UNIT 14 SEASONAL COMPONENT ANALYSIS

Structure

14.1 INTRODUCTION

In Unit 13, you have learnt that time series can be decomposed into four components: Trend (T), Cyclic (C), Seasonal (S) and Irregular Component (I). Our aim is to estimate T, C and S components and use them for forecasting. We have already described some methods for smoothing or filtering the time series, namely, the simple moving average method, weighted moving average method and exponential smoothing method in Unit 13. We have also described some methods for estimating Trend and Cyclic components, i.e., the method of least squares and the moving average method in Unit 13.

When time series data do not contain any trend and cyclic components but reflect seasonal variation, we have to estimate the seasonal component S by removing irregular components. In Sec. 14.2 of this unit, we discuss some methods for estimating the seasonal component (S), namely, the simple average method, the ratio to moving average method and the ratio to trend method. If the effect of seasonal variation is not removed from the time series data, the trend estimates are also affected. Therefore, we have to deseasonalise the data by dividing it by the corresponding seasonal indices. Once the data is free from seasonal effects, we estimate the trend equation using the method of least squares as explained in 14.3. In Sec. 14.4, we explain how to use data for forecasting purposes once we have estimated the trend, cyclic and seasonal components of the time series.

In the next unit, we shall discuss the stationary time series and explain the stationary processes, i.e., weak and strict stationary processes. We shall also discuss the autocovariance, autocorrelation function and correlogram of a stationary process.

Objectives

After studying this unit, you should be able to:

- explain the effect of seasonal variation in time series;
- apply the simple average method for estimating seasonal indices;
- apply the ratio to moving average method and ratio to trend method for estimation of seasonal indices;

- describe the method of estimation of trend component from deseasonalised time series data; and
- use trend (T), cyclical (C) and seasonal (S) components for forecasting purposes.

14.2 ESTIMATION OF SEASONAL COMPONENT

If the seasonal variation is substantial, we can express the variation in y_t by additive or multiplicative models:

Additive Model: $Y_t = T_t + C_t + S_t + I_t$ … (1)

Multiplicative Model: $Y_t = T_t \times C_t \times S_t \times I_t$ … (2)

In the additive model, the seasonal indices S_t are normalised so that their sum over months in a year is zero. In the multiplicative model, they are normalised. In both cases, the forecast of the yearly output is not affected by the seasonal effect S_t and we can work on yearly totals to estimate the trend. In such cases, we can estimate S_t by working on the annual data. However, in many cases, we may be interested in forecasting monthly (or quarterly) figures. This requires monthly (or quarterly) estimate of the seasonal index S_t.

In many cases it has been found that seasonal effects increase with increase in the mean level of time series. Under these circumstances, it may be more appropriate to use the multiplicative model. If seasonal effects remain constant, the additive model is more appropriate. The classical approach is to consider the multiplicative model and estimate seasonal effect (S_t) for forecasting purposes. In this unit, we use the classical multiplicative model. We describe two methods for estimating seasonal indices based on the ratios of time series observation (Y) and estimated trend and cycle effects:

$$\frac{Y}{(T_t \times C_t)} = \frac{T_t \times C_t \times S_t \times I_t}{(T_t \times C_t)} = S_t \times I_t \quad \ldots (3)$$

This ratio gives an estimate of $S_t \times I_t$. We shall estimate the seasonal indices (S_t) by smoothing out the irregular component (I_t).

We now describe the methods of estimating the seasonal indices S_t.

14.2.1 Simple Average Method

The method of simple average is the simplest of all methods. It is used to eliminate the seasonal effect from the given time series data. This method is based on the assumption that the data do not contain any trend and cyclic components and consists of eliminating irregular components by averaging the monthly (or quarterly or yearly) values over years. This assumption may or may not be true since most economic or business time series exhibit trends.

This method consists of the following steps:

Step 1: We arrange the data by years, months or quarters if data are collected on yearly, monthly or quarterly basis.

Step 2: After arranging the time series data, the average $\bar{y}_i$ is calculated for the i^{th} month of the year.

Step 3: After the monthly averages are calculated, we calculate the average of the monthly averages, that is,

$$\bar{\bar{y}} = \frac{\bar{y}_1 + \bar{y}_2 + ... + \bar{y}_{12}}{12}$$

Step 4: After calculating the average $\bar{\bar{y}}$, we express monthly averages $\bar{y}_i$ as the percentage of the average $\bar{\bar{y}}$. These percentages are known as seasonal indices. Thus,

$$\text{Seasonal index for the i}^{\text{th}} \text{ month} = \frac{\bar{y}_i}{\bar{\bar{y}}} \times 100, \quad \text{for } i = 1, 2, ..., 12. \ \ldots (4)$$

Let us consider an example to explain this method.

Example 1: Determine the monthly seasonal indices for the following data of production of a commodity for the years 2010, 2011 and 2012 using the method of simple averages.

Years / Months	Production in Tonnes		
	2010	2011	2012
January	120	150	160
February	110	140	150
March	100	130	140
April	140	160	160
May	150	160	150
June	150	150	170
July	160	170	160
August	130	120	130
September	110	1360	100
October	100	120	100
November	120	130	110
December	150	140	150

Solution: First of all, we arrange the data as shown in columns 1 to 4 of the table given below:

Table 1: Seasonal indices for the given time series

Months	**Production (in Tonnes)**			**Total**	**Monthly Avgs.**	**Seasonal Index**
	2010	**2011**	**2012**			
January	120	150	160	430	143.3	104.886
February	110	140	150	400	133.3	97.566
March	100	130	140	370	123.3	90.247
April	140	160	160	460	153.3	112.205
May	150	160	150	460	153.3	112.205
June	150	150	170	470	156.6	114.620
July	160	170	160	490	163.3	119.524
August	130	120	130	380	126.6	92.662
September	100	130	100	340	113.3	82.928
October	100	120	1000	320	106.6	78.024
November	120	130	110	360	120.0	87.832
December	150	140	150	440	146.6	107.301
Total				4920	1639.5	1200
Average				410	136.625	100

