DEPARTMENT OF STATISTICS,
GOVERNMENT ARTS COLLEGE, COIMBATORE – 641018

SUBJECT TITLE : TIME SERIES AND INDEX NUMBERS
SUBJECT CODE : 18BST23C
PREPARED BY : DR. P. VASANTHAMANI
MOBILE NUMBER : 9994575462

UNIT I

TIME SERIES

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

A time series is a set of values of some variable 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, it is used 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, y_t)** represents a time series data.
For example

Year 't'	1991	1992	1993	1994	1995	1996	1997
Sales(crores) 'y_t'	121	101	130	132	26	142	137

Uses of Time Series

Variables such as sales, production, profit, population etc., have different values at different points of time. Analysis of such series of values is important.

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

1. **It helps in understanding past behavior:** the observations at the past period of time indicates the condition which existed.

2. **It helps in accessing the present achievements:** If the past condition has existed what will be the present position? What is the actual position at present? What are the reasons for the changes? Questions like these can be answered with the help of timeseries analysis.

3. **It helps in planning future operations**: There are many methods inn statistics to estimate the value of a variable at a certain time in the future. It has been found that the forecasts by analysis of time series are more reliable.

4. **It facilitates comparison:** Relevant time series could be compared and vital inferences be drawn. For example, the production of motor cycles of two companies can be compared over a period of time.

5. **It forewarns:** As it predicts the future most reliably, future could be met with due preparedness. If the sales is in cloth shop is likely to fall, advertisement campaign can be tried to increase the sales, the services of certain staff ay be terminated, unnecessary godown facilities may be surrendered, etc. On the contrary if increased sales is expected stock may be increased, more sales personnel be employed, etc, In short losses, if any, could be minimized. Profits, if any, could be maximized.

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.

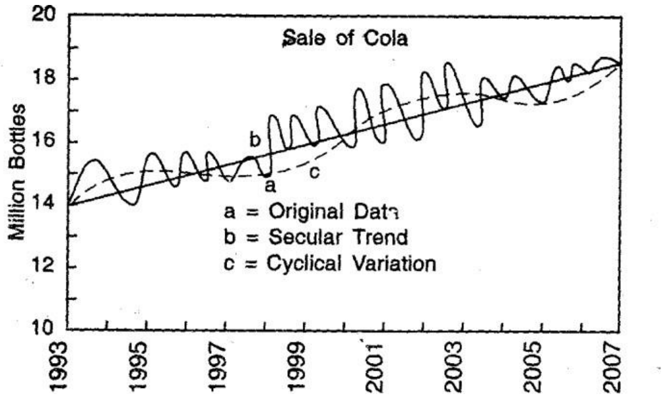
The fluctuations in a time series are of four different types

Long – term effect

1. Secular Trend

Short - term Variations

2. Seasonal Variations
3. Cyclical Variations
4. 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.

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.

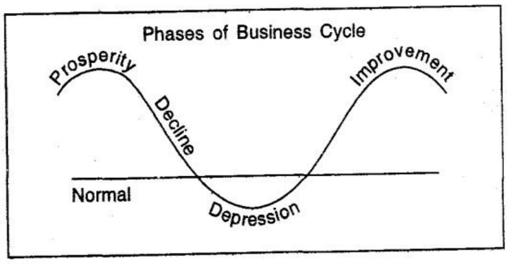
The most important factor causing seasonal variations is the climate changes in the climate and weather conditions such as rain fall, humidity, heat etc. act on different products and industries differently. For example during winter there is greater demand for woolen clothes, hot drinks etc. Where as 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 Diwali, Dussehra, 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.



Phases of Business Cycle

- Boom
- Decline
- Depression
- Improvement

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

If Y_t is the original data
T_t is the secular trend
S_t is the seasonal variation and
I_t is the irregular variation then,

1. Additive model

$$Y_t = T_t + S_t + C_t + I_t$$

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

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

2. Multiplicative model

$$Y_t = T_t \text{ x } S_t \text{ x } C_t \text{ x } I_t$$

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.

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.
the simple

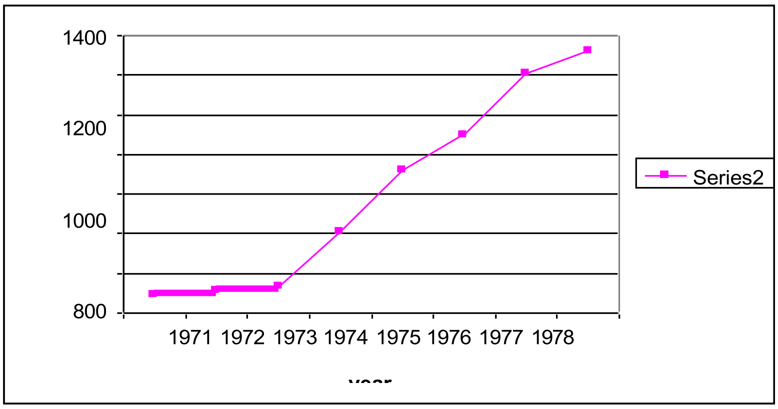
Measurement of Secular trend:

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.



