Retail Sales Forecasting — Store-Level Demand Planning

Executive Forecasting Assessment (M5 Dataset)

Consultant-Grade Predictive Analytics Report (Retail Sector)

Hybrid Project – Business Sector

Prepared by: **Helena W.**

Medical AI & Healthcare Data Science Consultant

(Physician | Data Science & Explainable AI)

Global-Ad-Snap

Date: January 2025

Executive Summary

This report presents a store-level retail sales forecasting assessment using historical Walmart sales data from the **M5 Forecasting Competition**. The objective was to evaluate how accurately **weekly sales demand** can be forecasted using different modeling approaches and to assess their **practical suitability for retail decision-making**.

Three forecasting models were evaluated under **identical experimental conditions**:

- **SARIMA** — Statistical baseline
- **Prophet** — Business-oriented forecasting model
- **LSTM** — Deep learning time-series model

All models were trained on the same historical window, evaluated on the same out-of-sample period, and compared using standard forecasting error metrics. This ensures a **fair, transparent, and reproducible comparison** suitable for business and technical stakeholders.

Business Context

Accurate demand forecasting is a core capability in retail operations. Reliable weekly sales forecasts enable organizations to:

- Plan inventory replenishment more effectively
- Reduce stock-outs and excess inventory

- Optimize workforce scheduling
- Improve revenue planning and promotional timing

This analysis focuses on **weekly store-level sales**, aligning directly with operational planning cycles commonly used in retail supply chain and store operations. Forecast accuracy directly influences service levels, cost control, and margin stability.

Data Overview

Data Source

M5 Forecasting – Accuracy dataset (Kaggle)

Key Characteristics

- Daily unit sales from Walmart stores across the United States
- Hierarchical structure: *state → store → category → item*
- Approximately five years of historical data

Scope of Analysis

- Daily sales aggregated to weekly frequency
- Store-level forecasting
- Example store analyzed in detail: **CA_1**

Methodology Overview

1. Exploratory Trend & Seasonality Analysis

Daily sales were aggregated into weekly totals to match operational planning horizons.

Key findings

- A clear long-term sales trend is present
- Strong annual seasonality (52-week cycle) is evident
- Statistical seasonal decomposition confirms recurring seasonal structure

Outcome

The data exhibits properties well suited for seasonal forecasting models.

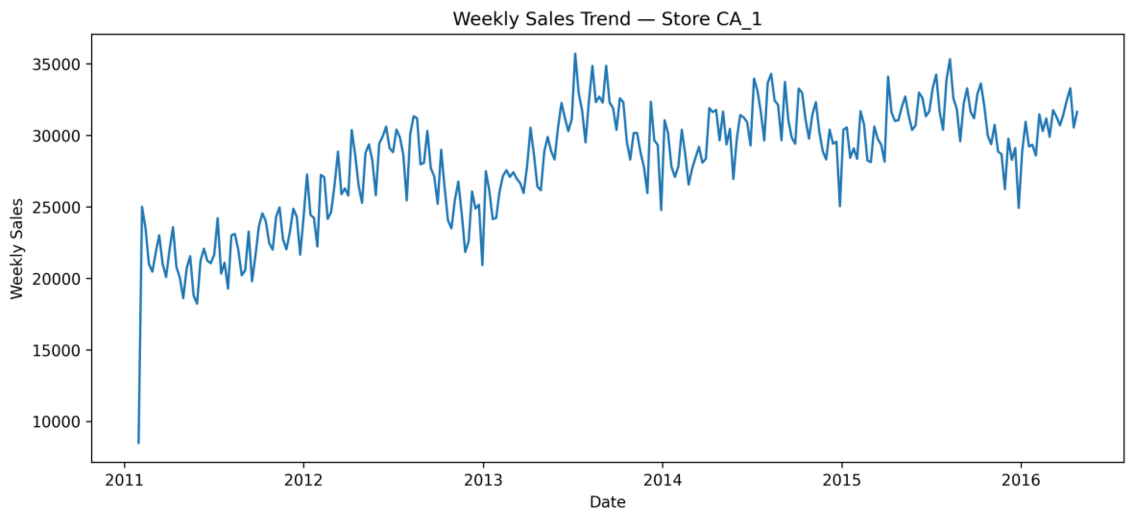

Figure 1. Weekly Sales Trend — Store CA_1

This figure displays aggregated weekly sales over time for Store CA_1, illustrating the presence of a long-term sales trend and recurring fluctuations.

