Comprehensive Sales Forecasting and Customer Segmentation Report

# Introduction

Sales forecasting and customer segmentation are crucial for businesses to predict future sales, optimize inventory, and enhance targeted marketing strategies. This report presents a detailed analysis of retail sales forecasting using time-series models and customer segmentation using clustering techniques. The study explores patterns in historical sales data and identifies key customer behaviors to drive data-driven business decisions.

## Objectives

The primary objectives of this analysis are:

* To analyze historical sales data and detect trends.
* To predict future sales using time-series forecasting.
* To identify key customer segments through clustering techniques.
* To evaluate customer purchase behavior and retention likelihood.
* To provide business recommendations based on data insights.

## Methodology

This project follows a structured approach, including data preparation, exploratory analysis, predictive modeling, and clustering.

### Data Preparation and Cleaning

* **Data Loading:** The dataset containing transaction details was imported.
* **Missing Values Handling:** Replaced missing values and removed duplicates.
* **Feature Selection:** Focused on relevant columns.
* **Datetime Conversion:** Transformed the Date column into a datetime format.

### Exploratory Data Analysis (EDA)

* Checking for Correlations
* Identifying Best-Selling Products
* Sales Trend Analysis
* Customer Behavior Analysis

### Predictive Modeling

* **Objective:** Predict whether a customer will purchase again.
* **Model Used:** Random Forest Classifier
* **Feature Importance Analysis:** Identified key attributes influencing customer retention.

### Clustering Analysis

* **Objective:** Group customers into segments based on their purchasing behavior.
* **Algorithm Used:** MiniBatchKMeans for efficient clustering.
* **Segmentation Insights:** Identified groups such as high-value customers, discount seekers, and infrequent buyers.

### Sales Forecasting

* **Objective:** Predict future sales trends using historical transaction data.
* **Model Used:** Facebook Prophet
* **Forecast Duration:** Next 30 days
* **Evaluation Metrics:** Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE).

## Tools Used

The following tools and libraries were used:

* Python
* Pandas
* Matplotlib & Seaborn
* Scikit-learn
* Facebook Prophet
* Jupyter Notebook

## Findings & Insights

### Predictive Model Accuracy

1. **MAE:** 244.1843669896603
2. **RMSE:** 1219.6870722750962

### Sales Forecast Insights

* Sales peak during the holiday seasons.
* A decline in sales was observed in mid-year months.
* Top-selling products contribute to 70% of total revenue.

### Customer Segmentation Findings

* High-Spending Customers.
* Discount Seekers.
* Occasional Buyers.

## Challenges & Limitations

* Data Gaps
* Limited Features
* Changing Market Conditions.

## Recommendations

1. **Inventory Management:** Adjust stock levels according to forecasted demand trends.
2. **Personalized Marketing:** Target high-value customers with loyalty programs.
3. **Discount Timing Strategy:** Launch promotions during slow months.
4. **Continuous Model Optimization:** Retrain forecasting models periodically.
5. **Expansion of Features:** Integrate external factors such as seasonality and competitor pricing.

## Conclusion

This study successfully implemented a predictive sales forecasting model and customer segmentation analysis to extract meaningful business insights. The findings can be leveraged to improve inventory planning, targeted marketing, and customer retention strategies.