Chapter Ten: Time Series Analysis

When you completed this chapter, you will be able to:

- ✓ understand the components of the classical time series model;
- ✓ apply Linear Regression Model to fit linear trend;
- ✓ use Moving Average and Exponential Smoothing to fit non-linear trend;
- ✓ comprehend the importance of deciding the moving average interval and smoothing coefficient;
- ✓ forecast using Linear Regression, Moving Average and Exponential Smoothing;
- ✓ use Moving Average Method to decompose the seasonal factors of a time series;
- ✓ understand the applications and limitations of various methods in forecasting.

Reference(s): Mason Chapter 18, Berenson Chapter 19, Owen Chapter 6 and 7.
Exercise(s): Seminars 23, 24 and 25, Mason Chapter18 Exercises 1, 7, 9, 27.

Many type of business, economic and scientific data are observations on a variable usually at ***equidistant*** *point in time.* A data set of this type is called a *time series*, and the variable is called a *time-series variable*.

e.g.1 the daily reading of Hang Sang Index,
the weekly sales of automobiles,
the monthly starts in new housing construction.

Definition

Time Series is a sequence of data values for a variable Y_1, Y_2, ..., Y_n that are separated by *equal time intervals. Y_t is the notation for the value of the time series at time t.*

The analysis of a time series is done ultimately *for the purpose of forecasting*. We think that **past and present observations** of a time series variable may be used to ***forecast*** **its future value**.

Components of a Time Series

Time Series analysis is a complicated topic, and there is a diversity of opinion as to how analysis should be performed.

The most accepted approaches is to view a time series as a composition of four components:

1. Trend / Secular Trend / Long Term Trend
2. Seasonal Variation
3. Cyclical Variation
4. Irregular Variation

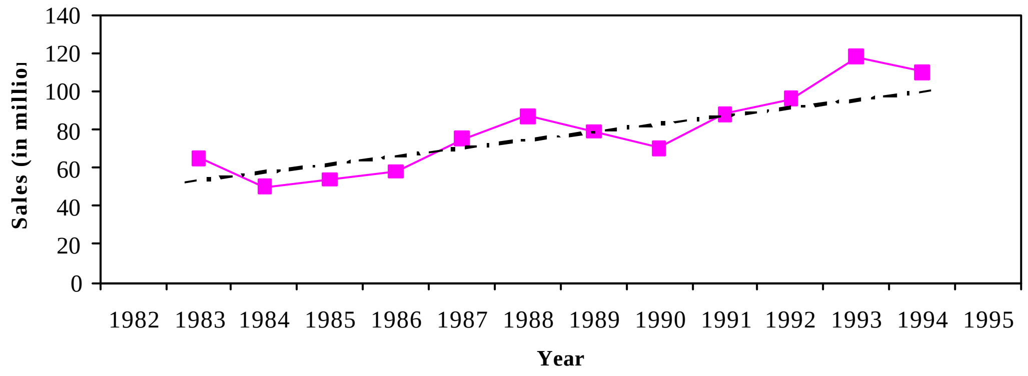
1. Trend / Secular Trend / Long Term Trend

Trend is the ***underlying movement*** in a series of figures over a time period. It is the long-term movements of the series that can be characterised by *steady* or only *slightly varied rate of change*.

e.g.2 a time series of the price of housing might vary upward and downward over short intervals of time, but would *trend gradually upward* because of long term inflation.

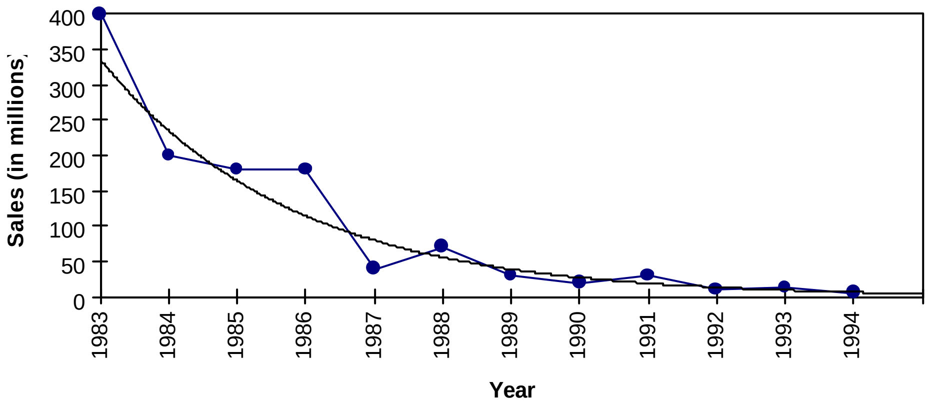
Trends can be represented by *straight lines (when the data, in figure 1, have a steady rate of change)* or by *smooth curve (when rate of change of the data , in figure 2, is slightly varied)*.

Figure 1. Time series for the data with steady rate of change



The data in figure 1 could show the sales of personal insurance policies that have become more popular.

Figure 2. Time series for the data with slightly varied rate of change



The data in figure 2 shows possibly represent sales of electric ovens, now largely replaced by the more popular microwave ovens.

