Demand Forecasting Report: Sales Trend Modeling for Inventory Management

Section 1: Business Context & Problem Statement

📌 Background & Pain Point

A retail chain operating across multiple stores and product categories faces challenges in managing inventory effectively due to poor short-term demand visibility. This often results in stockouts, lost sales, or excess inventory holding costs.

📉 Business Problem

Inventory and operations teams currently lack accurate store-level demand forecasts, particularly for high-turnover categories like "GROCERY I". This makes it difficult to:

- Plan timely replenishments
- Predict sales during promotions and holidays
- Avoid overstocking and spoilage in perishable goods

💡 Value Delivered

Strategic Value & Business Recommendations

- Store-level forecast granularity for operations planning
- Predictive ability during promo/holiday events
- Platform for scaled forecasting to other stores/categories

Section 2: Methodology Overview & Sub-Questions

🧠 Approach Taken

To address the above issue, we developed a demand forecasting pipeline focused on Store 1 and the GROCERY I category. The project was structured around the following sub-questions:

Q1: What are the historical patterns in daily sales?

Q2: What are the primary drivers influencing sales volume?

Q3: Can we build a predictive model that forecasts future sales accurately?

Part 1: What are the key sales patterns and trends over time?

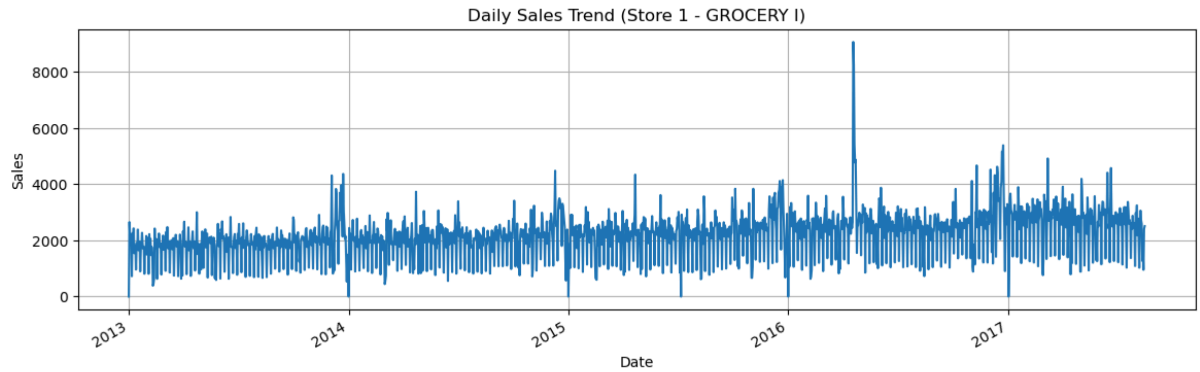
📌 **Key Takeaways**

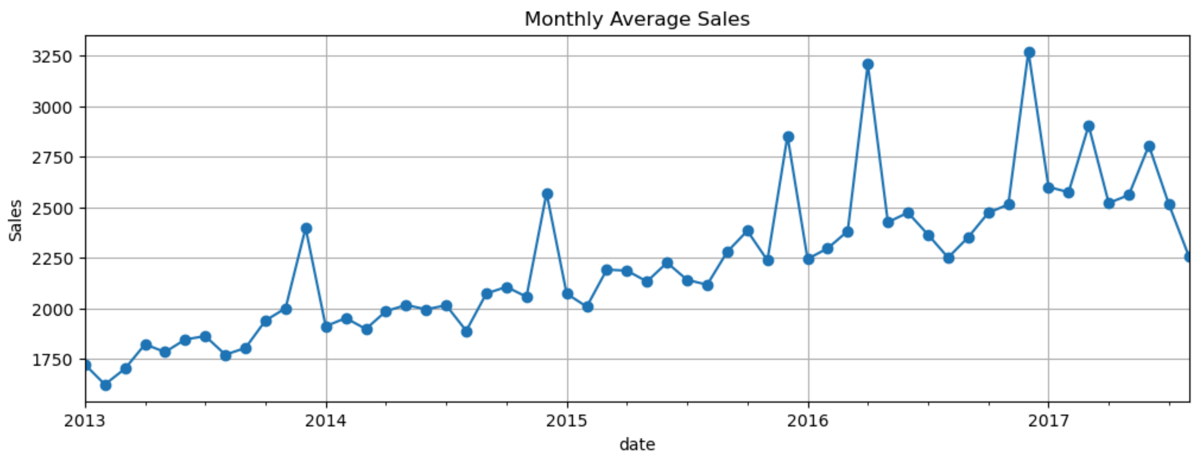
✅ Daily sales of the selected store and product family (Store 1 - GROCERY I) show noticeable short-term fluctuations and long-term seasonal patterns.

