Module 3 Modelling and Forecasting Methods

Introduction: Concept of Linear and Non Liner Forecasting model ,Concepts of Trend, Exponential Smoothing, Linear and Compound Growth model, Fitting of Logistic curve and their Applications, Moving Averages, Forecasting accuracy tests.

Probability models for time series: Concepts of AR, ARMA and ARIMA models.

INTRODUCTION

- A series of observations on a variable recorded after successive intervals of time is called a time series.
- The term ‘Time Series’ means a set of observations concurring any activity against different periods of time. The duration of time period may be hourly, daily, weekly, monthly or annually.

Examples of time series data:

- a) Profits earned by a company for each of the past five years.
- b) Workers employed by a company for each of the past 15 years.
- c) Number of students registered for CA examination in the institute for the past five years.
- d) The weekly wholesale price index for each of the past 30 week.

Importance of time series analysis

– Understand the past.

What happened over the last years, months?

– Forecast the future.

Government wants to know future of unemployment rate, percentage increase in cost of living etc.

For companies to predict the demand for their product etc.

Time-Series Components

1.Trend (T) 2. Seasonal variation (S)

3. Cyclical variation (C) 4. Random variation (R) or irregular

- Trend: the long-term patterns or movements in the data.
- Overall or persistent, long-term upward or downward pattern of movement.
- The trend of a time series is not always linear.
- Seasonal variation: Regular periodic fluctuations that occur within year.
- Examples:
- Consumption of heating oil, which is high in winter, and low in other seasons of year.
- Gasoline consumption, which is high in summer when most people go on vacation.

- Cyclical variations are similar to seasonal variations. Cycles are often irregular both in height of peak and duration.
- Examples:
- Long-term product demand cycles.
- Cycles in the monetary and financial sectors. (Important for economists!)
- Irregular component:
- Unpredictable, random, “residual” fluctuations
- Due to random variations of
  - Nature
  - Accidents or unusual events
- “Noise” in the time series

❑ Time Series Model

- **Addition Model:**

$$Y = T + S + C + I$$

Where:- Y = Original Data

T = Trend Value

S = Seasonal Fluctuation

C = Cyclical Fluctuation

I = Irregularity

- **Multiplication Model:**

$$Y = T \times S \times C \times I$$

or

$$Y = TSCI$$

Measurement of Secular trend:-

- The following methods are used for calculation of trend:
  - FREE HAND CURVE METHOD:
  - METHOD OF AVERAGES:
    - MOVING AVERAGE METHOD:
    - WEIGHTED MOVING AVERAGE
    - SEMI AVERAGE
  - LEAST SQUARE METHOD:

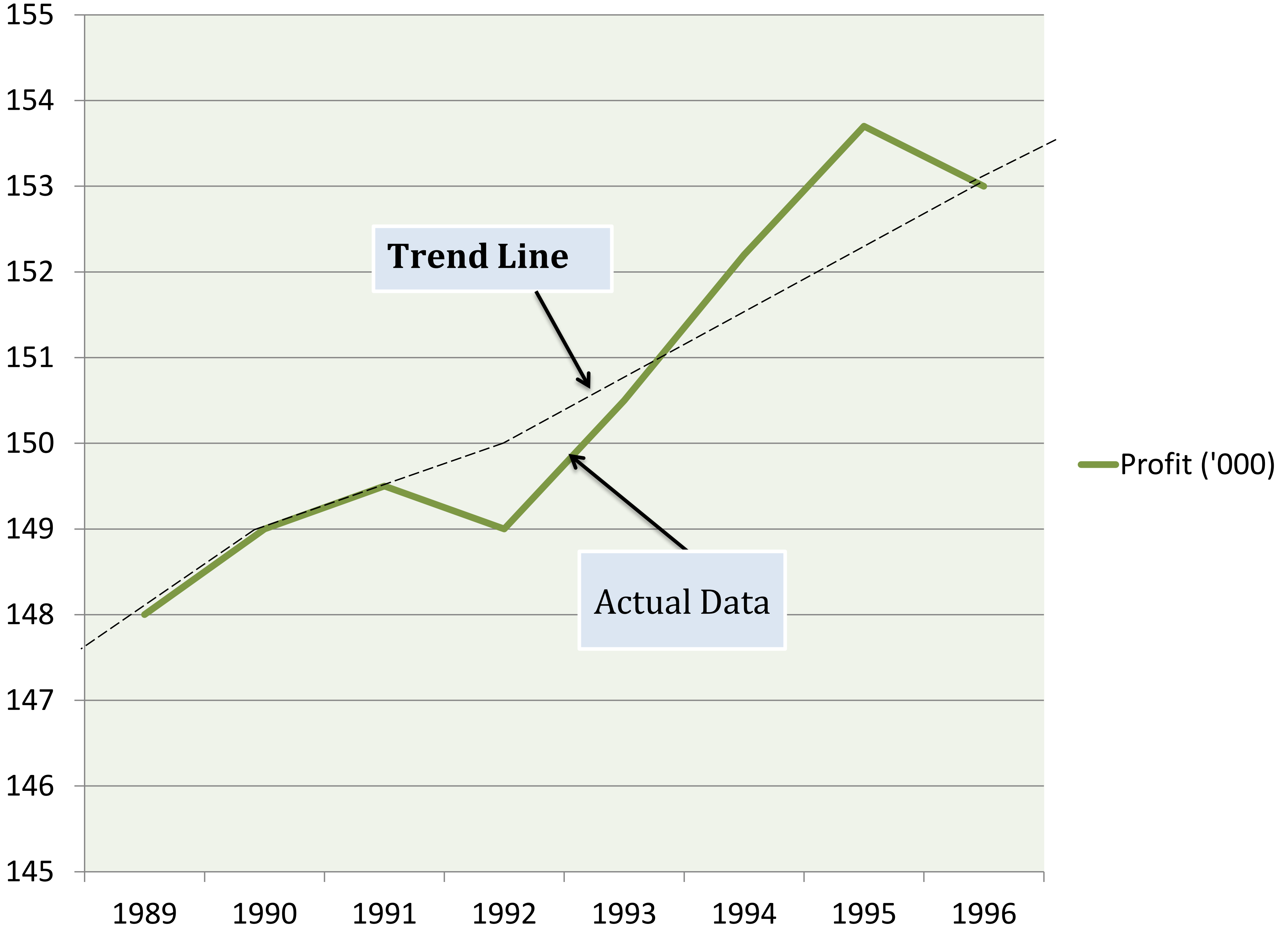
Free hand Curve Method:-

- In this method the data is denoted on graph paper. We take "Time" on 'x' axis and "Data" on the 'y' axis. On graph there will be a point for every point of time. We make a smooth hand curve with the help of this plotted points.

❑Example:

Draw a free hand curve on the basis of the following data:

Years	1989	1990	1991	1992	1993	1994	1995	1996
Profit (in '000)	148	149	149.5	149	150.5	152.2	153.7	153



Trend Line
Actual Data
Profit ('000)
155
154
153
152
151
150
149
148
147
146
145
1989
1990
1991
1992
1993
1994
1995
1996

Semi – Average Method:-

- In this method the given data are divided in two parts, preferable with the equal number of years.

- For example, if we are given data from 1991 to 2008, i.e., over a period of 18 years, the two equal parts will be first nine years, i.e.,1991 to 1999 and from 2000 to 2008. In case of odd number of years like, 9, 13, 17, etc.., two equal parts can be made simply by ignoring the middle year. For example, if data are given for 19 years from 1990 to 2007 the two equal parts would be from 1990 to 1998 and from 2000 to 2008 - the middle year 1999 will be ignored.

