Topic 1: Introduction to Time Series Analysis

STA 3050A: Time Series and Forecasting

1. OVERVIEW

Most of the data used in statistical analysis is collected at one point of time such type of data is called cross-sectional data. In cross-sectional data, we collect information about different individuals/subjects at the same point of time or during the same time. For example, data related to learners pursuing the MSCAST programme in July 2023 such as name, qualification, age, address, marks in graduation, etc., production of milk, import and export, information on the household income of New Delhi residents, etc. For such type of data, we just describe the status of the group at a point. For example, for the data related to the income of families living in New Delhi, we just find the average income, number of families below the poverty line, etc. but we do not find whether the income increasing or decreasing.

There are so many situations where we collect data over time. For example, in business, we observe daily sales, weekly interest rates, and daily closing stock prices. In meteorology, we observe daily high and low temperatures and hourly wind speeds. In agriculture, we record annual figures for crops and quarterly production, In the biological sciences, we observe the electrical activity of the heart at millisecond intervals, etc. Such types of data are called **time series** data. **A time series is a set of numeric data of a variable that is collected over time at regular intervals and arranged in chronological (time) order.**

In this topic, we shall discuss what a time series is and what are its components. In Sec. 2, we discuss what is time series with various examples. The components of a time series are described in Sec. 3. In Sec. 4, we explore different basic models of time series which show the relationships among the various components of a time series. To see better patterns of the time series, we describe the methods of smoothing or filtering such as simple and weighted moving averages, and exponential smoothing in Sec. 5. Sec. 6 and 7 are devoted to estimation of trend effects using the method of least squares (curve fitting) and moving average, respectively. In the next unit, you will learn various methods of estimating other components of a time series.

Expected Learning Outcomes

After studying this unit, you would be able to:

- explain what the time series is
- describe the components of time series
- explain the basic models of time series
- decompose the time series into different components for further analysis
- describe smoothing techniques for forecasting models, including, simple moving average, weighted moving average, and exponential smoothing
- explain various methods for the estimation of the trend.

2. INTRODUCTION TO TIMES SERIES

A time series is a collection of observations made sequentially through time. In other words, **the data on any characteristic collected with respect to time over a span of time periods is called a time series**. Normally, we shall assume that observations are available at equal intervals of time e.g., yearly, monthly, daily, hourly, etc. Some time series cover a period of several years. The time series data are collected in most of the fields, ranging from economics to engineering. For example, in business, the time series data gathered such as daily sales, daily closing stock prices, price of an item, in the meteorological department, daily high and low temperatures, hourly wind speed, in agriculture, annual figures for crops and yearly production, soil erosion, In the biological sciences, the electrical activity of the heart at millisecond intervals, brain monitoring (ECG), in ecology, the abundance of an animal species. In medicine, blood pressure tracking, weight tracking, cholesterol measurements, heart rate monitoring, etc.

There are two main goals in analysing a time series:

- First, one may want to describe or summarise the key features of the time series data

- Second, to predict what will happen in future based on past data (this is called forecasting).

For example, meteorologists forecast future weather conditions based on past observations, the milk production company forecasts the future demand of milk on the sales of milk on past days, business decision makers predict future sales, etc. Due to several special features of time series, we require different techniques to analyse and model the time series data for the forecast. Time series analysis is the art of extracting meaningful insights from time series data by exploring the series' structure and characteristics and identifying patterns that can then be utilized to forecast future events of the series.

Time series analysis assumes that data values of a time series variable are determined by four underlying environmental forces that operate both individually and collectively over time. They are trend (T), seasonal variations (S) cyclic variations (C) irregular (random) variations (I). These are called components of time series.

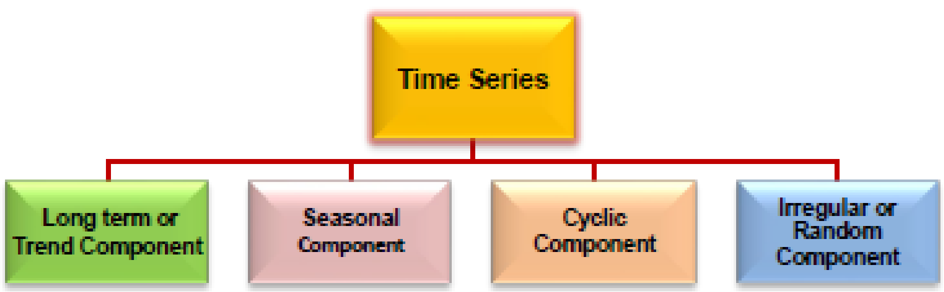
3. COMPONENTS OF TIME SERIES

The time series data do not remain constant over time while there is a variation in the values of the data. For example, the manager of a company collected the quarterly data of sales of a commodity for the period 2014-2020 which is given as follows:

Year	Quarter 1	Quarter 2	Quarter 3	Quarter 4
2014	18	8	12	9
2015	5	8	4	11
2016	4	10	14	18
2017	24	23	27	30
2018	35	32	30	38
2019	32	35	30	24

From the above data, we see that the sales of the commodity vary with time (quarterly and yearly). The variation occurs because of the effects of the various forces (such as seasons) at work, commonly known as components of time series.

