

SUBJECT NAME : APPLIED STATISTICS

SUBJECT CODE : CST54

UNIT : I

B.Sc. Statistics: Syllabus (CBCS)

PAPER - 8
APPLIED STATISTICS

Objective:

This course introduces the basic Statistical tools in time related Variables, economic variables. To enable the students understand index numbers and other Statistical tools applied to demographic and chorological data

UNIT - I

Time series - Concept - Components of time Series - Additive and multiplicative models - Measurement of trend - free hand method-semi average method-Moving average method - Least square method.

UNIT - II

Measurement of seasonal variations - Simple average method - Ratio to trend method - Ratio to moving average method - Link relative method - Variate Difference method.

UNIT - III

Index Numbers - uses, classification of index numbers - Problems in the construction of index numbers - Methods of constructing index numbers - Unweighted index numbers - weighted index numbers.

UNIT - IV

Quantity index numbers - Fixed and chain base index numbers - Optimum test for index numbers - Time reversal test - factor reversal test - cost of living index numbers.

UNIT - V

Demand Analysis Theory and analysis of consumer's demand Law of demand, Price elasticity of demand estimation of demand curves forms of demand functions - Demand and Supply utility and indifference maps determination of price and supply and demand

Books for Study:

1. Kapoor,V.K and Gupta,S.C (1978); Fundamentals of Applied Statistics, Sultan Chand & Sons.

Books for Reference:

1. Gupta, S.P (1999); Statistical Methods, Sultan & Sons, New Delhi.
2. Croxton, F.E & Cowdon, D.J. (1973); Applied general statistics, Prentice Hall
3. Mukhopadhyay P.(1999); Applied Statistics,New Central Book Agency Pvt.Ltd.,Calcutta.

APPLIED STATISTICS
REGISTER-V

D

UNIT I Time Series SUB. Code:
8ST54

Introduction:-

Arrangement of statistical data in chronological order i.e., in accordance with occurrence of time, is known as Time Series. Such series have a unique important place. Consumption in the field of Economic and Business Statistics since the series relating to Prices, Consumption and production of various commodities.

Money in circulation, bank deposits and bank clearings, sales and profits in a departmental store, agricultural and industrial production, national income and foreign exchange reserves etc., are all time series spread over a long period of time.

Definition:-

A time series may be defined as a collection of readings belonging to different time periods, of some economic variable or composite variables".

Mathematically, a time series is defined by the functional relationship.

$$U_t = f(t)$$

where U_t is the value of the phenomena

(or) variable under consideration at time

For example, (i) the population (U_t) of a country or a place in different years,

(ii) the number of births and deaths (U_t) in different months (t) of the year.

(iii) the sale (U_t) of a departmental store in different months (t) of the year.

(iv) the temperature (U_t) of a place on different days (t) of the week and so on.

Constitute time series.

Thus if the values of a phenomenon or variable at time t_1, t_2, \dots, t_n are U_1, U_2, \dots, U_n respectively, then the series.

$$t : t_1, t_2, \dots, t_n$$

$$U_t : U_1, U_2, \dots, U_n$$

Constitute a time series. Thus, a time series invariably gives a bivariate distribution, one of the two variables being time (t).

and the other being the value (U_t) of ⁽³⁾ the phenomenon at different points of time.

The values at t may be given yearly, monthly, weekly, daily or even hourly, usually but not always at equal intervals of time.

Components of time series:

The various forces at work, affecting the values of a phenomenon in a time series, can be broadly classified into the following four categories. Commonly known as the components of a time series, some or all of which are present in varying degrees.

- a) Secular Trend (or) Long-Term Movement
- b) Periodic Changes or Short-term Fluctuations
- c) Seasonal Variations (or) Cyclic Variations

c) Random or Irregular Movements

The value of a time series U_t at any time t is regarded as the resultant of the combined impact of above components.

In the following section we shall (6) briefly explain them one by one.

Trend:

By Secular trend or Simply trend we mean the general tendency of the data to increase or decrease during a long period of time. This is true of most of series of business or Economic Statistics.

For example:

An upward tendency would be seen in data pertaining to population, agricultural production, currency in circulation etc.

While a downward tendency will be noticed in data of births and deaths, epidemics, etc.

1. It may be clearly noted that trend is the general, smooth, long term, average tendency. It is not necessary that the increase or decline should be in the same direction throughout the given Period.

It may be possible that different tendencies of increase or decrease or stability are observed in different sections of time.

However, the overall tendency may be upward, downward or stable.

Such for example, the effect of population increase over a long period of time on the expansion of various sectors like agriculture, industry, education, textiles, etc., is a continuous but a gradual process. Similarly, the growth or decline in a number of economic time series is the interaction of forces like advances in production technology, large-scale production, improved marketing management and business organization, the invention and discovery of new natural resources and the exhaustion of the existing resources and so on.

2. It should not be inferred that all the series must show an upward (or) downward trend. We might come across certain series whose values fluctuate round a constant reading which does

not change with time e.g., the series of barometric readings or the temperature of a particular place.

3. Linear and Non-Linear (Curvi-linear)
Trend: If the time series values plotted on graph cluster more, or less, round a straight line, then the trend exhibited by the time series is termed as Linear. Otherwise Non-linear (curvi-linear). In a straight line trend, the time series values increase or decrease more or less by a constant absolute amount, i.e., the rate of growth (or decline) is constant.

Although, in practice, linear trend is commonly used, it is rarely obtained in economic and business data. In an economic and business phenomenon, the rate of growth or decline is not of constant nature throughout but varies considerably in different sectors of time.

4. The term 'long period of time' is a relative term and cannot be defined exactly. In some cases a period as small as a week may be fairly long while in some cases, a period as long as 2 years may not be enough.

For example, if the data of agricultural production for 24 months shows an increase it won't be termed as secular change over a period of 2 years whereas if the count of bacterial population of a culture every five minutes, for a week shows an increase, then we would regard it as a secular change.

