Unit IV

TIME SERIES

1. A Time Series is defined as a set of observations arranged in Chronological order.
2. An arrangement of Statistical data in accordance with time of occurrence or in a chronological order is called a time series.
3. The numerical or qualitative data which we get at different points of time – the set of observations – is known as time series.
4. Statistically, a Time Series is defined by the values $Y_1, Y_2, \ldots\ldots, Y_n$ of a variable Y at times $t_1, t_2, \ldots\ldots, t_n$. Here Y is a function of time t and Y_t denotes the value of the variable Y at time t.

Examples of Time Series:

i. The Annual Production of steel in India over the last 10 years
ii. The Monthly sales of sugar for the last 6 months
iii. The daily closing price of a share in the Exchange
iv. Hourly temperature recorded by the Meterological office in a city
v. Yearly Price or Quantity Index Number

According to Morris Hamburg- "*A time series is a set of observations arranged in Chronological Order*"

According to Kenny and Keeping– "*A set of data depending on the time is called time series*"

According to Croxton and Cowden– "*A time series consists of data arranged Chronological*"

According to Patterson – " *Time Series consists of Statistical data which are collected, recorded or observed over successive increments*"

USES OF TIME SERIES

The time series analysis is useful in every sphere. It is very important in economics, business, state administration and planning, science, astronomy, sociology, biology, research work etc., because of the following reasons.

i. It helps in understanding past behavior and it will help in estimating the future behavior
ii. It helps in planning and forecasting. It is very essential in business and economics. With the help of time series we can prepare plans for the future; short range and long range estimates are also possible
iii. Comparison between data of one period with that of another period is possible
iv. We can evaluate the progress in any field of economic and business activity. For example, With the change in price level, we can know the change in the purchasing power of money
v. Seasonal, Cyclical, Secular trend of data is useful not only in economics but also to the businessman.

Thus time series analysis can foretell accurately the course of future event.

COMPONENTS OF THE TIME SERIES

The four components of time series are:

i. Secular Trend

ii. Seasonal Variation

iii. Cyclical Variation

iv. Irregular Variation

I **Secular Trend:**

- A time series data may show upward trend or downward trend for a period of years
- This may be to factors like increase in population, change in technological progress, large scale shift in customer tastes, etc.
- For example, population increases over a period of years
- prices increase over a period of years
- production of goods in the country increases over a period of years.These are examples of upward trend.
- The sales of a commodity may decrease over a period of years because of better products coming to market.
- This is an example of declining trend or downward trend.
- The increase or decrease in ther movements of a time series is called a **Secular Trend**.

II Seasonal Variations :

- Seasonal Variations are short-term fluctuation in a time series which occur periodically in a year.
- This continues to repeat year after year.
- The major factors that are responsible for the repetitive pattern of seasonal variations are weather conditions and customs of people.
- Series of monthly and quarterly data are ordinarily used to examine these seasonal variations.
- One can observe that in each year more ice cream are sold during summer season and very less during winter season
- More woolen clothes are sold during winter and very little in summer
- The sales in departmental stores is at peak during Deepavali, Christmas, Pongal months. This reflects the shopping customs of consumers.

III Cyclical Variations:

- Cyclical Variations are recurrent upward or downward movements in a time series but the period of cycle is greater than a year.
- These variations are not regular as seasonal variation.
- There are different types of cycles of varying length and size.
- The ups and downs in business activities are the effects of cyclical variation.
- A business cycle showing these oscillatory movements has to pass through four phases – prosperity, recession, depression and recovery. In a business, these four phases are completed by passing one to another in this order.

IV. Irregular Variation:

- Irregular Variations are called as random variation.
- Random fluctuations in time series that are short in duration, erratic in nature and follow no regularity in the occurrence pattern.
- These variations are also referred to as residual variations since by definition they represent what is left out in a time series after Trend, Cyclical and Seasonal Variations.
- Irregular fluctuations result due to the occurrence of unforeseen events like floods, earthquakes, wars etc.