- **Example:**

Find the trend line from the following data by Semi – Average Method:-

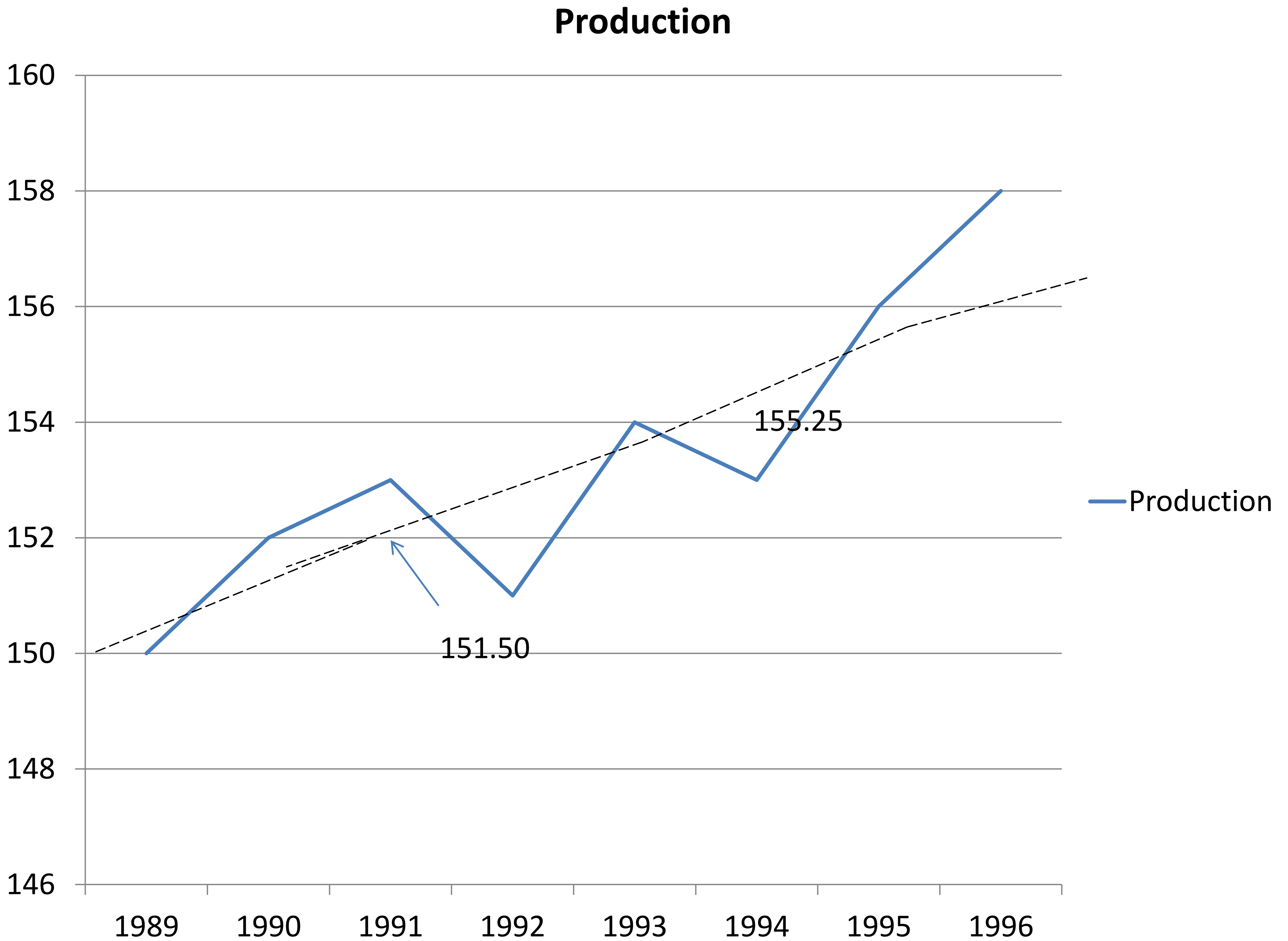
Year	1989	1990	1991	1992	1993	1994	1995	1996
Production (M.Ton.)	150	152	153	151	154	153	156	158

- There are total 8 trends. Now we distributed it in equal part. Now we calculated Average mean for every part.

$$\text{First Part} = \frac{150 + 152 + 153 + 151}{4} = 151.50$$

$$\text{Second Part} = \frac{154 + 153 + 156 + 158}{4} = 155.25$$

Year (1)	Production (2)	Arithmetic Mean (3)
1989	150	
1990	152	
1991	153	151.50
1992	151	
1993	154	
1994	153	
1995	156	155.25
1996	158	


