving average the more would be the reduction in the effect of random components but the more information is lost at the two ends of data.i.e. it reduces the curvature of curvi-linear trends.

When the trend is non linear, the moving averages would give biased rather than the actual trend values.

Suppose that the successive observations are taken at equal intervals of time, say, yearly are Y_1, Y_2, Y_3 , . . .

Moving Average when the period is Odd

Now by a three-yearly moving averages, we shall obtain average of first three consecutive years (beginning with the second year) and place it against time t=2; then the average of the next three consecutive years (beginning with the second year) and place it against time t=3, and so on. This is illustrated below:

Time (t)	Observations Y_t	Moving Total	Moving Average (3 Years)
1	Y_1		
2	Y_2	$Y_1 + Y_2 + Y_3$	⅓ $(Y_1 + Y_2 + Y_3)$
3	Y_3	$Y_2 + Y_3 +$	⅓ $(Y_2 + Y_3 +$

		Y_4	Y_4)
4	Y_4	$Y_3 + Y_4 + Y_5$	⅓ ($Y_3 + Y_4 + Y_5$)
5	Y_5		

It should be noted that for odd period moving average, it is not possible to get the moving averages for the first and the last periods.

Moving Average when the period is Even

For an even order moving average, two averaging processes are necessary in order to centre the moving average against periods rather than between periods. For example , for a four –yearly moving average we shall first obtain the average Y_1=1/4($Y_1 + Y_2 + Y_3 + Y_4$) of the first four years and place it in between t =2 and t=3 then the average Y_2 = 1/4($Y_2 + Y_3 + Y_4 + Y_5$) of the next four years is and place it in between t=3 and t=4 , and finally obtain the average ½($Y_1 + Y_2$) of the two averages and place it against time t=3. Thus the moving average is brought against time or period rather than between periods. The same procedure is repeated for further results. This is tabulated below:

T ime	Observations Y_t	Moving Average for	Centered Value

(t)		4-period	
1	Y_1		
2	Y_2		
	→	$\frac{1}{4}(Y_1 + Y_2 + Y_3 + Y_4) = \mathbf{A_1}$	
3	Y_3	→	$\frac{1}{2}(\mathbf{A_1} + \mathbf{A_2})$
	→	$\frac{1}{4}(Y_2 + Y_3 + Y_4 + Y_5) = \mathbf{A_2}$	
4	Y_4		

It should be noted that when the period of moving average is even, the computed average will correspond to the middle of the two middle most periods.

MERITS:

1. This method is easy to understand and easy to use because there are no mathematical complexities involved.
2. It is an objective method in the sense that anybody working on a problem with the method will get the same trend values. It is in this respect better than the free hand curve method.
3. It is a flexible method in the sense that if a few more observations are added, the entire calculations are not changed. This not with the case of semi-average method.

4. When the period of oscillatory movements is equal to the period of moving average, these movements are completely eliminated.
5. By the indirect use of this method, it is also possible to isolate seasonal, cyclical and random components.

DEMERITS:

1. It is not possible to calculate trend values for all the items of the series. Some information is always lost at its ends.
2. This method can determine accurate values of trend only if the oscillatory and the random fluctuations are uniform in terms of period and amplitude and the trend is, at least, approximately linear. However, these conditions are rarely met in practice. When the trend is not linear, the moving averages will not give correct values of the trend.
3. The trend values obtained by moving averages may not follow any mathematical pattern i.e. fails in setting up a functional relationship between the values of X(time) and Y(values) and thus, cannot be used for forecasting which perhaps is the main task of any time series analysis.
4. The selection of period of moving average is a difficult task and a great deal of care is needed to determine it.
5. Like arithmetic mean, the moving averages are too much affected by extreme values.

5. METHOD OF LEAST SQUARES

This is one of the most popular methods of fitting a mathematical trend. The fitted trend is termed as the best in the sense that the sum of squares of deviations of observations, from it, is minimized. This method of Least squares may be used either to fit linear trend or a non-linear trend (Parabolic and Exponential trend).