MEASUREMENT OF TREND

The following are some of the methods of finding trend in a time series

i. Free – Hand (Or Graphical) Method

ii. Semi – Average Method

iii. Moving Average Method

iv. Method of Least Squares

I. FREE – HAND (OR GRAPHICAL) METHOD:

This is the easiest, simplest and the most flexible method of estimating secular trend. The procedure of obtaining a straight line trend by this method is given below:

1. Plot the time series on a graph
2. Examine carefully the direction of the trend based on the plotted information (dots)
3. Draw a straight line which will best fit to the data according to personal judgement. The line now shows the direction of the trend.

<u>Merits of Graphical method</u>:

i. It is the simplest, easiest and quicker method.

ii. It saves time and labour

iii. It is adaptable and flexible, and it can be used to describe all types of trend i.e., linear and non-linear

iv. Experienced statisticians can draw a free-hand line more accurately than a mathematician; this can widely be used in applied situations

v. It will help to understand the character of time series and we can use the appropriate mathematical trend.

<u>Demerits of Graphical method</u>:

i. It is highly subjective. It is subject to personal bias. The results depend upon the judgment of the person who draws the line. There may be different curves for different persons.

ii. It seems to be very simple. But it requires more time for a careful job.

iii. If it is not drawn by experienced person, then it is dangerous to use for forecasting purpose.

iv. It does not help us to measure trend

II. SEMI – AVERAGE METHOD:

In this method, the original data is divided into two equal parts and average are calculated for both the parts. These are averages are called Semi – averages.

For example, we can divide the 10 years 1983 to 1992 into two equal parts; from 1983 to 1987 and 1988 to 1992.

If the period is odd number of years, the values of the middle year is omitted say, from 1983 to 1993 we must omit the year 1988.

We can draw the line by a straight line by joining the two points of averages. By extending the line downwards or upwards we can get the intermediate values or we can predict the future.

Example : 1 (For Odd Number of Items)

Determine the trend of the following by using semi-average. Estimate the sales for the year 1984

Year:	1975	1976	1977	1978	1979	1980	1981	1982	1983
Sales (Rs.00,000)	18	24	26	28	33	36	40	44	48

Solution:

Here the number of years is 9. The two middle parts will be 1975 to 1978 and 1980 to 1983. The value for the year 1979 will be ignored.

Year	Actual Value (Rs.00,000)		Trend Values
1975	18		(22.2 – 3.6) = 18.6
1976	24	$\overline{x}_1$ = 96/4 = 24	(24 – 1.8) = 22.2
			Center = 24
1977	26		(24 + 1.8) = 25.8
1978	28		(25.8 + 3.6) = 29.4
1979	33		
1980	36		(40.2 – 3.6) = 36.6
1981	40	$\overline{x}_2$ = 168/4 = 42	(42 - 1.8) = 40.2
			Center = 42
1982	44		(42 + 1.8) = 43.8
1983	48		(43.8 + 3.6) = 47.4
1984			(47.4 + 3.6) = 51

The value of 24 will be plotted in the middle of the years 1976 and 1977. Similarly the value 42 will be plotted against the middle of the years 1981 and 1982.

$$\text{Annual Increment} = (\overline{x}_2 - \overline{x}_1) / 5 = (42-24)/5 = 3.6$$

Again, 3.6/2 = 1.8 is the half yearly increment.

The figure is arrived at by dividing the difference between two semi averages by 5 not 4 because the averages are for the period 1976-77 and 1981-82 the difference of which is 5 years. This is so because in the calculation of semi averages we had ignored the year 1979.

Example : 2 (For Odd Number of Items) Determine the trend by using Semi-Average

Year :	1982	1983	1984	1985	1986	1987
Sales :	10	12	11	16	15	20

Solution:

Year	Sales		Trend Values
1982	10		$(11 - 2) = 9$
1983	12	$\overline{x}_1 = 33/3 = 11$	Center = 11
1984	11		$(11 + 2) = 13$
1985	16		$(17 - 2) = 15$
1986	15	$\overline{x}_2 = 51/3 = 17$	Center = 17
1987	20		$(17 + 2) = 19$

Annual Increment = $(\overline{x}_2 - \overline{x}_1)/3 = (17-11)/3 = 2$

Merits of Semi – Average Method:

i. It is simple and easier to understand than moving average and least square method

ii. As the line can be extended both ways, we can get the intermediate values and predict the future values.

iii. As it does not depend upon personal judgement, everyone who applies this method will get the same trend line unlike the former method.