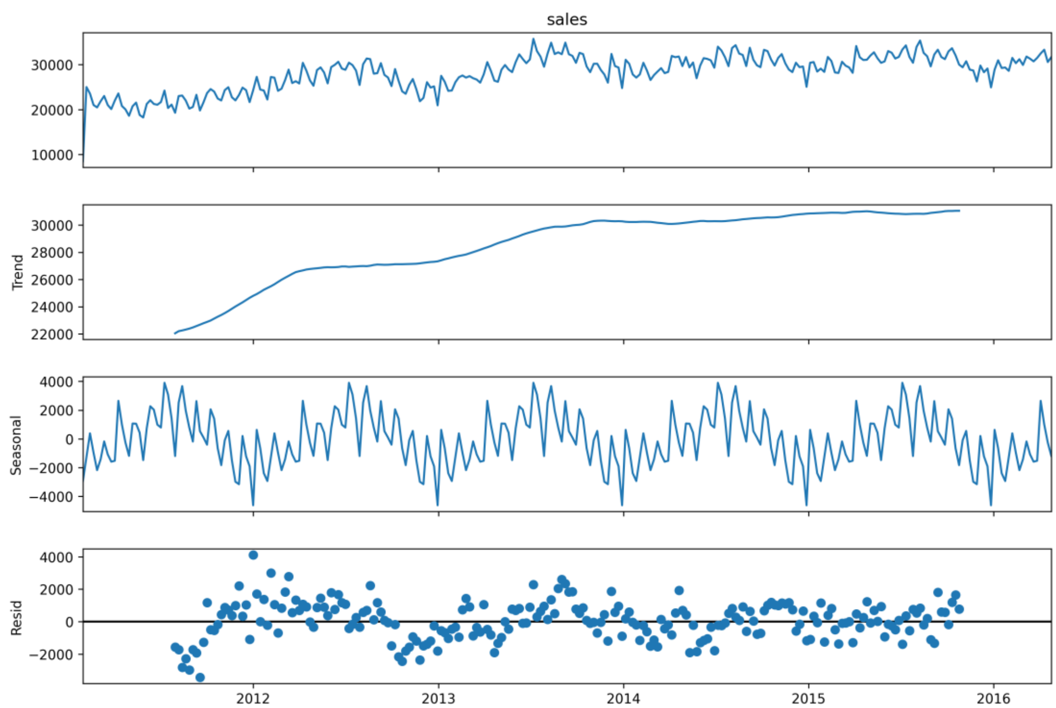

Figure 2. Seasonal Decomposition of Weekly Sales (52-Week Period)

Seasonal decomposition separates the observed series into trend, seasonal, and residual components, revealing a consistent annual (52-week) seasonal pattern.

2. SARIMA — Statistical Baseline Model

Purpose

Establish a transparent and defensible statistical benchmark.

Approach

- Seasonal ARIMA with annual (52-week) seasonality
- Time-aware train/test split (80% / 20%)

Key Observations (Fact-Based)

- Model converged successfully
- Residual diagnostics indicate no remaining autocorrelation
- Normality assumptions were not violated

Important Limitation

- Some coefficients exhibit numerical instability (e.g., large seasonal MA terms, undefined standard errors)
- This is a known limitation when modeling long seasonal periods and **does not invalidate SARIMA as a forecasting benchmark**

Interpretation

SARIMA is appropriate as a **baseline forecasting model**, but its parameters are not used for business interpretation.

