Churn Modelling Analysis

(python)

Github : https://github.com/KumarBoste/Churn_Modelling_Analysis

Project Overview

Customer retention is a critical challenge for businesses in the banking and financial services sector. Acquiring new customers is often more expensive than retaining existing ones. This project focuses on analyzing customer behavior to identify key factors that influence customer churn using Python-based exploratory data analysis (EDA).

The dataset contains customer demographic details, financial attributes, product usage information, and churn status. Through data visualization and analysis, the project aims to uncover meaningful patterns that can help businesses reduce churn and improve customer engagement.

Dataset Description

The dataset consists of the following key columns:

- **Demographics**: Geography, Gender, Age

- **Financial Attributes**: CreditScore, Balance, EstimatedSalary

- **Customer Engagement**: Tenure, NumOfProducts, HasCrCard, IsActiveMember

- **Target Variable**: Exited (0 = Retained, 1 = Churned)

Columns such as RowNumber, CustomerId, and Surname were treated as identifiers and excluded from analytical modeling.

Project Objective

- To understand customer churn behavior using exploratory data analysis

- To identify high-risk customer segments

- To analyze the impact of demographic, financial, and engagement factors on churn

- To generate actionable insights that support customer retention strategies
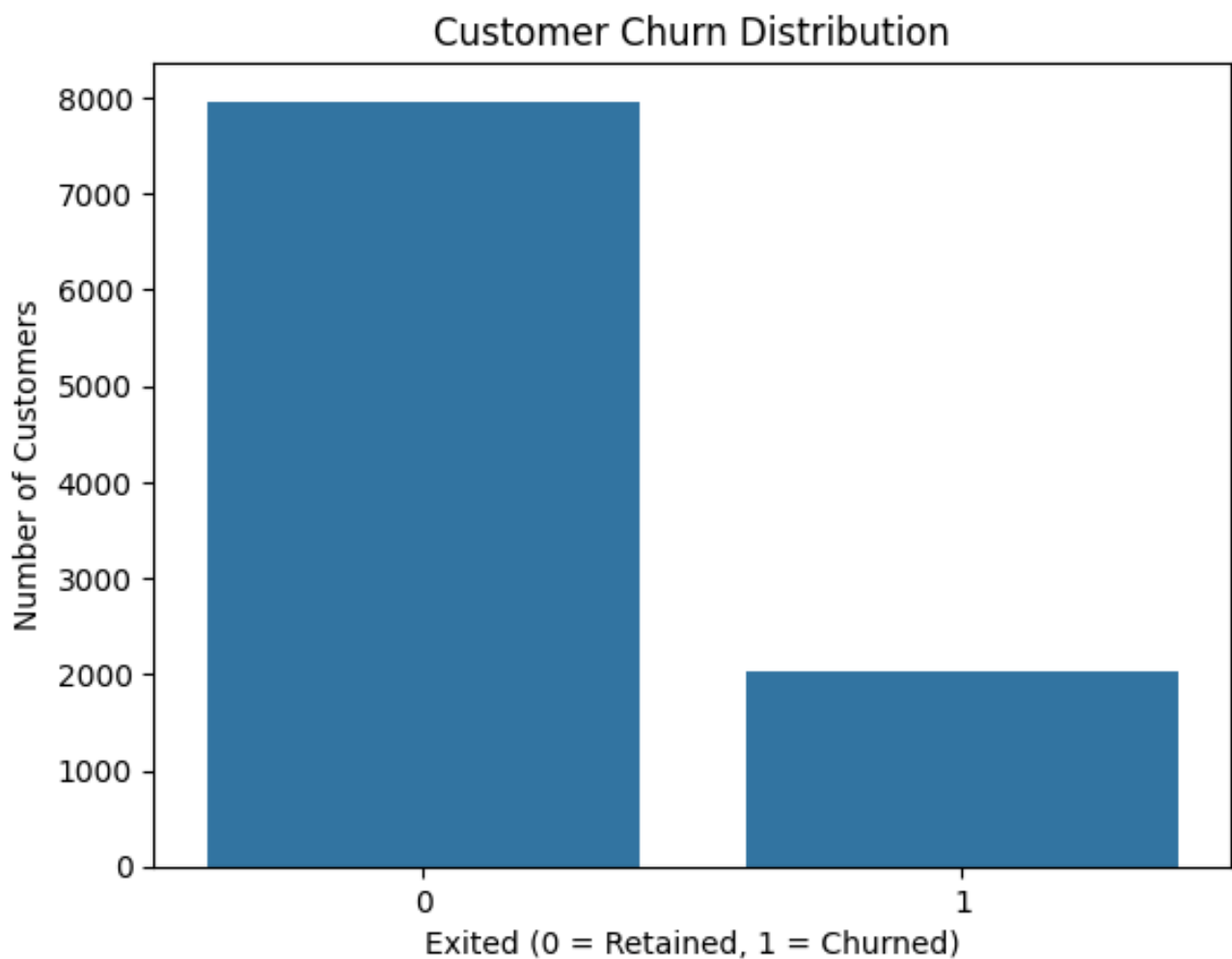
Problem Statement