Demerits of Semi – Average Method:

i. Under this method, it has an assumption of linear trend whether such a relationship exists or not

ii. It is affected by the limitation of arithmetic mean

iii. This method is not enough for forecasting the future trend or for removing trend from original data

Iii MOVING – AVERAGE METHOD:

Moving average method is a simple device of reducing fluctuations and obtaining trend values with a fair degree of accuracy. In this method the average value of number of years (or months, weeks or days) is taken as the trend value for the middle point of the period of moving average. The process of averaging smoothes the curve and reduces the fluctuation.

The first thing to be decided in the method is the period of the moving average. What it means is to take a decision about the number of consecutive items whose average would be calculated each time.

Suppose it has been decided that the period of the moving average would be 5 years or months or weeks or days (as the case may be) then the arithmetic average of the first 5 items (number 1, 2, 3, 4 and 5) would be placed against item No. 3 and then the arithmetic average of item Nos, 2,3,4,5 and6 would be placed against item No.4. This process would be repeated till the arithmetic average of the last 5 items has been calculated.

Odd Period of Moving Average:

Steps for calculating odd number of years (3,5,7,9), e.g., calculation of three yearly moving average include the following steps:

1. Compute the value of first three years (1,2,3) and place the three year total against the middle year (i.e., 2nd year)
2. Leave the first year's value and add up the values of the next three years i.e., 2,3,4 and place the three-year total against the middle year i.e., 3rd year

3. This process must be continued until the last year's value is taken for calculating moving average.
4. The three-yearly total must be divided by 3 and placed in the next column. This is the trend value of moving average.

The formula calculating3 yearly moving average is as follows :

(a+b+c)/3, (b+c+d)/3, (c+d+e)/3

The formula for 5 yearly moving average is as follows :

(a+b+c+d+e)/5, (b+c+d+e+f)/5, (c+d+e+f+g)/5

Even Period of Moving Average:

If the period of moving average is 4, 6 or 8, it is an even number. The four-yearly total cannot be place against any year as the median 2.5 is between the second and the third year. So the total should be placed in between the 2nd and 3rd years. We must centre the moving average in order to place the moving average against the year.

The steps are:

i. Compute the values of the first four years and place the total in between the 2nd and the 3rd years

ii. Leave the first year value and compute the value of the next four years and place the total in between the 3rd and 4th year

iii. This process must be continued until the last year is taken into account

iv. Compute the first two four-year totals and place it against the middle year (i.e., 3rd year)

v. Leave the first four year total and compute the next four-year totals and place in the 4^{th} year.

vi. This method must be continued until all the four-yearly totals are computed

vii. Divide the above totals by 8 (because it is the total of the two four-yearly totals) and put in the next columns. This is the trend values.

IV LEAST SQUARE METHOD:

When the trend is linear the trend equation may be represented by $y = a + bt$ and the values of a and b for the line $y = a + bt$ which minimizes the sum of squares of the vertical deviations of the actual (Observed) values from the straight-line, are the solutions to the so called normal equations :

$$\sum y = na + b\sum t \quad \text{--------------} \ (1)$$

$$\sum yt = a\sum t + b\sum t^2 \quad \text{--------------} \ (2)$$

When n is the number of paired observations.

The normal equations are obtained by multiplying $y = a + bt$ by the coefficient of a and b, i.e., by 1 and t throughout and summing up.

1 LINEAR TREND:

When the trend is linear the trend equation may be represented by $y = a + bt$ and the values of a and b for the line $y = a + bt$ which minimizes the sum of squares of the vertical deviations of the actual (Observed) values from the straight-line, are the solutions to the so called normal equations :

$$\sum y = na + b\sum t \quad \text{--------------} \quad ①$$

$$\sum yt = a\sum t + b\sum t^2 \quad \text{--------------} \quad ②$$

When n is the number of paired observations.

The normal equations are obtained by multiplying $y = a + bt$ by the coefficient of a and b, i.e., by 1 and t throughout and summing up.