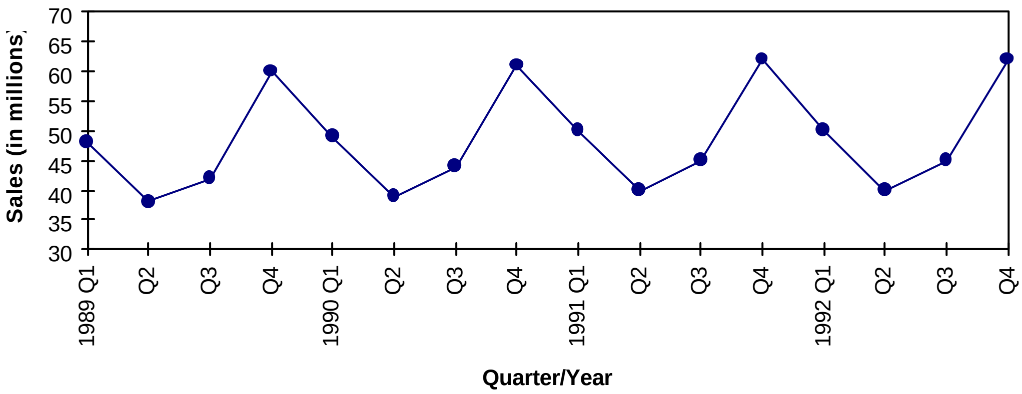
2. Seasonal Variations

This is the regular pattern in a time series that ***occurs again and again*** in the ***same months or quarters***.

e.g.3 fuel oil consumption rises in the winter and falls in the summer.

As their name indicates, seasonal variations in a time series are those variations that occur rather predictably *at a particular time each year* repeatedly year after year. Seasonal variations can be found in data recorded at *intervals of less than a **year***; quarterly, monthly or weekly data might well indicate variations of this type.

Figure 3. Time series of data with seasonal variations



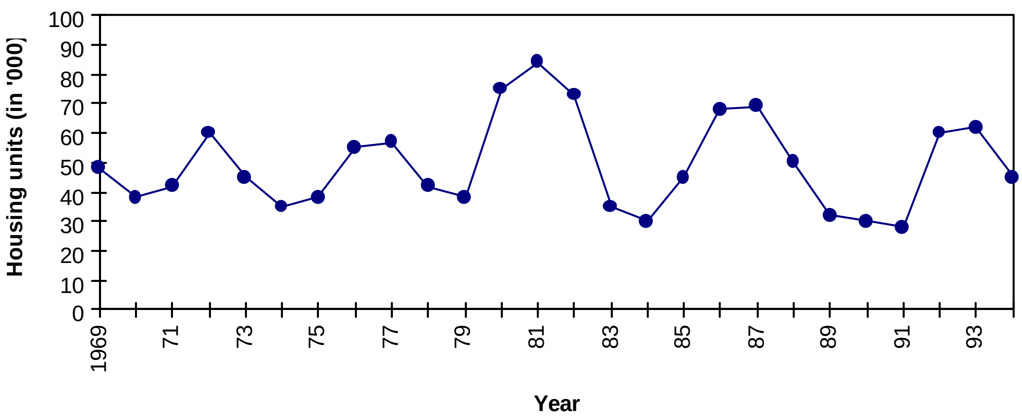
3. Cyclical Variations

These are movements in a time series that, like seasonal variations, are recurrent but that, unlike seasonal variations, *occur in cycles of longer than* ***one year***.

e.g.4 The most common example of cyclical fluctuation is the business cycle. Over time, there are years when the business cycle hits a peak above the trend line. At other times, business activity is likely to slump, hitting a low point below the trend line. The time between hitting peaks or falling to low points is at least one year, and it can be as many as 15 or 20 years, or even more.

The time series in the figure 4 below is an annual series that shows cyclical variation in the number of housing starts in the U.S. from 1969 through 1994.

Figure 4. Time series for the data with cyclical variations



Cycles are *not constant in* ***amplitude*** *and* ***duration***, and this lack of regular pattern makes their future occurrence *difficult to predict*.

Cyclical variations are often *caused by general economic conditions, Government policy changes, or shift in consumer tastes and purchasing habits.*

4. Irregular Variations / Unexplained Variations

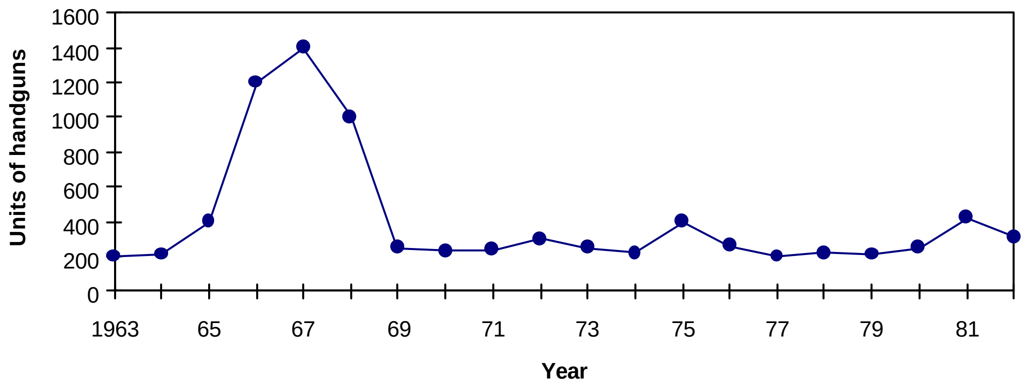
In many situations, the value of a variable may be completely ***unpredictable***, changing in a random manner. Irregular variations describe such movements. Essentially, they are the leftover movements in the time series when trend, seasonal variations, and cyclical variations have been identified. *Irregular variations cannot be predicted by using historical data, and they are not periodic in nature.* They are caused by such factors as changes in weather, wars, strikes, government legislation and elections.

e.g.5 a particular harsh winter will cause a variation in crop production and fuel oil usage.