The business is experiencing customer churn, leading to revenue loss and reduced customer lifetime value. However, the reasons behind customer churn are unclear. The challenge is to analyze historical customer data to determine:

- Which customer segments are more likely to churn

- What behavioral and financial factors contribute to churn

- How engagement and product usage affect customer retention

Data Visualization & Analysis
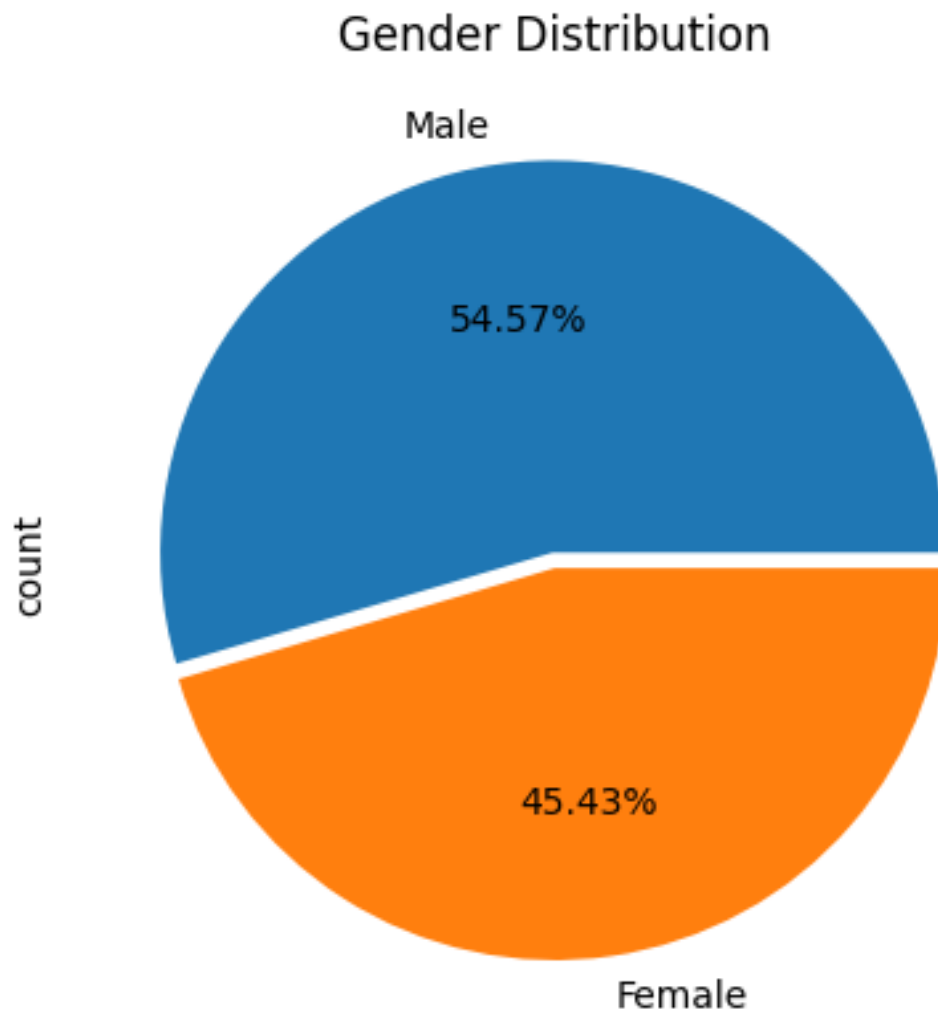
1. Customer Churn Distribution

The churn distribution analysis shows that while the majority of customers are retained, a significant portion has exited the bank. This indicates a clear need for churn prevention strategies.