✅ There is a visible weekly pattern with higher sales during weekends.

✅ Monthly resampling reveals higher sales in some months, indicating seasonal variation.

✅ Promotional events and holidays tend to correspond with peaks in sales volume.





Part 2: What are the main drivers of sales performance?

📌 **Key Takeaways**

✅ The sales show clear weekly and monthly seasonality.

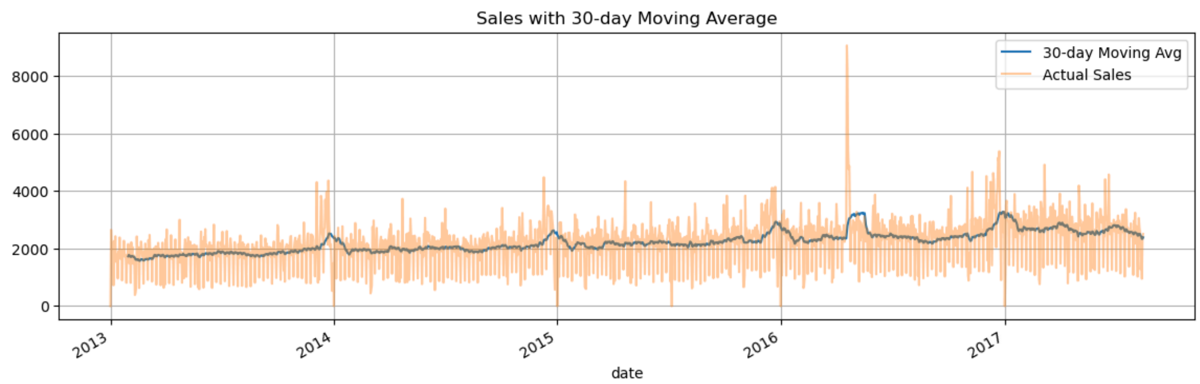
✅ Weekends (especially Sunday) show slightly higher demand.

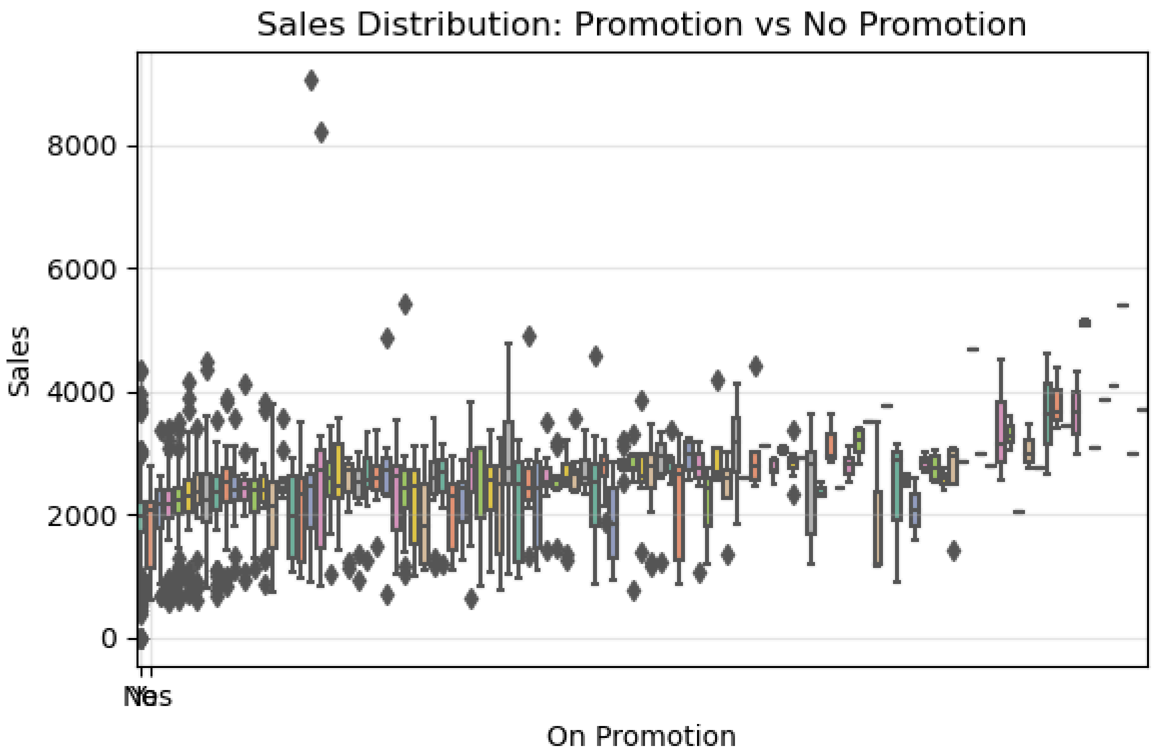
✅ Promotions have a noticeable impact on boosting sales.