Production
160
158
156
154
152
150
148
146
1989
1990
1991
1992
1993
1994
1995
1996
155.25
151.50
Production

❑ Moving Average Method:-

- It is one of the most popular method for calculating Long Term Trend. This method is also used for 'Seasonal fluctuation', 'cyclical fluctuation' & 'irregular fluctuation'. In this method we calculate the 'Moving Average for certain years.
- For example: If we calculating 'Three year's Moving Average' then according to this method:

$$= \frac{(1)+(2)+(3)}{3}, \quad \frac{(2)+(3)+(4)}{3}, \quad \frac{(3)+(4)+(5)}{3}, \quad \ldots\ldots\ldots\ldots\ldots$$

Where (1),(2),(3),………. are the various years of time series.

❑ **Example:** Find out the five year's moving Average:

Year	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996
Price	20	25	33	33	27	35	40	43	35	32	37	48	50	37	45

Year (1)	Price of sugar (Rs.) (2)	Five year's moving Total (3)	Five year's moving Average (Col 3/5) (4)
1982	20	-	-
1983	25	-	-
1984	33	135	27
1985	30	150	30
1986	27	165	33
1987	35	175	35
1988	40	180	36
1989	43	185	27
1990	35	187	37.4
1991	32	195	39
1992	37	202	40.4
1993	48	204	40.8
1994	50	217	43.4
1995	37	-	-
1996	45	-	-

❖ Least Square Method:-

- This method is most widely in practice. When this method is applied, a trend line is fitted to data in such a manner that the following two conditions are satisfied:-

- The sum of deviations of the actual values of y and computed values of y is zero.

$$\sum (Y - Y_c) = 0$$

- i.e., the sum of the squares of the deviation of the actual and computed values is least from this line. That is why method is called the method of least squares. The line obtained by this method is known as the line of `best fit`.

$$\sum (Y - Y_c)^2 \text{ is least}$$

The Method of least square can be used either to fit a straight line trend or a parabolic trend.

The straight line trend is represented by the equation:-

$$= Y_c = a + bx$$

Where, Y = Trend value to be computed

X = Unit of time (Independent Variable)

a = Constant to be Calculated

b = Constant to be calculated

- **Example:-**

Draw a straight line trend and estimate trend value for 1996:

Year	1991	1992	1993	1994	1995
Production	8	9	8	9	16

Solution:-

Year (1)	Deviation From 1990 X (2)	Y (3)	XY (4)	X^2 (5)	Trend $Y_c = a + bx$ (6)
1991	1	8	8	1	$5.2 + 1.6(1) = 6.8$
1992	2	9	18	4	$5.2 + 1.6(2) = 8.4$
1993	3	8	24	9	$5.2 + 1.6(3) = 10.0$
1994	4	9	36	16	$5.2 + 1.6(4) = 11.6$
1995	5	16	80	25	$5.2 + 1.6(5) = 13.2$
, N= 5	$\sum X = 15$	$\sum Y = 50$	$\sum XY = 166$	$\sum X^2 = 55$	

Now we calculate the value of two constant 'a' and 'b' with the help of two equation:-

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

Now we put the value of $\sum X, \sum Y, \sum XY, \sum X^2, \& N$:-

$50 = 5a + 15(b)$ .... (i)

$166 = 15a + 55(b)$ .... (ii)

Or $5a + 15b = 50$ .... (iii)

$15a + 55b = 166$ .... (iv)

Equation (iii) Multiply by 3 and subtracted by (iv)

$$-10b = -16$$
$$b = 1.6$$

Now we put the value of "b" in the equation (iii)

$$= \quad 5a + 15(1.6) = 50$$

$$5a = 26$$

$$a = \frac{26}{5} = 5.2$$

As according the value of 'a' and 'b' the trend line:-

$$Yc = a + bx$$

$$Y = 5.2 + 1.6X$$

Now we calculate the trend line for 1996:-

$$Y_{1996} = 5.2 + 1.6\ (6) = 14.8$$

Shifting The Trend Origin:-

- In above Example the trend equation is:

$$Y = 5.2 + 1.6x$$

Here the base year is 1993 that means actual base of these year will 1st July 1993. Now we change the base year in 1991. Now the base year is back 2 years unit than previous base year.

