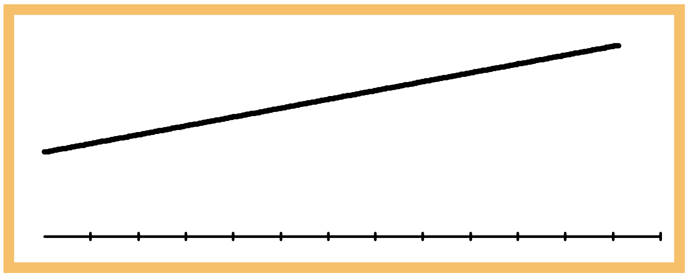
Chapter 5: Time Series

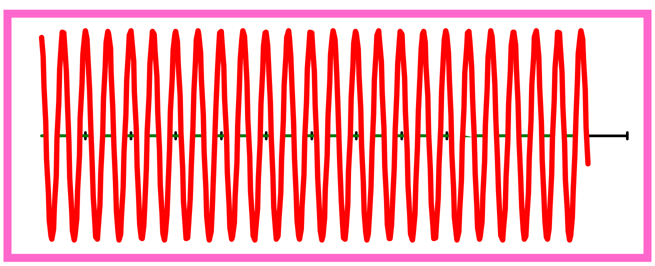
5.1 Components of Time Series

A time series usually consists of four components:

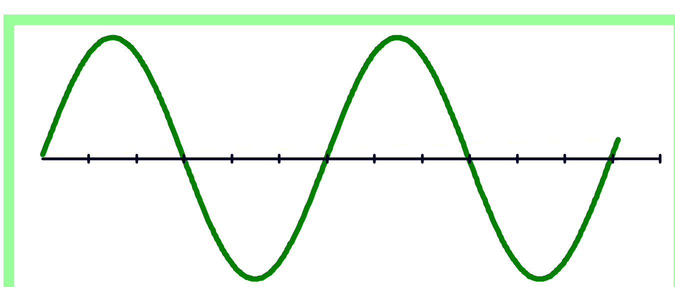
1. **Trend** (or **Secular Trend**), *T*
   The *trend* of a time series is the underlying long-term movement or tendency of the data. The duration of trend is more than one year and its fluctuation is due to factors which change slowly over a long stretch of time. The trend does not always show a linear pattern. However, the trend in time series is, in general, represented by a smooth graph.



2. **Seasonal Variations**, *S*
   Seasonal variation is the term used to describe patterns of change that recur over short period of time. It is a short-term cyclic movement of the data. The duration is usually less than one year. 'Season' in this case may mean a period of quarter of a month, or even a day as in the case of foreign exchange rate.