We are given the monthly production of a commodity for 3 years. For calculating the monthly seasonal indices, we first calculate the month-wise total production for the 3 years. Then we calculate the monthly averages for all 12 months. Note from Table 1 that for January, it is 143.3, for February, it is 133.3, …, etc. Next, we calculate the average of all monthly averages, i.e.

$$\bar{\bar{y}} = \frac{1}{12}\left(143.3 + 133.3 + \ldots + 146.6\right) = 136.625$$

Now we calculate the seasonal indices by taking the percentage of monthly averages $\bar{y}_i$ to the combined averages $\bar{\bar{y}}$, one at a time, for i = 1, 2, …, 12:

Seasonal Index for January $= \dfrac{143.3}{136.625} \times 100 = 104.886$

Seasonal Index for February $= \dfrac{133.3}{136.625} \times 100 = 97.566$

Seasonal Index for March $= \dfrac{123.3}{136.625} \times 100 = 90.247$

Seasonal Index for April $= \dfrac{153.3}{136.625} \times 100 = 112.205$

Seasonal Index for May $= \dfrac{153.3}{136.625} \times 100 = 112.205$

Seasonal Index for June $= \dfrac{156.6}{136.625} \times 100 = 114.620$

Seasonal Index for July $= \dfrac{163.3}{136.625} \times 100 = 119.524$

Seasonal Index for August $= \dfrac{126.6}{136.625} \times 100 = 92.662$

Seasonal Index for September $= \dfrac{113.3}{136.625} \times 100 = 82.928$

Seasonal Index for October $= \dfrac{106.6}{136.625} \times 100 = 78.024$

Seasonal Index for November $= \dfrac{120}{136.625} \times 100 = 87.832$

Seasonal Index for December $= \dfrac{146.6}{136.625} \times 100 = 107.301$

14.2.2 Ratio to Moving Average Method

In Sec. 14.2.1, we have discussed the simple average method for calculating seasonal indices. Now we discuss the most widely used method known as the **ratio to moving average method**. It is better because of its accuracy. Also the seasonal indices calculated using this method are free from all the three components, namely, trend (T), cyclic (C) and Irregular variations (I).

As you have learnt in Unit 13, the moving average eliminates periodic variations if the span of period of moving average is equal to the period of the oscillatory variation sought to be eliminated. Therefore, we have to choose the span of time for moving average to be equal to one cycle. For example, if a cycle is completed in 3 months, we calculate the moving average for 3 months. You may note that for some quarters, three values are available and for some quarters, four values are available. By taking the average over three or four values, we get seasonal indices S_i for i=1, 2, 3, 4. We usually normalise them so that their mean is 100 by dividing them by the mean of S_i and multiplying by 100. These normalised S_i have mean 100. Usually, not much difference exists between normalised and non-normalised seasonal indices S_i . When data are monthly, the same procedure will yield twelve monthly seasonal indices S_i . This method of estimating seasonal indices is known as the **ratio to moving average** method.

We have explained the ratio to moving average method for monthly time series data. The same method may be applied for any other periodic data such as quarterly, weekly data, etc. The steps for obtaining seasonal indices using this method are as follows:

Step 1: We arrange the data chronologically.

Step 2: If the cycle of oscillation is 1 year, we take the 12 months moving average of the 1st year, which will give estimates of the combined effects of trend and cyclic fluctuation. We enter the average value against the middle position, i.e., between the months of June and July.

Step 3: We discard the value for the month of January of the first year and include the value for the month of January of the subsequent year. Then we calculate the average of these 12 values and enter it against the middle position, i.e., between July and August. We repeat the process of taking moving averages MA (1) and entering the value in the middle position, till all the monthly data are exhausted.

Step 4: We calculate the centred moving average, i.e., MA(2), of the two values of the moving averages MA(1) and enter it against the first value, i.e., the month of July in the first year and subsequent values against the month of August, September, etc.

Step 5: After calculating the MA(1) and MA(2) values, we treat the original values (except the first 6 months in the beginning and the last 6 months at the end) as the percentage of the centred moving average values. For this we divide the original monthly values by the corresponding centred moving average, i.e., MA (2) values, and multiply the result by 100. We have now succeeded in eliminating the trend and cyclic variations from the original data. We now have to get rid of the data of irregular variations.

Step 6: We prepare another two-way table consisting of the month-wise percentage values calculated in Step 5, for all years. The purpose of this step is to average the percentages and to eliminate the irregular variations in the process of averaging.

Step 7: We find the median of the percentages or preliminary seasonal indices calculated in Step 5 month-wise and take the average of the month-wise median. Then we divide the median of each month by the average value and multiply it by 100. Generally, the sum of all medians is not 1200. Therefore, the average of all medians is not equal to 100. Hence, the seasonal indices are subjected to the same operation. We multiply the medians by the ratio of expected total of indices, i.e., 1200 to the actual total as follows:

$$\text{Seasonal Index} = \text{Median} \times \frac{1200}{\text{Total of Indices}}$$

The seasonal index for each month is given in the last column of the table. We calculate the sum of indices.

Let us consider an example to understand this method.