Now we will reduce the twice of the value of the 'b' from the value of 'a'.

Then the new value of 'a' = 5.2 – 2(1.6)

Now the trend equation on the basis of year 1991:

$$Y = 2.0 + 1.6x$$

Parabolic Curve:-

Many times the line which draw by "Least Square Method" is not prove 'Line of best fit' because it is not present actual long term trend So we distributed Time Series in sub-part and make following equation:-

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

- If this equation is increase up to second degree then it is "Parabola of second degree" and if it is increase up to third degree then it "Parabola of third degree". There are three constant 'a', 'b' and 'c'.

Its are calculated by following three equation:-

Parabola of second degree:-

$$\sum Y = Na + b\sum X + c\sum X^2$$
$$\sum XY = a\sum X + b\sum X^2 + c\sum X^3$$
$$\sum X^2Y = a\sum X^2 + b\sum X^3 + c\sum X^4$$

If we take the deviation from 'Mean year' then the all three equation are presented like this:

$$\sum Y = Na + C\sum X^2$$
$$\sum XY = b\sum X^2$$
$$\sum X^2Y = a\sum X^2 + c\sum X^4 +$$

❑ Example:

➢ Draw a parabola of second degree from the following data:-

Year	1992	1993	1994	1995	1996
Production (000)	5	7	4	9	10

Year	Production	Dev. From Middle Year (x)	xY	x^2	x^2Y	x^3	x^4	Trend Value $Y = a + bx + cx^2$
1992	5	-2	-10	4	20	-8	16	5.7
1993	7	-1	-7	1	7	-1	1	5.6
1994	4	0	0	0	0	0	0	6.3
1995	9	1	9	1	9	1	1	8.0
1996	10	2	20	4	40	8	16	10.5
	$\sum Y = 35$	$\sum X = 0$	$\sum XY = 12$	$\sum X^2 = 10$	$\sum X^2Y = 76$	$\sum X^3 = 0$	$\sum X^4 = 34$	

We take deviation from middle year so the equations are as below:

$$\sum Y = Na + \sum X^2$$

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

$$\sum X^2 Y = a\sum X^2 + c\sum X^4 +$$

Now we put the value of $\sum X, \sum Y, \sum XY, \sum X^2, \sum X^3, \sum X^4, \& N$

$35 = 5a + 10c$ .... (i)

$12 = 10b$ .... (ii)

$76 = 10a + 34c$ .... (iii)

From equation (ii) we get $b = \frac{12}{10} = 1.2$

Equation (ii) is multiply by 2 and subtracted from (iii):

$$10a + 34c = 76 \qquad \text{................} \qquad \text{(iv)}$$

$$10a + 20c = 70 \qquad \text{................} \qquad \text{(v)}$$

$$14c = 6 \text{ or } c = \frac{6}{14} = 0.43$$

Now we put the value of c in equation (i)

$$5a + 10\,(0.43) = 35$$

$$5a = 35\text{-}4.3 \quad = 5a = 30.7$$

$$a = 6.14$$

Now after putting the value of ‘a’, ‘b’ and ‘c’, Parabola of second degree is made that is:

$$\mathbf{Y = 6.34 + 1.2x + 0.43x^2}$$

Parabola of Third degree:-

- There are four constant 'a', 'b', 'c' and 'd' which are calculated by following equation. The main equation is $Y_c = a + bx + cx^2 + dx^3$. There are also four normal equation.