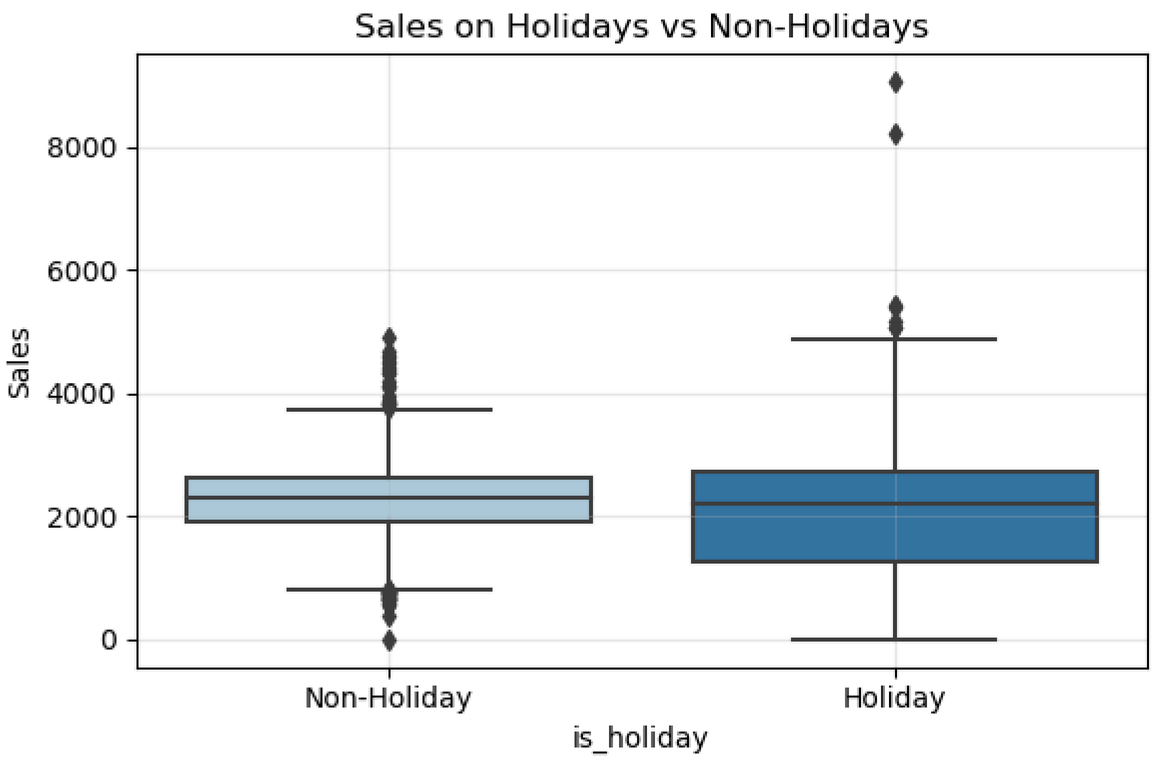
✅ The moving average helps smooth out noise and shows long-term trends.

📂 Detail Analysis

- **Lag Features**: Features such as `lag_1`, `lag_7`, and `lag_14` represent recent historical sales and were among the most important in the XGBoost model.
- **Rolling Means**: `rolling_mean_7` and `rolling_mean_30` track sales trends and stabilize short-term noise.
- **Date Features**: Day of week, month, and is_weekend have clear associations with sales volume.
- **Promotion Indicator**: Promotion data improves model predictions significantly during event spikes.