In the past when we analysed the time series, then we assumed that data values of a time series variable are determined by four underlying environmental forces that operate both individually and collectively over time. They are: **(i) Trend (ii) Seasonal (iii) Cyclic** and **(iv) Remaining variation** attributed to **Irregular fluctuations** (sometimes referred to as **Random component**).



This approach is not necessarily the best one and we shall discuss the modern approach in later units. Some or all the components are present in varying amounts and can be classified into the mentioned four categories. We shall now discuss these components in more detail one at a time.

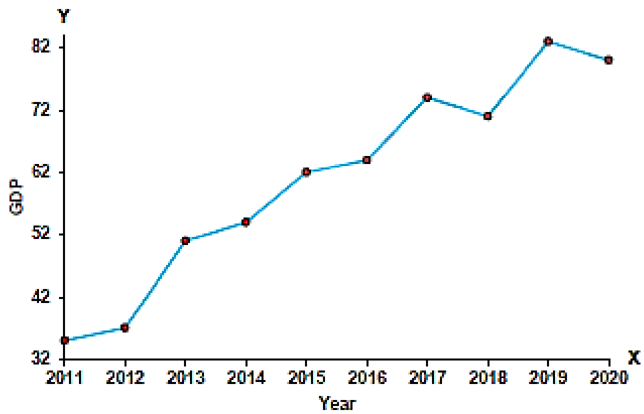
3.1. Trend Component

Usually, time series data show random variation, but over a long period of time, there may be a gradual shift in the mean level to a higher or a lower level. This gradual shift in the level of time series is known as the trend. In other words, **the general tendency of values of the data to increase or decrease during a long period of time is called the trend**.

When time series values are plotted on a graph and the values show an increasing or decreasing (on an average) pattern during a long period with reference to the time, then the time series is called the time series with a trend effect. The time series may show different types of trends.

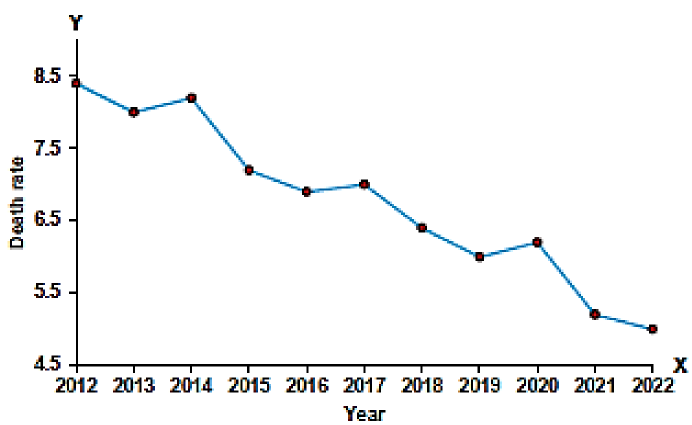
Time Series with Upward Trend

When a time series values are plotted on a graph and the values are increasing or showing an upward pattern with reference to the time then the time series is called the time series with an upward trend. For example, upward tendencies are seen in the data of population growth, currency in circulation, prices of petroleum products in India, number of passengers in the metro, literary rate, GDP of a country, etc. We plot a time series graph (Fig. 10.1) of the GDP of a country from 2011 to 2020 that shows an upward trend.



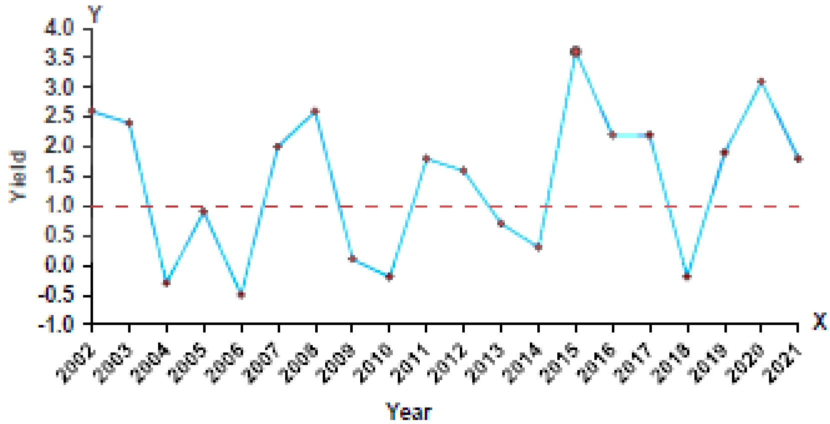
Time Series with Downward Trend

When a time series values are plotted on a graph and the values decrease or show a downward pattern (as shown in Fig. 10.2) with reference to the time then the time series is called the time series with a downward trend. For example, the downward trend is seen in the data of death rate, birth rate, number of landline phones, etc. The death rate of a country from 2012 to 2022 12 is showing a downward trend as shown below.