$$\sum Y = Na + b\sum X + c\sum X^2 + d\sum X^3$$

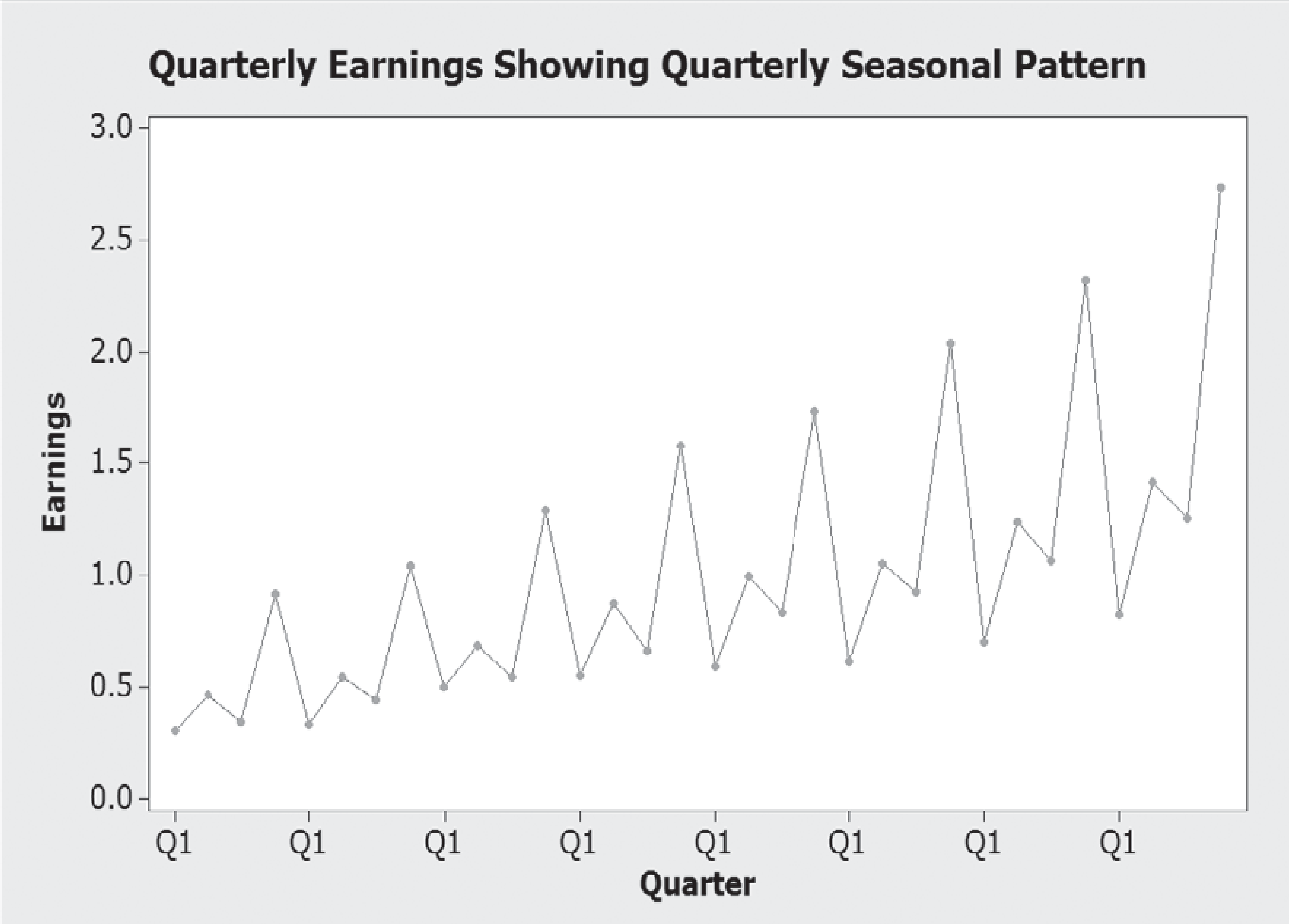
$$\sum XY = a\sum X + b\sum X^2 + c\sum X^3 + d\sum X^4$$

$$\sum X^2Y = a\sum X^2 + b\sum X^3 + c\sum X^4 + d\sum X^5$$

$$\sum X^3Y = a\sum X^3 + b\sum X^4 + c\sum X^5 + d\sum X^6$$

❑Methods Of Seasonal Variation:-

- Seasonal Average Method
- Link Relative Method
- Ratio to Trend Method
- Ratio to Moving Average Method