Insight: Customer churn exists at a meaningful level and should be addressed proactively rather than reactively.
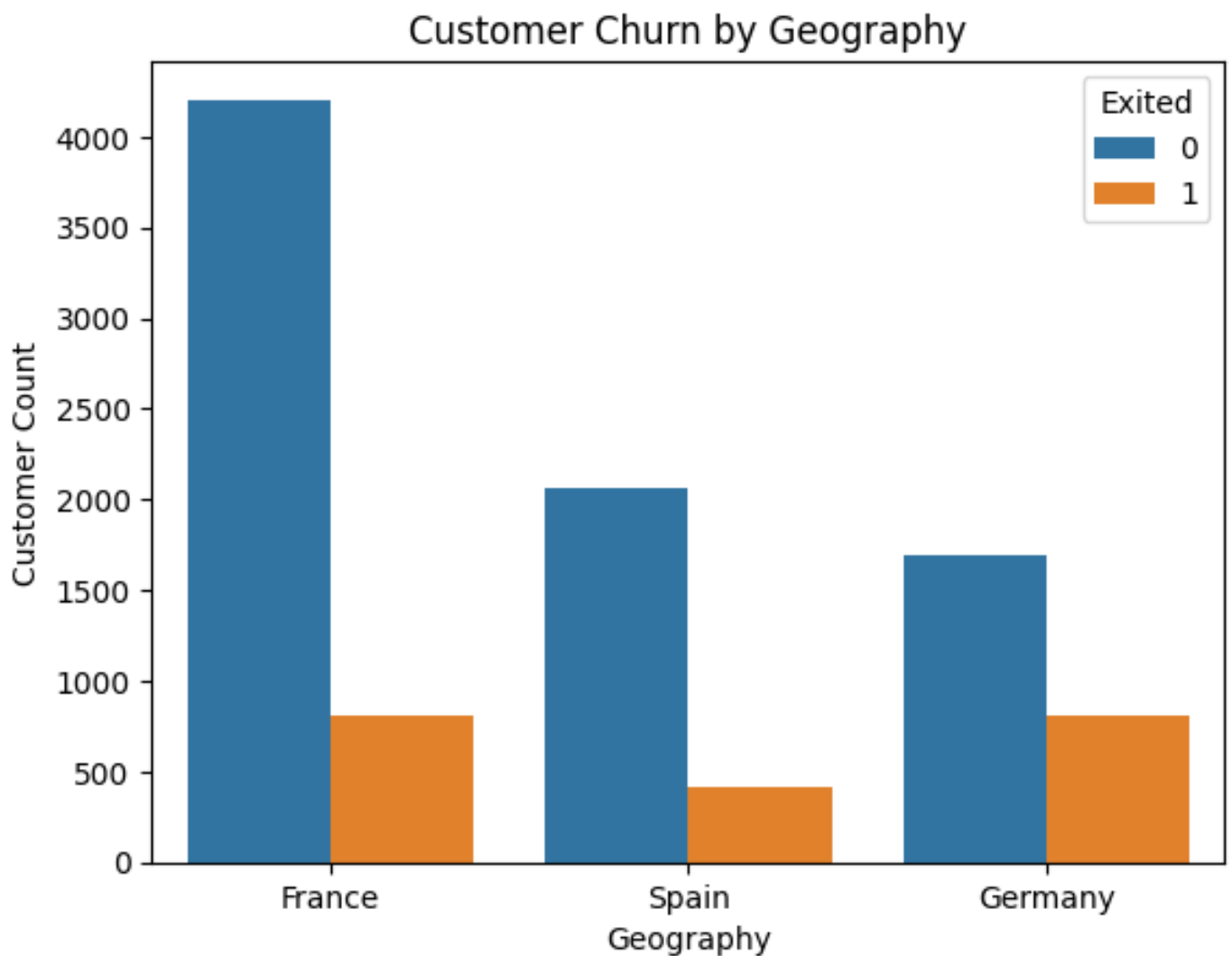
2. Gender Distribution

The near-even gender split suggests churn or engagement strategies should target both genders, rather than focusing on a single dominant group, as both represent significant portions of the customer base.