All of these unusual changes would be called irregular variations. *One can often determine the causes of irregular variations that are large.*

The following figure, figure 5, shows the time series of the U.S. imports of handguns for private use from 1963 to 1982. The high figures in the 1966-1968 period *correspond to a period of urban violence and political assassinations.*

Figure 5. Time series for data with irregular variations



Time series models usually fall into one of two categories, depending on whether their components are expressed as sums or products. The first one is called *Additive model*, assumes that the value of y_t equals the sum of the four components.

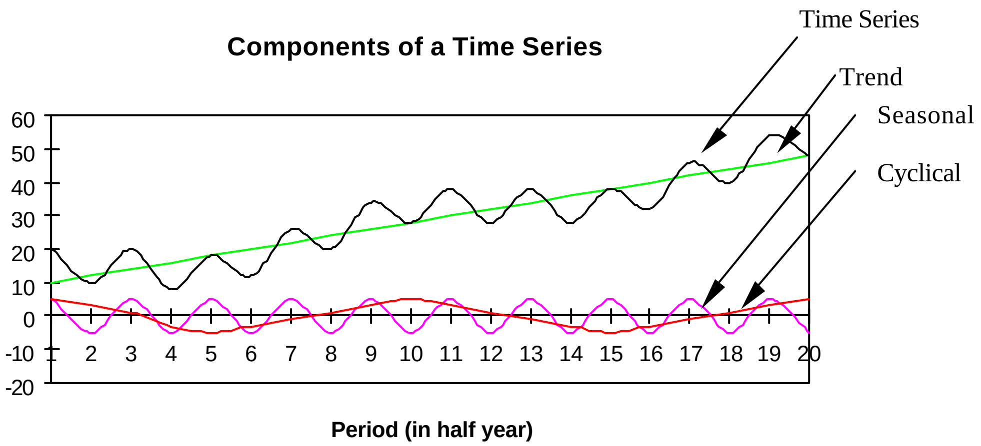
Additive model : $y_t = T + S + C + I$

By assuming that the components of a time series are additive, we are, in effect, assuming that these *components are **independent** of one another*. Thus, for example, trend cannot affect neither seasonal nor cyclical variation, nor can these components affect trend.

The other major type of relationship between the components expresses y_t in the form

Multiplicative model : $y_t = T \times S \times C \times I$

A model of this form assumes that the four components are ***related** to one another*.



Reasons for studying trends

1. *The study of trends allows us to describe a historical pattern.* There are many instances when we can use a past trend to evaluate the success of a pervious policy.

e.g.6 an university may evaluate the effectiveness of recruiting program by examining its past enrolment trends.

2. *Study of trend permits us to project past patterns, or trends, into the future.* Knowledge of the past can tell us a great deal about the future.

e.g.7 examining the growth rate of the world's population can help us to estimate the population for some future time, such that planning for education, medical, transportation, and etc, can be done.

3. In many situations, studying the secular trend of a time series *allows us to eliminate the trend component from the series*. This makes it easier for us to study the remaining three components of the time series.

e.g.8 if we want to determine the seasonal variation in ski sales, eliminating the trend component gives us a more accurate idea of the seasonal component.

e.g.9 The increases of pollutants in the environment follows an upward sloping curve similar to that in figure 6.

Another common example of curvilinear relationship is the life cycle of a new business product, illustrated in figure 7. When a new product is introduced, its sales volume is low. As the product gains recognition and success, unit sales grow at an increasingly rapid rate. After the product is firmly established, its unit sales grow at a stable rate. Finally, as the product reaches the end of its life cycle, unit's sales begin to decrease.

Figure 6. Trend of pollution increase

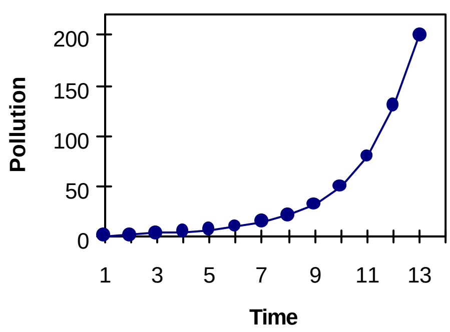
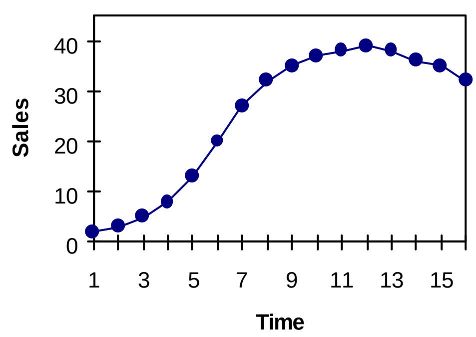


Figure 7. Typical product life cycle



Secular trend of a time series can be determined by using the smoothing techniques:
Regression (for linear trend), Moving Average and Exponential Smoothing (both for curved trend)

Fitting the linear trend line by the Least Squares Method

We can describe the general trend of many time series using a straight line. But we are faced with the problem of finding the ***best-fitting*** *line*.

Equation for a straight line : Y = m X + c
where *m is the slope, and c is the y-intercept of the straight line.*

So the equation for the best-fitting line should be in form of:

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

where : $\hat{Y}$ is the estimated value of the dependent variable
X is the independent variable (time in trend analysis)
a is the Y-intercept (the value of Y when X = 0)
b is the slope of the trend line

And we can use the *least squares method* to determine the value of a and b,

$$\boxed{b = \frac{\sum XY - n\overline{X}\overline{Y}}{\sum X^2 - n\overline{X}^2}} \qquad \text{----------} \quad ①$$

$$\boxed{a = \overline{Y} - b\overline{X}} \qquad \text{----------} \quad ②$$

where : Y represents the values of the dependent variable

X represents the values of the independent variable

$\overline{Y}$ is the mean of the values of the dependent variable

$\overline{X}$ is the mean of the values of the independent variable

n is the number of data point in the time series

a is the Y-intercept

b is the slope

e.g.10 Given the data set:

Year	1983	1984	1985	1986	1987	1988	1989
Sales (in millions)	100	120	130	150	150	170	180

find the equation of the best-fitting line.