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.

1. The sales in tonnes of a commodity varied from 2000 to 2011.
Fit a trend line by the method of semi-averages. Estimate the sales in 2012.

Year	Sales in Tonnes	Middle most year	Mean Sales
2000	280		
2001	300		
2002	280		
		2002.5	1650/6 = 275
2003	280		
2004	270		
2005	240		
2006	230		
2007	230		
2008	220		
		2008.5	1290/6 = 215
2009	200		
2010	210		
2011	200		

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

Year	Production In Tonnes	Middle most year	Mean Sales
2007	90		
2008	110	2008	330/3 = 110
2009	130		
2010	150	Omit	
2011	100		
2012	150	2012	450/3 = 150
2013	200		

3. Draw the trend line by the method of semi-averages.

Year	Net Profit In (Rs.Lakhs)	Middle most year	Mean Sales
2003	38		
2004	39		
		2004.5	161/4 = 40.25
2005	41		
2006	43		
2007	40		
2008	39		
		2008.5	139/4 = 34.75
2009	35		
2010	25		

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.

Case 1: Period of Moving Average is an ODD number such as 3 or 5 or 7 …

4. Using three year moving averages determine the trend and short-term fluctuations.

Year	Production ('000 tons) Y	3 Yearly Moving total	3 Yearly Moving Average (Or) Trend Y_t	Short term Fluctuations $Y - Y_t$
2003	21	-	-	-
2004	22	66	66/3 =22.00	0
2005	23	70	23.33	-0.33
2006	25	72	24.00	1
2007	24	71	23.67	0.33
2008	22	71	23.67	-1.67
2009	25	73	24.33	0.67
2010	26	78	26.00	0
2011	27	79	26.33	0.67
2012	26	-	-	-

5. Calculate five yearly moving averages determine the trend and short-term Fluctuations.

Year	No of Students Y	5 Yearly Moving total	5 Yearly Moving Average (Or) Trend Y_t	Short term Fluctuations $Y - Y_t$
1997	332	-	-	-
1998	311	-	-	-
1999	357	1794	358.8	-1.8
2000	392	1867	373.4	18.6
2001	402	1966	393.2	8.8
2002	405	2036	407.2	-2.2
2003	410	2049	409.8	0.2
2004	427	2085	417.0	10
2005	405	-	-	-
2006	438	-	-	-

6. Calculate the trend and short-term Fluctuations by seven yearly moving average method.

Year	Sales (Rs.Crores) Y	7 Yearly Moving total	7Yearly Moving Average (Or) Trend Y_t	Short term Fluctuations $Y - Y_t$
2001	35	-	-	-
2002	40	-	-	-
2003	37	-	-	-
2004	35	243	34.71	0.29
2005	34	235	33.57	0.43
2006	32	227	32.43	-0.43
2007	30	229	32.71	-2.71
2008	27	239	34.14	-7.14
2009	32	-	-	-
2010	39	-	-	-
2011	45	-	-	-

Case II: Period of Moving Average is an EVEN number such as 4 or 6 or 8 …

7. Using four yearly moving averages, calculate the trend values and short term fluctuations.

Year	Production Y	4 Yearly Moving totals	2 period Moving Totals	4 Yearly Centred Moving Averages (Or) Trend (Y_t)	Short term Fluctuations $Y - Y_t$
1991	464		-	-	-
1992	515		-	-	-
		1964			
1993	518		3966	3966/8 =495.75	22.25
		2002			
1994	467		4029	503.63	-36.63
		2027			
1995	502		4093	511.63	-9.63
		2066			
1996	540		4236	529.50	10.5
		2170			
1997	557		4424	553.00	4.0
		2254			
1998	571		4580	572.50	-1.5
		2326			
1999	586		-	-	-
2000	612		-	-	-

8. Calculate 6 yearly centered moving averages of the Earnings Per Share (EPS)of a company.

Year	EPS Y	6 Yearly Moving total	2 period Moving Totals	6 Yearly Centred Moving Averages Trend (Y_t)	Short term Fluctuations $Y - Y_t$
1995	10				
1996	12				
1997	13				
		78			
1998	15		162	162/12=13.5	1.5
		84			
1999	14		174	14.5	-0.5
		90			
2000	14		189	15.75	-1.75
		99			
2001	16		207	17.25	-1.25
		108			
2002	18		228	19.00	-1
		120			
2003	22		255	21.25	0.75
		135			
2004	24		279	23.25	0.75
		144			
2005	26		291	24.25	1.75
		147			
2006	29		297	24.75	4.25
		150			
2007	25		303	25.25	-0.25
		153			
2008	21				
2009	25				
2010	27				