Gender Distribution

Male

54.57%

count

45.43%

Female

Insight: The customer base is slightly male-dominated, with males accounting for 54.57% and females 45.43%, indicating a relatively balanced gender distribution overall.
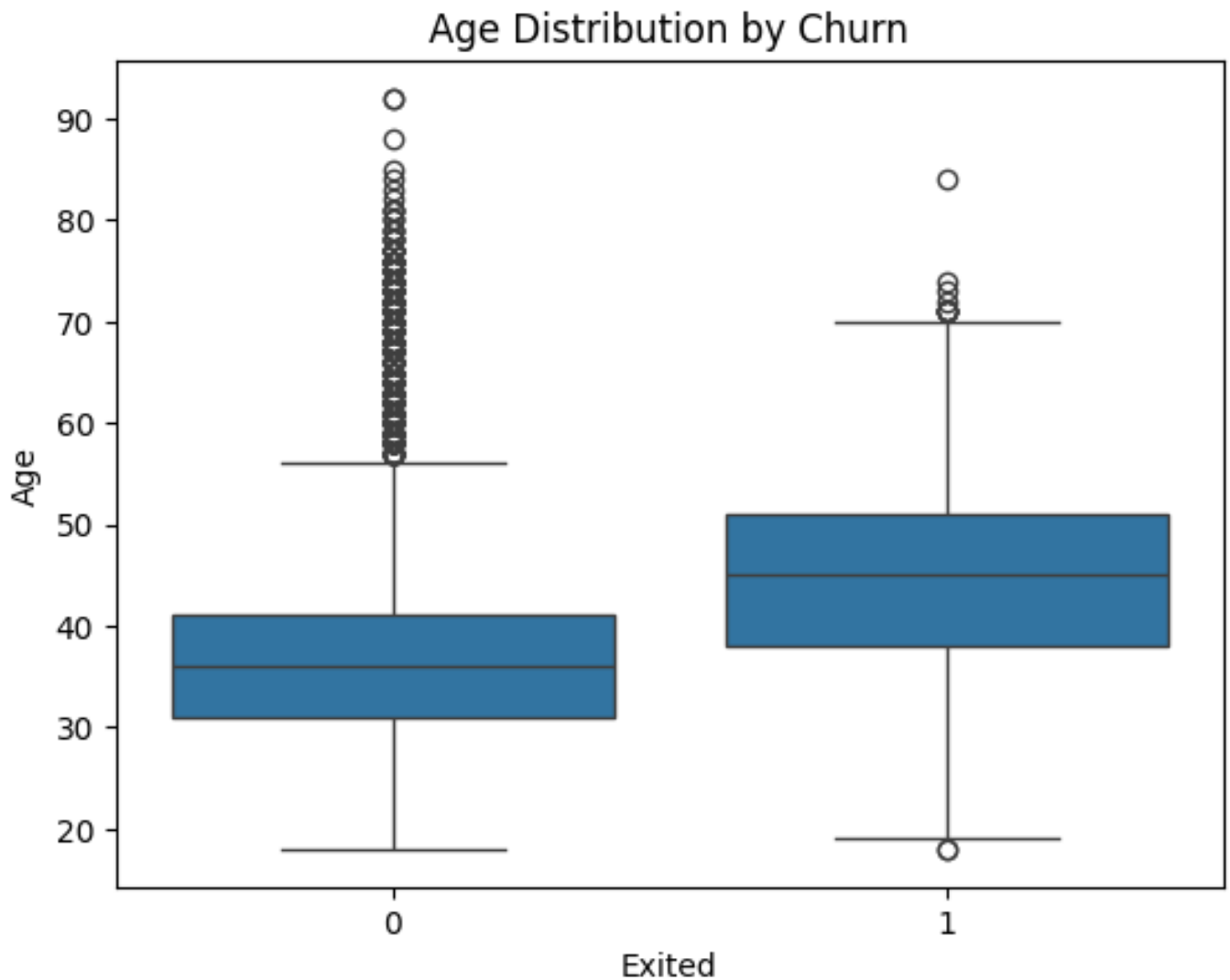
3. Churn by Geography

Different churn patterns were observed across geographical regions. Certain countries show a noticeably higher churn rate compared to others.

