Introduction to Time Series Analysis

STA 3050: Time Series and Forecasting

EXPECTED LEARNING OUTCOMES

After studying this topic, 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.

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 city 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 the city, 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.

OVERVIEW

- 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.**

OVERVIEW

- In this topic, we shall discuss what a time series is and what are its components.

  - we discuss what is time series with various examples.

  - The components of a time series are described.

  - 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.

  - estimation of trend effects using the method of least squares (curve fitting) and moving average, respectively.

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.

INTRODUCTION TO TIMES SERIES

• 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.

- Business: daily sales, daily closing stock prices, price of an item

- Meteorological department: daily high and low temperatures, hourly wind speed

- Agriculture: annual figures for crops and yearly production, soil erosion

- Biological sciences: the electrical activity of the heart at millisecond intervals, brain monitoring (ECG)

- Ecology: the abundance of an animal species

- Medicine: blood pressure tracking, weight tracking, cholesterol measurements, heart rate monitoring, etc.

INTRODUCTION TO TIMES SERIES

There are two main goals in analyzing a time series:

- First, one may want to describe or summarize 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

- 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.

COMPONENTS OF TIME SERIES

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

- 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.

- 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:

COMPONENTS OF TIME SERIES

• 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.
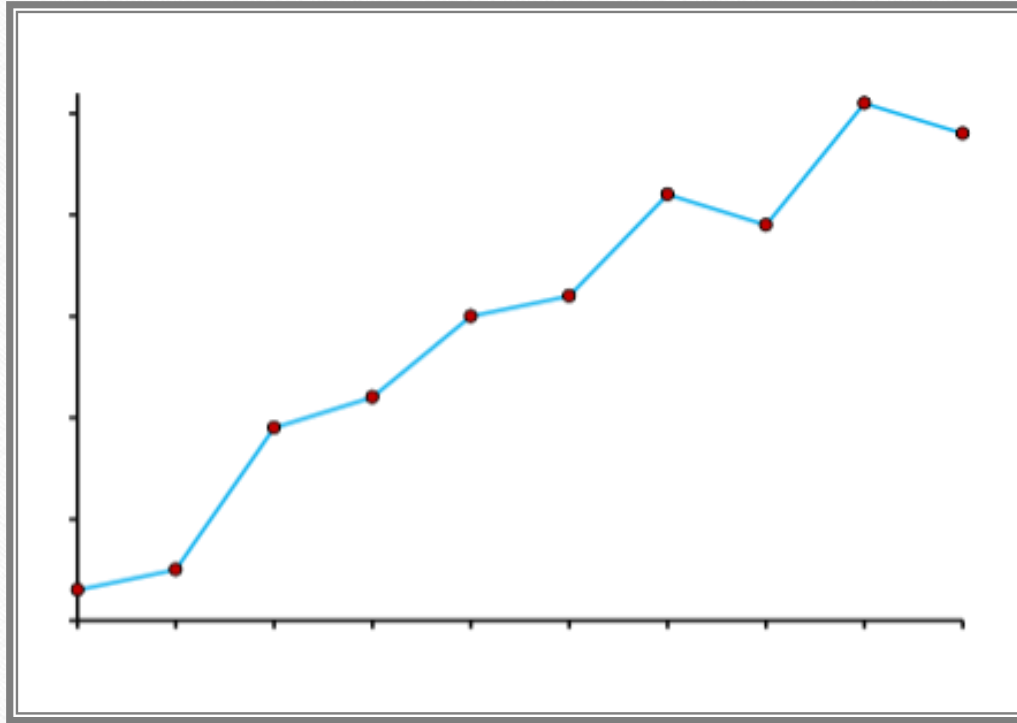
TREND COMPONENT

- 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.

TREND COMPONENT – 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
- number of passengers in the public transport
- literary rate
- GDP of a country, etc.

TREND COMPONENT – UPWARD TREND

We plot a time series graph of the GDP of a country from 2011 to 2020 that shows an upward trend.