Case 1: When the Number of Years is Odd

We can use this method when we are given odd number of years. It is easy and is widely used in practice. If the number of items is odd, we can follow the following steps:

1. Denote time as the t variable and value as y
2. Middle year is assumed as the period of origin and find out deviations
3. Square the time deviations and find t^2
4. Multiply the given value of y by the respective deviation of t and find the total $\sum yt$
5. Find out the values of y; get $\sum y$
6. The value so obtained are placed in the 2 equations:

i. $\sum y = na + b\sum t$

ii. $\sum yt = a\sum t + b\sum t^2$; find out the value of a and b

7. The calculated values of a and b are substituted and the trend value of y are found for various values of t

When the number of years is odd the calculation will be simplified by taking the mid year as origin and one year as unit and in that case $\sum x = 0$ and the two normal equations take the form

$\sum y = na$;

$\sum yt = b\sum t^2$

And hence, $a = \sum y / n$, $b = \sum xy / \sum x^2$

Example:

Fit a straight line trend to the following data :

Year	1965	1966	1967	1968	1969	1970	1971
Gross ex-factory Value (Rs. Crores)	672	824	967	1204	1464	1758	2057

And estimate the Gross ex-factory values (Rs Crores) for the year 1975

Solution:

Let the straight-line trend is represented by the equation $y = a + bt$. The values of a and b shall be determined by solving the normal equation $\sum y = na + \sum t$ and $\sum yt = a\sum t + b\sum t^2$

Here, since the number of years is odd the mid-year, i.e., year 1968 is taken as origin and one year as unit.

Year	Gross ex-factory value (in crores) (y)	t = Year - 1968	t^2	yt
1965	672	-3	9	-2016
1966	824	-2	4	-1648
1967	967	-1	1	-967
1968	1204	0	0	0
1969	1464	1	1	1464
1970	1758	2	4	3516
1971	2057	3	9	6171
Total	$\sum y = 8946$	$\sum t = 0$	$\sum t^2 = 28$	$\sum yt = 6520$

Here n = 7

From normal equations,

$8946 = 7a + bx0$ or $8946 = 7a$ or $a=1278$

$6520 = a \times 0 + b \times 28$ or $6250 = 28b$ or $b = 232.9$

Therefore, the trend equation is $y = 1278 + 232.9t$ with origin at 1968 and t unit = 1 year

The value of t for 1975 will be 7.

Hence the estimate for the year 1975 is

$Y = 1278 + 232.9 \times 7 = 1278 + 1630.3 = 2908.3$ (Rs. Crores)

Case 2: When the Number of Years is Even

When the number of years is even the origin is placed in the midway between the two middle years and the unit is taken to be ½ year instead of one year. With this change of origin and scale we have again

$\sum x = 0$ and hence $a = \sum y / n$ and $b = \sum yt / \sum t^2$

Example :

Fit a straight line trend to the following data:

Year	1951	1952	1953	1954	1955	1956
Electricity generated (Million kw Hours	101	107	113	121	136	148

Solution:

Let the straight line trend be represented by $y = a + bt$. Here number of years is even, so we take the origin at the mid-point of 1953 and 1954 and t unit = ½ year.

Year	Electricity generated (y)	t = (year – 1953.5)/0.5	t^2	yt
1951	101	-5	25	-505
1952	107	-3	9	-321
1953	113	-1	1	-113
1954	121	1	1	121
1955	136	3	9	408
1956	148	5	25	740
Total	$\sum y = 726$	$\sum t = 0$	$\sum t^2 = 70$	$\sum yt = 330$

Normal equations,

$$\sum y = na + b\sum t$$

Or $726 = 6a + b \times 0$ or $726 = 6a$ or $a = 121$

$$\sum yt = a\sum t + b\sum t^2 \text{ or } 330 = a \times 0 + b \times 70 \text{ or } 330 = 70b \text{ or } b = 4.7$$

Therefore, the trend equation is $y = 121 + 4.7\ t$

with origin at the mid-point of 1953 & 1954 and T unit = ½ year.

The trend value of 1953, $t = -1$

$$Y = 121 + 4.71(-1) = 116.29$$