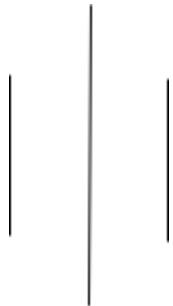# 📊 Store Sales Forecasting Dashboard Using SARIMAX with Promotion Impact

Submitted By:

Indra Jaiswal

Date:

February 3, 2026

mindrisers

Table of Content:

1. Introduction

Accurate sales forecasting is crucial for retail operations. It helps in inventory optimization, resource planning, and promotion strategy.

This project focuses on **daily store sales forecasting** using historical sales data and promotion information. We implemented **ARIMA, SARIMAX, and Prophet** to compare their effectiveness. The **SARIMAX model** was finally deployed in a **Streamlit dashboard** to provide an interactive interface for testing future promotion scenarios.

---

2. Objectives

- Forecast daily sales using historical data.

- Incorporate **promotion effects** to see their impact on sales.

- Provide an **interactive dashboard** to simulate multiple promotion strategies: No Promotion, Full Promotion, or Custom.

- Visualize **forecasted sales trends**, key KPIs, and total sales for business planning.

---

3. Dataset Description

Column	Description
id	Unique identifier for each record
date	Date of the sales record
store_nbr	Store number (1, 2, ..., n)
family	Product family/category (e.g., BEVERAGES, PRODUCE)
sales	Number of items sold (target variable)
onpromotion	Number of items on promotion (0 = no, 1 = yes)

4. Data Preprocessing

- date converted to datetime and set as index

- Missing values handled (filled with 0 where necessary)

- sales and onpromotion converted to **numeric types**

- Aggregated sales at **daily level** for modeling

- Checked for duplicates and chronological consistency

- Split into **training and test sets** using time-aware splitting

5. Methodology

5.1 Models Implemented

5.1.1 ARIMA

How it works in the project:

- ARIMA captures the **trend and autocorrelation** in daily sales.

- The model order (p=2, d=0, q=1) was selected via **grid search** using the training set.

- Applied **log transformation** to stabilize variance before fitting.

Results:

- Captures the overall trend but **cannot handle weekly seasonality or promotions**.

- Performance:

  o   MAE: 100,723

  o   RMSE: 133,142

- **Observation:** Forecasts follow the trend but **miss spikes due to promotions**.

5.1.2 SARIMAX

How it works in the project:

- Extends ARIMA to include **seasonality (weekly)** and **exogenous variable (onpromotion)**.

- Model order: order=(2,0,1) and seasonal_order=(1,1,1,7)

- Promotions directly influence the forecast: when onpromotion=1, the forecasted sales slightly increase.

- Fit on historical daily sales with onpromotion as exogenous input.

Results:

- MAE: 92,980

- RMSE: 85,758

- **Observation:**

  - Forecast captures **trend + weekly seasonality + promotional spikes**.

  - The effect of promotions is **quantifiable**, though moderate.

  - Dashboard scenario testing (No Promo vs Full Promo vs Custom) shows **small but visible differences in forecast totals** (e.g., total sales slightly higher with promotions).

Example Promotion Impact:

Scenario	Total Forecasted Sales
No Promotion	10,379,141
Full Promotion	10,379,851

5.1.3 Prophet

How it works in the project:

- Prophet models **trend and seasonality automatically**.

- Tested with and without holidays.

- Cannot directly handle promotions, so effect is **not included unless explicitly added as regressor**.

-

Results:

- Without holidays: MAE = 78,450, RMSE = 97,403

- With holidays: MAE = 90,308, RMSE = 109,711

- **Observation:**

  - Performs well on **overall trend and seasonality**, but **fails to reflect promotion spikes**.

6. Hyperparameter Tuning

ARIMA & SARIMAX: Grid search over (p,d,q) and seasonal (P,D,Q,s), selected using **AIC**.

Prophet: Grid search over changepoint_prior_scale, seasonality_prior_scale, and holidays_prior_scale.

7. Model Evaluation Metrics

- MAE (Mean Absolute Error)

- RMSE (Root Mean Squared Error)

- AIC (for ARIMA/SARIMAX)

- Time Series Cross-Validation

8. Results and Model Comparison

8.1 Test Set Performance

Model	MAE	RMSE
ARIMA	100,723	133,142
SARIMAX	92,980	85,758
Prophet (No Holidays)	78,451	97,403
Prophet + Holidays	90,308	109,711

8.2 Cross-Validation Results (MAE)

Model	CV MAE
ARIMA	53,152
SARIMAX	92,981
Prophet	75,562

Observation:

- SARIMAX is preferred because it **handles promotions and seasonality**, producing more actionable forecasts despite slightly higher CV MAE than Prophet.

- ARIMA cannot model seasonality or promotions effectively.

- Prophet fails to account for promotions without additional regressors.

9. Promotion & Holiday Impact Analysis

Promotion Scenarios in Dashboard:

Scenario	Description	Effect on Forecast
No Promotion	All days = 0	Baseline trend
Full Promotion	All days = 1	Slight increase in sales
Custom	User-defined 0/1	Sales rise on selected promotion days

Holiday Analysis:

- Prophet with holidays showed negligible improvement; holidays have minimal impact.

10. Final Model Selection

✅ **Selected Model: SARIMAX**

Reasons:

- Captures **trend, seasonality, and promotion effects**

- Provides **quantifiable impact of promotions**

- Produces **stable, interpretable forecasts** suitable for business deployment

11. Deployment

- SARIMAX model saved using joblib.

- **Streamlit dashboard features:**

    o Forecast horizon selection (7–90 days)

    o Promotion scenario simulation (No / Full / Custom)

- KPI metrics: Last sales, average sales, total forecast, promotion days

- Historical + forecast visualization

- Data table export to CSV

---

12. Business Impact

- Improved inventory planning and demand forecasting

- Reduced stockouts and overstock risk

- Provides actionable insights for promotion planning

- Interactive dashboard enables **scenario testing and data-driven decision-making**

---

13. Limitations

- External shocks (economic changes, pandemics) not captured

- Limited effect of promotions in some stores/families

- Forecasts may be less accurate for **new stores** or **new product families**

---

14. Future Scope

- Include **holidays, events, and weather** as regressors

- Multi-store and product-family level forecasting

- Use deep learning models (LSTM/GRU) for complex patterns

- Real-time data integration and automated retraining

- Ensemble modeling (SARIMAX + Prophet) for higher accuracy

15. Conclusion

- SARIMAX effectively models **trend, weekly seasonality, and promotion effects**

- Dashboard allows **interactive forecasting**, scenario testing, and KPI visualization

- Provides **actionable insights** for inventory management, promotion planning, and business strategy