TREND COMPONENT – DOWNWARD TREND

- When time series values are plotted on a graph and the values decrease or show a downward pattern 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.
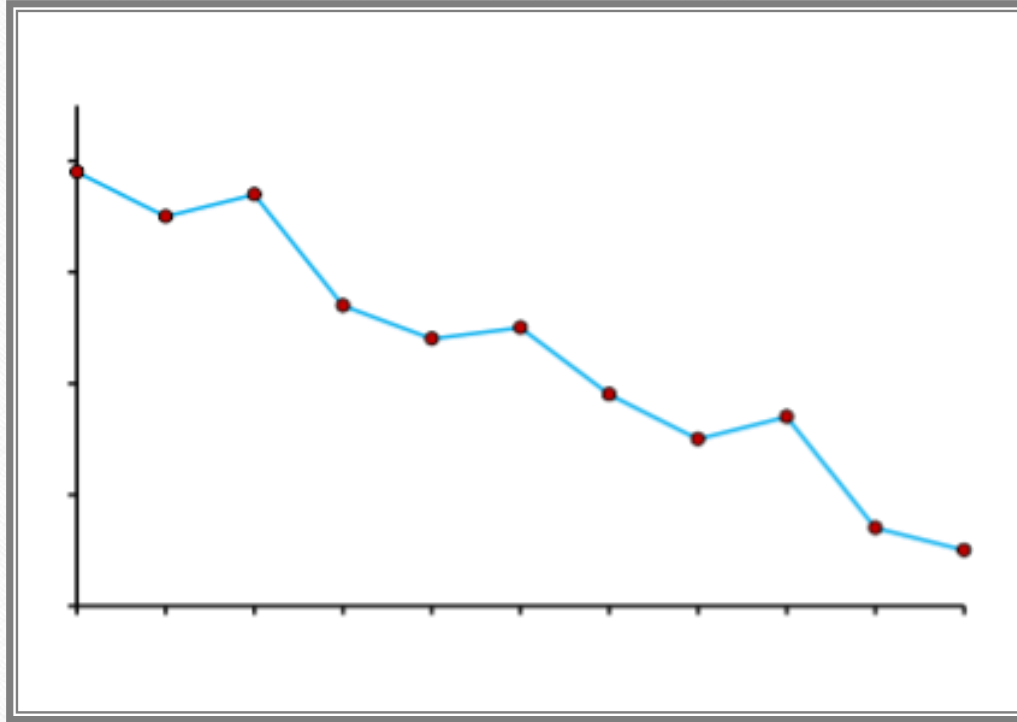
TREND COMPONENT – DOWNWARD TREND

The death rate of a country from 2012 to 2022 12 is showing a downward trend as shown below.

TREND COMPONENT – 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.
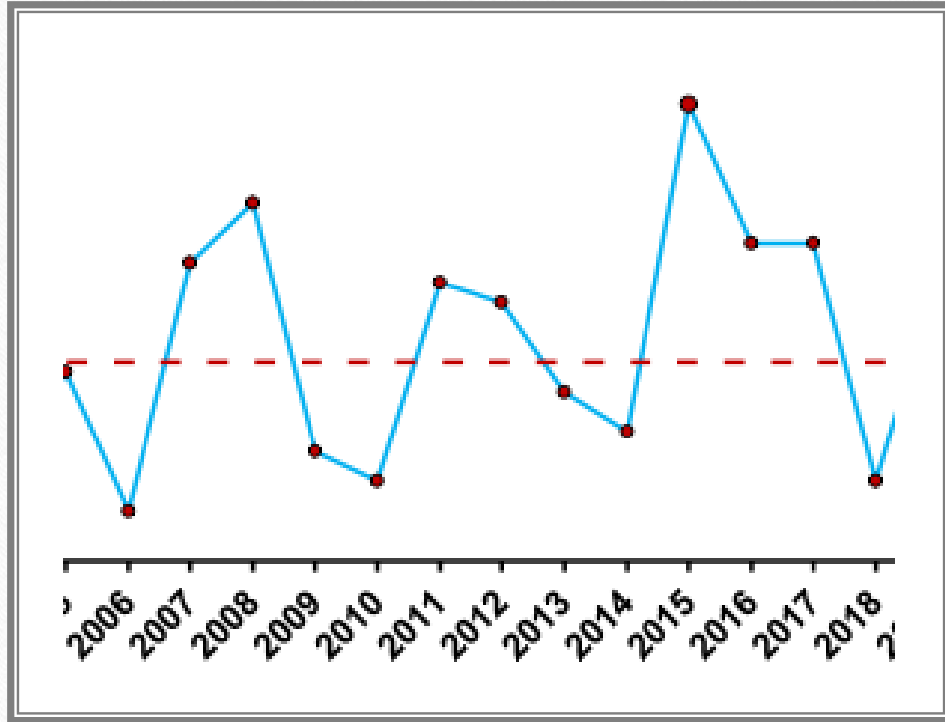
TREND COMPONENT – NO TREND

- For example, the yield of a crop in a particular area from 2002 to 2021 shows no trend as shown below.

TREND COMPONENT – NO TREND

• 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.