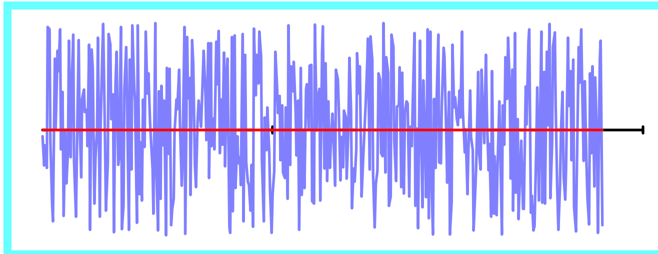


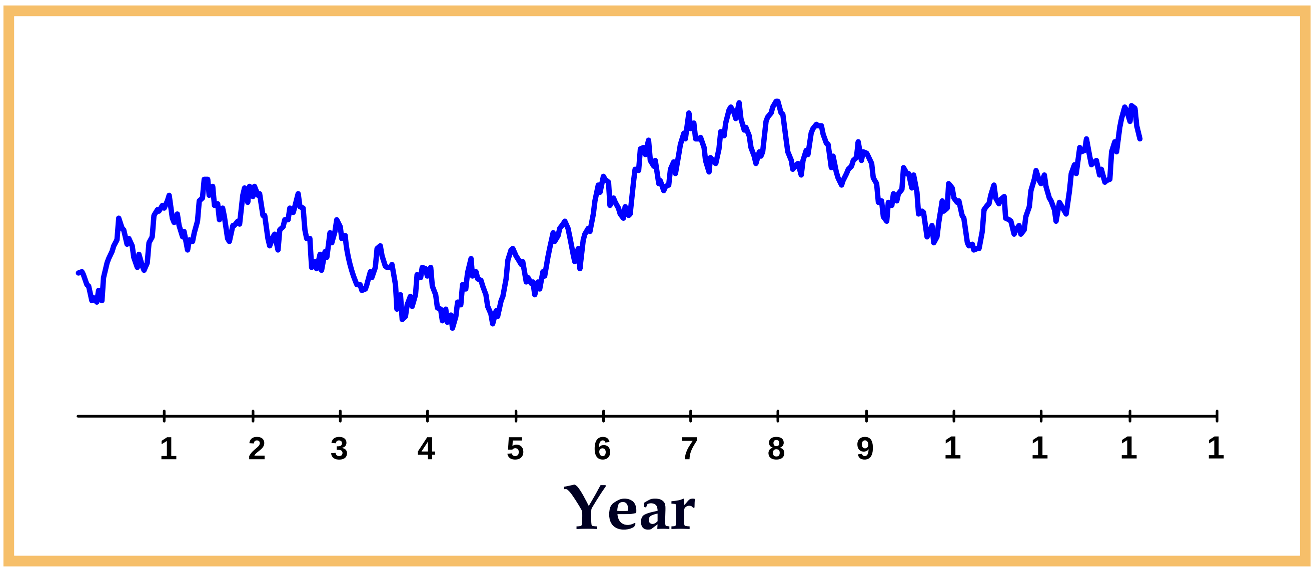
3. **Cyclical Variations**, *C*
   The *cyclical variations* are long-term cyclic movement of the data. The patterns of change occur repetitively over duration of more than one year. The long-term cyclic movement is due to the effect or influence of business or economic conditions which are irregular in length and amplitude.

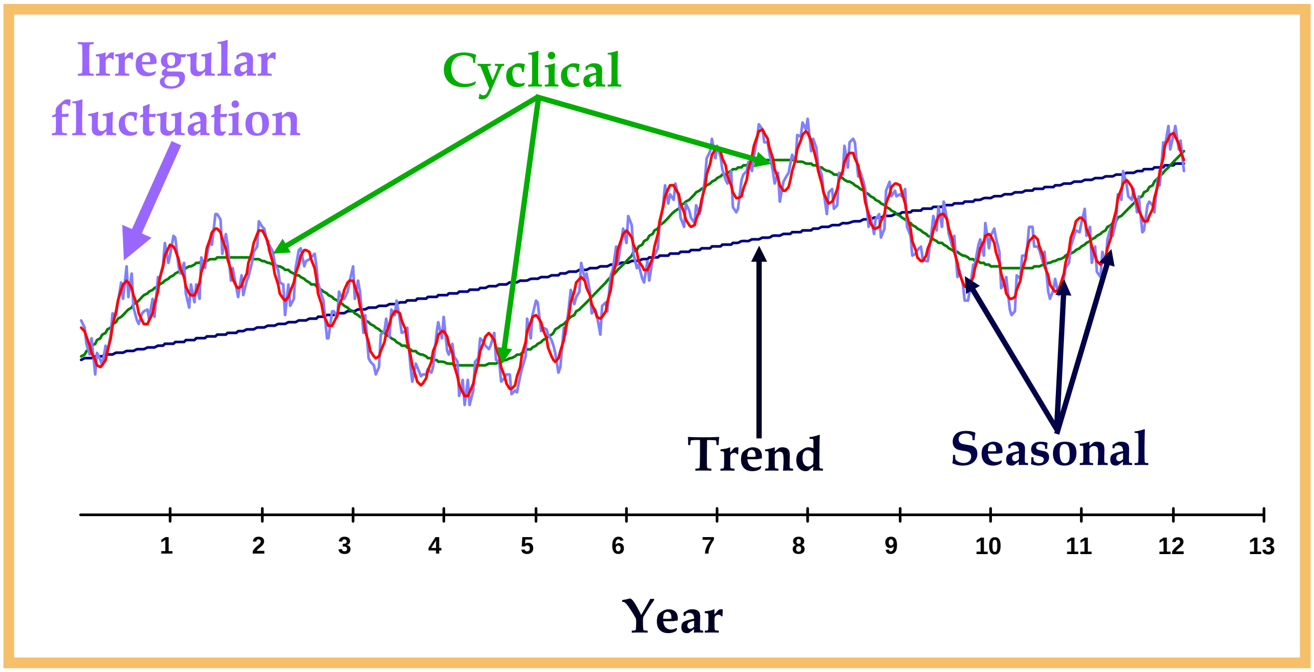


4. Irregular Variations, *I*

Irregular variations are random variations other than those that can be accounted for by the trend, seasonal, or cyclic variations. The changes occur in an unpredictable manner. Bad weather, illness, strikes and riots are examples of random factors that may occur at any time of the day.