FITTIG OF LINEAR TREND

Given the data ($Y_{t,}$ t) for n periods, where t denotes time period such as year, month, day, etc. We have to the values of the two constants, ‘a’ and ‘b’ of the linear trend equation:

$$Y_t = a + bt$$

Where the value of ‘a’ is merely the Y-intercept or the height of the line above origin. That is, when X=0, Y= a. The other constant ‘b’ represents the slope of the trend line. When b is positive, the slope is upwards, and when b is negative, the slope is downward.

This line is termed as the line of best fit because it is so fitted that the total distance of deviations of the given data from the line is minimum. The total of deviations is calculated by squaring the difference in trend value and actual value of variable. Thus, the term “Least Squares” is attached to this method.

using least square method, the normal equation for obtaining the values of a and b are :

$$\sum Y_t = na + b\sum_t$$

$$\sum tY_t = a\sum t + b\sum t^2$$

Let X = t – A, such that ∑X = 0, where A denotes the year of origin.

The above equations can also be written as

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

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

Since ∑x = 0 i.e. deviation from actual mean is zero

We can write

$$\mathbf{a = \sum Y / n}$$

$$\mathbf{b = \sum XY / \sum X^2}$$

FITTING OF PARABOLIC TREND

The mathematical form of a parabolic trend is given by:

$$Y_t = a + bt + ct^2$$

Here a, b and c are constants to be determined from the given data.

Using the method of least squares, the normal equations for the simultaneous solution of a, b and c are:

$$\sum Y = na + b\sum t + c\sum t^2$$

$$\sum tY = a\sum t + b\sum t^2 + c\sum t^3$$

$$\sum t^2Y = a\sum t^2 + b\sum t^3 + c\sum t^4$$

By selecting a suitable year of origin, i.e. define X = t – origin such that $\sum X = 0$, the computation work can be considerably simplified. Also note that if $\sum X = 0$, then $\sum X^3$ will also be equal to zero. Thus, the above equations can be rewritten as:

$$\sum Y = na + c\sum X^2 \quad \text{..........(1)}$$

$$\sum XY = b\sum X^2 \quad \text{..........(2)}$$

$$\sum X^2Y = a\sum X^2 + c\sum X^4 \text{..........(3)}$$

From equation (2), we get

$$\mathbf{b = \sum XY / \sum X^2}$$

From equation (1), w

$$a = \sum Y - c\sum X^2 /n$$

And from equati

$$c = n\sum X^2Y – (\sum X^2) (\sum Y) /n\sum X^4 – (\sum X^2)^2$$

or

$$c = \sum X^2Y - a\sum X^2 /\sum X^4$$

This are the three equations to find the value of constants a, b and c.

FITTING OF EXPONENTIAL TREND

The general form of an exponential trend is:

$$Y = a.b^t$$

Where ‘a’ and ‘b’ are constants to be determined from the observed data.

Taking logarithms of both side, we gave log Y = log a + log b.

This is linear equation in log Y and t can be fitted in a similar way as done in case of linear trend. Let A=log a and B =log b, then the above equation can be written as:

$$\log Y = A + Bt$$

The normal equations based on the principle of least squares are:

$$\sum \log Y = n A + B \sum t$$

$$\text{And } \sum t \log Y = n \sum t + B \sum t^2$$

By selecting a suitable origin, i.e. defining X = t – origin such that ∑X = 0, the computation work can be simplified. The values of A and B are given by:

$$\mathbf{A = \sum \log Y / n}$$

And

$$\mathbf{B = \sum X \log Y / \sum X^2}$$

Thus, the fitted trend equation can be written as:

$\log Y = A + BX$ **or**

$Y = \text{Antilog}\,[A + BX]$

$= \text{Antilog}\,[\log a + X \log b]$

$= \text{Antilog}\,[\log a.b^x]$

$= a.b^x$

MERITS:

1. Given the mathematical form of the trend to be fitted, the least squares method is an objective method.
2. Unlike the moving average method, it is possible to compute trend values for all the periods and predict the value for a period lying outside the observed data.
3. The results of the method of least squares are most satisfactory because the fitted trend satisfies the two most important properties, i.e. (1) $\sum(Y_0 - Y_t) = 0$ and (2) $\sum(Y_0 - Y_t)^2$ is minimum. Here Y_0 denotes the observed values and Y_t denotes the calculated trend value.

