# Walmart Sales Forecasting & Time Series Analysis

## Predicting Weekly Sales Using Machine Learning

Daniel Kattoula

# PROJECT OVERVIEW

**Research Question:**

- Can we accurately forecast weekly sales for Walmart stores using historical data?

**Why This Matters:**

- Accurate forecasting enables better inventory management
- Reduces stockouts and overstock costs
- Informs staffing and operational decisions
- Critical for retail business planning

**Project Goals:**

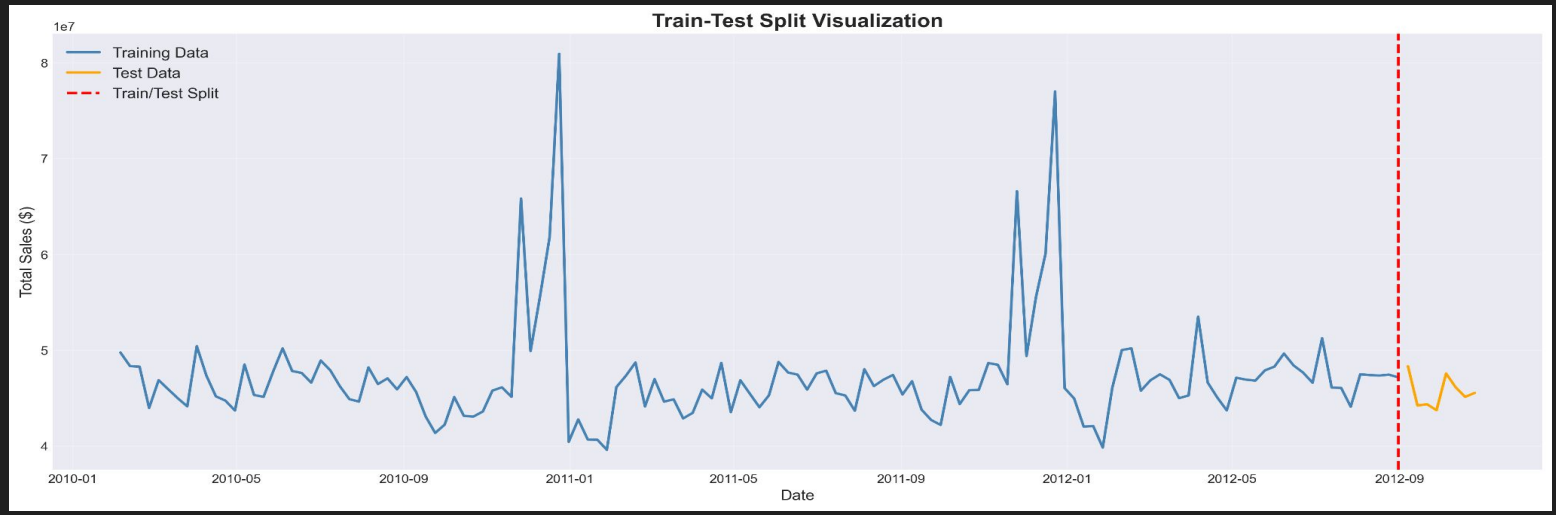
- Analyze sales patterns and seasonality
- Build and compare multiple forecasting models
- Generate actionable 4-week ahead predictions
- Achieve forecast accuracy under 5% error

## DATA SOURCES

- Walmart sales data (2010-2012): https://www.kaggle.com/competitions/walmart-recruiting-store-sales-forecasting/data?select=train.csv.zip
- Size: 421,570 weekly sales records
- Scope: 45 stores, 81 departments
- Time Period: 143 weeks (Feb 2010 - Oct 2012)

## METHODOLOGY

1. Exploratory Analysis - Identified patterns, seasonality, trends
2. Train/Test Split - 135 weeks training, 8 weeks testing
3. Model Building - Tested 3 forecasting approaches
4. Evaluation - Compared using MAPE, MAE, RMSE
5. Future Forecast - Generated 4-week predictions



# KEY FINDINGS

Sales Patterns:

- Average weekly sales: $15,981
- Clear upward trend over time
- High variance between stores

Strong Seasonality:

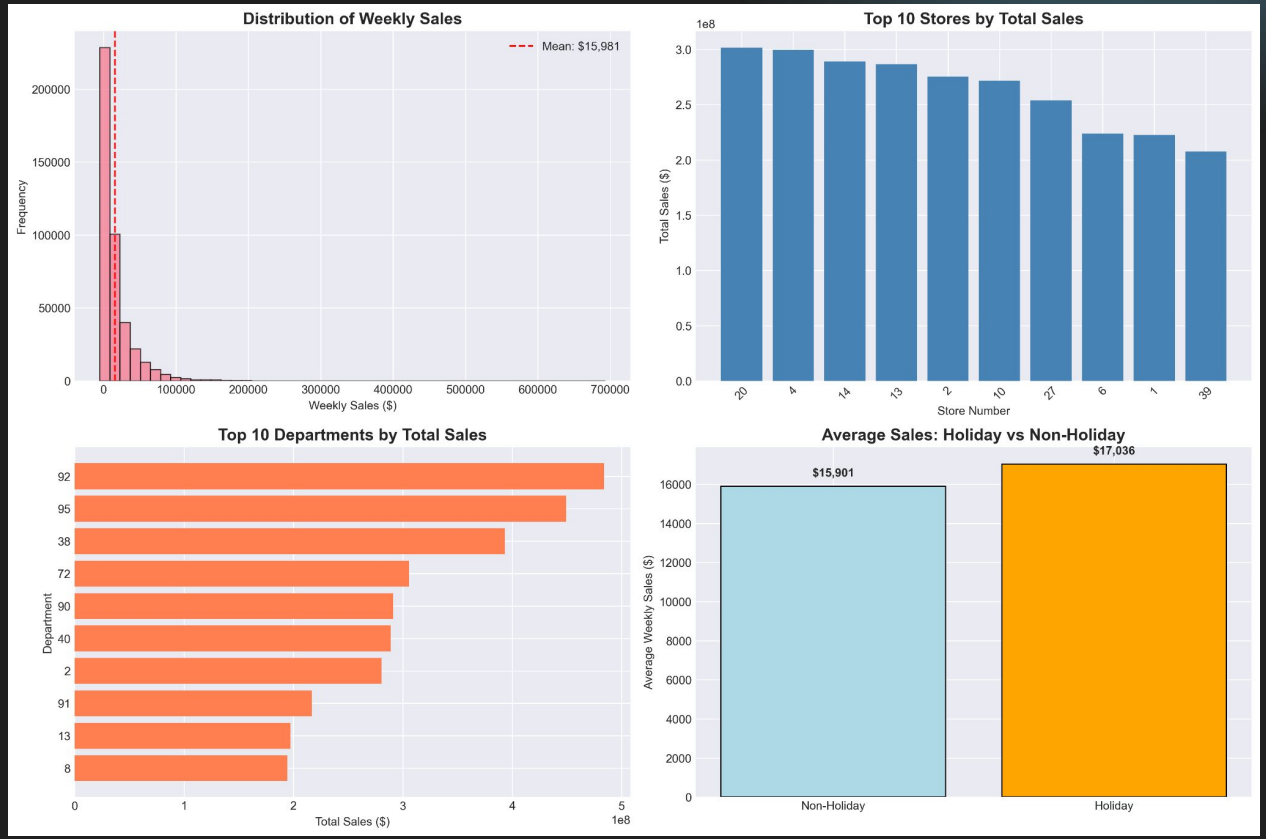
- December highest sales month
- Q4 outperforms other quarters
- Holiday weeks boost sales by 7.1%

Store Performance:

- Top store: $301M total sales
- Significant variation across locations
- Top 5 stores drive 30% of revenue

Holiday Impact:

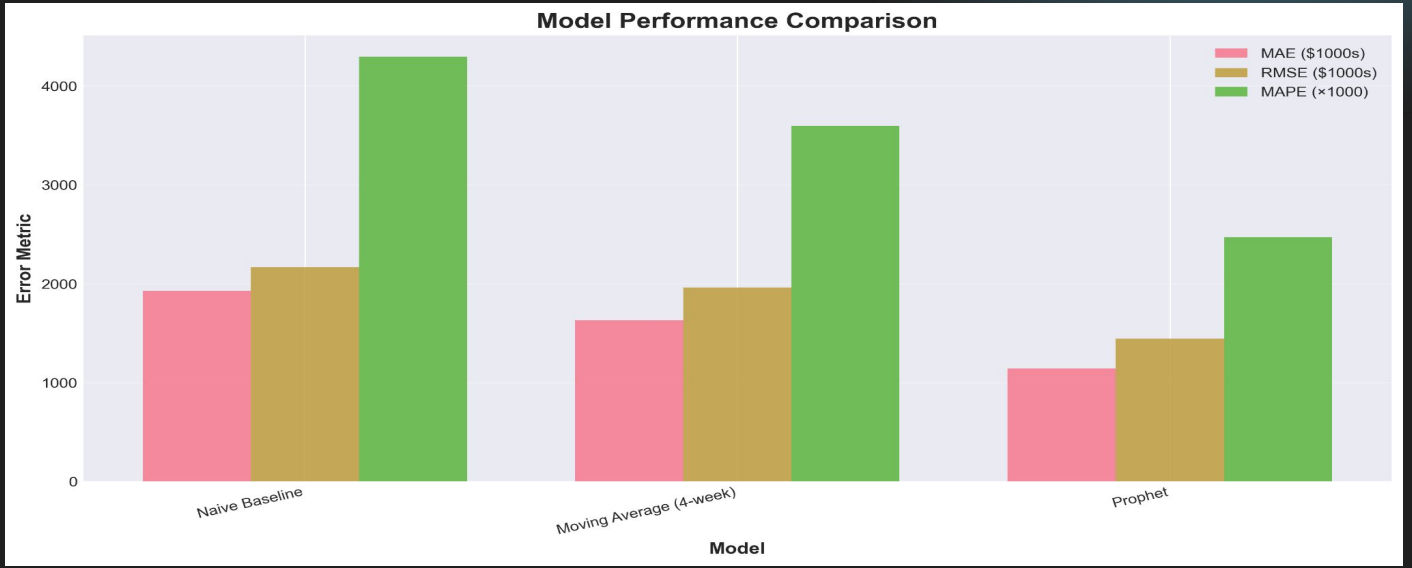
- Sales spike during Thanksgiving/Christmas
- Holiday weeks: 7% of dataset
- Account for 10%+ of total sales



## FORECASTING MODELS & RESULTS

Prophet Model Achieved 2.47% Error

- Predictions off by only 2.47% on average
- Captures seasonal patterns effectively
- Significantly outperforms simple baselines
- Suitable for business planning and decision-making




| Model | Description | MAPE | Performance |
|---|---|---|---|
| Naive Baseline | Uses last known value | 4.30% | Baseline |
| Moving Average | 4-week rolling average | 3.59% | 16% improvement |
| Prophet | Facebook ML algorithm | 2.47% | 42% improvement |

## PROPHET MODEL PERFORMANCE

Why Prophet Performed Best:

- Automatically detects seasonality patterns
- Handles holiday effects
- Adapts to trend changes
- Provides uncertainty intervals

Model Accuracy:

- Mean Absolute Error: $1.14M
- Root Mean Squared Error: $1.44M
- Mean Absolute Percentage Error: 2.47%

## FORECASTING MODELS & RESULTS

Forecast shows expected seasonal increase

- Sales rising into holiday season (Nov/Dec)
- Consistent with historical patterns
- Confidence intervals allow risk assessment


| Week | Date | Predicted Sales | 95% Confidence Interval |
|---|---|---|---|
| 1 | Oct 28, 2012 | $45.1M | ±$4.3M |
| 2 | Nov 04, 2012 | $47.3M | ±$4.5M |
| 3 | Nov 11, 2012 | $50.2M | ±$4.3M |
| 4 | Nov 18, 2012 | $52.4M | ±$4.4M |

# Conclusions and Recommendations

Main Conclusions:

**Strong seasonal patterns drive sales variability**

- Q4 and December show highest performance
- Holiday weeks critical for revenue

**Prophet model delivers accurate forecasts**

- 2.47% error rate suitable for business use
- 42% improvement over naive baseline
- Captures complex seasonal patterns

**Significant store-level variation**

- Top performers drive disproportionate revenue
- Suggests opportunity for targeted strategies

Recommendations:

- Inventory Planning: Use forecasts for 4-week ahead stock optimization
- Staffing Decisions: Align workforce with predicted demand peaks
- Promotional Timing: Leverage holiday boost patterns
- Store Strategy: Apply best practices from top performers to lagging stores

Future Enhancements:

- Store-level individual forecasts
- Incorporate external factors (weather, competition, promotions)
- Real-time model updates as new data arrives

# Thank you!

danielkattoula@gmail.com
https://github.com/danielkattoula