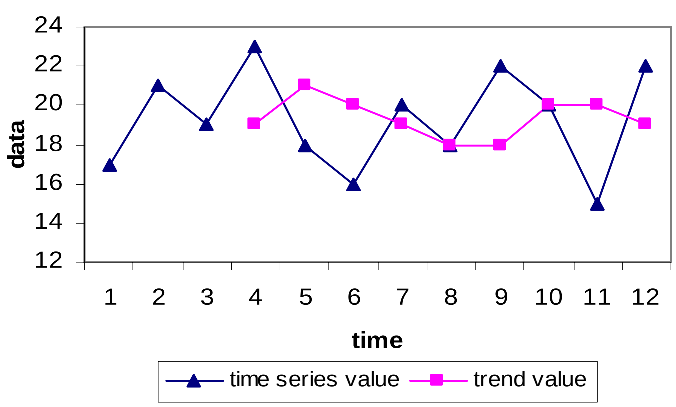


5.2 Standard Time Series Models

The main idea of constructing time series models is to study how various factors contribute to the ultimate formation of individual values of the time series. There are different types of time series models. However, the most common types are: (a) **the additive model** and (b) **the multiplicative model**. In this chapter, only the multiplicative model will be used

5.2.1 Graph of a Time Series

The graph in which a time series is plotted is known as **historigram**.



5.2.2 Time Series Multiplicative Model

The time series multiplicative model assumes that the value of time series *Y* is the product of the trend (*T*), the cyclic variation (*C*), the seasonal variation (*S*) and the irregular variation (*I*).

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

The time series multiplicative model implies that the various components of a time series are not independent of one another. In a simple time series treatment, the cyclic movement and the irregular variation are treated as residual variation.

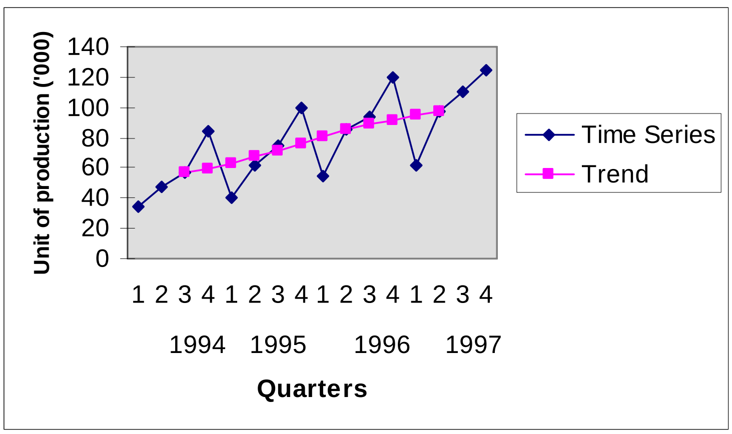
$$Y = T \times S \times R$$

Example 6.1
The following data shows the units of production of consumer goods for a manufacturing company over a period of four years.

	Units of production ('000)			
	Quarter 1	Quarter 2	Quarter 3	Quarter 4
1994	35	48	57	84
1995	40	62	75	100
1996	55	85	94	120
1997	62	97	110	125

Draw a time series to display the distribution of the data.

Solution



5.3 Time Series Analysis and the Extraction of Trend

In order to evaluate and understand the influence of each of the components on the time series, it is essential to separate each component from the time series.

Isolating trend and seasonal variations from a particular time series will help us in the study and understanding of trend and the influence of seasonal factors. Finally, a forecast may be made.

Trend can be isolated from time series by various methods. Two such methods will be discussed. They are:

(a) Moving Averages Method,
(b) Least Squares Regression Method.

5.3.1 The Method of Moving Averages

One of the most common methods of calculating the trend values is by using Moving averages. It helps to smooth the time series. The trend values may be obtained from either

(a) the moving averages or

(b) the centred moving averages.

This depends on whether the given number of data values is odd or even. The moving averages method uses the average of the most recent n data values in the time series,

$$Moving\ average = \frac{\Sigma(most\ recent\ n\ data\ values)}{n}$$

the term '*moving*' is used because every time a new observation becomes available for the time series, it replaces the oldest observation and a new average is computed. As a result, the average will change, or move, as new observations become available.