Periodic changes.

It would be observed that in many social and economic phenomena apart from the growth factor in a time series there are forces at work which prevent the smooth flow of the series in a particular direction and tend or repeat themselves over a period of time.

These forces do not act continuously but operate in a regular spasmodic manner. The resultant effect of such forces may be classified as:

(i) Seasonal Variations and

(ii) Cyclic Variations

(i) Seasonal Variations.

These variations in a time series are due to the rhythmic forces which operate in a regular and periodic manner over a span of less than a year & during a period of 12 months and have the same or almost same pattern year after year.

Thus seasonal variations in a time series will be there if the data are recorded quarterly (every three months), monthly, weekly, daily, hourly and so on.

Thus, in a time series data where only annual figures are given, there are no seasonal variations. Most of economic time series are influenced by seasonal swings, i.e. prices, production and consumption of commodities.

Sales and profits in a departmental store; bank clearings and bank deposits etc. are all affected by seasonal variations.

The seasonal variations may be attributed to the following two causes:

(i) Those resulting from natural forces

As the name suggests, the various seasons or weather conditions and climate changes play an important role in seasonal movements.

For instance the sale of umbrellas pick up very fast in rainy season, the demand for electric

fans goes up in summer season, (v)
the sale of ice and ice-cream
increases very much in summer; the
sales of woollens go up in winter - all
being affected by natural forces, viz.,
weather or seasons.

Likewise, the production of
certain commodities such as sugar,
rice, pulses, eggs, etc. depends on season.

Now the prices of agricultural
commodities always go down at the time
of harvest and then pick up gradually.
(ii) Those resulting from man-made conventions

These variations in a time series
within a period of 12 months are due to
habits, fashions, customs and conventions of
the people in the Society. For instance, the
sale of jewellery and ornaments goes up
in marriages; the sales and profits in
departmental stores go up considerably

during marriages, and festivals like ⁽¹⁾ Diwali, Dussehra (Durga Pooja), Christmas etc. Such variations operate in a regular sporadic manner and recur year after year.

The main objective of the measurement of seasonal variation is to isolate them from the trend and study their effects. A study of the seasonal patterns is extremely useful to businessmen, producers, sales-managers etc., in planning future operations in formulation of policy decisions regarding finance, purchase, production, inventory control, personnel requirements, selling and advertising programmes.

(2) Cyclic Variations:-

The oscillatory movements in a time series with period of oscillation more than one year are termed as cyclic fluctuations. One complete period is called a 'cycle'.

The cyclic movements in a time series are generally attributed to the so called 'Business Cycle', which may also be referred to as the 'four-phase cycle' composed of prosperity, recession, depression and recovery and normally lasts from seven to eleven years.

Eg., series relating to prices, production and wages etc., are affected by business cycles. Cyclic fluctuations, though more or less regular, are not periodic.

3. Irregular component:

A part from the regular variations almost all the series contain another called the random or irregular or residual fluctuations which are not accounted for by secular trend and seasonal and cyclic variation.

These fluctuations are purely random erratic, unforeseen, unpredictable

and are due to numerous non-recurring and irregular circumstances which are beyond the control of human hand but at the same time are a part of our system such as Earthquakes, wars, floods, famines, revolutions, epidemics etc. These isolated or irregular but powerful fluctuations due to floods, revolution, political upheaval, famines etc., are also called episodic fluctuations.

Analysis of Time series:

The main problem in a time series analysis are:

- (i) To identify the forces or components at work, the net effect of whose interaction is exhibited by the movement of a time series, and,
- (ii) To isolate, study analyse and measure them independently, i.e., by holding other things constant.

Mathematical models for Time Series

The following are the two models ⁽¹⁾ commonly used for the decomposition of a time series into its components.

(i) Decomposition by additive hypothesis

According to the additive model, a time series can be expressed as

$$U_t = T_t + S_t + C_t + R_t \quad \text{--- (1)}$$

where U_t is the time series value at time t . T_t represents the trend value, S_t , C_t and R_t represent the seasonal, cyclic and random fluctuations at time t .

Obviously, the term S_t will not appear in a series of annual data.

As such C_t (and S_t) will have +ve or -ve values according as whether we are in an above normal or below normal phase of the cycle (and year) and the total of positive and negative values for any cycle (and any year) will be zero.

3.3 Decomposition by Multiplicative Model

(15)

On the other hand if we have reason to assume that the various components in a time series operate proportionately to the general level of the series, the traditional classical multiplicative model is appropriate. According to the multiplicative model

$$U_t = T_t \cdot S_t \cdot C_t \cdot R_t \quad \text{--- (15)}$$

where S_t , C_t and R_t instead of assuming positive and negative values are indices fluctuating above or below unity and the geometric means of S_t in a year, C_t in a cycle and R_t in a long-term period are unity.

In a time series with both positive and negative values, the multiplicative model eqn (15) can not be applied unless the time series is translated by adding a suitable positive values.

The logarithmic values of the original time series, i.e.,

$$\log U_t = \log T_t + \log S_t + \log C_t + \log R_t$$

In practice, most of the series relating to economic data conform to the multiplicative model $\text{eqn } \text{D}$.

2. Mixed models

In addition to the addition and multiplication models discussed above the components in a time series may be combined in a large number of other ways. The different models, defined under different assumptions will yield different results.

Some of the mixed models resulting from combinations of additive and multiplicative models are given below.

$$U_t = T_t C_t + S_t R_t$$

$$U_t = T_t + S_t C_t R_t$$

$$U_t = T_t + S_t + C_t R_t$$

Uses of Time Series:

The time series analysis is of great importance not only to businessman or an Economist but also to people working in various disciplines in natural, Social and Physical Sciences. Some of its uses are enumerated below.

