Time Series Analysis

UAMMM0542

Credit-3
ESE -100 marks

Overview

1. Time series Analysis Overview
2. Exploratory Data Analysis for Time Series
3. Traditional Time Series Forecasting Methods
4. Machine Learning Approaches for Time Series Forecasting
5. Deep Learning Techniques for Time Series Forecasting
6. Applications of Time Series Analysis

Time series Analysis Overview

Unit 1

Outline

- Time series analysis and forecasting
- Understanding time series data
- Types of time series data
- Components of time series
- Residual Importance and applications of time series forecasting

Time series analysis and forecasting

- Time series analysis and forecasting are crucial for predicting future trends, behaviors, and behaviours based on historical data.
- Time Series Analysis is a way of studying the characteristics of the response variable concerning time as the independent variable.
- To estimate the target variable in predicting or forecasting, use the time variable as the reference point.
- TSA represents a series of time-based orders, it would be Years, Months, Weeks, Days, Horus, Minutes, and Seconds.
- It helps businesses make informed decisions, optimize resources, and mitigate risks by anticipating market demand, sales fluctuations, stock prices, and more.
- Additionally, it aids in planning, budgeting, and strategizing across various domains such as finance, economics, healthcare, climate science, and resource management, driving efficiency and competitiveness.

Importance of Time Series Analysis

1. **Predict Future Trends:** Time series analysis enables the prediction of future trends, allowing businesses to anticipate market demand, stock prices, and other key variables, facilitating proactive decision-making.
2. **Detect Patterns and Anomalies:** By examining sequential data points, time series analysis helps detect recurring patterns and anomalies, providing insights into underlying behaviors and potential outliers.
3. **Risk Mitigation:** By spotting potential risks, businesses can develop strategies to mitigate them, enhancing overall risk management.
4. **Strategic Planning:** Time series insights inform long-term strategic planning, guiding decision-making across finance, healthcare, and other sectors.
5. **Competitive Edge:** Time series analysis enables businesses to optimize resource allocation effectively, whether it's inventory, workforce, or financial assets. By staying ahead of market trends, responding to changes, and making data-driven decisions, businesses gain a competitive edge

Time series analysis

Time series analysis is a statistical technique used to analyze data points gathered at consistent intervals over a time span in order to detect patterns and trends. Understanding the fundamental framework of the data can assist in predicting future data points and making knowledgeable choices.

Objectives of Time Series Analysis

- To understand how time series works and what factors affect a certain variable(s) at different points in time.
- Time series analysis will provide the consequences and insights of the given dataset’s features that change over time.
- Supporting to derive the predicting the future values of the time series variable.
- Assumptions: TSA makes only one assumption, which is “stationary,” which means that the origin of time does not affect the properties of the process under the statistical factor.

Time Series Analysis

It is about **understanding the past** behavior of data over time.
It focuses on **identifying patterns**, trends, seasonality, or irregularities in the data.

Purpose:

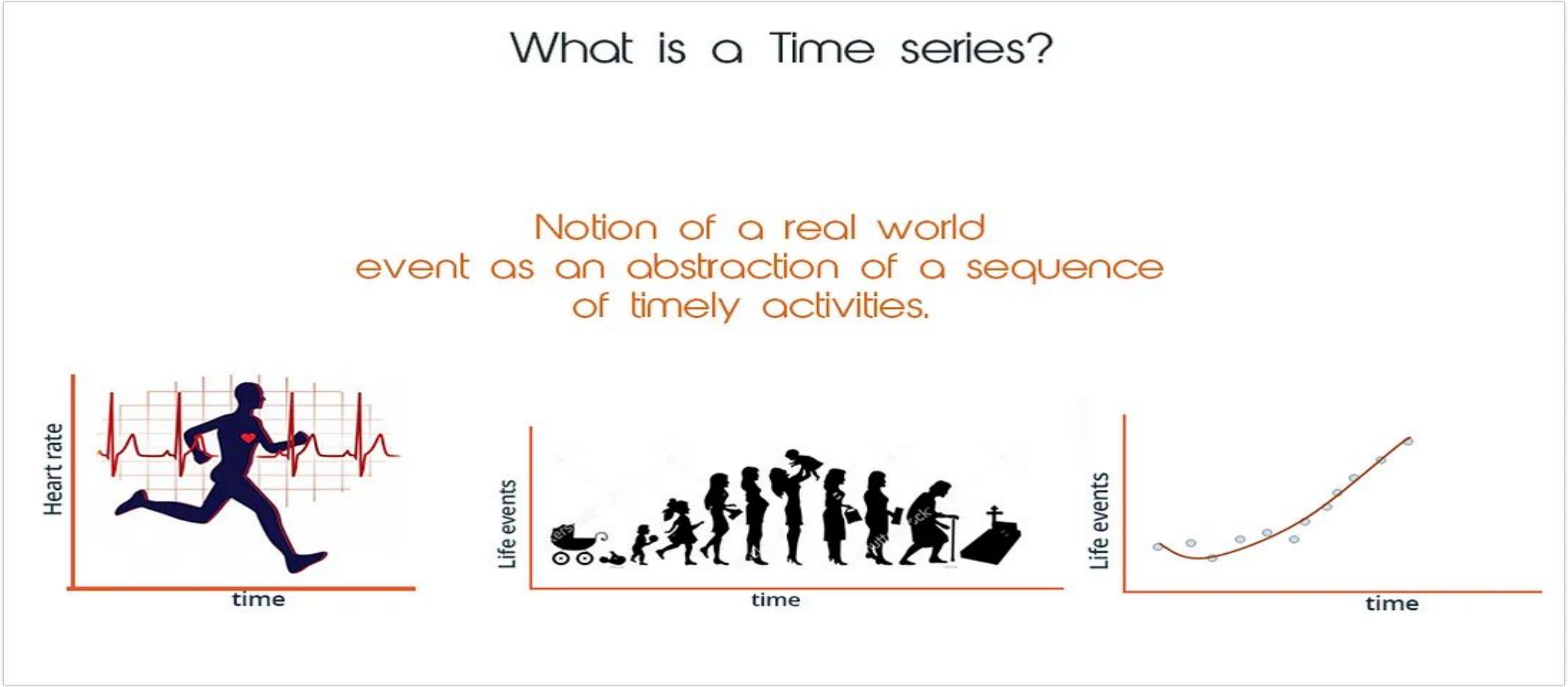
- To explore and explain the data
- To understand what happened and why
- Example: Looking at 5 years of sales data to find out if there's a seasonal pattern

Time Series Forecasting

It is about **predicting the future** values based on past data.
It uses the patterns found in analysis to make **future predictions**.