5.3.1.1 Moving averages with an *odd* number of data values

Example 6.2

The daily sales of canned food in retail shop over the past three weeks are given as follows:

	Daily sales (RM)				
	Mon	Tue	Wed	Thu	Fri
Week 1	140	150	100	150	220
Week 2	155	170	105	200	300
Week 3	175	190	130	225	325

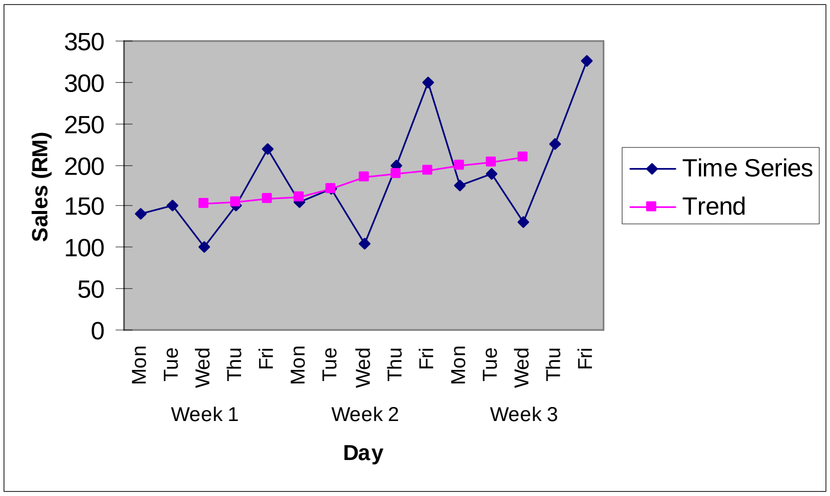
a) Construct a time series.

b) Using moving averages method, calculate the trend values.

c) Draw the trend line on the time series.

Solution

	Day	Sales, *Y*	Moving total of 5 items	Moving averages (Trend, *T*)
Week 1	Mon	140		
	Tue	150		
	Wed	100	760	152
	Thu	150	775	155
	Fri	220	795	159
Week 2	Mon	155	800	160
	Tue	170	850	170
	Wed	105	930	186
	Thu	200	950	190
	Fri	300	970	194
Week 3	Mon	175	995	199
	Tue	190	1020	204
	Wed	130	1045	209
	Thu	225		
	Fri	325		



5.3.1.2 Moving averages with an *even* number of data values – *The centered moving averages*

Since the number of data values is even, there is no exact median time point for each moving average, the mean of two moving averages is computed. The mean of two moving averages is the **centered moving average.**

Example 6.3

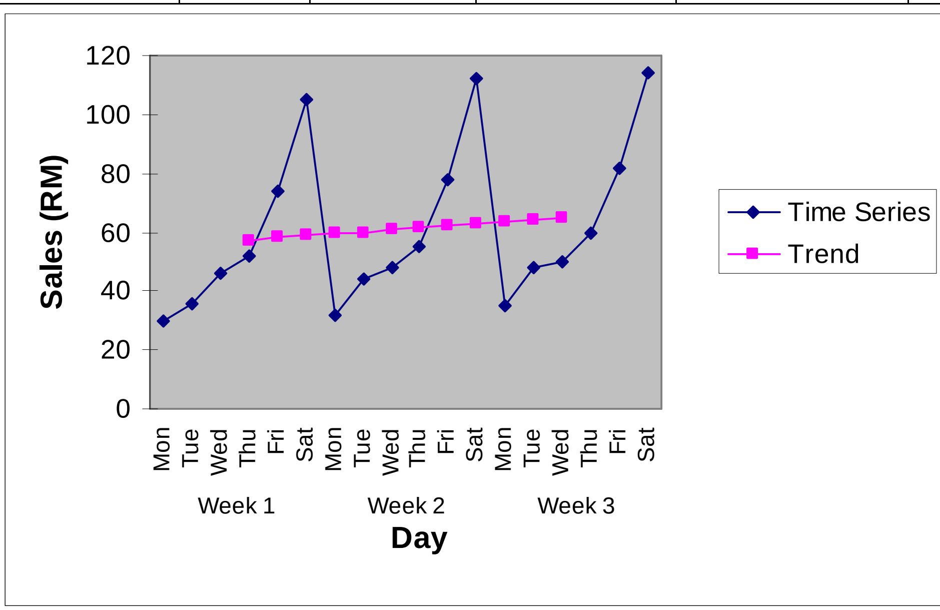
Consider the daily sales of a retail store as follows:

	Daily sales (RM)					
	Mon	Tue	Wed	Thu	Fri	Sat
Week 1	30	36	46	52	74	105
Week 2	32	44	48	55	78	112
Week 3	35	48	50	60	82	114

a) Construct a time series.
b) Using moving averages method, calculate the trend values.
c) Draw the trend line on the time series.

