Capstone Project Report

Walmart Sales Forecasting & Online Retail Analysis

Problem Statement 1

Analyze Walmart sales data and build forecasting models to address inventory challenges.

Project Objective

Generate insights and develop accurate sales forecasting models.

Data Description

Weekly sales, store/department info, CPI, fuel price, temperature, unemployment, holiday flag.

Data Preprocessing

Handled missing values, removed outliers, engineered features, encoded variables, scaled fields.

Algorithm Selection

Random Forest, ARIMA/SARIMA, Prophet, LSTM.

Model Evaluation

Used RMSE, MAE, MAPE with time-series split.

Inferences

Sales peak during holidays; temperature/CPI affect demand; promotions cause variance.

Future Scope

Automated forecasting system, dashboards, inventory optimization.

Problem Statement 2

Analyze online retail purchasing behavior and create customer segments.

Customer Insights

Repeat customers dominate revenue; seasonal purchase patterns; product trends discovered.

Customer Segmentation

Used RFM + KMeans to classify customers into groups.

Notebook Visual Outputs

Total Weekly Sales (capped) - All Stores

Correlation Matrix

Average Weekly Sales by Month (capped)

Sales vs Unemployment (sampled points + LOWESS)

Top 15 Stores by Mean Weekly Sales (capped)

Average Weekly Sales by Month (Seasonality View)

Weekly Sales vs Temperature

Top 10 Performing Stores (Mean Weekly Sales)

Store 20 - 12 Weeks Sales Forecast

Store 4 - 12 Weeks Sales Forecast

Store 14 - 12 Weeks Sales Forecast

Store 13 - 12 Weeks Sales Forecast

Store 2 - 12 Weeks Sales Forecast

Store 20 - 12 Weeks Sales Forecast

12 Weeks Sales Forecast - All Stores