1. It enables us to study the past behaviour of the phenomenon under consideration (12) i.e., to determine the type and nature of the variations in the data.
2. The segregation and study of the various components is of paramount importance to a businessman in the planning of future operations in the data.
3. It helps to compare the actual current performance of accomplishments with the expected ones (on the basis of the past performances) and analyse the causes of such variations, if any.
4. It enables us to compare the changes in the behaviour of the phenomenon in future which is very essential for business planning.
5. It helps us to compare the change in the values of different phenomenon at different times or places etc.

Measurement of Trend.

Trend can be studied and/or measured by the following methods.

- (i) Graphic Method (or) Trend by Inspection.

- (ii) Method of Semi Averages
- (iii) Method of Curve fitting by principles of least squares and
- (iv) Method of moving averages

(v) Graphic Method:

a) A free-hand smooth curve obtained on plotting the values y_t against t enables us to form an idea about the general 'trend' of the series. Smoothening of the curve eliminates other components viz. regular and irregular fluctuations.

b) The method is very subjective, i.e., the bias of the person handling the data plays a very important role and as such different trend curves will be obtained by different persons for the same set of data. As such trend by inspection should be attempted only by skilled and experienced statistician and this limits the utility and popularity of the method.

c) It does not enable us to measure trend.

(ii) Method of Semi-Averages

In this method, the whole data is divided into two parts with respect to time i.e., if we are given y_t for t from 1971-1983, i.e., over a period of 12 years, the two equal parts will be the data from 1971 to 1976, and 1977 to 1982.

In case of odd number of years the two parts are obtained by omitting the value corresponding to the middle year, i.e., for the data from 1971-81; the two parts would be the values for 1971-75 and 1977-81, the value corresponding to middle year, viz 1976 being omitted.

Next we compute the arithmetic mean for each part and plot these two averages against the mid-values of the respective periods covered by each part.

The line obtained on joining these two points is the required trend line and may be extended both ways to estimate intermediate or future values.

Merits:

1. Compared with graphic method, the obvious advantage of this method will be objectivity in the sense that everyone who applies it would get the same results.

2. It is readily comprehensible as compared to the method of least squares or the moving average method.

Limitations:

This method assumes linear relationship between the plotted points which may not exist. Moreover, the limitations of arithmetic mean as an average also stand in its way.

Example:

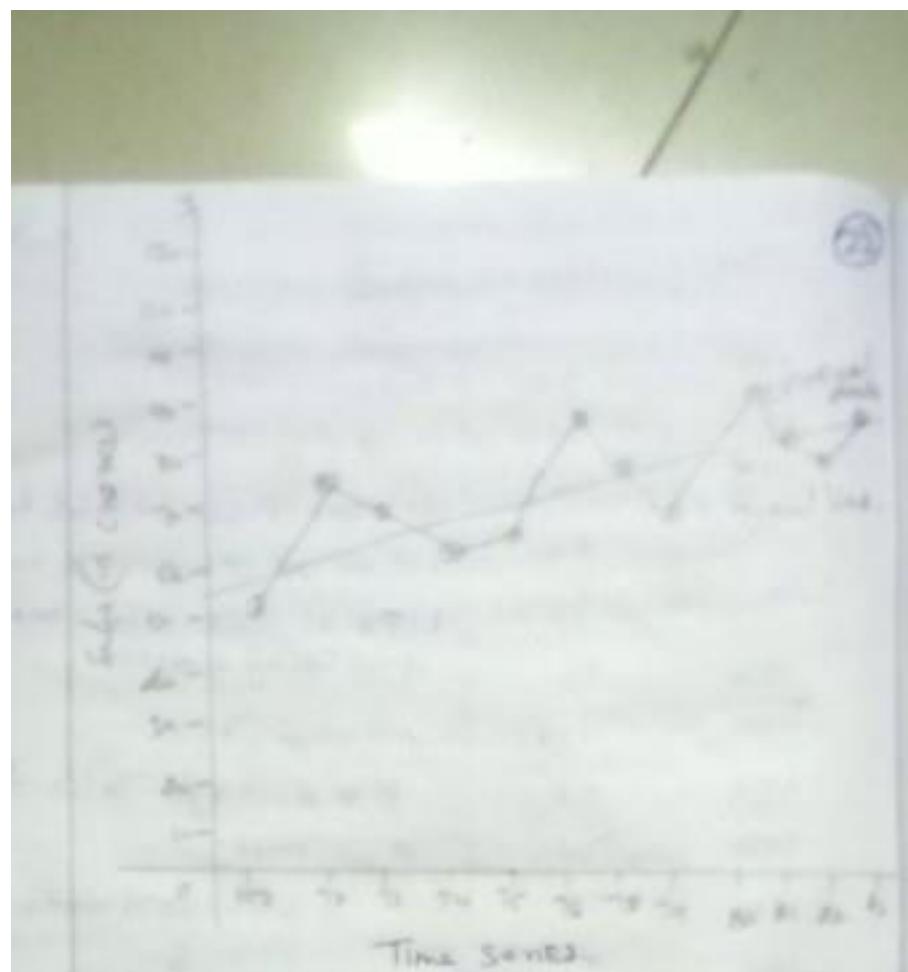
Fit a trend line to the following data by the method of semi averages.