Solution:

	Day	Sales, *Y*	Moving total of 6 items	Moving averages	Centred Moving Average, *T*
Week 1	Mon	30			
	Tue	36			
	Wed	46			
			343	57.17	
	Thu	52			57.34
			345	57.50	
	Fri	74			58.17
			353	58.83	
	Sat	105			59.00
			355	59.17	
Week 2	Mon	32			59.42
			358	59.69	
	Tue	44			60.00
			362	60.33	
	Wed	48			60.92
			369	61.50	
	Thu	55			61.75
			372	62.00	
	Fri	78			62.34
			376	62.67	
	Sat	112			62.84
			378	63.00	
Week 3	Mon	35			63.42
			383	63.83	
	Tue	48			64.17
			387	64.50	
	Wed	50			64.67
			389	64.83	
	Thu	60			
	Fri	82			
	Sat	114			



5.3.1.3 *Advantages and disadvantages of moving averages method*

1. It is most widely used.
2. It helps to smooth out time series.
3. It shows the trend as it is, either linear or nonlinear – it is representative of the data.
4. There are no trend values at the beginning and end time points.
5. When the number of items is even, a further centred moving average has to be obtained so that an appropriate time point is located.

5.3.2 The least squares regression method

The procedure is as follows:

1. Treat the time points of time series as independent variable, *X*. The time points are then coded accordingly as 1, 2, 3, … or …-2, -1, 0, 1, 2, …
2. Treat the values of the original data as dependent variable, *Y*.
3. Find the least squares regression line *y* on *x*, i.e. $\hat{Y} = a + bX$.
4. Trend values are obtained by substituting the value of *X* in $\hat{Y} = a + bX$.

Example 6.4

Using the least squares regression, compute the trend of the following data.

	Daily Sale (RM)				
	Mon	Tue	Wed	Thu	Fri
Week 1	140	150	100	150	220
Week 2	155	170	105	200	300
Week 3	175	190	130	225	325

Solution:

1. The time points are coded as below. (in *X*)

X	Y	X^2	XY	T
1	140	1	140	126.33
2	150	4	300	134.33
3	100	9	300	142.33
4	150	16	600	150.33
5	220	25	1100	158.33
6	155	36	930	166.33
7	170	49	1190	174.33
8	105	64	840	182.33
9	200	81	1800	190.33
10	300	100	3000	198.33
11	175	121	1925	206.33
12	190	144	2280	214.33
13	130	169	1690	222.33
14	225	196	3150	230.33
15	325	225	4875	238.33
$\Sigma X = 120$	$\Sigma Y = 2735$	$\Sigma X^2 = 1240$	$\Sigma XY = 24120$	

2. Let the equation of least squares regression line *Y* on *X* be

$$\hat{Y} = a + bX$$

$$b = \frac{S_{XY}}{S_{XX}} = \frac{\Sigma XY - \frac{\Sigma X \Sigma y}{n}}{\Sigma X^2 - \frac{(\Sigma X)^2}{n}} = \frac{24120 - \frac{(120)(2735)}{15}}{1240 - \frac{120^2}{15}} = 8$$

$$a = \overline{Y} - b\overline{X} = \frac{\Sigma Y}{n} - b\frac{\Sigma X}{n} = \frac{2735}{15} - 8 * \frac{120}{15} = 118.33$$