The first property implies that the position of fitted trend equation is such that the sum of deviations of observations above and below this equal to zero. The second property implies that the sums of squares of deviations of observations, about the trend equations, are minimum.

DEMERITS:

1. As compared with the moving average method, it is cumbersome method.
2. It is not flexible like the moving average method. If some observations are added, then the entire calculations are to be done once again.

3. It can predict or estimate values only in the immediate future or the past.
4. The computation of trend values, on the basis of this method, doesn’t take into account the other components of a time series and hence not reliable.
5. Since the choice of a particular trend is arbitrary, the method is not, strictly, objective.
6. This method cannot be used to fit growth curves, the pattern followed by the most of the economic and business time series.

MEASUREMENT OF SEASONAL VARIATIONS

The measurement of seasonal variations is done by isolating them from other components of a time series. There are four methods commonly used for the measurement of seasonal variations. These methods are:

1. Method of Simple Average
2. Ratio to Trend Method

3. Ratio to Moving Average Method
4. Method of link Relatives

1. METHOD OF SIMPLE AVERAGE

This is the easiest and the simplest method of studying seasonal variations. This method is used when the time series variable consists of only the seasonal and random components. The effect of taking average of data corresponding to the same period (say first quarter of each year) is to eliminate the effect of random component and thus, the resulting averages consist of only seasonal component. These averages are then converted into seasonal indices. It involves the following steps:

If figures are given on a monthly basis:

1. Average the raw data monthly year wise.
2. Find the sum of all the figures relating to a month. It means add all values of January for all the years. Repeat the process for all the months.
3. Find the average of monthly figures i.e., divide the monthly total by the number of years. For example if the data for 5 years (on monthly basis is available) there will be five figures for January. These have to be totaled and divided by five to get the average figures for January. Get such figures for all months. They may be X_1, X_2, X_3.......... X_{12}.
4. Obtain the average of monthly averages by dividing the sum of averages by 12 or

X1+X2+X3+...+X1212=X

5. Taking the average of monthly average as 100 find out the percentages of monthly averages. For the average of January (X_1) this percentage would be:
   average of januaryaverage of monthly averages×100

Or

X1X×100

If instead of the averages the monthly totals are taken into the account the result would be the same.

MERITS AND DEMERITS

This is a simplest method of measuring seasonal variations. However this method is based on the unrealistic assumption that the trend and cyclical variations are absent from the data.

2. RATIO TO TREND METHOD

This method is used when then cyclical variations are absent from the data, i.e. the time series variable Y consists of trend, seasonal and random components.

Using symbols, we can write **Y = T. S .R**

Various steps in the computation of seasonal indices are:

1. Obtain the trend values for each month or quarter, etc, by the method of least squares.
2. Divide the original values by the corresponding trend values. This would eliminate trend values from the data.

3. To get figures in percentages, the quotients are multiplied by 100.

Thus, we have three equations:

$$Y/T \times 100$$

$$T.S.R/T \times 100$$

$$S.R \times 100$$

MERITS AND DEMERITS

It is an objective method of measuring seasonal variations. However, it is very complicated and doesn’t work if cyclical variations are present.

3.RATIO TO MOVING AVERAGE METHOD

The ratio to moving average is the most commonly used method of measuring seasonal variations. This method assumes the presence of all the four components of a time series. Various steps in the computation of seasonal indices are as follows:

1. Compute the moving averages with period equal to the period of seasonal variations. This would eliminate the seasonal components and minimize the effect of random component. The resulting moving averages would consist of trend, cyclical and random components.