Year: 1971 1972 1973 1974 1975 1976

Banc clearances (in crores) } 53 77 76 66 69 74
(Ex) (Ex)

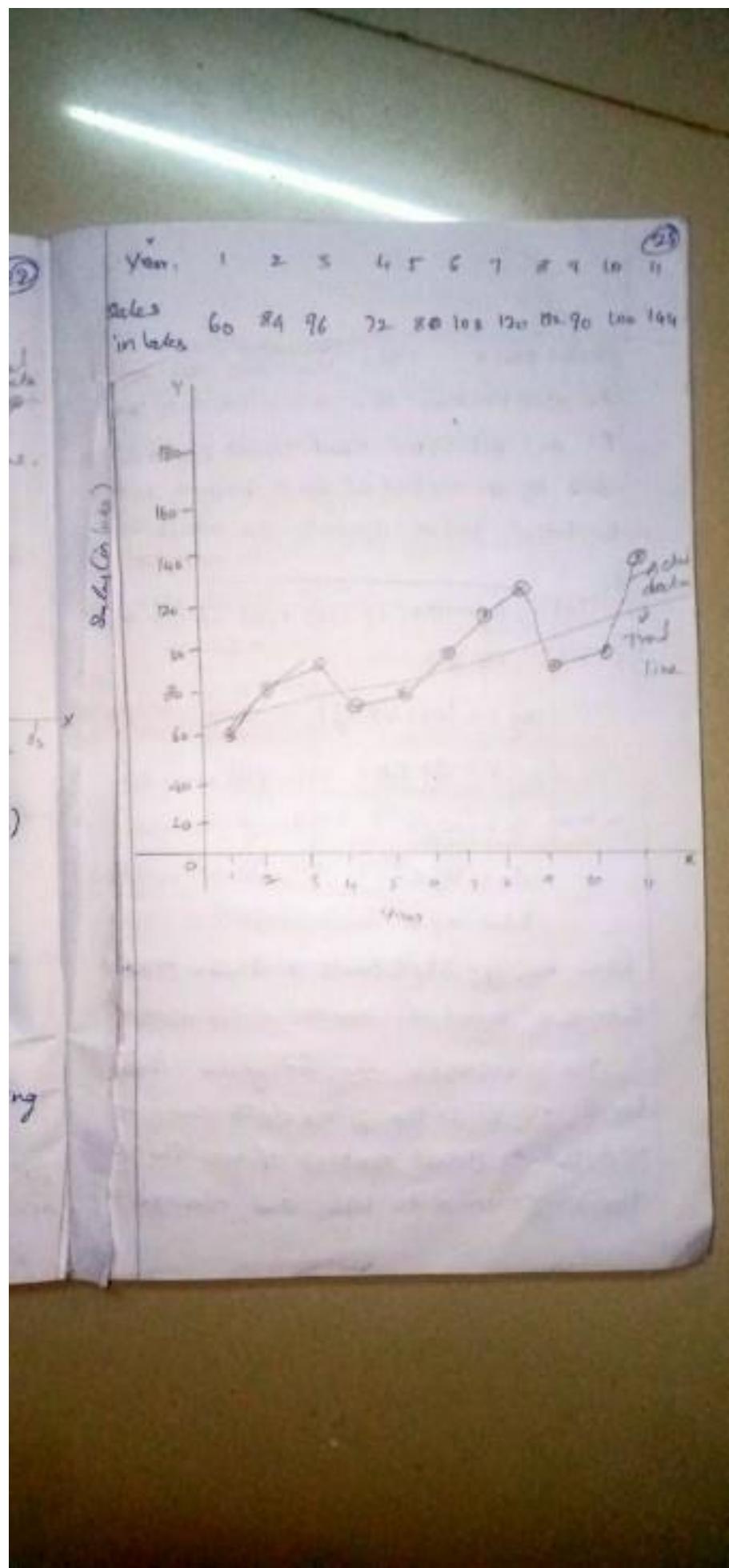
Year: 1977 1978 1979 1980 1981 1982 1983

Banc clearances: 105 87 79 104 99 92 101



1. Free hand smooth curve (Graphic method)
2. Method of semi average
3. Moving average method
4. Weighted moving average.
5. method of least square.

6. fit a freehand smooth curve representing the following data.



Method of curve fitting by principle of least squares.

The principle of least squares is the most popular and widely used method of fitting mathematical functions to a given set of data. The method yields very correct results if sufficiently good appraisal of the form of the function to be fitted.

A part from the usual arithmetic scales, semi-logarithmic or doubly-logarithmic scales may be used for the graphical representation of the data. The various types of curves that may be used to describe the given data in practice are.

(If U_t is the value of the variable corresponding to time t)

i) A straight line : $U_t = a + bt$

ii) Second degree parabola,

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

iii) k^{th} -degree polynomial :

$$U_t = a_0 + a_1 t + a_2 t^2 + \dots + a_n t^n$$

(2x)

(iv) Exponential curve: $U_t = ab^t$

$$\Rightarrow \log U_t = \log a + t \log b \\ = A + Bt \text{ (say)}$$

(v) Second degree curve fitted to logarithms

$$U_t = ab^t c^{t^2}$$

$$\Rightarrow \log U_t = \log a + t \log b + t^2 \log c \\ = A + Bt + Ct^2 \text{ (say).}$$

(vi) Growth curves:

(a) $U_t = a + bct$ (Modified Exponential curve)

(b) $U_t = ab^{ct}$ (Gompertz curve)

$$\Rightarrow \log U_t = \log a + ct \log b = A + Bc^t \text{ (say)}$$

$$(c) U_t = \frac{k}{1 + c^{at+bt}} \quad (\text{Logistic curve})$$

Note:

For deciding about the type of trend
to be fitted to a given set of data,

the following points may be helpful. (1)

(i) When the time series is found to be increasing or decreasing by equal absolute amounts, the straight line trend is used. In this case, the plotting of the data will give a straight line graph.

(ii) The logarithmic straight line exponential curve ($U_t = ab^t$) is used as an expression of the secular movement, when the series is increasing or decreasing by a constant percentage rather than a constant absolute amount. In this case, the data plotted on a semi-logarithmic scale will give a straight line graph.

(iii) Second degree curve fitted to logarithms may be tried for trend fitting if the data plotted on a semi-logarithmic scale is not a straight line.

graph but shows curvature, being
Concave either upward or downward. (29)

Fitting of straight line by least square
Method.

$$U_t = a + bt \quad \text{--- (1)}$$

Principle of least squares consists in minimizing
the sum of squares of the deviations between
the values of U_t and their estimated given
by eqn (1).