3. Substitute the values of X in the equation, $T = 118.33 + 8X$ to obtain the trend values respectively.

<u>An alternative coding of *X*</u>

X	*Y*	X^2	*XY*	*T*
-7	140	49	-980	126.33
-6	150	36	-900	134.22
-5	100	25	-500	142.33
-4	150	16	-600	150.33
-3	220	9	-660	158.33
-2	155	4	-310	166.33
-1	170	1	-170	174.33
0	105	0	0	182.33
1	200	1	200	190.33
2	300	4	600	196.33
3	175	9	525	206.33
4	190	16	760	214.33
5	130	25	650	222.33
6	225	36	1350	230.33
7	325	49	2275	238.33
$\Sigma X = 0$	$\Sigma Y = 2735$	$\Sigma X^2 = 280$	$\Sigma XY = 2240$	

$$\hat{Y} = a + bX$$

$$b = \frac{S_{XY}}{S_{XX}} = \frac{\Sigma XY - \frac{\Sigma X \Sigma y}{n}}{\Sigma X^2 - \frac{(\Sigma X)^2}{n}} = \frac{2240}{280} = 8$$

$$a = \overline{Y} - b\overline{X} = \frac{\Sigma Y}{n} - b\frac{\Sigma X}{n} = \frac{2735}{15} = 118.33$$

$$T = 118.33 + 8X$$

5.3.2.1 *Advantages and disadvantages of least squares regression method*

The main advantage of the least squares regression method is that it is representative of the data and the trend values of all the time points can be computed.

The disadvantage of this method is that it assumes the trend to be strictly linear and it is not appropriate for computation of trend values if there exists a strong seasonal fluctuation in the time series.

5.4 Seasonal Effects

To eliminate or remove any seasonal factors from a time series, we will deseasonalise the time series.

5.4.1 Deseasonalisation in multiplicative model of time series

The procedure for the computation is as follows:

1. Obtain the trend value using moving averages method (or other method).

2. Divide the original data of time series by the trend value and multiply by 100 to obtain the percentage deviation.

3. Find the average deviation for each season.

4. If the sum of the average deviation is not equal to the appropriate total percentage of the number of seasons (i.e. each season should have the value of 100 (%), therefore, if there are four seasons, we expect the total percentage to be 400), adjustment must be made accordingly.

5. To obtain the seasonally adjusted data, the original value is divided by the seasonal factor, i.e. $\frac{Y}{S}\times 100$.

Example 6.5

An export company recorded the following data concerning the export of vegetable for the various quarters of the year in the past three years:

	Quarterly Export (RM'000)			
	Quarter 1	Quarter 2	Quarter 3	Quarter 4
1994	55	35	25	55
1995	60	40	35	62.5
1996	75	50	40	65

a) Obtain the trend values using moving averages method and thus calculate the seasonal factor.
b) Plot the seasonally adjusted time series.

Solution:

a)

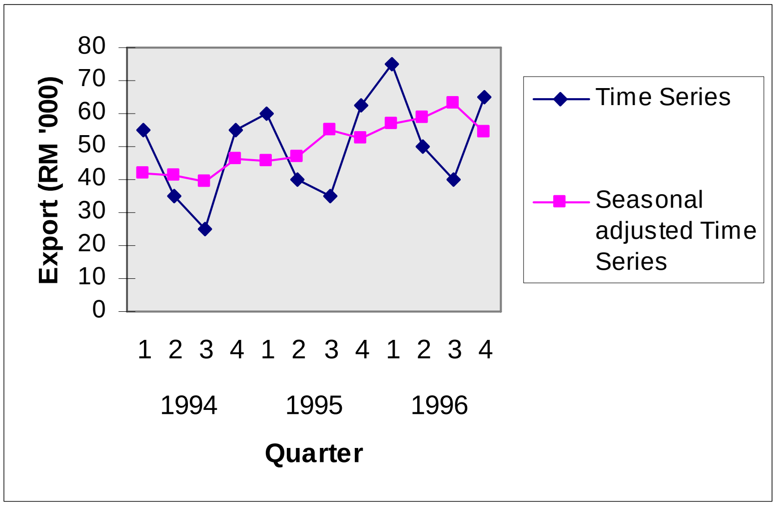
	Quarter	Original data, Y	Trend, T	Deviation $\frac{Y}{T}\times 100$	Seasonal factor, S	Seasonal adjusted time series, $\frac{Y}{S}\times 100$
1994	1	55			131.76	41.74
	2	35			85.24	41.06
	3	25	43.13	57.96	63.35	39.46
	4	55	44.38	123.93	119.65	45.97
1995	1	60	46.25	129.73	131.76	45.54
	2	40	48.44	82.58	85.24	46.93
	3	35	51.26	68.28	63.35	55.25
	4	62.5	54.39	114.91	119.65	52.24
1996	1	75	56.26	133.31	131.76	56.92
	2	50	57.19	87.43	85.24	58.66
	3	40			63.35	63.14
	4	65			119.65	54.33