Example 2: Apply the ratio to moving average method to ascertain seasonal indices for the following data:

Month \ Year	2009	2010	2011	2012
January	500	550	500	600
February	600	550	600	650
March	650	600	550	650
April	750	650	600	750
May	800	700	650	800
June	800	700	750	900
July	850	750	750	1000
August	900	750	850	1000
September	900	750	900	1050
October	950	800	1000	1100
November	1100	900	1100	1200
December	1100	1000	1200	1250

Solution: As described in Sec. 13.5 we shall first eliminate the effect of S_t from time series observations. If the data are quarterly and the period of seasonal effect is one year, then taking a moving average for 4 quarters will eliminate the effect of S_t. In this case, we have to calculate centred moving average by taking the average of two successive moving average values. Table 2 gives the original data and centred moving average values, denoted by MA(2), give an estimate of T × C. Then we calculate seasonal relatives S × I as percentages. Seasonal relatives in percentages are calculated by the ratio of y_t to MA (2) in percentages:

$$\text{Seasonal relative} = S \times I \times 100 = (y_t / MA(2)) \times 100$$

These are given for the years 2009-2012 in the following table in bold figures.

Table 2: Calculation of seasonal relatives for the given time series

Year (1)	Month (2)	Sales (3)	Moving Average MA(1) (4)	Centred Moving Average MA (2) (5)	Seasonal Relatives (3) / (5)
2009	January	500			
	February	600			
	March	650			
	April	750			
	May	800			
	June	800			
			825.0		
	July	850		827.00	**102.78**
			829.0		

	August	900		827.00	**108.827**
			825.0		
	September	900		823.00	**109.356**
			821.0		
	October	950		816.75	**116.314**
			812.5		
	November	1100		808.25	**136.096**
			804.0		
	December	1100		800.00	**137.500**
			796.0		
2010	January	550		791.75	**69.466**
			787.5		
	February	550		781.25	**70.4**
			775.0		
	March	600		768.75	**78.05**
			762.5		
	April	650		756.25	**85.95**
			750.0		
	May	700		741.75	**94.371**
			733.5		
	June	700		729.25	**95.989**
			725.0		
	July	750		723.00	**103.734**
			721.0		
	August	750		723.00	**103.734**
			725.0		
	September	750		723.00	**103.734**
			721.0		
	October	800		718.75	**111.3**
			716.5		
	November	900		714.50	**125.96**
			712.5		
	December	1000		714.50	**139.958**
			716.5		
2011	January	500		716.50	**69.78**
			716.5		
	February	600		720.75	**83.246**
			725.0		
	March	550		731.25	**75.213**
			737.5		
	April	600		745.75	**80.456**
			754.0		
	May	650		762.50	**85.246**
			771.0		
	June	750		779.25	**96.246**
			787.5		
	July	750		791.75	**94.727**
			796.0		

	August	850		798.00	**106.516**
			800.0		
	September	900		804.25	**111.905**
			808.5		
	October	1000		814.75	**122.737**
			821.0		
	November	1100		827.25	**132.97**
			833.5		
	December	1200		839.75	**142.899**
			846.0		
2012	January	600		856.25	**70.073**
			866.5		
	February	650		872.75	**74.477**
			879.0		
	March	650		885.25	**73.425**
			891.5		
	April	750		895.75	**83.728**
			900.0		
	May	800		904.25	**88.471**
			908.5		
	June	900		910.50	**98.846**
			912.5		
	July	1000			
	August	1000			
	September	1050			
	October	1100			
	November	1200			
	December	1250			

Once we have obtained S × I, we can take the average to eliminate the effect of irregularity I. This gives seasonal indices, S_i. Now we prepare a two way table which will include the percentage value of column (6) of Table 2 month-wise for every year as follows:

Table 3: Seasonal indices for the given time series

Year / Months	2009	2010	2011	2012	Median	Seasonal Indices
January	--	69.466	69.78	70.073	69.78	70.025
February		70.4	83.246	74.477	74.477	74.738
March		78.05	75.213	73.425	75.213	75.477
April		85.95	80.456	83.728	83.728	84.02
May		94.37	85.246	88.471	88.471	88.78
June		95.989	96.246	98.846	96.246	96.584
July	102.78	103.734	94.727		102.78	103.14
August	108.827	103.734	106.516		106.516	106.89
September	109.356	103.734	111.905		109.356	109.74
October	116.314	116.314	122.737		116.314	116.72
November	136.096	125.962	132.97		132.971	133.437
December	137.5	139.958	142.899		139.958	140.446
Total					1195.81	1200
Average					99.65	100

14.2.3 Ratio to Trend Method

This method provides seasonal indices free from trend and is an improved version of the simple average method as it assumes that seasonal variation for a given period is a constant fraction of the trend. The measurement of the seasonal indices by this method consists of the following steps:

Step 1: We obtain the trend values by the method of least squares for each period by establishing the trend by fitting a straight line or second degree parabola or a polynomial.

Step 2: To express the original time series values as percentages of the trend value, we divide each original value by the corresponding trend value and multiply it by 100. The indices so obtained are free from the trend.

Step 3: To obtain the seasonal indices free from the cyclic and irregular variations, we find average (mean or median) of ratio to trend values (or percentages values) for each season for any number of years. It is suggested that median be preferred instead of mean if some extreme values are present, which are not primarily due to seasonal effects. In this way, the irregular variation is removed. If there are a few abnormal values in the percentage values, the mean should be preferred to remove the randomness.

Step 4: If the seasonal periods are quarters, the sum of seasonal indices in the case of multiplicative model should be 400 and if the periods are in months, it should be 1200. After calculating the seasonal indices, the expected total is divided by actual total of the indices for quarterly and monthly data, respectively. Most of the time, the sum of all calculated seasonal indices is not exactly the same, as it is expected to be.

By following the above step-wise procedure, we calculate the seasonal indices using this method. This method is based on sound and logical footing and utilises complete information.

