Customer Churn Prediction & Retention Strategy

Comprehensive E-Commerce Churn Analysis & ML Model

By – Rupesh Varma

Introduction

- Objective: Analyze and predict customer churn for an e-commerce platform

- Approach:

- - Exploratory Data Analysis (EDA)

- - Feature Engineering

- - Machine Learning Models

- - Conclusion

Dataset Overview

- 5,630 customer records

- 20 features including: CustomerID, Churn, Tenure, PreferredLoginDevice, Payment Mode, Satisfaction Score

- Missing values handled for Tenure, OrderCount, and WarehouseToHome

- Target Variable: Churn (1 = Left, 0 = Retained)

Key Insights

- 16.84% churn rate, with most churn in first 18 months

- Female customers show slightly higher retention

- Satisfaction score strongly correlated with retention

- Spending behavior fluctuates after 30 months

Churn by Percentage

Overall Churn Percentage

Percentage of Customers who Left
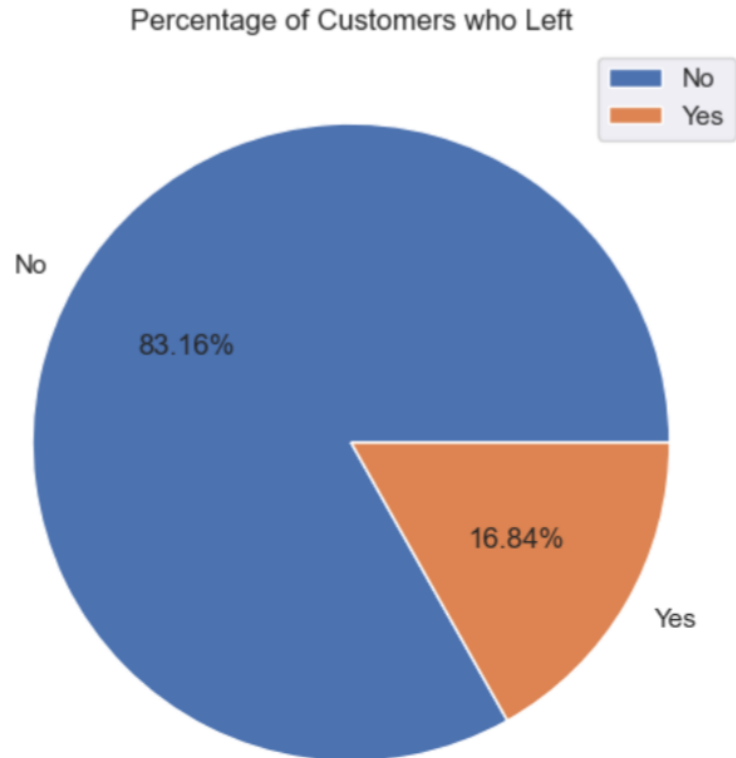
- The churn rate is **16.84%**, meaning a significant portion of customers are leaving.
- This highlights the need for effective **retention strategies** to maintain revenue stability.
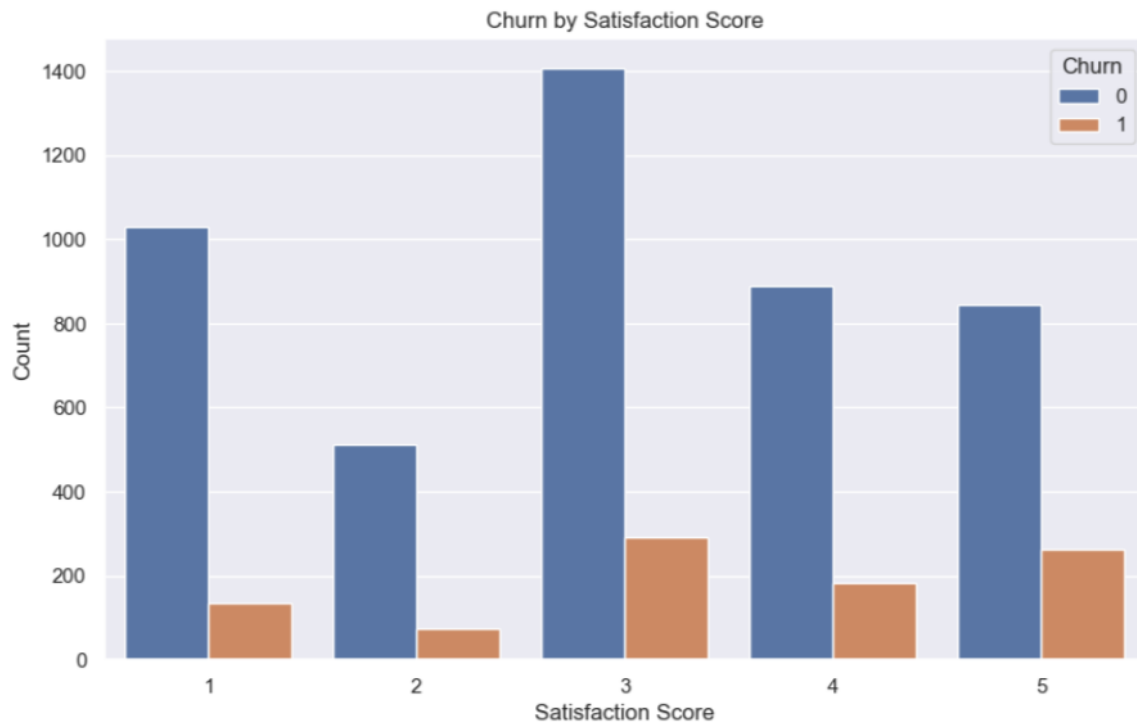
Churn by Tenure