TREND COMPONENT – NO TREND

- 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.
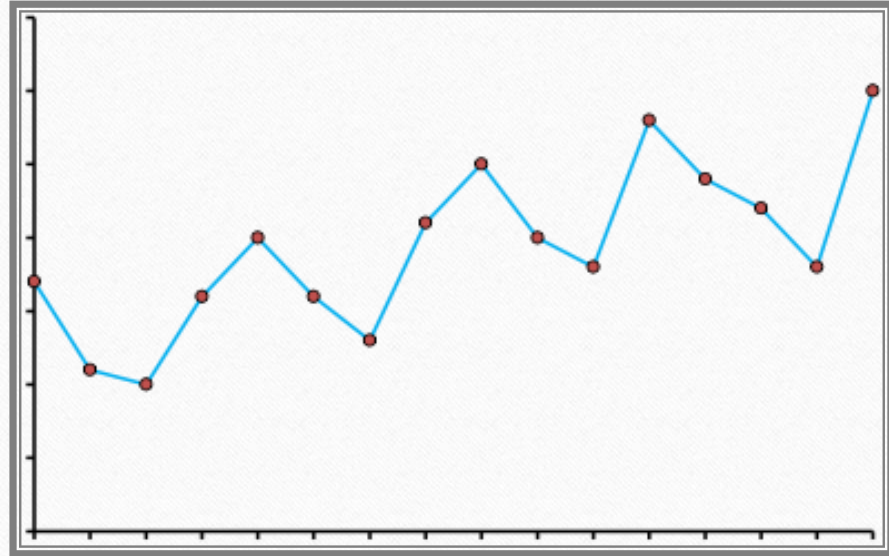
SEASONAL COMPONENT

- 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 behavior, with peak levels occurring during rush hours, moderate flow during the rest of the day, and light flow from midnight to early morning.

SEASONAL COMPONENT

- 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.

SEASONAL COMPONENT

- 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.

SEASONAL COMPONENT

- 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 some country sales of gold and clothes go up in marriage seasons and festivals.

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, 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.

CYCLIC COMPONENT

- 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.
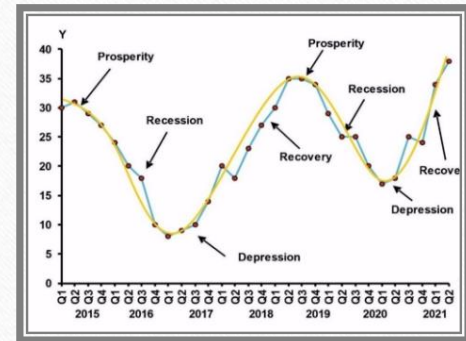
CYCLIC COMPONENT

- **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.
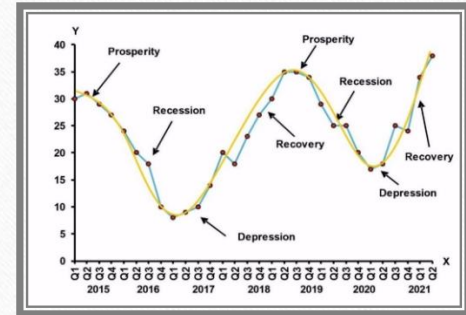
IRREGULAR COMPONENT

• 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.

IRREGULAR COMPONENT

• 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

IRREGULAR COMPONENT
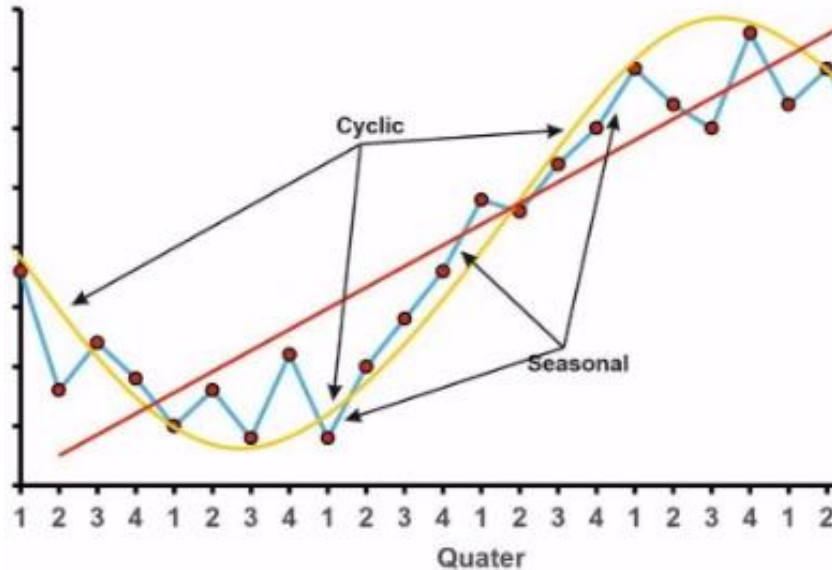
- We plot the time series data by taking sales on the Y-axis and the quarters on 16 the X-axis.

- 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.

IRREGULAR COMPONENT

• 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.