2. The original values, for each quarter (or month) are divided by the respective moving average figures and the ratio is expressed as a percentage, i.e. SR” = Y / M. A = TCSR / TCR’, where R’ and R” denote the changed random components.
3. Finally, the random component R” is eliminated by the method of simple averages.

MERITS AND DEMERITS

This method assumes that all the four components of a time series are present and, therefore, widely used for measuring seasonal variations. However, the seasonal variations are not completely eliminated if the cycles of these variations are not of regular nature. Further, some information is always lost at ends of the time series.

4. LINK RELATIVES METHOD

This method is based on the assumption that the trend is linear and cyclical variations are of uniform pattern. The link relatives are percentages of the current period (quarter or month) as compared with the previous period. With the computations of the link relatives and their average, the effect of cyclical and the random components is minimized. Further, the trend gets eliminated in the process of adjustment of chain relatives.

The following steps are involved in the computation of seasonal indices by this method:

1. Compute the Link Relative (L.R.) of each period by dividing the figure of that period with the figure of previous period. For example,
   Link relative of 3rd quarter=figure of 3rd quarter / figure of 2nd quarter ×100.
2. Obtain the average of link relatives of a given quarter (or month) of various years. A.M. or M_d can be used for this purpose. Theoretically, the later is preferable because the former gives undue importance to extreme items.
3. These averages are converted into chained relatives by assuming the chained relative of the first quarter (or month) equal to 100. The chained relative (C.R.) for the current period (quarter or month)
   = C.R. of the previous period ×L.R. of the current period / 100.

4. Compute the C.R. of the first quarter (or month) on the basis of the last quarter (or month). This is given by
   C.R. of the last quarter (month) × average L.R. of the first quarter (month) / 100
   a. This value, in general, is different from 100 due to long term trend in the data. The chained relatives, obtained above, are to be adjusted for the effect of this trend. The adjustment factor

d=14new C.R for 1st quater-100for quaterly data
d=112new C.R for 1st month-100for monthly data

b. On the assumption that the trend is linear d, 2d, 3d, etc, is respectively subtracted from the 2^{nd}, 3^{rd}, 4^{th}, etc quarter (or month).

5. Express the adjusted chained relatives as a percentage of their average to obtain seasonal indices.
6. Make sure that the sum of these indices is 400 for quarterly data and 1200 for monthly data.

MERITS AND DEMERITS

This method is less complicated than the ratio to moving average and the ratio to trend methods. However, this method is based upon the assumption of a linear trend which may not always hold true.

MEASUREMENT OF CYCLICAL VARIATIONS

A satisfactory method for the direct measurement of cyclical variations is not available. The main reason for this is that although these variations may be recurrent, these are seldom found to be of similar pattern having same period and amplitude of oscillations. Moreover, in most of the cases these variations are so intermixed with random variations that it is very difficult, if not impossible, to separate them.

The cyclical variations are often obtained, indirectly, as a residue after the elimination of other components. Various steps of the method are given as below:

1. Compute the trend values (T) and the seasonal indices(S) by appropriate methods. Here S is obtained as a fraction rather than the percentage.
2. Divide Y-values by the product of trend and seasonal index. This ratio would consist of cyclical and random component, i.e. C. R = Y / T. S
3. If there are no random variations in the time series, the cyclical variations are given by the step (2) above. Otherwise the random variations should be smoothened out by computing moving averages of C. R. values with appropriate period. Weighted moving average with suitable weights may also be used, if necessary, for this purpose.

MEASUREMENT OF RANDOM VARIATIONS

The random variations are also known as irregular variations. Because of their nature, it is very difficult to devise a formula for their direct computation. Like the cyclical variations, this component can also be obtained as a residue after eliminating the effects of other components.

BIBLIOGRAPHY:

BOOKS	AUTHORS
Business statistics	R S Bhardwaj
Fundamentals of Statistics	D N Elhance
Elements of business statistics	V K Puri
Mathematics and Statistics	Ajay Goel &Asha Goel