**Financial Time Series Analysis Report**

Assessment Notebook Link: <https://github.com/AbhiRishi96/ABB_assessment/blob/main/time_series_data_analysis_notebook.ipynb>

**Objective**

To analyze monthly cash flow data of an Automotive Manufacturing Company in India for the past 10 years, with the goal of:

1. Forecasting future cash flows for the next 12 months.

2. Detecting significant anomalies in historical data that could indicate financial risks or irregularities.

**Approach**

The analysis was structured into the following steps:

**1. Exploratory Data Analysis (EDA)**

• **Purpose**: Understand the dataset structure, trends, and seasonality.

• **Methods Used**:

• **Time Series Plot**: A line graph was used to visualize cash flow changes over time.

• **Boxplots by Year**: Highlighted yearly variability and potential outliers.

• **Key Findings**:

• Observed potential seasonal patterns and variability in cash flows over the years.

• Some years showed significant deviations, prompting further anomaly analysis.

**2. Data Preprocessing**

• **Purpose**: Prepare the data for analysis by ensuring cleanliness and consistency.

• **Steps Taken**:

• **Handling Missing Values**: Missing values were interpolated using linear interpolation.

• *Why*: Ensures continuity in time series and avoids biases from omitted data.

• **Stationarity Check (ADF Test)**:

• Used the Augmented Dickey-Fuller test to check if the time series is stationary.

• *Findings*: The series was non-stationary; differencing was applied to remove trends.

• **Differencing**: Created a differenced series to remove trends for ARIMA modeling.

• **Implications**:

• Ensured the data was in a form suitable for time series analysis and forecasting.

• Addressed non-stationarity, a requirement for ARIMA-based models.

**3. Forecasting**

To predict future cash flows, we applied two forecasting models: **ARIMA** and **Prophet**.

**ARIMA (Autoregressive Integrated Moving Average)**

• **Why Used**:

• A traditional time series model that performs well for short-term forecasting.

• Suitable for stationary data and captures trends and autocorrelation.

• **Approach**:

• Stationarized the data using differencing.

• Tuned the ARIMA hyperparameters (p, d, q) using ACF and PACF plots.

• **Implications**:

• Provided a simple, interpretable forecast for the next 12 months.

• Accuracy evaluated using **Mean Absolute Error (MAE)**.

**Prophet (Facebook’s Time Series Tool)**

• **Why Used**:

• Handles non-linear trends and seasonality better than ARIMA.

• Easily incorporates holiday effects or special events.

• Robust to missing data and outliers.

• **Approach**:

• Fit the model on raw cash flow data (no differencing required).

• Forecasted for the next 12 months and plotted confidence intervals.

• **Implications**:

• Produced forecasts that accounted for both trends and seasonal patterns.

• Compared accuracy against ARIMA using MAE.

**Model Comparison:**

• **Metrics Used**:

• Mean Absolute Error (MAE): Measured how close predictions were to actual values.

• **Findings**:

• Both models produced reasonable forecasts, with Prophet slightly outperforming ARIMA for capturing seasonality.

**4. Anomaly Detection**

• **Purpose**: Identify irregularities or outliers in historical cash flow data that may represent financial risks.

• **Method Used**: Isolation Forest

• A tree-based anomaly detection algorithm.

• Flagged approximately 5% of data points (as specified by contamination=0.05).

• **Implications**:

• Detected points where cash flows were abnormally high or low compared to the rest of the data.

• Flagged anomalies likely indicate:

• Business-specific events (e.g., unexpected expenses or revenue spikes).

• Data errors or external economic disruptions.

• **Visualization**:

• Red points overlaid on the cash flow time series highlighted flagged anomalies.

**Findings**

1. **Seasonal and Trend Patterns**:

• Clear seasonal trends were observed in cash flows, with periodic increases and decreases over the years.

• Both ARIMA and Prophet effectively modeled these patterns.

2. **Forecasting**:

• Prophet produced more reliable forecasts due to its ability to handle non-linear trends and seasonality.

3. **Anomaly Detection**:

• Several anomalies were identified using Isolation Forest, warranting further investigation into specific periods.

**Recommendations**

1. **Investigate Anomalies**:

• Examine flagged anomalies to understand their causes (e.g., data entry errors, unexpected market changes).

• Adjust financial planning to account for potential risks.

2. **Forecasting Improvements**:

• Consider incorporating external macroeconomic indicators (e.g., inflation, interest rates) to enhance forecast accuracy.

• Explore advanced models like SARIMA for more complex seasonality or LSTMs for long-term dependencies.

3. **Regular Updates**:

• Periodically retrain forecasting models with updated data to improve reliability.

• Monitor cash flow trends for early detection of irregularities.

**Conclusion**

This analysis provided valuable insights into the company’s cash flow patterns, forecasted future trends, and identified potential risks through anomaly detection. By leveraging the results, the company can make informed decisions, mitigate financial risks, and plan more effectively for the future.