<u>Time Series Analysis – Scenario based Questions with Solutions</u>

Scenario 1: Identifying Trend
You are given a dataset showing daily temperature for the past year. You notice the values generally increase over time. How do you confirm the presence of a trend?

Answer:

- First, we need to plot the temperature values over time to visually check if they consistently increase.
- Then we need to apply a moving average (e.g., 7-day or 30-day) to smooth short-term fluctuations and see the long-term direction.
- Use statistical decomposition to separate the trend component from the raw data.
- Calculate the slope of a fitted line (linear regression); a positive slope indicates an increasing trend.
- Check if average temperature values increase across months or quarters when grouped.
- Verify if first differences (today minus yesterday) are mostly positive, supporting an upward pattern.

Scenario 2: Detecting Seasonality
You are analyzing monthly electricity usage and notice similar spikes every summer. How do you confirm seasonality?

Answer:

- First, we need to plot the time series to visually check if spikes occur at the same time every year (e.g., every summer).
- Then we need to **group data by month** (e.g., Jan, Feb, …, Dec) and compare average usage for each month; repeated high values in specific months indicate seasonality.
- **Use time series decomposition** to extract the **seasonal component**.
    - If the seasonal pattern repeats consistently every 12 months, seasonality is confirmed.
- **Create a seasonal plot** by overlaying usage from multiple years.

- If the peaks align around the same months, it shows seasonal behavior.

- **Check autocorrelation (ACF plot)**.
  - Strong spikes at multiples of 12 lags suggest annual seasonality.
- **Compute month-to-month patterns** to see if similar behaviors repeat every year.

Scenario 3: Making Data Stationary
Your time series model gives poor results. The data shows trend and seasonality. How do you prepare it for ARIMA?

Answer:

- First, we need to **check stationarity** using ADF test or by plotting the data.
- **Remove trend** using:
  - Differencing (e.g., y(t) − y(t-1))
  - Or subtracting a moving average.
- Then we need to **remove seasonality** using:
  - Seasonal differencing (e.g., y(t) − y(t-12) for monthly data)
  - Or decomposition to isolate and subtract the seasonal component.
- **Apply transformations** such as log or square-root to stabilize variance if needed.
- **Recheck stationarity** using ADF test to ensure the transformed data is stationary.
- **Use the stationary data** (after differencing/transformations) as input to the ARIMA model.
- **Determine ARIMA parameters** (p, d, q) using ACF and PACF plots once stationarity is achieved.

Scenario 4: Choosing Model for No Seasonality
Your time series data (e.g., daily website visits) has no visible seasonality but has a clear trend. Which model would you choose?

Answer:

- Since the data has **trend** but **no seasonality**, we need to avoid models that require seasonal components (like SARIMA).
- Now, we need to use a **non-seasonal ARIMA** model to capture:

- Autoregressive patterns (AR)
- Differencing for trend (I)
- Moving average components (MA)

- Apply **first differencing** to remove the trend before modelling.
- Choose **ARIMA(p, d, q)** where:
  - **d = 1** usually removes the trend effectively.
- You may also consider **Exponential Smoothing (Holt's Trend Model)** if:
  - The trend is smooth and deterministic.
- Final choice:
  - **ARIMA** when data shows autocorrelation patterns.
  - **Holt's Linear Trend** when the trend is strong but random noise is minimal.

Scenario 5: Evaluating Forecast Performance
You forecast sales for 3 months and want to check how accurate it is. What steps do you follow?

Answer:

- We need to **compare actual vs predicted values** for the 3-month period to see how close the forecasts are.
- Then we need to **calculate error values** (Actual − Forecast) to understand where predictions deviate.
- Compute **standard accuracy metrics**, such as:
  - **MAE** (Mean Absolute Error) – average size of errors.
  - **MSE** or **RMSE** – penalizes larger errors more strongly.
  - **MAPE** – percentage error, easy to interpret.
- **Visualize actual vs forecast** on a line plot to quickly see how well predictions align.
- **Check residuals**:
  - Residuals should be small and randomly scattered.
  - No trend or pattern indicates a good model.

- **Compare accuracy across models** (if multiple models were tested) to choose the best performer.
- **Interpret results**:
    - Low errors = good forecast
    - High errors = model needs improvement.

Scenario 6: Sudden Spike in Data

You're analyzing daily water consumption. One day, the value suddenly jumps very high. What will you do?

Answer:

- Firstly, we need to **verify the data point** to check if the spike is due to a data entry error or sensor malfunction.
- **Compare the spike with historical values** to confirm that it is unusually high.
- **Check external factors** (e.g., repairs, leakage, seasonal events) that might explain the sudden increase.
- **Plot the time series** to visually confirm that the spike is an outlier.
- **Use statistical methods** (e.g., z-score, IQR) to formally detect the outlier.
- **Decide how to handle it**:
    - Keep it if it represents a real event.
    - Remove or replace it (e.g., with mean/median) if it is erroneous.
- **Document the anomaly** and note its impact on forecasting or trend analysis.
- **Re-run the model** after handling the spike to ensure stable performance.

Scenario 7: Sales Goes Up Every December

You're analyzing a store's monthly sales. Sales always increase in December. What does this indicate?

Answer:

- The repeated increase happening **every year in the same month** indicates **seasonality**.

- December sales likely rise due to **holiday shopping, festivals, bonuses, or year-end events**.
- Since this pattern repeats consistently, it is considered a **predictable seasonal effect**.
- This seasonal spike should be **captured in forecasting models** like SARIMA or Seasonal Exponential Smoothing.
- When analysing trends, the December rise should be treated as part of the **seasonal component**, not as a random event.
- Forecasts for December should expect **higher values** based on past patterns.

Scenario 8: Forecasting with Missing Values

You are forecasting stock prices, but a few days of data are missing. What should you do?

Answer:

- Firstly, we need to **identify the missing dates** and confirm how many data points are missing.
- **Check the pattern** of missingness (random or continuous gaps).
- **Fill missing values appropriately**, because forecasting models cannot handle gaps:
  - Use **forward fill** (previous day's value).
  - Or **interpolation** (linear or time-based) for smoother filling.
- **Avoid deleting rows** unless very few and not critical, because stock data is sequential.
- **Re-check the time index** to ensure continuity after filling.
- **Proceed with forecasting** only after the series is complete and consistent.
- **Document the imputation method** to maintain transparency in analysis.