In other words, we have to find
 a and b such that for given values of U_t
corresponding to n different values of t ,

$$Z = \sum_t (U_t - a - bt)^2$$

is minimum. For a maxima or minima of Z ,
for variation in a and b , we should have

$$\begin{aligned} \frac{\partial Z}{\partial a} &= 0 = -2 \sum (U_t - a - bt) \\ \frac{\partial Z}{\partial b} &= 0 = -2 \sum t(U_t - a - bt) \\ \sum U_t &= na + bt \\ \sum tU_t &= a \sum t + b \sum t^2 \end{aligned} \quad \text{--- (2)}$$

Which are the normal equations for
estimating a and b . (2)

The values of ΣU_t , ΣtU_t , Σt , Σt^2
are obtained from the given data and the
equations (2) can now be solved for
 a and b . With these values of a and b ,
the line $a + bt$ gives the desired trend
line.

Note:

It has been shown that the solution
of normal equations (2) provides a
minima of Z . The proof is, however,
beyond the scope of this book.

Fitting of second Degree Parabola.

$$U_t = a + bt + ct^2 \quad \text{--- (3)}$$

Proceeding similarly as in the case of a
straight line, the normal equation of for
estimating a , b and c are give by

$$\begin{aligned}\sum U_t &= n\alpha + b\sum t + c\sum t^2 \\ \sum tU_t &= \alpha\sum t + b\sum t^2 + c\sum t^3 \\ \sum t^2 U_t &= \alpha\sum t^2 + b\sum t^3 + c\sum t^4\end{aligned} \quad \text{--- (4)}$$

Fitting of Exponential Curve:

$$U_t = ab^t \quad \text{--- (5)}$$

$$\Rightarrow \log U_t = \log a + t \log b$$

$$\Rightarrow y = A + Bt \quad (\text{say}) \quad \text{--- (6)}$$

where $y = \log U_t$.

$$A = \log a$$

$$B = \log b$$

Eqn (6) is a straight line in t and y
and thus the normal equations for
estimating A and B are

$$\begin{aligned}\sum y &= nA + B\sum t \\ \sum ty &= A\sum t + B\sum t^2\end{aligned} \quad \text{--- (7)}$$

These equations can be solved for A
and B and finally on using eqn (5), we get

$$a = \text{antilog}(A)$$

$$b = q \text{antilog}(B)$$

Merits and Limitations of Trend Fitting by principle of Least Squares: 20

Merits:

The method of least squares is the most popular and widely used method of fitting mathematical functions to a given set of observations. It has the advantages:

1. Because of its mathematical or analytical character, this method completely eliminates the element of subjective judgement or personal bias on the part of the investigator.

2. Unlike the method of moving averages, this method enables us to compute the trend values for all the given time periods in the series.

3. The trend equation can be used to estimate or predict the value of the variable for any period t in

future or even in the intermediate periods of the given series and the forecast values are also quite reliable.

4. The curve fitting by the principle of least squares is the only technique which enables us to obtain the rate of growth per annum, for yearly data, if linear trend is fitted.

Demerits:

1. The method is quite tedious and time-consuming as compared with other methods. It is rather difficult for a non-mathematical person to understand and use.

2. The addition of even a single new observation necessitates all calculation to be done afresh.

3. Future predictions or forecasts based on this method are based only on the long term variations, i.e., trend and completely ignore the cyclical, seasonal

and irregular fluctuations.

(ii)

4. The most serious limitation of the method is the determination of the type of the trend curve to be fitted, viz., whether we should fit a linear or a parabolic trend or some other more complicated trend curve.

5. It cannot be used to fit growth curves like Modified Exponential curve, Gompertz curve and Logistic curve to which most of the economic and business time series data conform.

Example: In a certain industry, the production of a certain commodity (in '000 tons) during the year 1973-83 is given in the following table.

Year:	1973	1974	1975	1976	1977
Production ('000 of ton)	66.6	84.6	88.6	78.0	96.8
	1978	1979	1980	1981	1982
	105.2	95.2	111.6	82.3	117.0

- (i) Graph the data (2)
- (ii) Obtain the Least Square line fitting the data and construct the graph of the trend line.
- (iii) Compute the trend values for the years 1973-83 and estimate the production of commodity during the years 1984 and 1985, if the present trend continues.
- (iv) Eliminate the trend.

Solution:

Here: $n = 11$.

2., odd and therefore, we shift the origin to the middle time period. 1978

$$\text{Let } t = x - 1978$$

Year (x)	Production (U _t)	t	tU _t	t ²	Trend Values U _t \hat{U}_t (24)
1973	66.6	-5	-333.0	25	75.74
1974	84.6	-4	-338.4	16	79.69
1975	88.6	-3	-265.8	9	83.64
1976	78.0	-2	-156.0	4	87.69
1977	96.8	-1	-96.8	1	91.54
1978	105.2	0	0	0	95.49
1979	93.2	1	93.2	1	99.44
1980	111.6	2	223.2	4	103.39
1981	98.3	3	264.9	9	107.39
1982	117.0	4	468.0	16	111.29
1983	115.2	5	576.0	25	115.24
Total	1050.4	0	434.1	110	

The least square line of U_t can be

$$U_t = a + bt$$

The normal equation for estimating a and b

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

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

$$\alpha = \frac{\sum U_t}{n}$$

$$= \frac{1050.4}{11}$$

$$\alpha = 95.49$$

$$b = \frac{\sum tU_t}{\sum t^2}$$

$$= \frac{434.19110}{285}$$

$$b = 3.95$$

Hence the least square line fitting the data is

$$U_t = 95.49 + 3.95t$$

Trend values for the years 1973 to 1983 are obtained on putting $t = -5, -4, \dots, 5$ respectively.