Time Series with No Trend

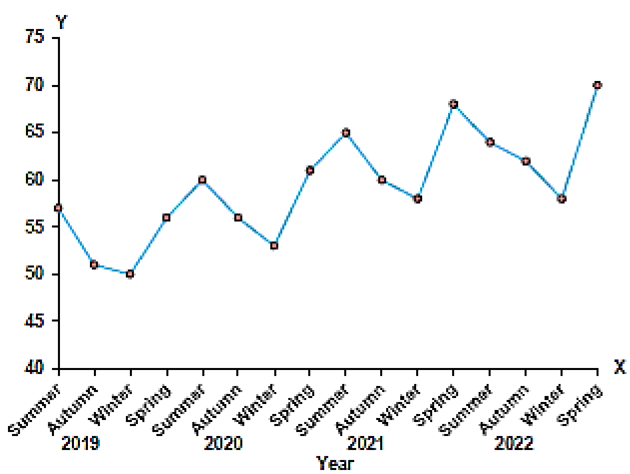
It is to be noted that all time series do not show an increasing or decreasing trend. In some cases, the values of time series fluctuate around a constant reading and do not show any trend with respect to time. Therefore, if a time series data is plotted on a graph paper and does not show any trend that is there is neither an upward nor a downward trend reflected in the time series plot then this kind of time series is called a time series with no trend. For example, the yield of a crop in a particular area from 2002 to 2021 shows no trend as shown below.



This should be clearly understood that a trend is general, smooth, long term and the average tendency of a time series data. The increase or decrease may not necessarily be in the same direction throughout the given period. The tendency of a time series may be found in either the form of a linear or a nonlinear (curvy linear) trend. If the time series data is plotted and the points on the graph cluster more or less around a straight line, then the tendency shown by the data is called a linear trend in time series. Similarly, if the points plotted on the graph do not cluster more or less around a straight line, then the tendency shown by the data is called a nonlinear or curvilinear trend. Trends are also known as **long-term variations**. The long-term or long period of time is a relative term which cannot be defined. In some cases, a period of one week may be long while in some cases a period of 2 years may not be enough. Some of the more important causes of long-term trend movements in a time series include population growth, urbanisation, technological improvements, economic advancements and developments, and consumer shifts in habits and attitudes.

3.2. Seasonal Component

In a time series, the variations which occur due to the rhythmic or natural forces and operate in a regular and periodic manner over a span of **less than** or **equal** to **one** year are termed as **seasonal variations**. We generally think of seasonal movement in time series as occurring yearly, but it can also represent any regularly repeating pattern that is less than one year in duration. For example, daily traffic volume data show within-day seasonal behaviour, with peak levels occurring during rush hours, moderate flow during the rest of the day, and light flow from midnight to early morning. Thus, in a time series, seasonal variation may exist if data are recorded quarterly, monthly, daily and so on. Even though the data may be recorded over a span of three months, one month, a week or a day, the amplitudes of the seasonal variation may be different. Most of the time series data of economic or business fields show the seasonal pattern. For example, the number of farming units (such as ploughs and tractors) sold quarterly for the period 2019 to 2022 shows a seasonal effect as shown below.



The seasonal pattern existing in a time series may be either due to natural forces or man-made conventions.

Seasonal Variations due to Natural Forces

Variations in time series that arise due to changes in seasons or weather conditions and climatic changes are known as seasonal variations due to natural forces. For example, sales of umbrellas and raincoat increase very fast 14 in the rainy season, the demand for air conditioners goes up in the summer season, and the sale of woollens go up in winter all being operated by natural forces.

Seasonal Variation due to Man-Made Conventions

Variations in time series that arise due to changes in fashions, habits, tastes, and customs of people in any society are called seasonal variations due to man-made conventions. For example, in our country sales of gold and clothes go up in marriage seasons and festivals.

3.3. Cyclic Component

Apart from seasonal effects, some time series exhibit variation due to some other physical causes, which is called cyclic variation. Cyclic variations are wave-like movements in a time series (as shown in Fig. 10.5), which can vary greatly in both duration and amplitude. Cyclical variations are recurrent upward or downward movements in a time series, but the period of a cycle is **greater than a year** whereas the period is less than one year in seasonal variation. Cyclic and seasonal variations are seen as similar, but they are quite different. If the variations are not of a fixed period, then they are cyclic and if the period is constant and associated with some aspect of the season, then the pattern is seasonal. In general, the average length of cycles is longer than the length of a seasonal pattern, and the magnitude of cycles tends to be more variable than the magnitude of seasonal variations.

The cyclic variation in a time series is usually called the **"Business cycle"** and comprises four phases of business i. e. prosperity (boom), recession, depression, and recovery.

Prosperity (boom)

The prosperity of any business is its profit. During a period of boom, businessmen and industrialists invest more and the economy surpasses the level of full employment and the level of production increases. These incentives make them produce more and therefore profit more.