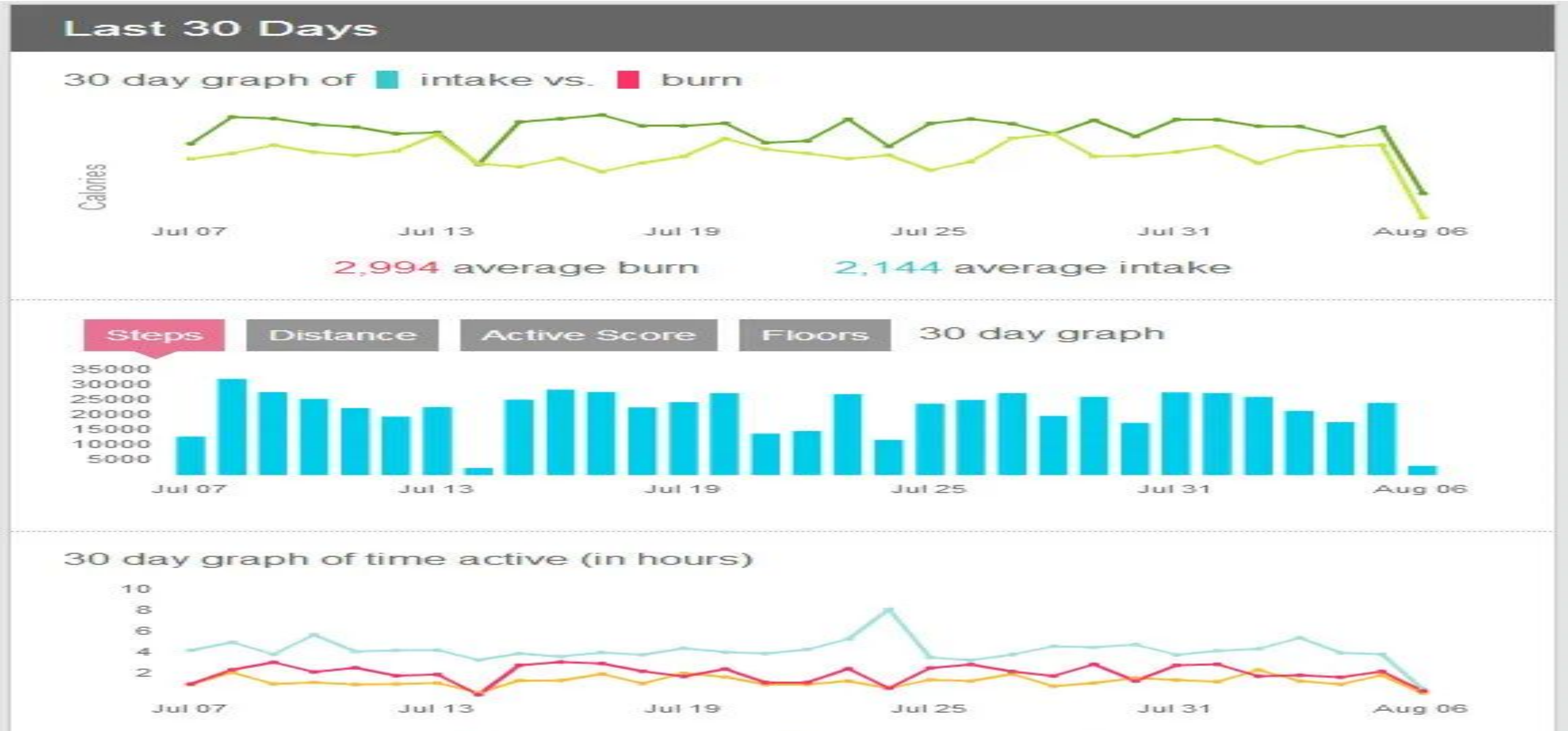
Purpose:

- To estimate or project future outcomes
- Example: Using the past 5 years of sales data to predict sales for the next 6 months



A time series is a sequence of data points collected, recorded, or measured at successive, evenly-spaced time intervals.

Ex.- Fitness device market is built around buy people to help track fitness related data to monitor effectiveness of their fitness exercises.



Ex.- Sales growth of a product over period of time is a good indicator of sales performance of a product manufacturing company.

Types of time series data

1. **Continuous time series data**: Continuous time series data is collected continuously over time, in which data points are taken at regular intervals.

Examples of this type of data include temperature, stock prices, and sales data.

Continuous **time series data is useful for analyzing** trends and patterns over time.

2. **Discrete time series data**: Discrete time series data is collected at irregular intervals, with data points taken at specific points in time.

Examples of this type of data include patient data in a hospital and weather data collected at specific times of the day.

Discrete *time series data* is useful for analyzing *specific events* and their impact on the data.

Ex. A doctor measures a patient's **body temperature** only when the patient reports feeling unwell:

10:00 AM: 98.7°F

2:30 PM: 101.2°F

7:00 PM: 100.5°F

3. **Longitudinal time series data**: Longitudinal *time series data* is collected from the same group of individuals over a long period of time.

Examples of this type of data include medical data and data collected from social surveys. Longitudinal *time series data* is useful for analyzing changes in *individual behavior* or characteristics over time.

Ex.- A hospital conducts a 5-year study to monitor blood pressure in a group of 100 diabetic patients.

Each patient's blood pressure is measured once every month.

The same individuals are followed over time.

4. **Cross-sectional time series data**: Cross-sectional time series data is collected at different points in time from different groups of individuals.

Examples of this type of data include census data and data collected from opinion polls.

Cross-sectional *time series data* is useful for analyzing changes in behavior or characteristics across different groups of individuals over time.

Ex. A labor ministry surveys employment rates in different Indian states every year for 5 years.