Customer Churn by Tenure Range

Customer Churn by Tenure Range
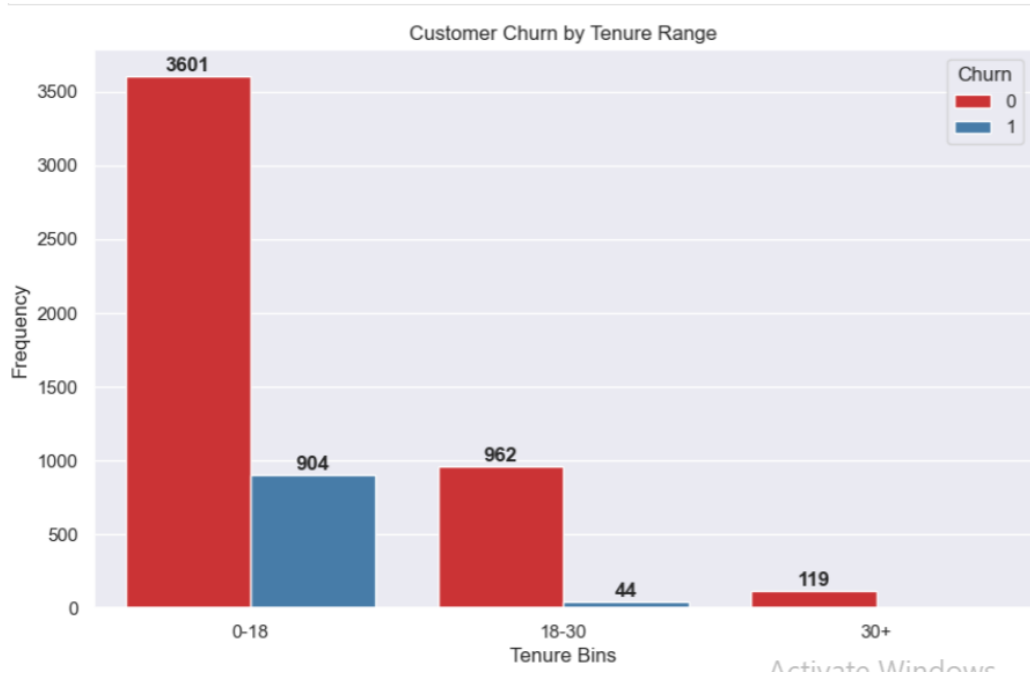
1. Customers with lower tenure have significantly higher churn rates.
2. This indicates the importance of early engagement and onboarding programs.
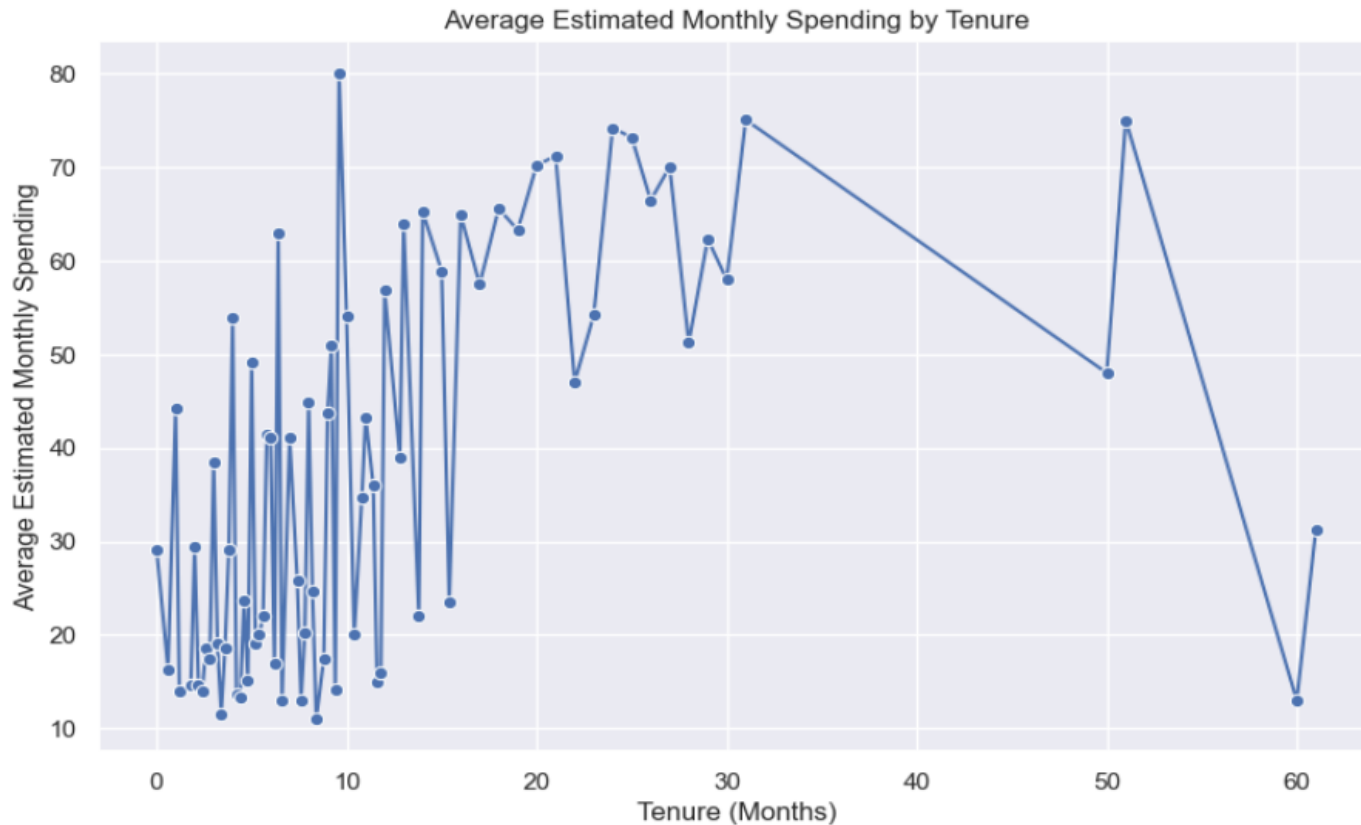
Churn by Satisfaction Score

Churn by Satisfaction Score

- Dissatisfied customers are much more likely to churn, highlighting the need for feedback systems and proactive support.

Monthly Spending by Tenure

Average Estimated Monthly Spending by Tenure

- Spending patterns increase with tenure, meaning retaining long-term customers directly impacts revenue growth.

Model Performance

Model	Accuracy	Precision (Churn=1)	Recall (Churn=1)	F1-Score
Logistic Regression	0.89	0.72	0.53	0.61
Decision Tree	0.95	0.82	0.87	0.84
Random Forest	0.97	0.96	0.84	0.89

Key Takeaway: **Random Forest** selected as final model for its**, high accuracy (97%)** and **balanced recall (84%).**

Conclusion

▶ **Churn Prediction Success:** Achieved **97% model accuracy** with Random Forest.

▶ **Key Finding:** Highest churn observed in the first 18 months — critical period for retention efforts.

▶ **Customer Insights:** Satisfaction score and spending behavior are strong indicators of churn risk.

▶ **Actionable Outcome:** Use churn predictions to personalize engagement, improve onboarding, and design loyalty programs.

▶ **Business Impact:** Enables proactive customer retention, potentially reducing churn by double digits.