Recession

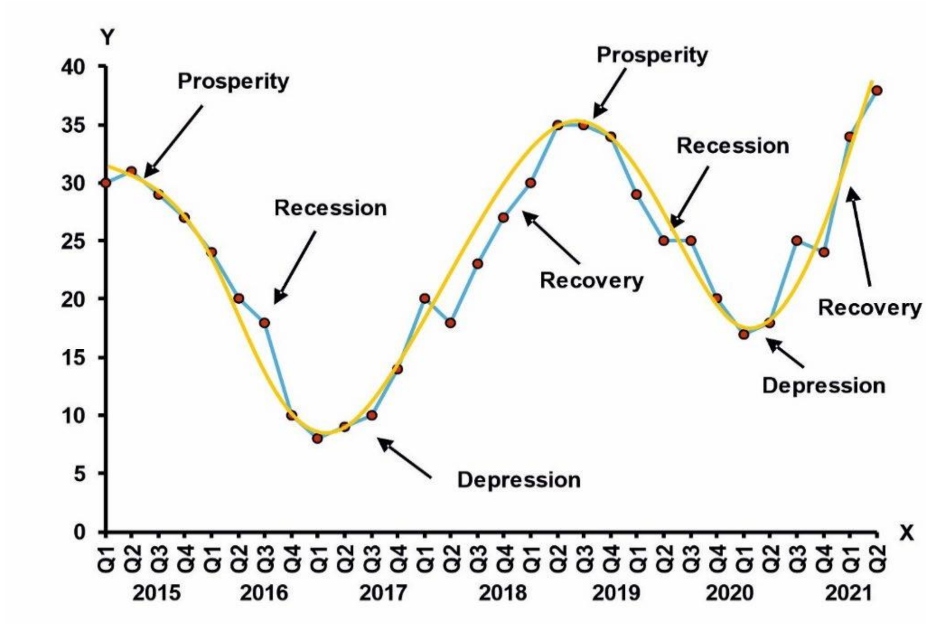
When there is excessive expansion, then it results in diseconomies that make it difficult to keep up with large-scale production. Additionally, it causes greater prices, rising salaries, and additional shortages. In an economic cycle, this is referred to as a recession.

Depression

In this phase of the economic cycle, output, income, and employment all start to drop rapidly. Also, investments decrease, and businesses are demoralized. Thus, it leads to pessimism which leads to deflation and depression.

Recovery

The depressive phase does not last forever. After some time, there is a cooling down and the improvement of trade begins. During the recovery period, old debts are repaid, and the units which are weaker are settled. As a result, the unemployment rate is gradually declining over time and income is generated.



3.4. Irregular Component

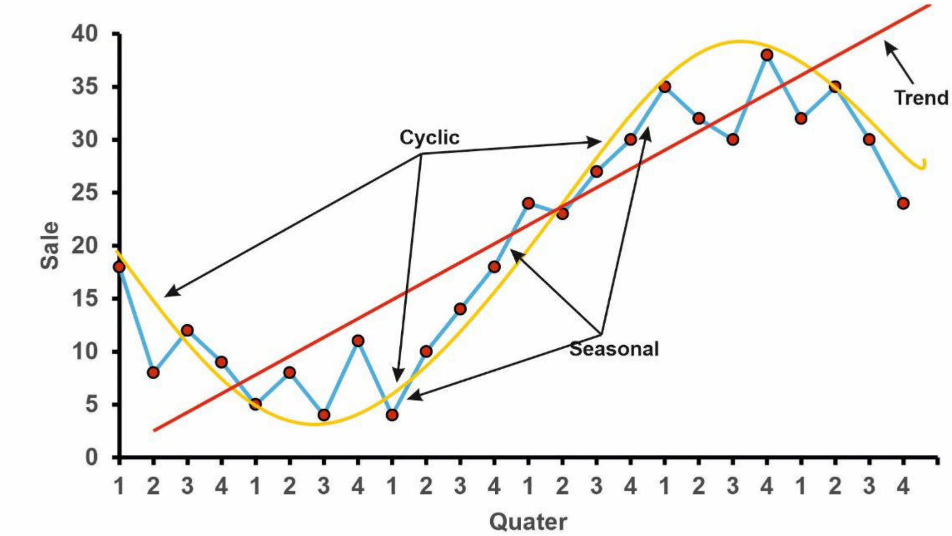
Apart from these regular variations, long-term and short-term variation, random or irregular factors which are not accounted for trend, seasonal or cyclic variations, exists in almost all time series. The variations in a time series which do not repeat in a definite pattern are called irregular variations or irregular component of a time series. The irregular variations in a time series may be either due to unforeseen one-off events such as natural disasters (floods, droughts, fires) or man-made disasters (strikes, boycotts, accidents, war, riots).

Since occurrences of irregular variations are totally unpredictable and follow no specific pattern, therefore, we cannot think of their time of occurrence, direction, and magnitude. In the latter units, we shall try to explain it by probability models such as autoregressive (AR) and moving average (MA) models, etc.