Example 3: Compute the seasonal indices for the following time series of sales (in thousand `) of a commodity by the ratio to trend method:

Year	I Quarter	II Quarter	III Quarter	IV Quarter
2008	800	920	880	820
2009	540	760	680	620
2010	400	580	540	480
2011	340	520	500	440
2012	300	400	360	340

Solution: We are given the time series data for 5 years of quarterly sales of a commodity. To compute the seasonal indices, we first determine the trend value for the yearly averages (Y) by fitting a linear trend by the method of least squares. The following table is constructed for fitting the straight line:

$$Y = a + b\left(X - \bar{X}\right) = a + bU$$

where U = $\left(X - \bar{X}\right)$

Table 4: Trend values for the yearly averages

Year X	Yearly Total	Yearly Average (y)	U = X−2010	U×Y	U^2	Trend Value
2008	3400	850	−2	−1700	4	800
2009	2600	650	−1	−650	1	680
2010	2000	500	0	0	0	560
2011	1800	450	1	450	1	440
2012	1400	350	2	700	4	320
Total		2800	0	−1200	10	2800

For the straight line Y = a + b X, the normal equations for estimating a and b are:

$$\sum Y = na + b\sum U$$

$$\sum UY = a\sum U + b\sum U^2$$

Now we put the values of $\sum Y$, $\sum U$, $\sum UY$, $\sum U^2$ in the normal equations:

$$5a = 2800 \Rightarrow a = 560$$

$$10b = -1200 \Rightarrow b = -120$$

On putting the optimum value of a and b in the equation of the straight line Y = a + bX, we get the trend line for the given time series data as:

$$Y = 560 - 120\,U$$

Thus, the trend values for each value of U are obtained as follows:

$$U = -2,\quad Y = 560 - 120\,(-2) = 800$$

$$U = -1,\quad Y = 560 - 120\,(-1) = 680$$

$$U = 0,\quad Y = 560 - 120\,(\,0\,) = 560$$

$$U = 1,\quad Y = 560 - 120\,(\,1\,) = 440$$

$$U = 2,\quad Y = 560 - 120\,(\,2\,) = 320$$

Since yearly decline in the trend value is −120, the quarterly increment would be

$$\text{Quarterly increment} = \frac{-120}{4} = -30$$

Now we determine the quarterly trend values as follows:

For 2008, the trend value for the middle quarter, i.e., half of the second quarter and half of the third quarter is 800. Since the quarterly increment is −30, we obtain the trend value for the 2nd quarter as 800 − (−15) and for the 3rd quarter as 800 + (−15). Thus, these are 815 and 785, respectively. Consequently, the trend value for the first quarter is 815 − (− 30) = 845 and for the 4th quarter, it is 785 + (−30) = 755. Similarly, we can get the trend values for other years as we have obtained for all the quarters of the year 2008. After calculating the trend values, we also calculate the seasonal indices for each quarter of every year, which estimates the trend component from the data. This is shown in Table 5.

Table 5: Calculations for seasonal indices

Trend Values					Seasonal Indices (Given value as % of trend values)				Total
Year	1st Qtr	2nd Qtr	3rd Qtr	4th Qtr	1st Qtr	2nd Qtr	3rd Qtr	4th Qtr	
2008	845	815	785	755	94.67	112.88	112.1	108.61	428.26
2009	725	695	665	635	74.48	109.35	102.25	97.64	383.72
2010	605	575	545	515	66.11	100.87	99.08	93.2	359.26
2011	485	455	425	395	70.1	114.28	117.64	111.39	413.41
2012	365	335	305	275	82.19	119.4	118.03	123.63	443.25
Total					387.55	556.78	549.1	534.47	2027.9
Average (A. Mean)					77.51	111.356	109.82	106.892	405.08
Adjusted Seasonal Index					76.445	109.824	108.301	105.42	399.99

The average yearly seasonal indices obtained above are adjusted to a total of 400 because the total of the seasonal indices for each quarter is 405.578, which is greater than 400. So we multiply each quarterly index by the ratio

$$K = \frac{400}{\text{Total of Indices}} = \frac{400}{405.08} = 0.986.$$

The adjusted seasonal indices for each quarter are given in the last row of the table.

You may now like to solve the following problems to check your understanding of the three methods explained so far.

E1) Determine the seasonal indices for the data given below for the average quarterly prices of a commodity for four years:

Years	Quarter I	Quarter II	Quarter III	Quarter IV
2009	554	590	616	653
2010	472	501	521	552
2011	501	531	553	595
2012	403	448	460	480

E2) Calculate the seasonal indices for the following data of production of a commodity (in hundred tons) of a firm using the ratio to trend method:

Years	Quarter I	Quarter II	Quarter III	Quarter IV
2001	470	531	500	480
2002	340	450	410	380
2003	270	360	340	310
2004	240	330	320	290
2005	220	270	250	240

E3) Apply the ratio to moving average method for calculating the seasonal indices for the time series data given in Example 8 of Unit 13.

14.3 ESTIMATION OF TREND FROM DESEASONALISED DATA

In Unit 13, you have learnt the estimation of trend. However, when a substantial seasonal component is present in the data, it is advisable to first remove the seasonal effect from the data. Otherwise, the trend estimates are also affected by seasonal effects, which makes the estimation unreliable. Hence, after estimating the seasonal indices, we deseasonalise the data values by dividing them by the corresponding seasonal indices (S_t). Thus, the deseasonalised values are given by

$$\text{Deseasonalised } Z_t = \frac{y_t}{S_t} = T_t \times C_t \times I_t$$

Once the data are made free from the seasonal effect, we estimate the trend line by the method of least squares as explained in Sec.13.4 of Unit 13. Thus, we have a reasonably good estimate of T_t, C_t and S_t.