Estimate for 1984. Taking $t = 1984$ in

$$t = 1984 - 1978$$

$$= 6.$$

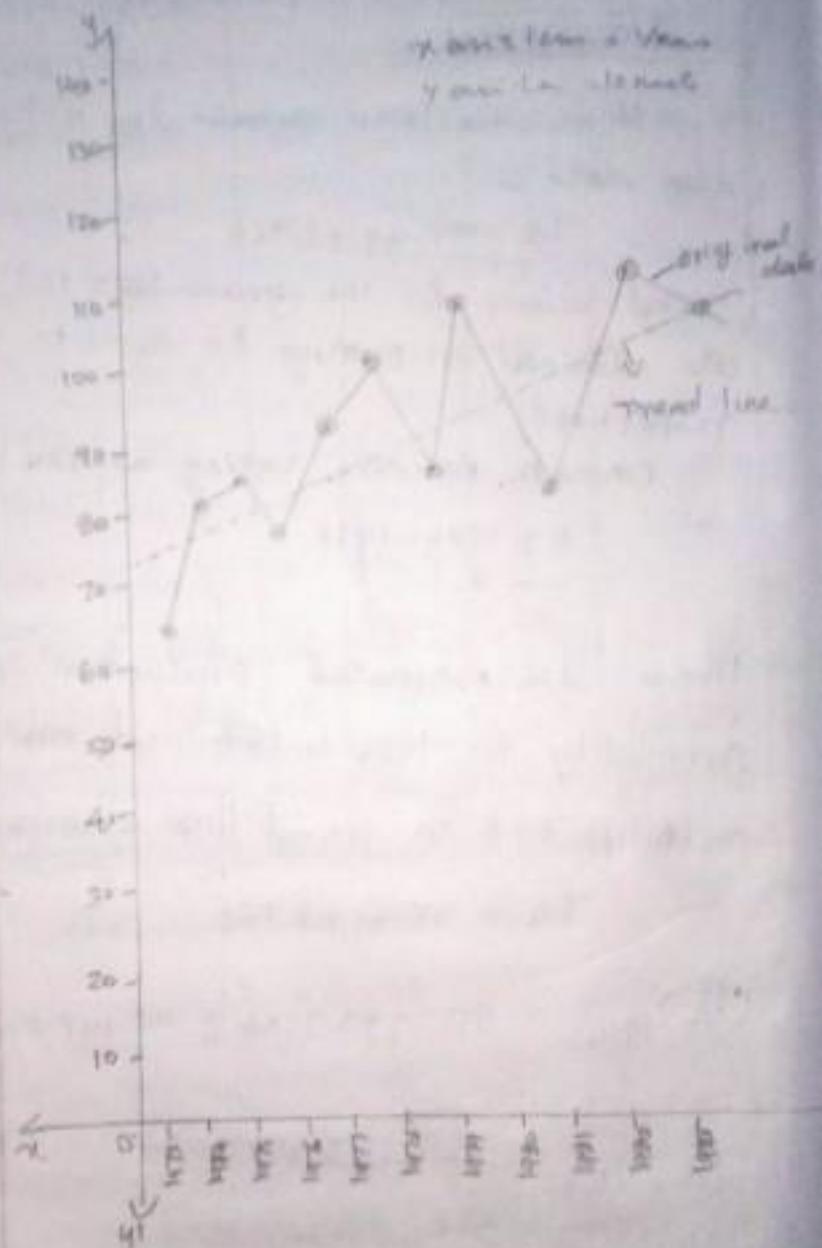
Hence the estimated production of the commodity for 1984 is obtained on putting $t = 6$ in trend line equation.

$$U_t = 95.49 + 3.95t$$

$$(U_t)_{1984} = 95.49 + 3.95 \times 6 \Rightarrow 119.14 \text{ (Ans)}$$

$$\text{Ans} \\ (U_t)_{1985} = 95.49 + 3.95 \times 7 \Rightarrow 123.14 \text{ (Ans)}$$

The graph of the original data and
the trend line is given below. (3)



Assuming multiplicative model, the trend values are eliminated on dividing the given value (U_t) by the trend values U_t .

If $(U_t - U_t)$ The resulting values contain short-term (seasonal and cyclic) variation and irregular variations.

ELIMINATION OF TREND

YEAR	Trend eliminated values based on	
	Additive model ($U_t - U_t$)	Multiplicative model (U_t/U_t)
1973	$66.6 - 74.74 = -8.14$	$66.6/74.74 = 0.88$
1974	$84.9 - 79.69 = 5.21$	$84.9/79.69 = 1.065$
1975	$58.6 - 53.64 = 4.96$	1.059
1976	-9.59	0.891
1977	4.26	1.102
1978	9.71	0.937
1979	6.24	1.079
1980	8.21	0.928
1981	-9.04	1.051
1982	45.71	0.999
1983	-0.04	

Home work for students:

1. Fit a straight line trend by the method of least squares of the following data
Assuming that the same rate of change continues, what would be predicted earnings for the year 1985?

Year: 1976 1977 1978 1979 1980 1981 82

Sales 176 80 130 144 133 120 174 190
Watch ad: 76 80 130 144 133 120 174 190

Ans: $n = 9$

$t = 2x - 3959.$

$a = 181.5$

$b = 7.33$

$y_t = 181.5 + 7.33t.$

Example:

2. Below are give the figures of Production (in thousand quintals) of sugar factory.

Year: 1973 1975 1976 1977 1978 1979 1980

Production 77 88 94 85 91 98 100

(i) Fit a straight line by the 'Least Squares method' and tabulate the trend values

(ii) Eliminate the trend. What component

of the five series are thus left over (2)
 Q) What is the monthly increase in the
 production of sugar.

Solution.