Advantages:

- This method is simple to under stand 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

The straight line trend equation be Y = a + bX

The Normal Equations to find 'a' and 'b'

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

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

1. Fit a linear trend equation by the method of least squares and estimate the net profit in 2003?

Year X	Net Profit (Rs. Crores) Y = y	x = X - $\overline{X}$ $\overline{X}$ = 1998 X – 1998	xy	x^2	Trend Y_t
1995	32	-3	-96	9	21.92
1996	36	-2	-72	4	33.71
1997	44	-1	-44	1	45.50
1998	37	0	0	0	57.29
1999	71	1	71	1	69.08
2000	72	2	144	4	80.87
2001	109	3	327	9	92.66
	$\sum y = 401$	$\sum x = 0$	$\sum xy = 330$	$\sum x^2 = 28$	$\sum y_t = 401.03$

Let the equation of straight line be Y = a + bX ------------ 1

$$\sum Y = Na + b\sum X \quad \rightarrow \quad a = \frac{\sum y}{N} = 410/7 = 57.29$$

401 = 7a + b(0)

a = 410/7 = 57.29

$$\sum XY = a\sum X + b\sum X^2 \rightarrow b = \frac{\sum xy}{\sum x^2} = 330/28 = 11.79$$

330 = a(0) + 28b

b = 330/28 = 11.79

Sub a = 57.29 and b= 11.79 in 1

Y = a + bX

Y = 57.29+11.79x

Linear Trend equation Yt = 57.29 11.79x where x =X- Mid X

1. For X = 1995
   Y_t = 57.29+11.79(1995 - 1998)
   Y_t = 57.29+11.79(-3) = 57.29-35.37
   Y_t = 21.92

2. For X = 1996
   Y_t = 57.29+11.79(1996 - 1998)
   Y_t = 57.29+11.79(- 2) = 57.29-23.58
   Y_t = 33.71

3. For X = 1997
   Y_t = 57.29+11.79(1997 - 1998)
   Y_t = 57.29+11.79(- 1) = 57.29-11.79
   Y_t = 45.5

4. For X = 1998
   Y_t = 57.29+11.79(1998 - 1998)
   Y_t = 57.29+11.79(0) = 57.29-0
   Y_t = 57.29

5. For X = 1999
   Y_t = 57.29+11.79(1999 - 1998)
   Y_t = 57.29+11.79(1) = 57.29+11.79
   Y_t = 69.08

6. For X = 2000
   Y_t = 57.29+11.79(2000 - 1998)
   Y_t = 57.29+11.79(2) = 57.29+23.58
   Y_t = 80.87

7. For X = 2001
   Y_t = 57.29+11.79(2001 - 1998)
   Y_t = 57.29+11.79(3) = 57.29+35.37
   Y_t = 92.66

8. For X = 2003
   Y_t = 57.29+11.79(2003 - 1998)
   Y_t = 57.29+11.79(5) = 57.29+58.95
   Y_t = 116.24

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.

Fitting of a parabolic trend by the method of least squares

Let the second degree parabolic trend curve be

$$Y = a + bx + cx^2$$

The normal equations to find a, b and c are

$$\Sigma Y = n a + b \Sigma x + c \Sigma x^2$$

$$\Sigma (x Y) = a \Sigma x + b \Sigma(x^2) + c \Sigma(x^3)$$

$$\Sigma (x^2 Y) = a \Sigma(x^2) + b \Sigma(x^3) + c \Sigma(x^4)$$

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. *(B.Com. (H), DU; M. Com., M.D. Univ.)*

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

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

Year	*Prices (Rs.) Y*	*X*	X^2	X^3	X^4	*XY*	X^2Y	*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
$N = 6$	$\Sigma Y = 848$	$\Sigma X = 3$	$\Sigma X^2 = 19$	$\Sigma X^3 = 27$	$\Sigma X^4 = 115$	$\Sigma XY = 771$	$\Sigma X^2Y = 3{,}099$	$\Sigma Y_C = 848.002$