We have not described the estimation of cyclic effect (C_t) separately. A cycle in the time series means a business cycle, which normally exceeds one year in duration. Note that hardly any time series possess strict cycles because cycles are never regular in periodicity and amplitude. This is why the business cycles are the most difficult types of economic fluctuation to measure. To construct meaningful typical cycle indices of curves similar to those that have been developed for trends and seasons is impossible. The successive cycles vary widely in time, amplitude and patterns and are inextricably mixed with irregular factors. Since it is very difficult to distinguish cyclic effects from long term trend effect, and in most cases we assume that either they are not present or they are estimated along with trend, we estimate $T_t C_t$ jointly by the least squares method.

Once we have estimated all components, we shall use them for forecasting purposes as described in the next section.

Example 4: Use the following data and calculate the deseasonalised values T × C × I. Use these values to estimate the trend line:

Years	Quarter I	Quarter II	Quarter II	Quarter IV
2003	289	241	273	232
2004	336	294	363	274
2005	297	270	263	198
2006	291	209	243	187

Solution: For calculating the deseasonalised values, we first calculate the seasonal indices for the given data. The seasonal indices are given in Table 6. Deseasonalised values are also calculated by dividing the time series values by the corresponding indices S_t.

Table 6: Calculations for obtaining trend values

Year X_t		Y_t	S_t	Deseasonalised values $Z_t = (Y_t/S_t)\times 100$	$(X_t - \bar{X}_t)$	$(X_t - \bar{X}_t)^2$	$Z_t(X_t - \bar{X}_t)$
2003	Q_1	289	113.13	255.46	1.875	3.515625	478.9875
	Q_2	241	94.82	254.17	1.625	2.640625	413.02625
	Q_3	273	107.43	254.12	1.375	1.890625	349.415
	Q_4	232	84.62	274.17	1.125	1.265625	308.44125

2004	Q_1	336	113.13	297.00	0.875	0.765625	259.875
	Q_2	294	94.82	310.06	0.625	0.390625	193.7875
	Q_3	363	107.43	337.89	0.375	0.140625	126.70875
	Q_4	274	84.62	323.80	0.125	0.015625	040.475
2005	Q_1	297	113.13	262.53	–0.125	0.015625	–032.81625
	Q_2	270	94.82	284.75	–0.375	0.140625	–106.78125
	Q_3	263	107.43	244.81	–0.625	0.390625	–153.00625
	Q_4	198	84.62	233.99	–0.875	0.765625	–204.74125
2006	Q_1	291	113.13	257.23	–1.125	1.265625	–289.38375
	Q_2	209	94.82	220.42	–1.375	1.890625	–303.0775
	Q_3	243	107.43	226.19	–1.625	2.640625	–367.55875
	Q_4	187	84.62	220.99	–1.875	3.515625	–414.35625
Total				**4257.58**		**21.25**	**298.995**

Once the data are deseasonalised, we apply the method of least squares to estimate the trend equation. The following values are calculated (as given in the above table):

$$\overline{X}_t = 2004.5, \quad \sum\left(X_t - \overline{X}_t\right)^2 = 21.25 \quad \sum Z_t\left(X_t - \overline{X}_t\right) = 298.995$$

$$\hat{a} = \overline{Z} = 266.10 \qquad \hat{b} = \frac{\sum Z_t\left(X_t - \overline{X}_t\right)}{\sum\left(X_t - \overline{X}_t\right)^2} = 14.07$$

The fitted trend equation is:

$$Y_t = 266.10 + 14.07\left(X_t - \overline{X}_t\right)$$

14.4 FORECASTING

Forecasting is one of the main purposes of time series analysis. It is always a risky task as one has to assume that the process will remain the same in future as in the past, at least for the period for which we are forecasting. This assumption is not very realistic and we shall assume that at least for short term forecasting, the process remains the same as in the past.

If a time series plot shows that there is no seasonal effect, or on theoretical basis there is no reason for having a seasonal component (S_t), we can estimate the trend component (T_t) and use the trend equation for forecasting.

If it is empirically observed that there is a significant seasonal effect (S) and on theoretical ground also there is a valid reason for the presence of such a component, we have to take the seasonal effect into account while estimating and forecasting. If data are collected on monthly basis and the period of seasonality is one year, we estimate twelve seasonal indices, one for each month. If data is quarterly, we estimate four seasonal indices. After deseasonalising the data, we fit the trend equation. Then we project the trend for the period for which we have to forecast. Next we adjust it for seasonal effect by multiplying it by the corresponding seasonal index. This gives the final forecast value which has been corrected for the seasonal effect. These steps can be described as follows:

Step 1: Calculate the moving average of suitable order. The order of moving average is taken as the period of the seasonal effect.

Step 2: Calculate the ratio of data to moving average values so that this ratio contains the seasonal component (S_t) and the irregular component (I_t), i.e., $S_t \times I_t$.

Step 3: Determine the seasonal indices by averaging $S_t \times I_t$ for respective seasons. This gives the estimates of the seasonal component (S_t).

Step 4: Obtain deseasonalised values by dividing data values by the corresponding seasonal component (S_t).

Step 5: Fit the trend equation to deseasonalised data. Compute deseasonalised forecast value from the trend equation.

Step 6: Adjust the forecast value for seasonal effect by multiplying it by the corresponding S_t. If the additive model is used, instead of multiplying or dividing we add or subtract the corresponding values. Seasonal effects are normalised so that their sum is equal to zero.

Example 5: Using estimated seasonal indices and the fitted trend equation of Example 4, forecast the value for the first quarter (Q_1) of 2007.

