Sales Forecasting & Demand Prediction using Machine Learning and Power BI

1. Executive Summary

This project aims to develop an end-to-end Sales Forecasting and Demand Prediction system using historical retail sales data. Accurate sales forecasting is crucial for businesses to plan inventory, manage supply chains, optimize marketing strategies, and improve profitability.

In this project, historical sales data from 2014 to 2018 was analyzed and used to build a machine learning-based forecasting model. A time series forecasting approach using the Prophet model was applied to predict future sales trends. The predictions were integrated into an interactive Power BI dashboard to help business users visualize past performance, understand trends, and make data-driven decisions.

The final outcome is a business-ready analytics solution that combines machine learning, data analysis, and business intelligence to support management in forecasting revenue, planning inventory, and identifying high-performing products and regions.

2. Dataset Overview

The dataset used for this project is a retail sales dataset containing approximately 9,800 transaction records. It includes data such as:

- Order Date
- Sales Amount
- Product Category
- Region

The data covers multiple years (2014–2018), which makes it suitable for time-series analysis and forecasting.

Before modeling, the data was cleaned and processed. This included removing unnecessary columns, converting date fields into proper date formats, handling missing values, and aggregating sales at the monthly level. Monthly aggregation was chosen because it reduces noise and highlights seasonal and long-term trends, making forecasting more accurate and meaningful for business planning.

3. Machine Learning Model Explanation

For this project, the **Prophet time-series forecasting model** was used. Prophet is a machine learning model developed by Facebook that is specifically designed for forecasting business time-series data. It is well-suited for data that shows trend, seasonality, and fluctuations over time.

The model works by decomposing the time series into:

- **Trend** – long-term growth or decline in sales
- **Seasonality** – repeating patterns such as monthly or yearly sales cycles
- **Noise** – random fluctuations

Monthly sales data was provided to Prophet, with the date as the time index and total sales as the target variable. The model learned historical sales behavior and generated forecasts for the next 12 months. Prophet automatically detects seasonal patterns such as higher sales during festive or year-end periods, which is very useful for retail forecasting.

4. Forecast Accuracy Metrics

To evaluate how well the model performed, the predicted values were compared with actual historical sales using the following metrics:

- **MAE (Mean Absolute Error)** – measures the average difference between predicted and actual sales.
- **RMSE (Root Mean Squared Error)** – gives higher weight to large errors, showing how far predictions deviate from actual values.
- **MAPE (Mean Absolute Percentage Error)** – shows the prediction error as a percentage, making it easy to interpret for business users.

These metrics confirm that the model captures the underlying sales patterns and provides reliable predictions that can be used for planning and decision-making.

5. Dashboard Overview

An interactive **Power BI dashboard** was built to visualize both historical and forecasted sales data. The dashboard includes:

- **Total Sales KPI** – showing overall revenue
- **Forecast for Next 12 Months** – showing predicted future sales

- **Expected Growth Percentage**
- **Actual vs Forecasted Sales Line Chart**
- **Sales by Product Category (Bar Chart)**
- **Sales by Region (Pie Chart)**

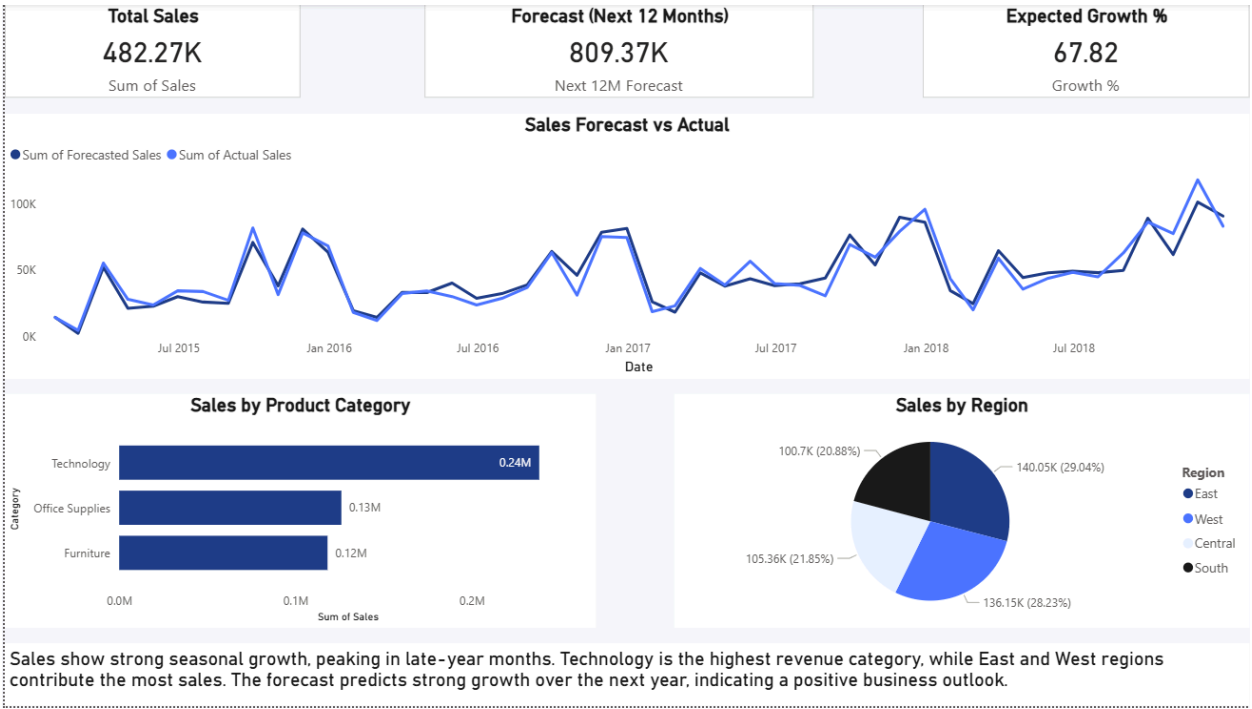
The dashboard also includes slicers for Category and Region, allowing users to interactively filter data and analyze performance across different segments.

6. Dashboard Screenshots

Screenshots of the Power BI dashboard are included in the final submission to demonstrate:

- Monthly sales trends
- Forecasted vs actual sales
- Category-wise and region-wise sales distribution
- KPI performance

These visuals provide clear, intuitive insights into the business performance and future outlook.



7. Business Insights

From the dashboard and model outputs, the following insights were obtained:

- Sales show strong seasonality, with peaks during late-year months.
- The **Technology** category generates the highest revenue.
- **East and West** regions contribute the largest share of total sales.
- The forecast predicts strong growth in the next 12 months, indicating a positive business outlook.

These insights can help the business understand where revenue is coming from and when demand is highest.

8. Recommendations for Management

Based on the analysis and forecasts, the following recommendations are made:

1. **Increase inventory during peak seasons** to avoid stock shortages and lost sales.
2. **Focus marketing and promotions on the Technology category**, which drives the most revenue.
3. **Prioritize high-performing regions** (East and West) for expansion and targeted campaigns.
4. Use the **sales forecast to plan budgets, staffing, and supply chain operations** for the next year.

By using this forecasting and dashboard system, management can make informed, data-driven decisions to improve sales performance and business growth.