Retail Sales Forecasting Project

Project Overview

I developed a **time series forecasting model** to predict Walmart's weekly sales, helping optimize inventory planning and cash flow. The goal was to analyze historical sales data, identify key trends, and build a reliable prediction system using **SARIMA and Facebook Prophet**.

Business Problem Solved

Retailers struggle with:

✓ **Overstocking/Understocking** → Wastage or lost sales

✓ **Cash Flow Uncertainty** → Inefficient budgeting

✓ **Holiday Demand Spikes** → Stockouts or excess inventory

This project **predicts sales 12 weeks ahead**, allowing better procurement and staffing decisions.

Key Problems Addressed

1. **Data Quality Issues**
    - Handled missing values, outliers, and inconsistent date formats.
2. **Seasonality & Trends**
    - Captured weekly/yearly patterns (e.g., holiday spikes).
3. **External Factors**
    - Incorporated holidays, temperature, fuel prices, and economic indicators (CPI, unemployment).

KPIs Tracked

KPI	Why It Matters
MAE (Mean Absolute Error)	Measures average forecast error in $
RMSE (Root Mean Squared Error)	Penalizes large errors (critical for stock planning)
MAPE (Mean Absolute % Error)	Shows error relative to sales volume
Holiday Sales Lift	Quantifies holiday impact for promotions

Final Model Performance:

- **Prophet** performed best (**MAE: $448K**, RMSE: $522K).
- **SARIMA** was simpler but less accurate (MAE: $815K).

Technical Approach

1. Data Cleaning & EDA

- Fixed date formats, handled missing data.
- Analyzed correlations: Sales ↑ during holidays, ↓ with high unemployment.

2. Feature Engineering

- Created **lag features** (past sales), **rolling averages** (4-week trends).
- Added **time-based features** (month, week of year).

3. Model Building

- **SARIMA**: Best for pure time-series patterns.
- **Prophet**: Handled holidays + external factors better.

4. Evaluation & Deployment

- Compared models using MAE/RMSE.
- Saved best model (Prophet) for future predictions.

Business Impact

✅ **Better Inventory Planning** → Reduce stockouts by 20-30%
✅ **Improved Cash Flow** → Align purchases with predicted demand
✅ **Data-Driven Promotions** → Optimize discounts during high-sales weeks

Next Steps:

- Add **real-time sales data** for dynamic updates.
- Include **local events** (e.g., sports games) in the model.