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

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

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

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

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

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

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

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

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 $X = 3$,

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

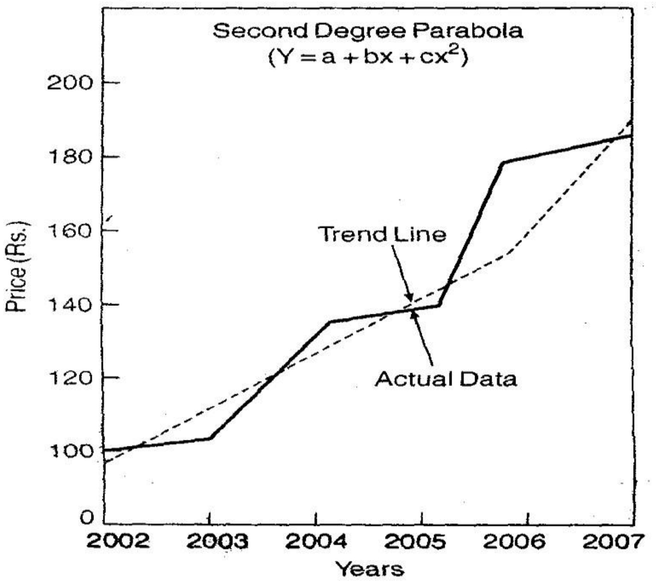
Price for the year 2008

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:



Fitting of exponential trend

The equation of the exponential trend is given by $y = a\,b^x$

$$y = a\,b^x$$

Taking log on both sides $\log y = \log a + x \log b$

$$Y = A + B\,x$$

Where $Y = \log y$, $A = \log(a)$ and $B = \log(b)$

The constants A and B are obtained by solving the equations

$$\Sigma Y = n\,A + B\,\Sigma x \qquad \text{--------------} \quad \text{I}$$

$$\Sigma\, xY = A\,\Sigma x + B\,\Sigma x^2 \qquad \text{---------------} \quad \text{II}$$

We get 'a' as Antilog (A) and 'b' as antilog (B)

Example

You are given the population figures of India as follows:

Year (x)	1911	1921	1931	1941	1951	1961	1971
Population in crores	25.0	25.1	27.9	31.9	36.1	43.9	54.7

Fit an exponential trend $y = a\,b^x$ to the above data by the method of least squares and find the trend values. Estimate the population in 1981, 2001 and 2011.

The equation of the exponential trend is given by $y = a\, b^x$

$$y = a\, b^x$$

Taking log on both sides $\log y = \log a + x \log b$

$$Y = A + B\,x$$

Where $Y = \log y$, $A = \log(a)$ and $B = \log(b)$

The constants A and B are obtained by solving the equations

$$\Sigma Y = n\,A + B\,\Sigma x \quad \text{--------------} \quad \mathrm{I}$$

$$\Sigma\, xY = A\,\Sigma x + B\,\Sigma x^2 \quad \text{---------------} \quad \mathrm{II}$$

Year(X)	Population (crores) y	x = (X -1941)/10	Y=logy	x^2	xY
1911	25.0	-3	1.3979	9	-4.1937
1921	25.1	-2	1.3997	4	-2.7994
1931	27.9	-1	1.4456	1	-1.4456
1941	31.9	0	1.5038	0	0
1951	36.1	1	1.5575	1	1.5575
1961	43.9	2	1.6425	4	3.2850
1971	54.7	3	1.7380	9	5.2140
Total		0	10.6850	28	1.6178

$\mathrm{I} \rightarrow 10.6850 = 7A + B(0) \rightarrow 7A = 10.6860 \rightarrow A = 10.6850 / 7 \rightarrow A = 1.5264$

$\mathrm{II} \rightarrow 1.6178 = A(0) + 28\,B \rightarrow 28B = 1.6178 \rightarrow B = 1.6178/28 \rightarrow B = 0.0577$

a = Antilog (A) = Antilog (1.15264) = 33.60

b = Antilog (B) = Antilog (0.0577) = 1.142

substituting the vales of a and b we get the exponential trend fitted to the given data is

$$y = 33.60 * (1.142)^{x}$$

To obtain the trend values y for different x, we use the linear trend

$Y = A + Bx \rightarrow Y = 1.5264 + 0.0577x$ and

finally the trend values are obtained as y = antilog(Y)

Year	x	0.0577x	Y= 1.5264 + 0.0577x	Trend y_t= Antilog(y)
1911	-3	-0.1731	1.3533	22.56
1921	-2	-0.1154	1.4160	25.76
1931	-1	-0.0577	1.4687	29.43
1941	0	0	1.5264	33.50
1951	1	0.0577	1.5841	38.38
1961	2	0.1154	1.6418	43.83
1971	3	0.1731	1.6995	50.06
1981	4	0.2308	1.7572	57.18
2001	6	0.3462	1.8726	74.57
2011	7	0.4039	1.9303	85.17

Hence, by assuming exponential trend the estimated population for 1981,2001 and 2011 is 57.18crores, 74.57 crores and 85.17 crores resp.

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 cyclical, seasonal and irregular variations is ignored.
- This method can not be used to fit the growth curves like Gompertz