

WALMART PROJECT

Presentated by Mansi Singh

INTRODUCTION

Effective inventory management is vital for retail stores to align supply with demand across multiple outlets. This project analyzes and forecasts weekly sales for a nationwide retail chain using data on sales, holidays, temperature, fuel prices, consumer price index (CPI), and unemployment rates.

We explore how unemployment, seasonal trends, temperature, and CPI affect sales through statistical analysis and exploratory data analysis (EDA). Insights from this analysis will help optimize inventory and sales strategies.

Additionally, we use predictive modeling to forecast sales for each store over the next 12 weeks, enabling the retail chain to make informed decisions and enhance inventory management. This project aims to improve supply-demand matching, reduce stockouts, and boost customer satisfaction.

Company Overview

Walmart Inc. is a global retail corporation that operates a chain of hypermarkets, discount department stores, and grocery stores. Renowned for its wide range of products at competitive prices, Walmart serves millions of customers worldwide.

Analysis and Forecasting of Retail Store Sales

Today, I will be presenting my project on analyzing and forecasting sales for Walmart, a nationwide retail chain. Our goal is to optimize inventory management and enhance sales performance across multiple outlets.

Objectives:

1. Understand the Impact of Various Factors on Sales:

- **Unemployment Rates:** Determine how changes in unemployment affect weekly sales and identify stores most affected.
- **Seasonal Trends:** Analyze sales data to identify seasonal patterns and understand the underlying reasons.
- **Temperature:** Assess the influence of temperature variations on weekly sales.
- **Consumer Price Index (CPI):** Examine how fluctuations in CPI impact sales across different stores.
- **Store Performance:** Identify top-performing stores and analyze the performance gap between the best and worst-performing stores.

2. Forecast Future Sales:

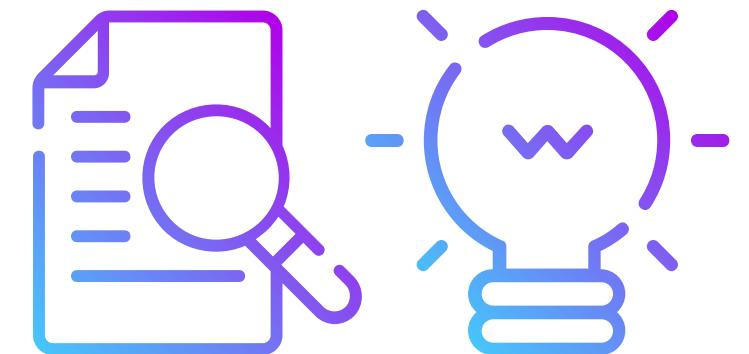
- Use predictive modeling techniques to forecast sales for each store over the next 12 weeks, enabling better inventory planning and management.



Dataset Overview

The dataset contains **6435** rows and **8** columns representing weekly sales data from multiple outlets of a retail store. The columns include:

- **Store:** Store number
- **Date:** Week of sales
- **Weekly_Sales:** Sales for the given store in that week
- **Holiday_Flag:** Indicates if it is a holiday week
- **Temperature:** Temperature on the day of the sale
- **Fuel_Price:** Cost of fuel in the region
- **CPI:** Consumer Price Index
- **Unemployment:** Unemployment rate



Insights from Statistical Analysis and Exploratory Data Analysis (EDA)

Impact of Unemployment Rate on Weekly Sales

- Analysis: Conduct a correlation analysis between Weekly_Sales and Unemployment for each store.
- Insight: Identify stores with high negative correlations, indicating that higher unemployment rates lead to lower weekly sales.
- Result: Stores with the highest negative correlations suffer the most from higher unemployment rates.

Seasonal Trend in Weekly Sales

- Analysis: Perform time series decomposition to detect seasonal patterns in the sales data. Additionally, visualize sales trends over time.
- Insight: Identify periods with consistently high or low sales. For example, increased sales during holiday seasons or summer months.
- Result: Seasonal trends could be attributed to holiday shopping sprees or seasonal promotions.

Effect of Temperature on Weekly Sales

- Analysis: Examine the relationship between Temperature and Weekly_Sales using scatter plots and correlation coefficients.
- Insight: Determine if higher or lower temperatures are associated with increased or decreased sales.
- Result: Specific temperature ranges that positively or negatively affect sales can be identified.

Impact of Consumer Price Index on Weekly Sales

- Analysis: Analyze the correlation between CPI and Weekly_Sales for each store.
- Insight: Understand how changes in the consumer price index affect purchasing power and sales.
- Result: Stores showing high sensitivity to changes in CPI could be identified.

Top Performing Stores

- Analysis: Calculate total and average weekly sales for each store over the entire period.
- Insight: Rank stores based on their total and average sales.
- Result: List the top-performing stores historically.

Worst Performing Store

- Analysis: Similar to the top-performing stores, identify the store with the lowest total and average weekly sales.
- Insight: Compare the sales performance of the best and worst-performing stores.
- Result: Highlight the significance of the difference in performance, which could point to management, location, or other operational issues.

Predictive Modeling for Sales Forecasting

To forecast the sales for each store for the next 12 weeks, we can use time series forecasting methods. Here's a high-level outline of the approach:

Data Preparation

- Handling Missing Values: Impute missing values using techniques like forward fill, backward fill, or interpolation.
- Outlier Detection: Identify and handle outliers using methods like Z-score analysis or IQR.

Model Selection

- ARIMA (AutoRegressive Integrated Moving Average): Suitable for univariate time series data with potential seasonality and trend.
- SARIMA (Seasonal ARIMA): An extension of ARIMA that accounts for seasonality explicitly.
- Prophet: A model by Facebook that is robust to missing data and shifts in the trend and can handle seasonality well.
- LSTM (Long Short-Term Memory) Networks: A type of recurrent neural network capable of learning long-term dependencies, suitable for complex time series.

Model Training and Evaluation

- Train-Test Split: Split the data into training and testing sets. Use the training set to train the model and the testing set to evaluate its performance.
- Cross-Validation: Use cross-validation techniques to ensure the model generalizes well to unseen data.
- Evaluation Metrics: Use metrics like Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE) to evaluate model performance.

Forecasting

- Generate Forecasts: Use the trained model to forecast sales for the next 12 weeks.
- Confidence Intervals: Provide confidence intervals for the forecasts to quantify uncertainty

Conclusion

By performing statistical analysis, EDA, and predictive modeling, we can gain valuable insights into the factors affecting sales and forecast future sales. This information helps the retail store in making data-driven decisions to manage inventory effectively, plan promotions, and optimize operations across different outlets.



**THANK
YOU**