X	Y	XY	X^2
1983	100	198300	3932289
1984	120	238080	3936256
1985	130	258050	3940225
1986	150	297900	3944196
1987	150	298050	3948169
1988	170	337960	3952144
1989	180	358020	3956121
13902	1000	1986360	27609400

by: $$b = \frac{\sum XY - n\overline{X}\overline{Y}}{\sum X^2 - n\overline{X}^2}$$

$$= \frac{1986360 - 7(13902/7)(1000/7)}{27609400 - 7(13902/7)^2} \qquad = \frac{360}{28}$$

$$= 12.8571$$

$$a = \overline{Y} - b\overline{X}$$
$$= (1000/7) - 12.8571\ (13902/7)$$
$$= -25391.4289$$

therefore, the best-fitting equation is : $\hat{Y} = -25391.4289 + 12.8571X$

Translating or Coding Time

Suppose our time series consists of only three points, 1986, 1987 and 1988. If we had to place these numbers into equations ① and ②, we would find *the* ***resultant calculations tedious***. Fortunately, we can convert these traditional measures of time to a form that simplifies the computation, we called this process *coding*.

We can *transform the values 1986, 1987, and 1988 into corresponding values of -1, 0, and 1*, where 0 represents the mean (1987), -1 represents the first year (1986 - 1987 = -1), and 1 represents the last year (1988 - 1987 = 1).

We need to consider two cases when we are coding time values. The first is a time series *with an odd number of elements*. The second is a time series *with an even number of elements*.

1. The time series with an *odd number of elements*:

X	X- $\overline{X}$	Coded Time
1983	1983 - 1986 =	**-3**
1984	1984 - 1986 =	**-2**
1985	1985 - 1986 =	**-1**
1986	1986 - 1986 =	**0**
1987	1987 - 1986 =	**1**
1988	1988 - 1986 =	**2**
1989	1989 - 1986 =	**3**
$\sum X = 13902$		
$\overline{X} = \sum X/n$ = 13902/7= 1986		

2. The time series with an *even number of elements*:

X	X- $\overline{X}$	(X- $\overline{X}$) x 2	Coded Time
1984	1984 - 1986.5	-2.5 x 2 =	**-5**
1985	1985 - 1986.5	-1.5 x 2 =	**-3**
1986	1986 - 1986.5	-0.5 x 2 =	**-1**
1987	1987 - 1986.5	0.5 x 2 =	**1**
1988	1988 - 1986.5	1.5 x 2 =	**3**
1989	1989 - 1986.5	2.5 x 2 =	**5**
$\sum X = 11919$			
$\overline{X} = \sum X/n$ = 11919/6 = 1986.5			

There are two reasons for this translation of time:

First, it *eliminates the need to square numbers as large as 1983, 1984*, and so on.

Second, this method also sets the mean year, $\overline{X}$, equal to ***zero*** and allows us to *equations ① and ②,*

$$b = \frac{\sum XY - n\overline{X}\overline{Y}}{\sum X^2 - n\overline{X}^2} \quad \text{----------} \quad ①$$

$$= \frac{\sum xY - n\overline{x}\overline{Y}}{\sum x^2 - n\overline{x}^2}$$

{$\overline{x}$ (the coded variable) substitute for $\overline{X}$, and x substitute for X}

$$b = \frac{\sum xY - n(0)(\bar{Y})}{\sum x^2 - n(0^2)}$$ {replace $\bar{x}$ by 0}

$$\boxed{b = \frac{\sum xY}{\sum x^2}}$$ ---------- ③

$a = \bar{Y} - b\bar{X}$ ---------- ②

$a = \bar{Y} - b\bar{x}$ { $\bar{x}$ substituted for $\bar{X}$ }

$a = \bar{Y} - b(0)$ { replace $\bar{x}$ by 0 }

$$\boxed{\mathbf{a = \bar{Y}}}$$ ---------- ④

e.g.11 Given the data set :

Year	1983	1984	1985	1986	1987	1988	1989
Sales (in millions)	100	120	130	150	150	170	180

find the equation of the best-fitting line.

X	$X - \bar{X}$	**Coded Time (x)**	Y	xY	x^2
1983	1983 - 1986 =	**-3**	100	-300	9
1984	1984 - 1986 =	**-2**	120	-240	4
1985	1985 - 1986 =	**-1**	130	-130	1
1986	1986 - 1986 =	**0**	150	0	0
1987	1987 - 1986 =	**1**	150	150	1
1988	1988 - 1986 =	**2**	170	340	4
1989	1989 - 1986 =	**3**	180	540	9
$\sum$ = 13902			1000	360	28

$\bar{X} = \sum X/n$
$= 13902/7$
$= 1986$

$$b = \frac{\sum xY}{\sum x^2}$$ ---------- ③

$= 360 / 28$

$= 12.8571$

$a = \bar{Y}$ ---------- ④

$= 1000 / 7$

$= 142.8571$

therefore, the equation is $\hat{Y} = 142.8571 + 12.8571\,x$

Forecasts based on the regression equation:

Suppose we want to forecast the sales in 1990. To make a prediction, we must *first translate 1990 into a coded variable x by subtracting the mean year*, 1986:

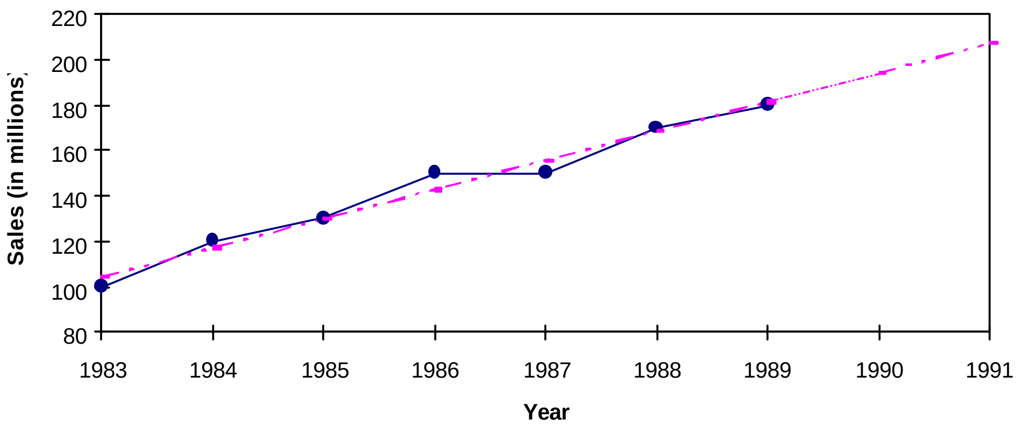
$$\begin{aligned} x &= X - \bar{X} \\ &= 1990 - 1986 \\ &= 4 \end{aligned}$$

this value, x = 4, is then substituted into the equation

$$\begin{aligned} \hat{Y} &= 142.8571 + 12.8571\ (4) \\ &= 194.2855\ \text{(millions)} \end{aligned}$$

the forecast sales in 1990 is 194.2855 millions.

Figure 8. Time series and trend line for the sales data



Class Exercise 1

Given a time series:

Year	1987	1988	1989	1990	1991	1992
Sales (in '000)	23	34	52	65	78	90

determine the equation of the best fitting line, using coded time, and hence to forecast the sales in 1993 and 1994.

Fitting the curved trend line by the Moving Average

Some time series may have irregular or random fluctuation to the extent that trends are difficult describe. In such cases it may be easier to estimate the movement of the time series if the effects of these fluctuations are removed from the data.

Methods for removing these effects are usually referred to as ***smoothing*** *techniques*. The most commonly used smoothing technique is called as *Methods of Moving Average*.

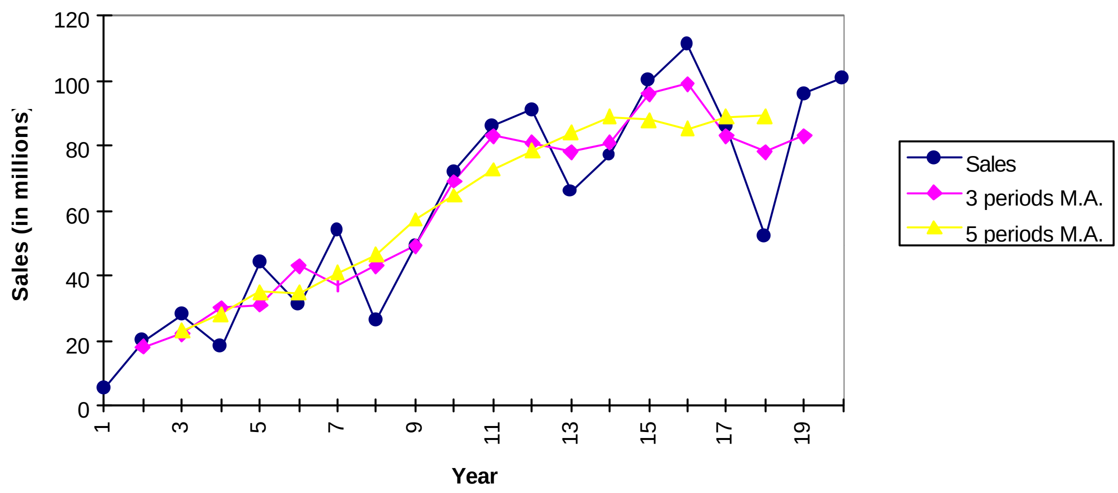
A moving average is actually *a* ***series of average***, where each average is the mean value of the time series over a fixed interval of time, and where all possible average of this time length are included in the analysis.

e.g.12 Given the past 20 years' annual sales of the EVI Co.

t (year)	1	2	3	4	5	6	7	8	9	10
Sales (in millions)	5	20	28	18	44	31	54	26	49	72
t (year)	11	12	13	14	15	16	17	18	19	20
Sales (in millions)	86	91	66	77	100	111	86	52	96	101

t (year)	**Sales (in millions)**	**3 period Moving Average**	**5 period Moving Average**
1	5	-	-
2	20	(5+20+28)/3 = 17.66	-
3	28	(20+28+18)/3 = 22	(5+20+28+18+44)/5 = 23
4	18	30	(20+28+18+44+31)/5 = 28.2
5	44	31	35
6	31	43	34.6
7	54	37	40.8
8	26	43	46.4
9	49	49	57.4
10	72	69	64.8
11	86	83	72.8
12	91	81	78.4
13	66	78	84
14	77	81	89
15	100	96	88
16	111	99	85.2
17	86	83	(100+111+86+52+96)/5 = 89
18	52	(86+52+96)/3 = 78	(111+86+52+96+101)/5 = 89.2
19	96	(52+96+101)/3 = 83	-
20	101	-	-

Figure 9. Time series and trend lines



One problem with the use of a moving average is the *choice of its length*, denoted by n, the number of consecutive values used in the averages. The *larger the value of n*, the *more the moving average smoothes out* the original data. But *as n increases so does (n-1), which is the total number of observations at the beginning and end of the data for which no moving-average value can be determined*.