Year	Production	t	t ¹	t ²	Total value	Million of Pounds
1973	77	-4	-304	16	85.82	-6.52
1974	82	-3	-276	9	86.66	+1.84
1975	94	-2	-91	4	87.63	+1.87
1976	85	0	0	0	88.60	-3.00
1977	91	1	91	1	90.17	+1.81
1978	83	2	196	4	91.04	+1.91
1979	90	3	650	9	92.65	+5.65
Total	623	-1	159	45	692.97	

Let the trend equation be $y_t = a + bt$

Normal equation for estimating a and b

Q/W

$$\sum U_t = n a + b \sum t$$

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

$$a = \frac{\sum U_t}{n} = 62.3/7 = 88.80$$

$$(a = 88.80)$$

$$b = \frac{\sum t U_t}{\sum t^2}$$

$$= 159/45$$

$$(b = 1.37)$$

∴ Trend equation is $U_t = 88.8 + 1.37t$

∴ Substituting the values of t , $U_2 = 88.8$
etc.

i.e. trend values for 1975 is

$$U_6 = 88.8 + 1.37(-4)$$
$$= 83.32$$

Hence trend values for 1974 is

$$U_5 = 88.8 + 1.37(-2)$$
$$= 86.06 \text{, and so on.}$$

(ii) Assuming additive model for the time series, the trend values are eliminated by subtracting them from the given values, as shown in the table.

The resulting values show the short-term fluctuations which change with a period more than one year.

(iii) Yearly increase in the production of sugar, as provided by linear trend

$$U_t = a + bt \quad \therefore b = 1.37$$

$$\therefore \text{Monthly Increase} = 1.37/12$$
$$= 0.114 \text{ thousand quintal}$$

Moving Average method:

(67)

It consists in measurement of trend by smoothing out the fluctuation of the data by means of a moving average.

In this method, the main problem which is of paramount importance lies in determining the extent (or the period) of the moving average which will completely eliminate the oscillatory movements affecting the series.

It has been established mathematically, that if the fluctuations are regular and period then the moving average completely eliminates the oscillatory movements provided

(i) the extent of moving average is exactly equal to the period of oscillation

(ii) the trend is linear since different cycles vary in amplitude and period in such case the appropriate

period of moving average should be equal to or somewhat greater than the mean period of the cycles in the data.

Moving average method is very flexible in the sense that the addition of a few more figures to the data simply results in some more trend values; the previous calculation are not affected at all.

The moving average method has the following drawbacks,

- (i) It does not provide trend values for all the terms e.g., for a moving average of extent $2k+1$, we have to forego the trend values for the first k and the last k terms of the series.
- (ii) It cannot be used for forecasting or predicting future trend which is the main objective of trend analysis.

To summarise moving average method
gives a correct picture of the long term trend of the series if

- i) The trend is linear or approximately linear,
- ii) oscillatory movement affecting the data are regular in period and amplitude.

Example:

A study of demand (d_t) for the past 12 years ($t=1, 2, \dots, 12$) has indicated the following:

$$d_t = 100; t = 1, 2, \dots, 5$$

$$= 2; t = 6$$

$$= 100; t = 7, 8, \dots, 12$$

Compute a 5-year moving average

Solution :

84

t	d _t	6 yearly moving total	6 yearly moving avg
1	100	-	-
2	100	-	-
3	100	600	100
4	100	420	54
5	100	420	54
6	90	420	54
7	100	420	54
8	100	420	54
9	100	500	100
10	100	600	100
11	100	- -	-
12	100	- -	-

Home work:

Calculate three yearly moving average method

Year : 1980 1981 1982 1983 1984 1985 1986 1987

Sales: 3 4 8 6 7 11 9 10
(in million)

1988 1989

14 12

Example

Work out the centered 4 Yearly moving average
for the following data:

(4)

Year:	1971	1972	1973	1974	1975	1976	1977
Tonnage of goods	2204	2100	2360	2480	2424	2634	2904
1978	1979	1980	1981	1982			
3098	3172	2952	3248	3172			

Solu:

Year (1)	Tonnage of goods carried (2)	Column of differen. (3)	Four yearly moving total (not centered) (4)	2 Period moving total in (5)	Four yearly moving total centered trend value over 8 (6)
1971 2204					
1972 2500	2424 - 2204 = 220	9744		19708	19708/8 = 2
1973 2360	2634 - 2500 = 134	9964			2463.5
1974 2480	3172 - 2360 = 812	10098		20062	20062/8 = 2507.75
1975 2424	3098 - 2480 = 618	10642		20740	2592.5
1976 2634				21702	2712.75
1977 2904	744	11060		22862	2853.5
1978 3098	318	11808		25934	309.75
1979 3172	944	12124		24596	3074.5
1980 2952	74	12470		26014	3126.75
1981 3248		12544			
1982 3172					