Sample data:

Year	State	Employment Rate (%)
2020	Maharashtra	72%
2020	Karnataka	68%
2020	Bihar	60%
2021	Maharashtra	74%
2021	Karnataka	69%
2021	Bihar	61%

- Each year, **data is collected from multiple states** (cross-section),
- Data is collected **repeatedly over time** (time series).

Components of Time Series Data

Time series data means data that is collected over time — like daily temperatures, monthly sales, or yearly rainfall.

This kind of data usually has different parts (called **components**) that show patterns in how the data changes.

By studying these parts, we can understand how the data behaves over time and make better predictions for the future.

- Trends: Long-term increases, decreases, or stationary movement
- Seasonality: Predictable patterns at fixed intervals
- Cycles: Fluctuations without a consistent period
- Noise: Residual unexplained variability

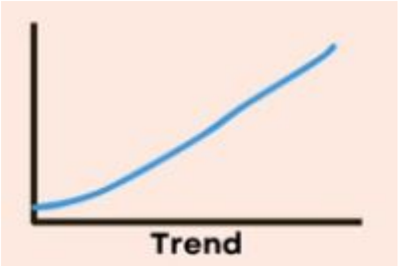
Trends

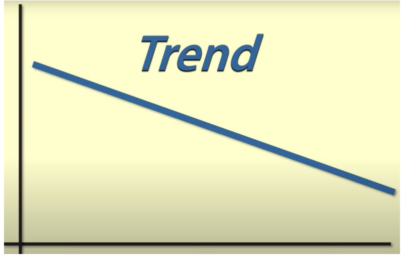
Trends show the general direction of the data, and whether it is increasing, decreasing, or remaining stationary over an extended period of time.

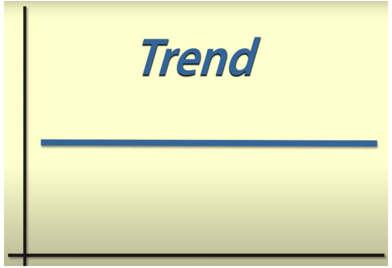
Trends indicate the long-term movement in the data and can reveal overall growth or decline.

Types of Trends:

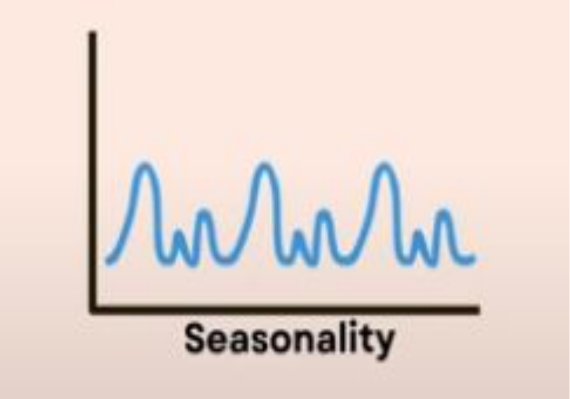
- **Upward Trend:** A consistent increase in values over time.
  Example: The steady rise in global temperatures due to climate change.
- **Downward Trend:** A continuous decline in values over time.
  Example: The decreasing landline telephone subscriptions as mobile phone usage increases.
- **Stable (No Trend):** Data remains relatively constant with no significant increase or decrease.
  Example: Daily energy consumption in an industrial plant with steady operations.







Seasonality



Seasonality is about repeating the specific patterns or deviations in the data that are observed at frequent time intervals, for example, daily, weekly, monthly, or yearly cycles.

Examples of Seasonality:

- **Retail Sales:** Increased shopping during festivals, Black Friday, & Christmas.
- **Weather Patterns:** Higher electricity consumption in summers due to air conditioning.
- **Website Traffic:** More visits to online education platforms at the start of a semester.

Characteristics of Seasonality:

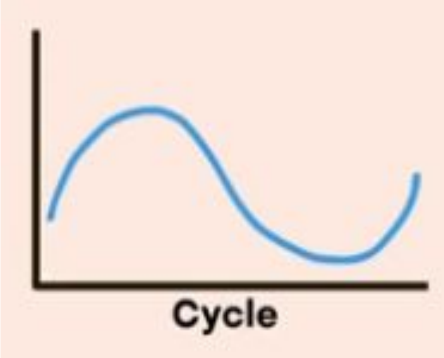
- **Fixed & Predictable:** Unlike cyclic variations, seasonal patterns repeat over fixed time periods.
- **Influences Business Planning:** Companies use seasonal trends to plan inventory, marketing campaigns, and staffing.
- **Can be Removed for Analysis:** Seasonal Adjusted Models help in isolating seasonality to better analyze the underlying trend.