1. Obtain the trend values using moving averages method (or least squares regression method).
2. Obtain the deviation from the trend by calculating $\frac{Y}{T}\times 100$.
3. Find the average deviation for each quarter:

	Deviations, $\frac{Y}{T}\times 100$			
	Quarter			
	1	2	3	4
1994 1995 1996	- 129.73 133.31	- 82.58 87.43	57.96 68.28 -	123.93 114.91 -
Total deviations	263.04	170.01	126.24	238.84
Average deviation Adjustment	131.52 0.24	85.01 0.23	63.12 0.23	119.42 0.23
Adjusted seasonal factor (S)	131.76	85.24	63.35	119.65

4. Since the sum of the average deviations is not 400 (sum = 399.07) and it is found to have a difference of 0.93 unit, therefore adjustment has to be made accordingly, i.e. divide 0.93 by 4 = 0.2325. To make the adjustment, the deviation in the first quarter should be added 0.24 (since this is the largest deviation) and 0.23 should be added to the other average deviations. The sum of the adjusted seasonal factor is now 400.
5. In order to deseasonalise the data, we calculate $\frac{Y}{S}\times 100$.

b)



5.4.1 Interpretation of the values of the seasonal factors

In the multiplicative model of time series, the seasonal factors are expressed in terms of percentages which are also known as the **seasonal indices**. When the seasonal indices are less than 100, they show that the seasonal factors are below the trend. For example, the seasonal factor in quarter 2 is 85.24 which means that the trend is deflated by 14.76% (i.e. 100 – 85.24), and the seasonal factor in quarter 1 (131.76) shows that the trend is inflated by 31.76%.

5.5 Forecasting

Not only does a time series enable us to study the general pattern of the movement of the past data, it also allows us to look into the general pattern of the movement of the data in the near future.

Since the forecast of a future movement is based on past data, the forecast so obtained should be treated with caution, for unknown factors could be present in the future.

Forecasting in multiplicative model of time series

In order to make a forecast, the following steps may be taken:

1. Project the trend into the future time point based on the calculated trend values.

2. Multiply the appropriate seasonal factor to the projected trend value to make the forecast, i.e. $Y_p = T_p * S$, where

   Y_p = projected value of time series.
   T_p = projected trend value.
   S = seasonal factor.

Note: Residual variation, R, is ignored in simple forecasting.

Example 6.6

Using the data in Example 6.5, project the trend and hence make a forecast for the 4 quarters of 1997 based on multiplicative model.

Solution:

Steps:

1. Find the range of the trend values.
   i.e. Range = 57.19 - 43.13 =14.06

2. Find the average change per time point
   i.e. Average change per time point = 14.06/ 7 = 2.01

   There are 8 trend values but there are 7 time points in between. Change per time point could be considered as a 'jump' from one trend value to the next. Hence from the lowest trend value to the highest trend value requires 7 'jumps', i.e. 7 changes in time points.

3. Projected trend values are:

   i.e. $T_p(1996, Q3) =$ 57.19 + 1(2.01) = 59.20

   $T_p(1996, Q4) =$ 57.19 + 2(2.01) = 61.21

   $T_p(1997, Q1) =$ 57.19 + 3(2.01) = 63.22

   $T_p(1997, Q2) =$ 57.19 + 4(2.01) = 65.23

   $T_p(1997, Q3) =$ 57.19 + 5(2.01) = 67.24

   $T_p(1997, Q4) =$ 57.19 + 6(2.01) = 69.25

4. The seasonal factors are: (as calculated previously)

	Quarter			
	1	2	3	4
Seasonal factor	131.76	85.24	63.35	119.65

5. Multiply the projected trend value by the appropriate seasonal factor and then divide by 100 to obtain the forecast, i.e. $Y_p = \frac{T_p \times S}{100}$, therefore, the predictions are:

   $Y_p(1996, Q3) =$ 59.20 X 63.35/100 = 37.50

   $Y_p(1996, Q4) =$ 61.21X119.65/100 = 73.24

   $Y_p(1997, Q1) =$ 63.22X 131.76/100 = 83.30

$Y_p(1997, Q2) =$ 65.23X 85.24/100 = 55.60

$Y_p(1997, Q3) =$ 67.24X 63.35/100 = 42.60

$Y_p(1997, Q4) =$ 69.25X119.65/100 = 82.86