Solution: We have fitted the trend equation in Example 4 as:

$$Y_t = 266.10 + 14.07\left(X_t - \overline{X}_t\right)$$

The projected trend value for Q_1 of 2007 is:

$$\hat{Y}_t = 266.10 + 14.07\ (2007.25 - 2004.5)$$

where $\overline{X}_t = 2004.5$

or $\hat{Y}_t = 266.10 + 14.07 \times 2.75$

$= 304.79$

The season corrected forecast = $(\hat{Y}_t \times S_t)$ /100

$= (304.79 \times 84.62)$ /100

$= 257.9$

You may now like to solve the following exercises to assess your understanding of forecasting.

E4) The following table gives the sales figures (in thousands) of television sets for 16 quarters over four years, coded as 1, 2, 3 and 4:

Year \ Quarter	Q_1	Q_2	Q_3	Q_4
1	480	410	600	650
2	580	520	680	740
3	600	560	750	780
4	630	590	800	840

i) Calculate the four quarters centred moving average values.

ii) Compute the seasonal indices for the four quarters.

iii) Obtain deseasonalised values and estimate the trend line.

iv) Obtain the season adjusted forecast value (Q_3) of the fifth year.

E5) The following data give production of a certain brand of motor vehicles. Determine indices using the ratio to moving average method for August and September, after calculating the centred moving average for twelve months.

Production (in thousand units)

Year	Jan	Feb	Mar	Apr	May	Jun	July	Aug	Sep	Oct	Nov	Dec
1985	7.92	7.81	7.91	7.03	7.25	7.17	5.01	3.90	4.64	7.03	6.88	6.14
1986	4.86	4.48	5.26	5.48	6.42	6.82	4.98	2.45	4.51	6.38	7.55	7.59

E6) Given below are the data of production of a company (in lakhs of units) for the years 1973 to 1979:

Year	1973	1974	1975	1976	1977	1978	1979
Production	15	14	18	20	17	24	27

i) Compute the linear trend by the method of least squares.

ii) Compute the trend values for each year.

Let us now summarise the concepts that we have discussed in this unit.

14.5 SUMMARY

1. When time series data do not contain any trend and cyclic components but reflect only seasonal variation, we have to estimate the seasonal component S by removing irregular components.

2. If the effect of seasonal variation is not removed from the time series data, the trend estimates are also affected. Therefore, we have to deseasonalise the data by dividing it by the corresponding seasonal indices. Once the data is free from the seasonal effects, we estimate the trend equation using the method of least squares.

3. If a cycle is completed in 3 months, we calculate the moving average for 3 months. By taking the average over all available values, we get seasonal indices S_i for i =1, 2, 3, 4.

4. We usually normalise the seasonal indices so that their mean is 100 by dividing them by mean of S_i for all i and multiplying by 100. The normalised seasonal indices have mean 100. Usually, there is not much difference between normalised and non-normalised seasonal indices.

5. The ratio to trend method provides seasonal indices free from trend and is an improved version of the simple average method as it assumes that seasonal variation for a given period is a constant fraction of the trend.

6. When substantial seasonal component is present in the data, it is advisable to first remove the seasonal effect from the data. Otherwise, the trend estimates are also affected by seasonal effects, which make the estimation unreliable. Hence, after estimating the seasonal indices, we deseasonalise the data values by dividing it by corresponding seasonal indices (S_t). Thus, the deseasonalised values are given by

$$\text{Deseasonalised } Z_t = \frac{y_t}{S_t} = T_t \times C_t \times I_t$$

7. A cycle in the time series means a business cycle, which normally exceeds one year in length. Note that hardly any time series possess

strict cycles because cycles are never regular in periodicity and amplitude. This is why the business cycles are the most difficult type of economic fluctuations to measure. To construct meaningful typical cycle indices of curves similar to those that have been developed for trends and seasons is impossible. The successive cycles vary widely in time, amplitude and patterns and are inextricably mixed with irregular factors. Since it is very difficult to distinguish cyclic effects from long term trend effect, and in most cases we assume that either they are not present or they are estimated along with trend, we estimate T_t C_t jointly by the least squares method.

14.6 SOLUTIONS/ANSWERS

E1) For determining the seasonal indices we first arrange the given data as follows:

Table 7: Seasonal Indices for quarterly price of a commodity

Years Quarters	2009	2010	201	2012	Total	Quarterly Avgs	Seasonal Index
I	554	472	301	403	1930	482.5	91.51
II	590	501	531	448	2070	517.5	98.21
III	616	521	553	460	2150	537.5	102.01
IV	653	552	595	480	2280	570.0	108.18
Total					**8430**	**2107.5**	**400**
Average					**2107.5**	**526.9**	**100**

We are given quarterly prices of a commodity for 4 years. For calculating the quarterly variation indices, we first calculate the quarter-wise total of the production amount for the 4 years. Then we calculate the quarterly averages for all 4 quarters. For Quarter I, it is 482.5, for Quarter II, 517.5, for Quarter III, 537.5 and for Quarter IV, 570.

After calculating the quarterly averages, we calculate the average of all quarterly averages:

$$\bar{\bar{y}} = \frac{1}{4}\left(482.5 + 517.5 + 537.5 + 570\right) = 526.9$$

Now we calculate the seasonal indices, one at a time, for i = 1, 2, 3, 4, as follows:

$$\text{Seasonal index for Quarter I } = \frac{482.5}{526.9} \times 100 = 91.57$$

$$\text{Seasonal index for Quarter II } = \frac{517.5}{526.9} \times 100 = 98.21$$

$$\text{Seasonal index for Quarter III } = \frac{537.5}{526.9} \times 100 = 102.01$$

$$\text{Seasonal index for Quarter IV } = \frac{570}{526.9} \times 100 = 108.18$$

E2) We are given the time series data for 5 years of quarterly sales of a commodity. To compute the seasonal indices, we first determine the trend value for the yearly averages (Y) by fitting a linear trend by the method of least squares. The following table is constructed for fitting the straight line Y = a + bX:

Table 8: Calculation of trend values for yearly averages

Year X	Yearly Total	Yearly Average (y)	U = X−2003	U×Y	U^2	Trend Value
2001	1980	495	−2	−990	4	470
2002	1580	395	−1	−395	1	410
2003	1280	320	0	0	0	350
2004	1180	295	1	295	1	290
2005	980	245	2	490	4	230
Total		1750	0	−600	10	1750

For the line $Y = a + b\left(X - \bar{X}\right) = a + bU$, the normal equations for estimating a and b are

$$\sum Y = na + b\sum U$$

$$\sum UY = a\sum U + bU^2$$

Now we put the values of $\sum Y$, $\sum U$, $\sum UY$, $\sum Y^2$ in the normal equations:

$$5a = 1750 \Rightarrow a = 350$$

$$10b = -600 \Rightarrow b = -60$$

Therefore, by putting the optimum value of a and b in the equation of the line Y = a + bU, we get the trend line for the given time series data as:

$$Y = 350 - 60\,U$$

Thus, the trend values for different values of U are as follows:

$$U = -2,\quad Y = 350 - 60(-2) = 470$$

$$U = -1,\quad Y = 350 - 60(-1) = 410$$

$$U = 0,\quad Y = 350 - 60(0) = 350$$

$$U = 1,\quad Y = 350 - 60(1) = 290$$

$$U = 2,\quad Y = 350 - 60(2) = 230$$

As the yearly decline in the trend value is −60, the quarterly decline is

$$\text{Quarterly increment } = \frac{-60}{4} = -15$$

We now determine the quarterly trend values as follows:

For 2001, the trend value for the middle quarter, i.e., half of the second quarter and half of the third quarter is 400. Since the quarterly increment is 15, we obtain the trend values for 2^{nd} quarter as 470 – (−7.5) and for 3^{rd} quarter as 470 +(−7.5), i.e., 477.5 and 462.5, respectively. Consequently the trend value for the first quarter is 477.5–(–15) = 492.5 and for the 4^{th} quarter, it is 462.5+(−15) =447.5. We can get the trend values for other years in the same way as for all quarters of the year 2001. After calculating the trend values, we also calculate the seasonal indices for each quarter of every year, which estimate the trend component from the data. These are given in the following table:

Table 9: Calculations for seasonal indices

Trend Values					Seasonal Indices (Given value as % of trend values)			
Year	**Qtr I**	**Qtr II**	**Qtr III**	**Qtr IV**	**Qtr I**	**Qtr II**	**Qtr III**	**Qtr IV**
2001	492.5	477.5	462.5	447.5	95.43	110.99	108.1	107.26
2002	432.5	417.5	402.5	387.5	78.61	107.78	101.86	98.06
2003	372.5	357.5	342.5	327.5	72.48	100.7	99.27	94.65
2004	312.5	297.5	282.5	267.5	76.8	110.92	113.27	108.4
2005	252.5	237.5	222.5	207.5	87.13	113.68	112.36	115.66
Total					410.45	544.07	534.86	524.03
Average (A. Mean)					82.09	108.81	106.97	104.80
Adjusted Seasonal Index					81.539	108.081	106.253	104.097

The average yearly seasonal indices obtained above are adjusted to a total of 400 because the total of the seasonal indices for each quarter is 402.67, which is greater than 400. So we multiply each quarterly index by the ratio

$$K = \frac{400}{\text{Total of Indices}} = \frac{400}{402.67} = 0.9933$$

The adjusted seasonal indices for each quarter are given in the last row of the table.

E3) As described in Sec. 13.5 we shall first eliminate the effect of S_t from time series observations. Since the data are quarterly and the period of seasonal effect is one year, taking a moving average for 4 quarters will eliminate the effect of S_t. In this case, we have to calculate centred moving average by taking average of two successive moving average values. The following table gives the original data and centred MA (2) values as:

Table 10: Calculation for MA(1), MA(2) and seasonal relatives

Year (1)	Quarter (2)	Sales (3)	Moving Average (4)	Centered Mov. Average (5)	Percentage Values
2001	Qtr 1	935			
	Qtr 2	1215			
			1162.50		
	Qtr 3	1045		1169.375	89.364
			1176.25		
	Qtr 4	1455		1188.75	122.397
			1201.25		
2002	Qtr 1	990		1239.375	79.878
			1277.50		
	Qtr 2	1315		1281.25	102.634
			1285.00		
	Qtr 3	1350		1332.50	101.313
			1380.00		
	Qtr 4	1485		1442.50	102.946
			1505.00		
2003	Qtr 1	1370		1520.00	90.131
			1535.00		

	Qtr 2	1815		1559.125	116.411
			1583.25		
	Qtr 3	1470		1619.75	90.754
			1656.25		
	Qtr 4	1680		1600.00	105.00
			1543.75		
2004	Qtr 1	1160		1510.625	76.789
			1477.50		
	Qtr 2	1365		1448.625	94.227
			1419.75		
	Qtr 3	1205		1340.50	89.89
			1261.25		
	Qtr 4	1445		1275.00	113.33
			1288.75		
2005	Qtr 1	1030		1287.50	80.00
			1286.25		
	Qtr 2	1475		1303.75	113.135
			1321.25		
	Qtr 3	1195		1340.625	89.137
			1360.00		
	Qtr 4	1585		1341.875	118.118
			1323.75		
2006	Qtr 1	1185		1361.875	87.012
			1400.00		
	Qtr 2	1330		1470.00	90.476
			1540.00		
	Qtr 3	1500		1568.125	95.655
			1596.25		
	Qtr 4	2145		1695.00	126.548
			1793.75		
2007	Qtr 1	1410		1845.625	76.396
			1897.50		
	Qtr 2	2120		1845.625	109.95
			1958.75		
	Qtr 3	1915		1928.125	94.948
			2075.00		
	Qtr 4	2390		2016.875	115.007
			2081.25		
2008	Qtr 1	1875		2078.125	89.82
			2093.75		
	Qtr 2	2145		2145.00	100.00
			2196.25		
	Qtr 3	1965		2195.00	89.52
			2193.75		
	Qtr 4	2800		2190.00	127.853
			2186.25		
2009	Qtr 1	1865		2182.50	85.452
			2178.75		
	Qtr 2	2115		2099.375	100.744
			2020.00		
	Qtr 3	1935			
	Qtr 4	2165			