Quarterly Earnings Showing Quarterly Seasonal Pattern
Earnings
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Q1
Q1
Q1
Q1
Q1
Q1
Q1
Q1
Quarter

Seasonal Average Method

- Seasonal Averages $= \dfrac{\text{Total of Seasonal Values}}{\text{No. Of Years}}$
- General Averages $= \dfrac{\text{Total of Seasonal Averages}}{\text{No. Of Seasons}}$
- Seasonal Index $= \dfrac{\text{Seasonal Average}}{\text{General Average}}$

EXAMPLE:-

- From the following data calculate quarterly seasonal indices assuming the absence of any type of trend:

Year	I	II	III	IV
1989	-	-	127	134
1990	130	122	122	132
1991	120	120	118	128
1992	126	116	121	130
1993	127	118	-	-

Solution:-

Calculation of quarterly seasonal indices

Year	I	II	III	IV	Total
1989	-	-	127	134	
1990	130	122	122	132	
1991	120	120	118	128	
1992	126	116	121	130	
1993	127	118	-	-	
Total	503	476	488	524	
Average	125.75	119	122	131	497.75
Quarterly Turnover seasonal indices 124.44 = 100	101.05	95.6	98.04	105.03	

- General Average $= \dfrac{497.75}{4} = 124.44$

  Quarterly Seasonal variation index $= \dfrac{125.75}{124.44} \times 100$

  So as on we calculate the other seasonal indices

Link Relative Method:

- In this Method the following steps are taken for calculating the seasonal variation indices
- We calculate the link relatives of seasonal figures.

$$\text{Link Relative: } \frac{\text{Current Season's Figure}}{\text{Previous Season's Figure}} \times 100$$

- We calculate the average of link relative for each season.
- Convert These Averages to chain relatives on the basis of the first seasons.

- Calculate the chain relatives of the first season on the base of the last seasons. There will be some difference between the chain relatives of the first seasons and the chain relatives calculated by the pervious Method.
- This difference will be due to effect of long term changes.
- For correction the chain relatives of the first season calculated by 1st method is deducted from the chain relative calculated by the second method.
- Then Express the corrected chain relatives as percentage of their averages.

Ratio To Moving Average Method:

- In this method seasonal variation indices are calculated in following steps:
- We calculate the 12 monthly or 4 quarterly moving average.
- We use following formula for calculating the moving average Ratio:

$$\textbf{Moving Average Ratio} = \frac{\text{Original Data}}{\text{Moving Average}} \times 100$$

Then we calculate the seasonal variation indices on the basis of average of seasonal variation.