Forecasts based on the Moving Average:

Moving Average may also be used to obtain simple ***short-term*** *forecasts* of the future values of the time series.

To make these forecasts, we calculate the moving average as before, however, for forecasting, *the mean of the n periods becomes the forecast for the period* ***immediately following*** *the n periods in the time series*.

Therefore, by 3 period moving average, the forecast sales in year 21 is the mean value of the last three years, that is 83 millions. Or by 5 period moving average, the forecast sales in year 21 is the mean value of the last five years, that is 89.2 millions.

Although moving average provides easy-to-calculate forecast, it makes an important assumption that the patterns of the past will hold for the future. However, if there are ***rapidly changing trends***, the moving average time series tends to lag behind the original time series, resulting in poor forecast.

Disadvantage of Moving Average Method:

- Trend value is not ***completed***. Trend values of the first and last several periods are lost.
- Moving average is calculated over a ***limited number*** of periods only, and all earlier data are ignored.
- A ***suitable period*** must be selected as a basis for calculation. The suggested period should be the regular interval of any cyclical or seasonal pattern in the series. However, some series show no such apparent pattern, then the choice of the period will be extremely difficult.

Fitting the curved trend line by Exponential Smoothing Techniques

Exponential smoothing is used to smooth out the fluctuations in a time series and to demonstrate the long-term movement in the series.

For doing the exponential smoothing techniques, the **first trend value is set to equal to the value of the data for the *first period***,

$$E_1 = Y_1$$

then all subsequent trend values are calculated as follows:

Trend for the current period
= a (actual value for current period) + (1- a) (trend value for last period)

$$E_i = \alpha Y_i + (1-\alpha)E_{i-1}$$

Where α is called the smoothing constant, and could be any value between 0 and 1;
E_i is the trend value at period i;
Y_i is the actual value at period i.

Unlike moving average, all individual past data affect the current trend value. However, the contribution of any individual value in the past to the current trend value becomes less at each successive time period.

The speed at which the effect of historical data is reduced depends on the size of the smoothing constant α. For values of α near to 1, then the effect of any past value fall quickly. Larger value of the smoothing constant, more weight is given to current period than to past periods.

When the data under consideration is subject to large random fluctuations, then ***small*** smoothing constant should be used. It would even out the random fluctuation quickly, since ***more*** weight is given to the previous trend than to any current item that might be abnormal.

Forecasts based on Exponential Smoothing Techniques

To use the exponential smoothing for the purpose of forecasting, the current trend value, say time period i, is used as the forecast of the next period, i +1.

$$\hat{Y}_{i+1} = E_i$$

Therefore, the formula can be revised as:

$$\hat{Y}_{i+1} = \alpha Y_i + (1-\alpha)\hat{Y}_i \qquad \text{where } 0 \le \alpha \le 1$$

e.g.13 Data given in the accompanying table represent the annual sales dollars (in millions) for a food processing company for the year 1995 – 1999.

Year	1995	1996	1997	1998	1999
Sales	23	30	21	40	45

a) Using a smoothing constant of 0.25, exponentially smooth the series.
b) Using a smoothing constant of 0.75, exponentially smooth the series
c) For both smoothing constants, what are the respective exponentially smoothed forecast for the trend in 2000?
d) Which smoothing constant, either 0.25 or 0.75, has greater smoothing effect? Why?

Sol] a)

$E_1 = Y_1 = 23$

$E_2 = \alpha Y_2 + (1-\alpha)E_1$ = (0.25) (30) + (1-0.25) (23) = 24.75

$E_3 = \alpha Y_3 + (1-\alpha)E_2$ = (0.25) (21) + (1-0.25) (24.75) = 23.81

$E_4 = \alpha Y_4 + (1-\alpha)E_3$ = (0.25) (40) + (1-0.25) (23.81) = 27.86

$E_5 = \alpha Y_5 + (1-\alpha)E_4$ = (0.25) (45) + (1-0.25) (27.86) = 32.15

b)

$E_1 = Y_1 = 23$

$E_2 = \alpha Y_2 + (1-\alpha)E_1$ = (0.75) (30) + (1-0.75) (23) = 28.25

$E_3 = \alpha Y_3 + (1-\alpha)E_2$ = (0.75) (21) + (1-0.75) (28.25) = 22.81

$E_4 = \alpha Y_4 + (1-\alpha)E_3$ = (0.75) (40) + (1-0.75) (22.81) = 35.70

$E_5 = \alpha Y_5 + (1-\alpha)E_4$ = (0.75) (44) + (1-0.75) (35.70) = 41.93

c) For α = 0.25, forecasted trend in 2000 = $\hat{Y}_6 = E_5 = 32.15$

For α = 0.75, forecasted trend in 2000 = $\hat{Y}_6 = E_5 = 41.93$

d) α = 0.25 has greater smoothing effect, since it gives more weight to the past data, and less weight to the current data which is highly fluctuated.

Extraction of the seasonal component through moving averages

For Additive Model ($Y_t = T + S + C + I$) :

1. **The first step in computing a seasonal index is to calculate *the 4-quarterly moving total* for the time series.**

   We use 4 period moving average, since the 4-quarter moving average *will remove seasonal variations*.