Cycles

A cyclic pattern shows changes that go up and down over time, but these do not happen on a regular schedule. These changes are often caused by large-scale influences, like economic cycles or shifts in business dynamics.

Unlike seasonal changes, which occur at specific times each year, cyclic variations do not stick to a fixed timeline.

Instead, they rely on outside factors such as how the market is doing, economic changes, or significant global events. These outside factors can make the pattern hard to predict because they do not follow a set path.



Examples of Cyclic Patterns:

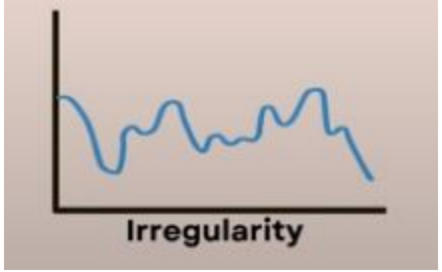
- Stock Market: The rise and fall of stock prices based on economic cycles.
- Housing Market: Booms and recessions in real estate, influenced by interest rates and inflation.
- Business Performance: Revenue fluctuations based on industry-wide economic changes.

Residuals (Irregular Variations or Noise)

Random noise or residuals are the unforeseen fluctuations in time series data that are not explicitly explainable by trend, seasonality, or cyclic patterns.

Examples of Residuals:

- Stock Market Crashes: Unexpected economic crises causing sharp market drops.
- Natural Disasters: Earthquakes or hurricanes disrupting supply chains and production.
- Viral Social Media Trends: Sudden spikes in traffic due to a viral video or news.



How to Handle Residuals?

- Smoothing Techniques: Methods like moving averages help reduce random fluctuations.
- Anomaly Detection Models: Machine learning models identify outliers caused by unpredictable events.
- Advanced Forecasting Models: Deep learning models (like LSTMs) can capture complex patterns beyond simple trend analysis.

How to Analyze Time Series?

To perform the time series analysis, we have to follow the following steps:

- Collecting the data and cleaning it
- Preparing Visualization with respect to time vs key feature
- Observing the stationarity of the series
- Developing charts to understand its nature.
- Model building – AR, MA, ARMA and ARIMA
- Extracting insights from prediction

1. **Data Cleansing**: Time series analysis techniques such as smoothing and seasonality adjustments help remove noise and outliers, making the data more reliable and interpretable.

2. **Understanding Data**: Models like ARIMA or exponential smoothing provide insight into the data's underlying structure. Autocorrelations and stationarity measures can help understand the data's true nature.

3. **Forecasting**: One of the primary uses of time series analysis is to predict future values based on historical data. Forecasting is invaluable for business planning, stock market analysis, and other applications.

4. **Identifying Trends and Seasonality**: Time series analysis can uncover underlying patterns, trends, and seasonality in data that might not be apparent through simple observation.

5. **Visualizations**: Through time series decomposition and other techniques, it's possible to create meaningful visualizations that clearly show trends, cycles, and irregularities in the data.

6. **Efficiency**: With time series analysis, less data can sometimes be more. Focusing on critical metrics and periods can often derive valuable insights without getting bogged down in overly complex models or datasets.

7. **Risk Assessment**: Volatility and other risk factors can be modeled over time, aiding financial and operational decision-making processes.
   (**Volatility** refers to how much a value (like a stock price, currency, or demand) **fluctuates** over time.
   High volatility = lots of ups and downs (more risk).
   Low volatility = more stable (less risk).)

ARIMA and Exponential Smoothing

These are **forecasting models** used in time series analysis.