As we have discussed all four components which affect individually as well as jointly to the time series. Now, let us take an example of quarterly data of sales of a commodity for the period 2014-2019 given in the following table:

Year	Quarter 1	Quarter 2	Quarter 3	Quarter 4
2014	18	8	12	9
2015	5	8	4	11
2016	4	10	14	18
2017	24	23	27	30
2018	35	32	30	38
2019	32	35	30	24

We plot the time series data by taking sales on the Y-axis and the quarters on 16 the X-axis. We get the time series plot as shown below.



The plot shows more clearly the presence of different components in the time series data. The plot also shows seasonal as well as cyclic effects. If we draw a free-hand line to show the approximate movement of a curve around the line, then this line shows the presence of a long-term linear trend.

All time series need not necessarily exhibit all four components. For example, the time series data of annual production of a yield does not have seasonal variations and similarly, a time series for the annual rainfall does not contain cyclical variations.

Before moving to the next section, you can try the following Self Assessment Question for better understanding.

SAQ 1: What is time series? Describe its components.

4. BASIC MODELS OF TIME SERIES

In the previous section, we discussed different types of factors which affect the time series. In this section, we shall discuss the commonly used mathematical models which explain the time series data reasonably well. The basic models show the functional relationship among the various components of a time series. While discussing these models, we shall use the notation Y_t for the value of the time series at the time t. The following are the basic time series models:

4.1. Additive Model

The additive model is one of the most widely used models. The additive model assumes that at any time t the time series value Y_t is the sum of all four components. According to the additive model, the value of a time series variable can be expressed as

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

where, T_t, C_t, S_t and I_t are the trend, cyclic, seasonal and irregular variations at time t, respectively. In this model, it is assumed that cyclic effects remain constant for all cycles and the seasonal effects remain constant during any year or corresponding period. In this model, it is also assumed that irregular variation is an independent identically normally distributed variable with mean 0, i.e. the irregular variation effect remains constant throughout. Obviously, the additive model implies that seasonal variations in different years, cyclic variations in different cycles and irregular variations in different trends show equal absolute effects irrespective of the trend value.

4.2 Multiplicative Model

In the additive model, we have assumed that a value of a time series variable is the sum of the trend, cyclic, seasonal, and irregular components but this model is based on the assumption that cyclic effects remain constant for all cycles and the seasonal effects remain constant during any year or corresponding period. However, there are several situations where the seasonal variations exhibit an increase or decrease trend over time. When seasonal variations exhibit any change over time in terms of an increasing or decreasing trend, we can try the multiplicative model. In other words, if the various components in a time series operate proportionately to the general level of the series, the multiplicative model is appropriate. The multiplicative model is formed with the assumption that the time series value Yt at time t is the product of the trend, cyclic, seasonal and irregular components of the series. Symbolically, the multiplicative model can be described as.

$$Y_t = T_t \times C_t \times S_t \times I_t$$

where, T_t, C_t, S_t and I_t denote the trend, cyclic, seasonal and irregular variations. The multiplicative model is really a multiplicative version of the additive model. This model is found appropriate for many business and economic data. For example, the time series of the production of electricity, the time series of the number of passengers who opted the air travelling, the time series of the consumption of soft drinks, etc.

You can try the following Self Assessment Question for better understanding.

SAQ 2: What is the difference between additive and multiplicative models?

5. SMOOTHING TIME SERIES

For the estimation of trend and seasonal effects, it is very important to smooth out (or filter out) the effect of irregular fluctuations of time series so that the effects of trend and seasonal components can be easily estimated. Smoothing helps us to see better patterns of the time series such as trend. Generally smoothing smooths out the irregular roughness to see a clearer trend. Smoothing techniques are based on averaging values over multiple periods to reduce irregular fluctuations. Since these techniques **"smoothing out"** the short term/irregular fluctuations from the time-series data, therefore, they are called smoothing techniques. After smoothing out the short-term fluctuation, we can estimate and forecast and trend effect. They manage the data in the sense that we can estimate time series components directly from the data without a predetermined structure. In this section, we shall discuss two simple and important smoothing methods for the time series data namely moving averages and exponential smoothing methods.

5.1. Simple Moving Average

The moving average (MA) is the simplest method for smoothing time series data. A moving average removes irregular fluctuations and short-term fluctuations from a time series. This method is based on averaging each observation of a series with its surrounding observations, that is, past and future observations in chronological order. In this method, we find the simple moving averages of time series data over m span/period of time, and these averages are called **m-period moving averages.** These averages are smoothed versions of the original time series. In this method, we put the average on the middle value of the set of observations, therefore, we explain the moving average for old and even periods as follows:

When m is odd

In some situations, the data may show seasonal effects over an **odd period** of time, e.g., 5 days, 7 months, etc. It means that after every 5 days or 7 months, data behave in a similar manner. Therefore, to remove the seasonal effects from the time series, we calculate the moving average of an odd span/period of time, in our example, it is 5 days or 7 months. For odd periods, the method consists of the following steps:

Step 1: We calculate the average of the first m values of the time series and place it against the middle value, i.e., $\frac{m+1}{2}$th observation. For example, if $m = 3$, we compute the average of the first three observations of the time series as $MA_1 = \frac{y_1+y_2+y_3}{3}$ where y_1, y_2 and y_3 are the first three observations of the time series. and place it against the $\frac{m+1}{2}$th observation, i.e., $\frac{3+1}{2}$th = 2nd observation y_2.