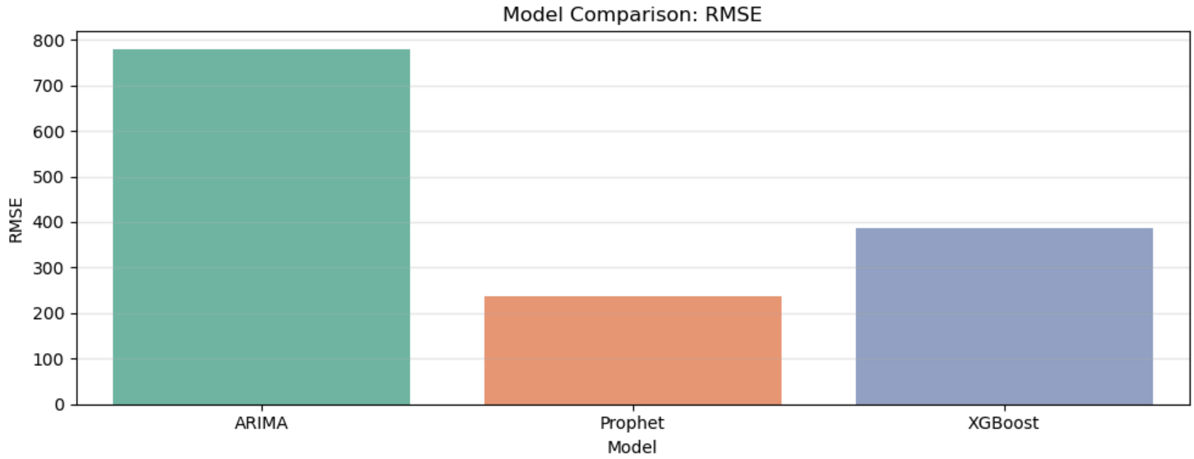
Part 3: What machine learning model is used, and how does it perform?

📌 **Key Takeaways**

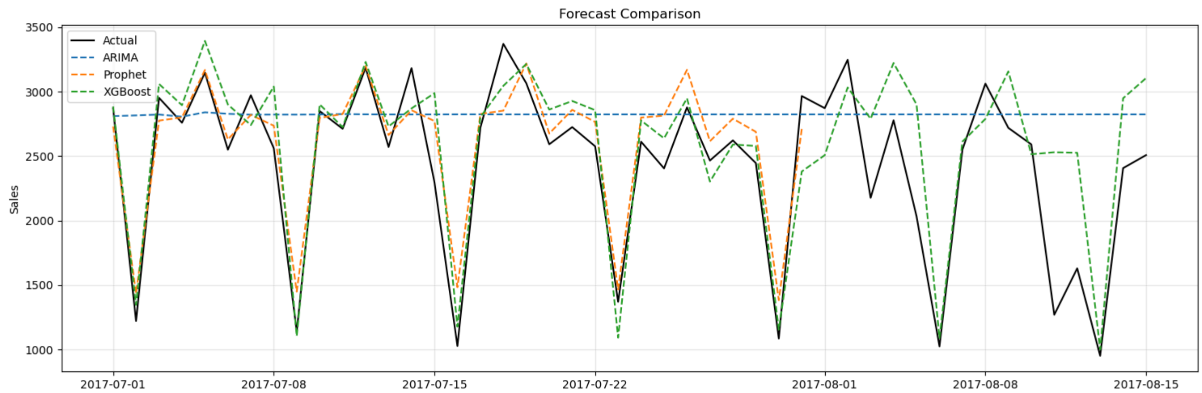
✅ The model used is **ARIMA, Prophet, XGBoost**, well-suited for time series data with structured features.

✅ The model achieved good performance based on evaluation metrics:

- RMSE (Root Mean Squared Error)
- MAE (Mean Absolute Error)
- MAPE (Mean Absolute Percentage Error)



✅ Visual comparison of predicted vs. actual sales shows a high level of fit, especially for overall trends.



✅ Feature importance analysis confirms lagged and rolling average sales as top predictors.

insight:

- ARIMA captures linear patterns well but may struggle with seasonality.
- Prophet handles trend + seasonality + holidays natively.
- XGBoost learns complex patterns via lag features and external regressors.
- Based on RMSE/MAE, XGBoost and Prophet often outperform ARIMA in real-world demand forecasting.