   We total the values for the quarters during the first year, that is 4240. *A moving total is associated with the middle data point in the set of values from which it was calculated*. Since the first total of 4240 was calculated from four data points, we place it opposite the midpoint of those quarters, so it falls in column (2).

2. **Compute the 4-quarterly moving average by dividing each of the 4-quarterly total by 4.**

   We divide the value in column (2) *by 4*.

3. **Centre the 4-quarterly moving average.**

   The moving averages in column (3) all fall halfway between the quarters. We could like to have moving average associated with each quarter. In order to centre our moving averages, we *associate with each quarter the average of the two 4-quarter moving averages falling just above and just below*.

4. **Calculate the difference between the actual value and the moving average value for each quarter in the time series having a 4-quarter moving average entry.**

   It allows us to recover the seasonal components for the quarters, in column (5).

5. **To collect all the difference, and arrange them by quarter.**

   The seasonal indices for each of the 4 quarters are calculated by *averaging the difference for each quarter*.

6. **Adjusts the seasonal indices slightly**

   For additive model, *the sum of the 4-quarterly indices should be zero*. However, then indices sum, of the example, is -5.6. Therefore we have to adjust the seasonal indices of the 4 quarters by adding (-5.6/4 = -1.4) 1.4 to each of the index to make the sum zero.

e.g.14 Given the time series :

Beef Usage (in Kg)	1989	1990	1991
Quarter I	1240	1460	1680
Quarter II	1020	1300	1670
Quarter III	830	1050	1440
Quarter IV	1150	1320	1800

If the time series is assumed to be *an additive model*, determine the seasonal index of the 4 quarters.

Year	Quarter	1 Usage (Kg)	2 4-quarter total	3 4-quarter average	4 centred average	5(= 1 - 4) variation
1989	I	1240			-	-
	II	1020			-	-
			4240	1060		
	III	830			**1087.5**	-257.5
			4460	1115		
	IV	1150			**1150**	0
			4740	1185		
1990	I	1460			**1212.5**	247.5
			4960	1240		
	II	1300			**1261.25**	38.75
			5130	1282.5		
	III	1050			**1310**	-260
			5350	1337.5		
	IV	1320			**1383.75**	-63.75
			5720	1430		
1991	I	1680			**1478.75**	201.25
			6110	1527.5		
	II	1670			**1587.5**	82.5
			6590	**1647.5**		
	III	1440			-	-
	IV	1800			-	-

	Quarter I	Quarter II	Quarter III	Quarter IV		
1989 1990 1991	- 247.5 201.3	- 38.8 82.5	-257.5 -260.0 -	0.0 -63.8 -		
Total Average	448.8 224.40	121.3 60.65	-517.5 -258.75	-63.8 -31.90	(-5.6)	(-1.4)
Adjustment Adjusted Seasonal Factor	1.4 **225.80**	1.4 **62.05**	1.4 **-257.35**	1.4 **-30.50**		

Therefore, the seasonal index of the 4 quarters are 225.80, 62.05, -257.35 and -30.50 respectively.

Forecasts based on moving average with seasonal variations

Since the given data are from 1989-I to 1991-IV, and moving average can only be used to obtain simple *short-term forecasts*, we can *only forecast the beef usage for one-quarter ahead (1992-I).*

The forecast usage without seasonal effect is 1647.50, and the seasonal index of quarter I is 225.8, *so the forecast usage for 1992-I = 1647.50 + 225.80 = 1873.30 Kg*

For Multiplicative Model (Y_t = T x S x C x I) :

1. **The first step in computing a seasonal index is to calculate the 4-quarterly moving total for the time series.**

2. **Compute the 4-quarterly moving average by dividing each of the 4-quarterly total by 4.**

3. **Centre the 4-quarterly moving average.**

4. **Calculate the percentage of the actual value to the moving average value for each quarter in the time series having a 4-quarter moving average entry.**

5. **To collect all the percentages, and arrange them by quarter.**

6. **Adjusts the seasonal indices slightly**

   For multiplicative model, the base for an index is 100 %, thus, the *four quarterly indices should total 400%.* However, the index sum, of the example, is 397.1%. Therefore we have to adjust the seasonal indices of the 4 quarters by dividing (397.1% / 4 =) 99.3 % to each of the index to make the sum 400 %.

e.g.15 Determine the seasonal index for the time series in e.g.14, if the time series is assumed to be a *multiplicative model*.

Year	Quarter	1 Usage (Kg)	2 4-quarter total	3 4-quarter average	4 centred average	5(= 1/4 x100) variation
1989	I	1240			-	-
	II	1020			-	-
			4240	1060		
	III	830			**1087.5**	76.3
			4460	1115		
	IV	1150			**1150**	100.0
			4740	1185		
1990	I	1460			**1212.5**	120.4
			4960	1240		
	II	1300			**1261.25**	103.1
			5130	1282.5		
	III	1050			**1310**	80.2
			5350	1337.5		
	IV	1320			**1383.75**	95.4
			5720	1430		
1991	I	1680			**1478.75**	113.6
			6110	1527.5		
	II	1670			**1587.5**	105.2
			6590	**1647.5**		
	III	1440			-	-
	IV	1800			-	-

	Quarter I	Quarter II	Quarter III	Quarter IV		
1989 1990 1991	- 120.4 113.6	- 103.1 105.2	76.3 80.2 -	100.0 95.4 -		
Total Average	234.0 117.00	208.3 104.15	156.5 78.25	195.4 97.70	397.10	99.275
Adjustment Adjusted Seasonal Factor	99.275 **117 ¸ 0.99275 = 117.85**	99.275 **104.2 ÷ 0.99275 = 104.91**	99.275 **78.3 ÷ 0.99275 = 78.82**	99.275 **97.7 ÷ 0.99275 = 98.41**		