Step 2: We discard the first observation and include the next observation. Then we take the average of m values again. For example, if $m = 3$, we discard the first observation and determine the average of the second, third and fourth observations and place it against the middle of the second, third and fourth observations, that is, in front of 3rd observation.

Step 3: We repeat this process until all data are exhausted. These steps provide us with a new time series of m-period moving averages which will be smoother than the original time series.

When m is even

Sometimes, there may be a seasonal effect over an even period, e.g., quarterly or yearly, etc. This means that after every fourth quarter or 12 months, data behave in a similar way. Since there is no centre of these four observations, therefore, the simple moving averages for even periods need to be centred. So, it is also known as centred moving averages. If we take a centred moving average with $m = 4$ then it will filter (or eliminate) the effect of the season by a quarter. This method consists of the following steps:

Step 1: We calculate the average of the first m values of the time series and place it against the middle value, i.e., $\frac{m}{2}$ .For example, if $m = 4$, we compute the average of the first four observations and place it against the middle of the second and third observations.

Step 2: We discard the first observation and include the next observation and then take the average of the next m observations again. For example, if $m = 4$, we discard the first observation, find the average of the second, third, fourth and fifth observations and place it against the middle of the third and fourth observations. We repeat this process until all data are exhausted as shown in Example 2.

Step 3: Since the period (m) is even and there is no exact midpoint, so we cannot put the MA corresponding to any year. Therefore, we determine the centred moving average. To determine the first centred moving average, we compute the average of the first two moving averages and place it against the middle value of the two moving averages. For example, if $m = 4$, we calculate the average of the first two moving averages and place it against the middle of the first and second moving averages, i.e., against the third observation.

Step 4: We repeat this process until all data are exhausted.

Note: The moving average eliminates periodic variations if the span of the period of the moving average (m) is equal to the period of the oscillatory variation. Therefore, we should choose m which constitutes a cycle, for example, if a cycle is completed in 3 months, we should calculate exactly 3 monthly moving averages. If there is a variation in the span of cycles, for example, suppose, the first cycle is completed in 3 months, the second in 3 months and the third in 5 months and so on then we should use the average of these time spans as m.

Let us take an example to explain the procedure of the moving average method.

Example 1: The following table shows the number of fire insurance claims received by an insurance company in each four-month period from 2018 to 2021:

Year	**2018**			**2019**			**2020**			**2021**		
Period	I	II	III	I	II	III	I	II	III	I	II	III
No. of Claims	17	13	15	19	17	19	22	14	20	23	19	20

Calculate and plot the three-period and five-period moving average series for the number of fire insurance claims. Compare these two moving average series.

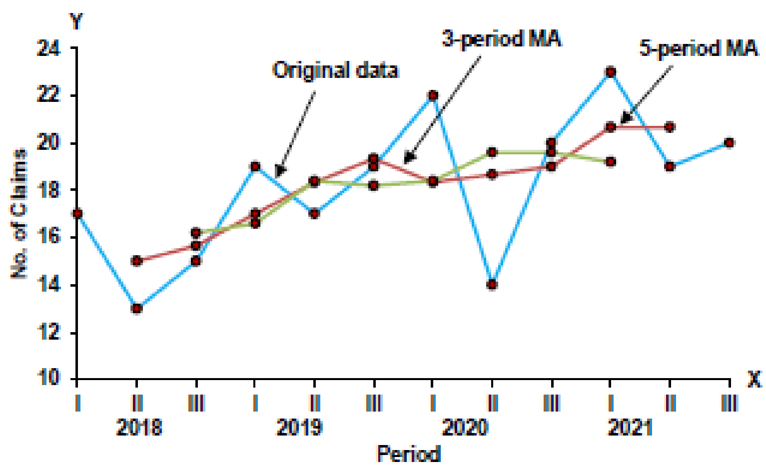
Solution: Since $m = 3$ is odd, therefore, we take the average of the first three observations and put the same in the middle of these observations, that is, in 20 front of the second observation. The first value of $m = 3$ years moving average $\frac{17+13+15}{3} = 15$ and we put this value in front of II period of 2018. The second value of moving averages is obtained by discarding the first observation i.e. 17 and including the next observation, that is, 19, this gives an average of 13, 15 and 19, which is equal to $\frac{13+15+19}{3} = 15.67$. And we put it against the third observation, that is III period of 2018. We repeat the same procedure of calculating three-period moving averages until all data are exhausted. Similarly, we can find the five- period moving averages by following the same procedure of three-period moving averages, that is, we calculate the average of the first five observations 17, 13, 15, 19 and 17 and put the same in front of the third period. We calculate the rest of the moving averages in the following table:

Year	Period	No. of Claims	3-period MA	5-period MA
2018	I	17	-	-
	II	13	$\frac{17+15+15}{3} = 15.00$	-
	III	15	$\frac{13+15+19}{3} = 15.67$	$\frac{17+15+15+19+17}{5} = 16.20$

Year	Period	No. of Claims	3-period MA	5-period MA
2019	I	19	$\frac{15+19+17}{3} = 17.17$	$\frac{13+15+19+17+19}{5} = 16.60$
	II	17	$\frac{19+17+19}{3} = 18.33$	$\frac{15+19+17+19+22}{5} = 18.40$
	III	19	$\frac{17+19+22}{3} = 19.33$	$\frac{19+17+19+22+14}{5} = 18.20$
2020	I	22	$\frac{19+22+14}{3} = 18.33$	$\frac{17+19+22+14+20}{5} = 18.40$
	II	14	$\frac{22+14+20}{3} = 18.67$	$\frac{19+22+14+20+23}{5} = 19.60$
	III	20	$\frac{14+20+23}{3} = 19.00$	$\frac{22+14+20+23+19}{5} = 19.60$
2021	I	23	$\frac{20+23+19}{3} = 20.67$	$\frac{14+20+23+19+20}{5} = 19.20$
	II	19	$\frac{23+19+20}{3} = 20.67$	-
	III	20	-	-

From the above table, we observed that the original series varies between 13 and 22 whereas moving averages vary between 15 and 20.67 (3-period) and 16.20 to 19.60 (5-period) which are much smoother than the original series. The moving averages fluctuate less than the fluctuation of the original observations which are calculated as they smooth (or filter out) the effect of seasonal/irregular components. This helps us to appreciate the effect of the trend more clearly.

We now plot the original observations with both the 3-period and 5-period MA values by taking them on the Y-axis and the period on the X- axis as shown below:



From a comparison of the line plots of the 3-period and 5-period moving average values, we can see that there is less fluctuation (greater smoothing) in the 5-period moving average series than in the 3-period moving average series.

Therefore, we can conclude that the term, m, for the moving average affects the degree of smoothing:

- A shorter period (m) of the moving average produces a more jagged moving average curve.
- A longer period (m) of the moving average produces a smoother moving average curve.

Now the problem is what should be the value of m. If m is increased then the series becomes much smoother and it may also smooth out the effect of cyclical and seasonal components, which are our main interest of study. To take away seasonality from a series, we would use a moving average with a length equal to the seasonal span. Thus, in the smoothed series, each smoothed value has been averaged across all seasons. Sometimes 3-year, 5- year or 7-year moving averages are used to expose the combined trend and cyclical movement of time series.

The moving average can also be utilised as a forecasting model with some simple steps. In simple moving average, we put the moving average at the centre of the set of m observations as you have seen in Example 1 and we have put the average of 17, 13 and 15 which equals 15 against the Period II of 2018. It means we forecast the value of the Period II of 2018. And for that, we use both the past (observation of period I) and future (observation of period III) of a given time point. In that sense, we cannot use moving averages as such forecasting because, at the time of forecasting, the future is typically unknown. Hence, for the purpose of forecasting, we use **trailing moving averages.** In this method, we take an average of past consecutive m observations of the series as follows:

$$y'_{t+1} = \frac{y_t + y_{t-1} + \ldots + y_{t-m+1}}{m}$$

where y'_{t+1} *is the forecast value of* y_{t+1} . For example, if $m = 3$, we compute the first forecast value by taking the average of the first three values as

$$y'_4 = \frac{y_3 + y_2 + y_1}{3}$$

Furthermore, only the first forecast value is constructed by averaging only the actual values of the series. As we move to the second forecast, the actual values are replaced with the previously forecasted values. For instance, the second forecast value is defined by the following expression:

$$y'_{t+2} = \frac{y'_{t+1} + y_t + y_{t-1} + \ldots + y_{t-m+2}}{m}$$

For example, if $m = 3$, we compute the second forecast value as:

$$y_5{}' = \frac{y_4{}' + y_3 + y_2}{3}$$

If the forecast value for y t is y'_t, then we can calculate the forecast error as

$$e_t = y_t - y'_t$$

5.2 Weighted Moving Average

The simple moving average described in the previous sub-section is not generally recommended for measuring trend, although it can be useful for removing seasonal variation. In simple moving averages, we give equal importance to all observations, therefore, we assign equal weights but sometimes we may observe that a certain time period has more importance in comparison to the others. For example, a forecaster might believe that the previous month's value is two times as important in forecasting as other months. Therefore, a moving average in which some time periods are weighted differently than others is called a weighted moving average (WMA). Generally, more weights are assigned to the recent observations and less to past observations in the weighted moving average. We assign weights in such a way that all weights are positive. If wi denotes the weight assigned to the i^{th} observation, $(i = 1,2, \ldots, m)$ then we can compute the first weighted moving average (WMA) as follows:

$$WMA_t = \frac{w_1 y_1 + w_2 y_2 + \ldots + w_m y_m}{w_1 + w_2 + \ldots + w_m} \text{ such that } w_i \geq 0$$

The procedure for getting the smooth trend values by the weighted moving average is the same as that for the simple moving average in spite of weights. Therefore, we find the weighted average instead of the simple average in the weighted moving average.