Ratio To Trend Method:-

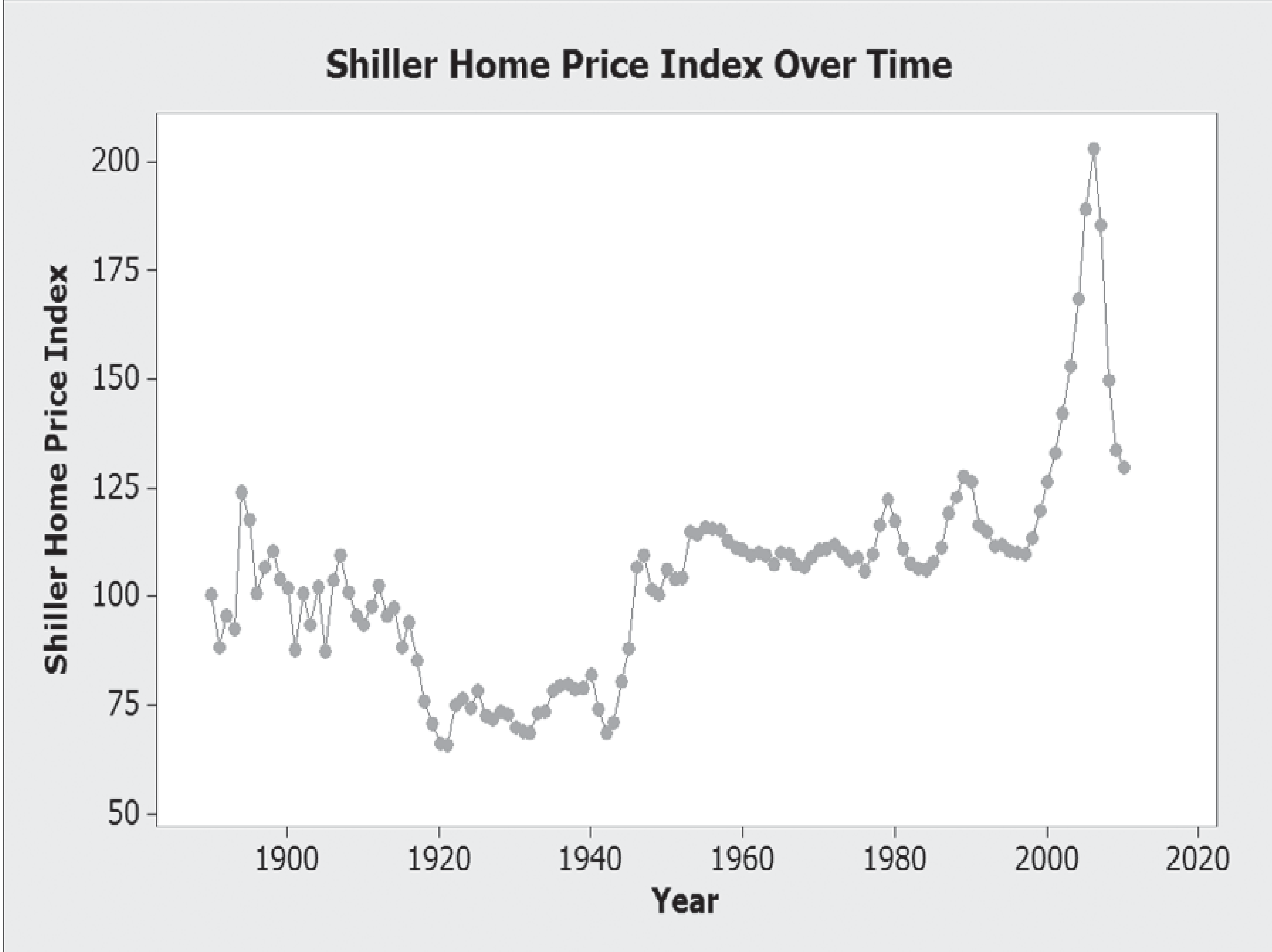
- This method based on Multiple model of Time Series. In It We use the following Steps:
- We calculate the trend value for various time duration (Monthly or Quarterly) with the help of Least Square method
- Then we express the all original data as the percentage of trend on the basis of the following formula.

$$= \frac{\text{Original Data}}{\text{Trend Value}} \times 100$$

Rest of Process are as same as moving Average Method

Methods Of Cyclical Variation:-

- Residual Method
- References cycle analysis method
- Direct Method
- Harmonic Analysis Method


Shiller Home Price Index Over Time
Shiller Home Price Index
200
175
150
125
100
75
50
1900
1920
1940
1960
1980
2000
2020
Year

Residual Method:-

- Cyclical variations are calculated by Residual Method . This method is based on the multiple model of the time Series. The process is as below:

- **(a) When yearly data are given:**

In class of yearly data there are not any seasonal variations so original data are effect by three components:

- Trend Value
- Cyclical
- Irregular

(b) When monthly or quarterly data are given:

- First we calculate the seasonal variation indices according to moving average ratio method.

- At last we express the cyclical and irregular variation as the Trend Ratio & Seasonal variation Indices

Measurement of Irregular Variations

- The irregular components in a time series represent the residue of fluctuations after trend cycle and seasonal movements have been accounted for. Thus if the original data is divided by T,S and C ; we get I i.e. . In Practice the cycle itself is so erratic and is so interwoven with irregular movement that is impossible to separate them.