Forecasts based on moving average with seasonal variations

The forecast usage without seasonal effect is 1647.50, and the seasonal index for quarter I is 117.8, *so the forecast usage for 1992-I = 1647.50* **x** *117.85% = 1941.58 Kg*

Deseasonalized the time series

After determining the seasonal indices, for both additive and multiplicative model, we can remove the seasonal effect from the original time series by deseasonalising the time series. Therefore, the deseasonalised time series for e.g.15 is:

Year	Quarter	1 Usage (Kg)	2 Seasonal index	(1 / 2 =)3 **Deseasonalized time series**
1989	I	1240	117.85	**1052.18**
	II	1020	104.91	**972.26**
	III	830	78.82	**1053.03**
	IV	1150	98.41	**1168.58**
1990	I	1460	117.85	**1238.86**
	II	1300	104.91	**1239.16**
	III	1050	78.82	**1332.15**
	IV	1320	98.41	**1341.33**
1991	I	1680	117.85	**1425.54**
	II	1670	104.91	**1591.84**
	III	1440	78.82	**1826.95**
	IV	1800	98.41	**1829.08**

Class Exercise 2

Given a time series as follows:

Sales (in '000)	1990	1991	1992
Quarter I	240	450	520
Quarter II	120	270	370
Quarter III	350	420	550
Quarter IV	620	680	800

If the time series is assumed to be ***multiplicative*** model,

i) determine the seasonal index for the 4 quarters, and

ii) deseasonalised the time series, and

iii) forecast the sales in 1993 quarter I.

Class : ____________ **Name :** __________________ **No. :** __________

Class Exercise 1 (Solution)

X	$X - \bar{X}$	$(X - \bar{X}) \times 2$	**Coded Time**	Y	xY	x^2
1987	1987 - 1989.5	-2.5 x 2 =	**-5**	23	-115	25
1988	1988 - 1989.5	-1.5 x 2 =	**-3**	34	-102	9
1989	1989 - 1989.5	-0.5 x 2 =	**-1**	52	-52	1
1990	1990 - 1989.5	0.5 x 2 =	**1**	65	65	1
1991	1991 - 1989.5	1.5 x 2 =	**3**	78	234	9
1992	1992 - 1989.5	2.5 x 2 =	**5**	90	450	25
$\sum X$ = 11937				342	480	70

$\bar{X} = \sum X/n$
$= 11937/6$
$= 1989.5$

$$b = \frac{\sum xY}{\sum x^2} \qquad \text{----------} \quad ③$$

$= 480/70$
$= 6.8571$

$$\mathbf{a} = \bar{Y} \qquad \text{----------} \quad ④$$

$= 342/6$
$= 57$

therefore, the equation is $\hat{Y} = 57 + 6.8571\,x$

Forecast sales in 1993 :

$x = (X - \bar{X}) \times 2 = (1993 - 1989.5) \times 2 = 7$

$\hat{Y} = 57 + 6.8571\,(7)$
$= 105$ ('000)

Forecast sales in 1994 :

$x = (X - \bar{X}) \times 2 = (1994 - 1989.5) \times 2 = 9$

$\hat{Y} = 57 + 6.8571\,(9)$
$= 118.71$('000)

Class : ____________ **Name :** ________________ **No. :** __________

Class Exercise 2 (Solution)

Year	Quarter	1 Sales ('000)	2 4-quarter total	3 4-quarter average	4 centred average	5 (= 1/4 x100) variation
1990	I	240			-	-
	II	120			-	-
			1330	332.5		
	III	350			**358.75**	97.56
			1540	385		
	IV	620			**403.75**	153.56
			1690	422.5		
1991	I	450			**431.25**	104.35
			1760	440		
	II	270			**447.5**	60.34
			1820	455		
	III	420			**463.75**	90.57
			1890	472.5		
	IV	680			**485**	140.21
			1990	497.5		
1992	I	520			**513.75**	101.22
			2120	530		
	II	370			**545**	67.89
			2240	**560**		
	III	550			-	-
	IV	800			-	-

Class Exercise 2 (Solution) (cont.)

	Quarter I	Quarter II	Quarter III	Quarter IV		
1990 1991 1992	- 104.35 101.22	- 60.34 67.89	97.56 90.57 -	153.56 140.21 -		
Total Average	205.57 102.79	128.23 64.12	188.13 94.07	293.77 146.89	407.87	101.97
Adjustment Adjusted Seasonal Factor	101.97 **102.79/.10197 = 100.80**	101.97 **64.12/.10197 = 62.88**	101.97 **94.07/.10197 = 92.25**	101.97 **146.89/.10197 = 144.05**		

the seasonal indices of the 4 quarters are 100.80, 62.88, 92.25 and 144.05 respectively.

Year	Quarter	1 Sales ('000)	2 Seasonal index	(1 / 2 =)3 **Deseasonalized time series**
1990	I	240	100.80	**238.10**
	II	120	62.88	**190.84**
	III	350	92.25	**379.40**
	IV	620	144.05	**430.41**
1991	I	450	100.80	**446.43**
	II	270	62.88	**429.39**
	III	420	92.25	**455.28**
	IV	680	144.05	**472.06**
1992	I	520	100.80	**515.87**
	II	370	62.88	**588.42**
	III	550	92.25	**596.21**
	IV	800	144.05	**555.36**

Forecast usage in 1993 quarter I :

The forecast usage without seasonal effect is 560,

and the seasonal index for quarter I is 100.80,

so the forecast usage for 1993-I = 560 ***x*** *100.80% = 564.48 ('000)*