- **ARIMA (AutoRegressive Integrated Moving Average):**
    - It captures **trend**, **seasonality**, and **autocorrelation** in the data.
    - It's especially good when the data needs to be made **stationary** first.
    - Helps understand how current values relate to **past values and past errors**.
- **Exponential Smoothing:**
    - Uses **weighted averages** of past data.
    - More recent data points are given more importance.
    - Good for capturing **level**, **trend**, and **seasonal patterns**.

  Both models help **identify hidden patterns and structures** in time series data so we can make better predictions.

Autocorrelation

- Measures how **current values relate to previous values** in the series.
- If data points are highly auto correlated, it means past values strongly influence future values.
  Example:
  If sales are high this month and tend to be high next month too → high autocorrelation.

Stationarity

- A time series is **stationary** if its **mean, variance, and autocorrelation are constant over time**.
- Many forecasting models (like ARIMA) **require** the data to be stationary.

Residual in Time Series Forecasting

A residual is the difference between the actual observed value and the predicted (forecasted) value of a time series.

$$\text{Residual} = \text{Actual Value} - \text{Predicted Value}$$

It represents the error or unexplained variation after modeling the trend, seasonality, and other components.

Importance of Residuals

1. Model Accuracy Evaluation

Residuals help assess how well the model fits the data.

If residuals are small and randomly distributed (white noise), the model is likely good.

2. Diagnostic Checking

Plotting residuals can show patterns, which indicate model inadequacy.

E.g. trends or cycles in residuals imply the model missed some structure.

3. Model Comparison

Comparing residual statistics (e.g., RMSE, MAE) helps choose the best forecasting model.

4. Stationarity Check

Residuals should ideally be stationary (no trend/seasonality); this supports the validity of certain models like ARIMA.

5. Detecting Outliers and Change Points

Sudden large residuals may indicate outliers, anomalies, or structural breaks in the time series.

Applications of time series forecasting

1. Economics and Finance: In the financial realm, time series analysis is used for forecasting stock prices, currency exchange rates, and economic indicators. It plays a significant role in risk assessment, portfolio management, and trading strategies.
2. Meteorology and Climate Science: Time series analysis is indispensable in predicting weather conditions, climate trends, and natural disasters. It enables meteorologists to issue weather forecasts and monitor long-term climate changes.
3. Sales and Demand Forecasting: Businesses rely on time series analysis to predict product sales and customer demand. Accurate forecasts aid in inventory management, production planning, and marketing strategies.
4. Healthcare and Epidemiology: In healthcare, time series analysis helps predict disease outbreaks and healthcare resource requirements. It's a critical tool for tracking the spread of diseases, making it especially relevant in the context of the COVID-19 pandemic.

5. Transportation and Traffic Management: Time series analysis is used to predict traffic patterns, public transportation ridership, and optimize traffic signal timing. This enhances transportation efficiency and reduces congestion.

6. Energy Consumption: Utilities employ time series analysis to forecast energy consumption patterns, allowing for better energy production and distribution planning. This leads to cost savings and resource optimization.

7. Retail and E-Commerce:

Time series analysis is widely used in retail to forecast future sales trends and consumer buying behavior. It helps retailers plan promotions, manage pricing strategies, and optimize stock levels. E-commerce platforms also rely on time series analysis to predict website traffic, customer engagement, and product demand during peak seasons, ensuring effective resource allocation.

8. Financial Fraud Detection:

In the finance industry, time series analysis can be applied to detect anomalies in transaction patterns that may indicate fraudulent activities. By analyzing the time-based sequences of transactions, financial institutions can identify irregularities or unusual behavior in accounts, improving fraud detection systems and enhancing security.

Reference Questions

1. What is time series analysis? Explain its importance.
2. How is time series forecasting different from time series analysis?
3. Describe the main goals of time series analysis.
4. What is time series data? How is it different from cross-sectional data?
5. What are the different types of time series data?
6. Explain the role of time series forecasting in real-world applications (e.g., business, healthcare, finance).
7. What are the main components of a time series?
8. Explain the difference between trend and seasonality.
9. How do cyclical and seasonality components differ?
10. How to analyze time series data?
11. What is a residual in time series forecasting?
12. Why is analyzing residuals important for model accuracy?
13. Describe applications of time series forecasting.
14. How does forecasting help in decision-making?
15. Mention two models used in time series forecasting and their uses.