Centred moving average values, denoted by MA(2), give an estimate of $T_t \times C_t$. We calculate seasonal relatives $S_t \times I_t$ in percentages as follows:

$$S_t \times I_t \times 100 = (Y_t / MA(2)) \times 100$$

These are given for the years 1973−1976 in the above table in bold figures. Now we prepare a two way table, which includes the quarterly percentage moving average values year-wise as follows:

Table 11: Calculations for seasonal indices

Years / Qtrs	2001	2002	2003	2004	2005	2006	2007	2008	2009	Median	Indices
Qtr 1		79.87	90.13	76.78	80.00	87.01	76.39	89.82	85.45	82.72	84.56
Qtr 2		102.63	116.44	94.22	113.13	90.47	109.95	100.00	100.74	101.68	103.95
Qtr 3	89.364	101.31	90.75	89.89	89.13	95.65	94.95	89.52		90.32	92.33
Qtr 4	122.39	102.94	105.00	113.33	118.12	126.55	115.01	127.85		116.56	119.15
Total										391.29	400
Average											100

Once we have obtained S × I, we can take the average to eliminate the effect of irregularity I. This gives seasonal indices S_i. We note that for some quarters, three values are available and for some quarters, four values are available. By taking the average over three or four values, we get seasonal indices S_i for i = 1, 2, 3, 4. We usually normalise them so that their mean is 100. These normalised values of S_i have mean 100. These are obtained in the table given above. Usually not much difference exists between normalised and non-normalised seasonal indices. When the data are collected on monthly basis, the same procedure will yield twelve monthly seasonal indices. This method of estimating seasonal indices is known as **ratio-to-moving average** method.

E4) From the given data for sales of television sets, we have the following table:

Table 12: Calculations for seasonal indices and deseasonalised values

Year	Quarter	Sales (y)	Centred MA(MA(1))	S × I × 100 = y/MA	Seasonal Index S_i	Deseasonalised Value (Z_t)
1	1	480	-	-	93.00	516
	2	410	-	-	83.67	490
	3	600	547.5	109.6	109.14	550
	4	650	573.8	113.3	114.13	570
2	5	580	597.5	97.1	93.06	623
	6	520	618.8	84.0	83.67	621
	7	680	632.5	107.5	109.14	623
	8	740	640.0	115.6	114.13	648

3	9	600	653.8	91.8	93.06	645
	10	560	667.5	83.9	83.67	669
	11	750	676.3	110.9	109.14	687
	12	780	683.8	114.1	114.13	683
4	13	630	693.8	90.8	93.06	677
	14	590	707.5	83.4	83.67	705
	15	800	-	-	109.14	733
	16	840	-	-	114.13	736
	$\bar{X} = 8.5$					

From the above table we get

$$\bar{X}_t = 8.50, \quad \sum \left(X_t - \bar{X}_t\right)^2 = 340.00$$

and $$\sum Z_t \left(X_t - \bar{X}_t\right) = 4870.5$$

$$\hat{a} = \bar{Z} = 636 \qquad \hat{b} = \frac{\sum Z_t \left(X_t - \bar{X}_t\right)}{\sum \left(X_t - \bar{X}_t\right)^2} = 14.325$$

The trend equation is

$$\hat{y} = 636 + 14.325\,(X_t - 8.50)$$

For Q_3 of fifth year, $X_t = 19.00$. The forecast for trend value is

$$\hat{y} = 636 + 14.325\,(19.00 - 8.50)$$

$$= 786.4125$$

Season corrected forecast = (786.4125 × 109.14) /100

= 858.29

E5) Values obtained from the data are as follows:

Table 13: Calculation of seasonal indices

	July	Aug.	Sep.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	Apr.	May	June
Centred MA	6.43	6.16	5.91	5.74	5.64	5.59	5.58	5.51	5.45	5.41	5.42	5.50
$(y_t/MA)x100$	77.92	63.31	78.51	122.47	121.99	109.84	87.10	81.31	96.51	101.29	118.45	124.00
Normalised S_i	79.06	64.23	79.65	124.26	123.77	111.49	88.37	82.49	97.92	102.77	120.18	125.80

The required seasonal indices for August and September are:

$S_{Aug} = 64.23, \quad S_{Sep} = 79.65$

The seasonal indices have been normalised so that their average is 100.

E6) From the given information, we have calculated:

$$\hat{a} = \overline{Z} = 19.28$$

We also have $\sum\left(X_t - \overline{X}_t\right)^2 = 28$ $\quad \sum Z_t\left(X_t - \overline{X}_t\right) = 55$

$$\hat{b} = \frac{\sum Z_t\left(X_t - \overline{X}_t\right)}{\sum\left(X_t - \overline{X}_t\right)^2} = \frac{55}{28} = 1.96$$

$$\hat{Y} = 19.28 + 1.96\left(X_t - \overline{X}_t\right)$$

The estimated trend values are:

Years	1973	1974	1975	1976	177	1978	1979
$\hat{y}_t$	13.40	15.36	17.32	19.28	21.24	23.20	25.16