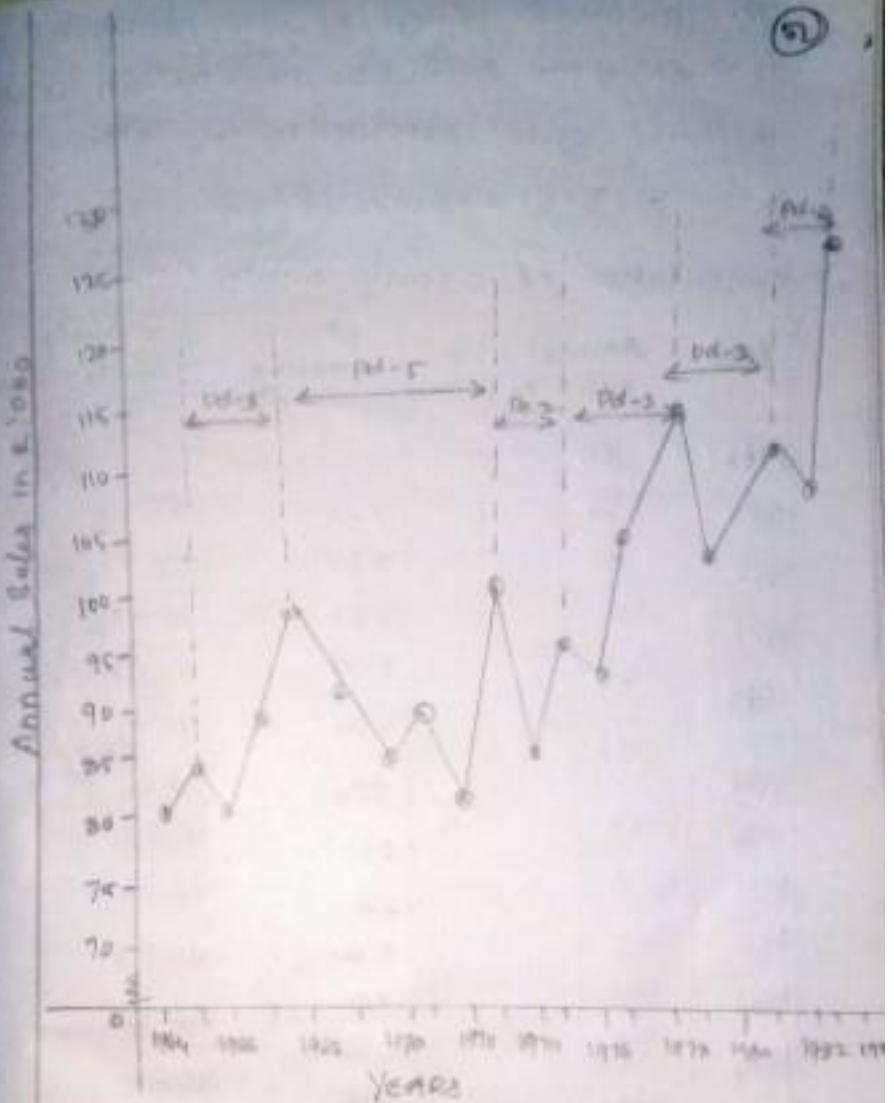
Example:

Find the trend of annual sales of a ^{trading} organization by Moving Average Method.

Year : 1964	1965	1966	1967	1968	1969	
Annual Sales	80	84	80	88	98	92
1970	1071	1072	1073	1074	1075	
84	88	80	100	84	96	
1976	1077	1078	1079	1080	1081	
92	104	116	112	102	84	
1982	1083					
103	126					

Sales

We know that the appropriate period for the moving average is the period of cyclic variation. The given data does not reveal a regular cycle of any fixed period. If we examine the data carefully we have the peaks at the following points.



Year 1965 1969 1973 1975 1978 1981 1983

Peak 84 92 100 96 116 114 124
Value
Period 3 5 2 3 3 2

Thus the data exhibits 6 cycles with
Varying periods 3, 5, 2, 3, 3 and 2 respectively.

The appropriate period of the moving average is given by the arithmetic mean of periods of different cycles exhibited by the data.

$$= \frac{3+5+2+5+3+2}{6} = 3.$$

Computation of 3 yearly moving average.

Year	Annual Sales (Rs in 1000)	3 yearly moving totals	3 yearly moving average $A = \frac{3+3}{2}$
1964	90		
1965	84	244	81.33
1966	80	252	84.00
1967	93	266	88.67
1968	92	272	92.67
	92	274	91.33
1969		264	88.00
1970	84	268	86.67
1971	88	276	92.00
1972	80	272	90.67
1973	1000	280	93.33
	84	272	90.67
1974	96	292	96.67
1975	92	312	104.00
1976		332	110.67
1977	104	336	108.00
1978	116	328	109.33
1979	112	324	108.00
1980	114	348	116.00
1981	108		
1982	126		
1983			

Example:

53

For the following series of observations verify that the 4-year centered moving average is equivalent to a 5-year weighted moving average weights 1, 2, 2, 2, 1 respectively.

Year: 1973	1974	1975	1976	1977	1978	1979	1980
Annual Sales 1973 → 2	6	1	5	3	7	2	6
1974 → 1981	1982	1983					
4	2	3					

Solu:-

Computation of 4-Yearly moving Average

Year (1)	Annual Sales (1000Rs) (2)	4 Years moving total not centred (3)	4 Year moving average (not centred) (4)	4 Years moving average centred (5)	4 Year moving average (not centred) (6)
1973	2				
1974	6	14	3.50	7.25	3.625
1975	1	15	3.75	7.75	3.875
1976	5	16	4.00	8.25	4.125
1977	3	17	4.25	8.75	4.375
1978	7	18	4.50	9.25	4.625
1979	2	19	4.75	9.75	4.875
1980	6	20	5.00	10.25	5.125
1981	4	21	5.25		
1982	8				
1983	3				

The weighted average obtained on dividing the weighted totals by the sum of the weights is, by using the formula

Weighted Moving Average

$$= \frac{\sum wY}{\sum w} \quad \text{where } \sum w = 1+2+3+4+5$$

Year (ii)	Annual Sales (Rs.)	5 year weighted moving total (iii)	5 year weighted moving average
1975	2	$2 \times 1 + 6 \times 2 + 5 \times 3 + 3 \times 4 + 2 \times 5 = 24$	3.625
1974	6	$6 \times 1 + 5 \times 2 + 3 \times 3 + 2 \times 4 + 1 \times 5 = 31$	3.875
1975	1	$1 \times 1 + 5 \times 2 + 3 \times 3 + 2 \times 4 + 1 \times 5 = 32$	4.000 - 12.5
1976	5	35	4.375
1977	3	37	4.625
1978	7	39	4.875
1979	2	49	5.125
1980	6		
1981	4		
1982	3		
1983	3		