Let us take an example to illustrate this method.

Example 2: Compute 4-month weighted moving average for the number of fire insurance claims data given in Example 1 using weights 1,1,2 and 4.

Solution: Here, the weights are given as

$$w_1 = 1, w_2 = 1, w_3 = 2, w_4 = 4$$

Thus, we can find the first weighted moving average as

$$WMA_1 = \frac{w_1 y_1 + w_2 y_2 + \ldots + w_m y_m}{w_1 + w_2 + \ldots + w_m} = \frac{1 \times 17 + 1 \times 13 + 2 \times 15 + 4 \times 19}{1 + 1 + 2 + 4} = 17.00$$

In a similar way, you compute the rest weighted moving averages. Since period (4) is even, therefore, we put the first WMA in the middle of the second and third observations. For that, we create blank rows after each observation. Also, we determine the centred moving average by averaging the first two WMAs and placing it against the middle value, i.e., we calculate the average of the first two WMAs and place it against the middle of the first and second moving averages, i.e., against the third observation. We show these in the following table:

Year	Period	No. of Claims	4-period WMA	Centred 4-period WMA
2018	I	17		
	II	13		
			$\frac{1\times 17+1\times 13+2\times 15+4\times 19}{1+1+2+4}=17.00$	
	III	15		$\frac{17.00+16.75}{2}=16.88$
			$\frac{1\times 13+1\times 15+2\times 19+4\times 17}{1+1+2+4}=16.75$	
2019	I	19		$\frac{16.75+18.00}{2}=17.38$
			$\frac{1\times 15+1\times 19+2\times 17+4\times 19}{1+1+2+4}=18.00$	
	II	17		$\frac{18.00+20.25}{2}=19.13$
			$\frac{1\times 19+1\times 17+2\times 19+4\times 22}{1+1+2+4}=20.25$	
	III	19		$\frac{20.25+17.00}{2}=18.63$
			$\frac{1\times 17+1\times 19+2\times 22+4\times 14}{1+1+2+4}=17.00$	
2020	I	22		$\frac{17.00+16.75}{2}=17.81$
			$\frac{1\times 19+1\times 22+2\times 14+4\times 20}{1+1+2+4}=18.63$	
	II	14		$\frac{18.63+21.00}{2}=19.81$
			$\frac{1\times 22+1\times 14+2\times 20+4\times 23}{1+1+2+4}=21.00$	
	III	20		$\frac{21.00+19.50}{2}=20.25$
			$\frac{1\times 14+1\times 20+2\times 23+4\times 19}{1+1+2+4}=19.50$	
2021	I	23		$\frac{19.50+20.13}{2}=19.81$
			$\frac{1\times 20+1\times 23+2\times 19+4\times 20}{1+1+2+4}=20.13$	
	II	19		
	III	20		

After understanding the moving average as a smoothing and forecast technique, we now discuss the merits and demerits of the moving average.

Merits

1. It is simple as compared to the other method (the exponential method discussed in the next section).

2. A moving average time series is a smoother series than the original time series values. It has removed the effect of short-term fluctuations (i.e. seasonal and irregular fluctuations) from the original observations by averaging over these short-term fluctuations.

3. If the moving average period coincides with the period of cyclical variations, the moving average method eliminates the regular cyclical fluctuations. Even if the variations are not eliminated, it reduces their intensity.
4. The method is very flexible in the sense that the addition of a few more figures to the data simply results in a few more trend values without affecting the previous calculations.
5. The method is suitable for determining trend when a linear trend is present in the time series.

Demerits

1. As seen from the moving average calculations, its primary drawback is a loss of information (data values) at both ends of the original time series. However, this is not a significant drawback if the time series is long, say 50 time periods or more.
2. If a series consists of a non-linear trend, a moving average will not reveal the trend present.
3. The selection of the extent or period of the moving average is quite difficult particularly when the time series data exhibit cycles which are not regular in period and amplitude. The effect of an inappropriate selection of the extent or period is that moving averages may generate cycles or other movements which were not present in the original data.
4. The moving averages are greatly affected by extreme values. To overcome this somewhat, a weighted moving average with appropriate weights is used.

Before moving to the next method of smoothing, i.e., exponential smoothing, you can try the following Self Assessment Question for better understanding.

SAQ 3: The marketing manager of an electricity company recorded the following quarterly demand levels for electricity (in 1000 megawatts) in a city from 2020 to 2022.

Season	**2020**	**2021**	**2022**
Summer	70	101	146
Monsoon	52	64	92
Winter	22	24	38
Spring	31	45	49

1) Calculate the quarterly moving averages of the electricity demand.
2) Calculate the quarterly weighted moving averages taking weights as 1, 2, 3 and 4.
3) Plot both the quarterly simple and weighted moving averages together with the original data on the same axis. Also, compare both methods.