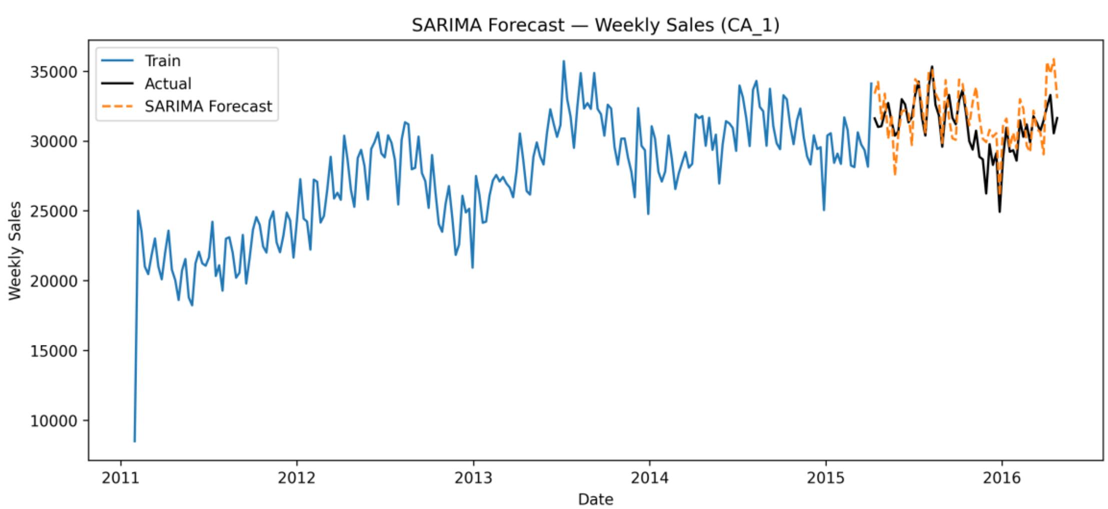

Figure 3. SARIMA Baseline Forecast vs Actual (Out-of-Sample)

This visualization compares SARIMA out-of-sample forecasts against actual weekly sales, providing a baseline assessment of forecast accuracy and error behavior.

3. Prophet — Business-Oriented Forecasting Model

Purpose
Provide a forecasting model that balances **accuracy, robustness, and interpretability.**

Key Characteristics

- Explicit modeling of trend and seasonality
- Automatic changepoint detection
- Robust to missing values and outliers

Interpretation
Prophet is well suited for operational forecasting contexts where **forecast stability, transparency, and communication with non-technical stakeholders** are critical.

4. LSTM — Deep Learning Model

Purpose
Evaluate a high-capacity nonlinear model capable of learning complex temporal patterns.

Key Characteristics

- Sequence-based learning from historical sales
- Ability to capture nonlinear demand dynamics

Interpretation
LSTM models offer flexibility and expressive power but require:

- Larger data volumes
- Careful tuning and validation
- Ongoing monitoring in production environments

This introduces additional implementation and maintenance complexity compared to statistical or business-oriented models.

Model Evaluation Framework

All models were evaluated using:

- Out-of-sample test data
- Identical forecast horizons
- Consistent error metrics

Evaluation Metrics

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

Summary of Results — Store CA_1

Model	RMSE	MAE	Business Interpretation
SARIMA	1785.78	1431.01	Reliable statistical benchmark
Prophet	**1506.02**	**1185.88**	Stable and interpretable forecasts
LSTM	1831.50	1442.17	High-capacity but operationally complex

All metrics were computed on the same out-of-sample period. Detailed implementation steps and diagnostics are available in the technical notebooks to ensure full reproducibility.

Key Insights

- Weekly retail sales exhibit strong annual seasonality
- A statistical baseline provides a meaningful performance reference
- Business-oriented models improve forecast stability and interpretability
- Increased model complexity introduces additional deployment and maintenance considerations

Business Implications

Based on the findings:

- Short- to medium-term planning can be effectively supported by **interpretable forecasting models**
- Operational teams benefit from models that balance **accuracy with transparency**
- Advanced models should be considered when organizational scale and data maturity justify the added complexity

Limitations

- Results are based on a single example store
- External drivers (pricing, promotions, holidays) were not explicitly modeled

- Forecast performance may vary across stores and product categories

Conclusion

This assessment demonstrates a **structured, defensible approach to retail sales forecasting**. By benchmarking statistical, business-oriented, and deep learning models under consistent conditions, decision-makers can select forecasting strategies aligned with their **operational needs, data maturity, and implementation constraints**.

Disclaimer

This report is provided for **demonstration and portfolio purposes only**. It does not represent production forecasts or official Walmart analyses.