Customer Churn by Geography

Insight: Geography plays an important role in customer behavior. Region-specific strategies and localized offers may help reduce churn.
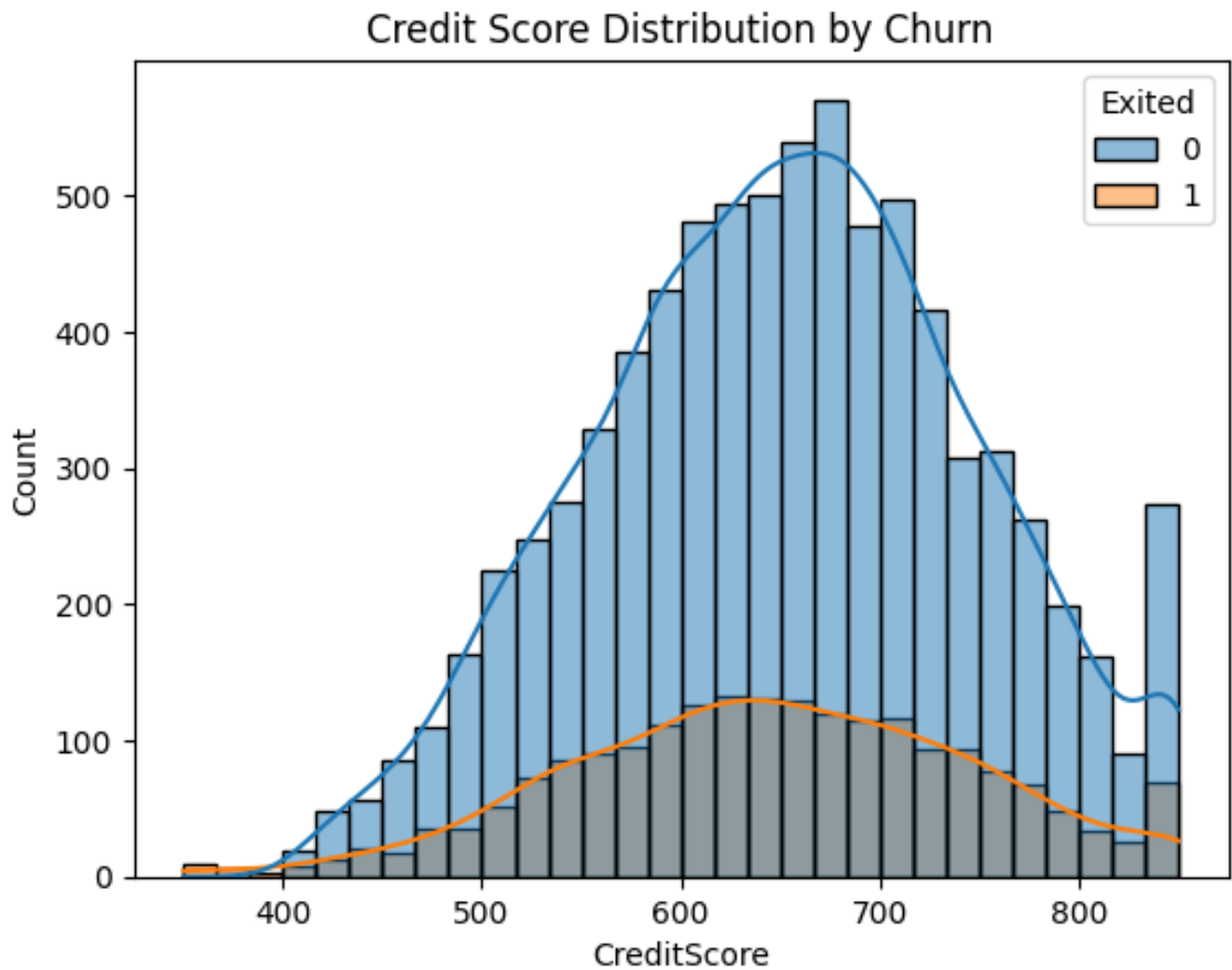
4. Age vs Churn

The age-based analysis reveals that churned customers tend to be older compared to retained customers.

Age Distribution by Churn

Insight: Middle-aged and senior customers are more likely to churn, possibly due to changing financial needs or dissatisfaction with services.
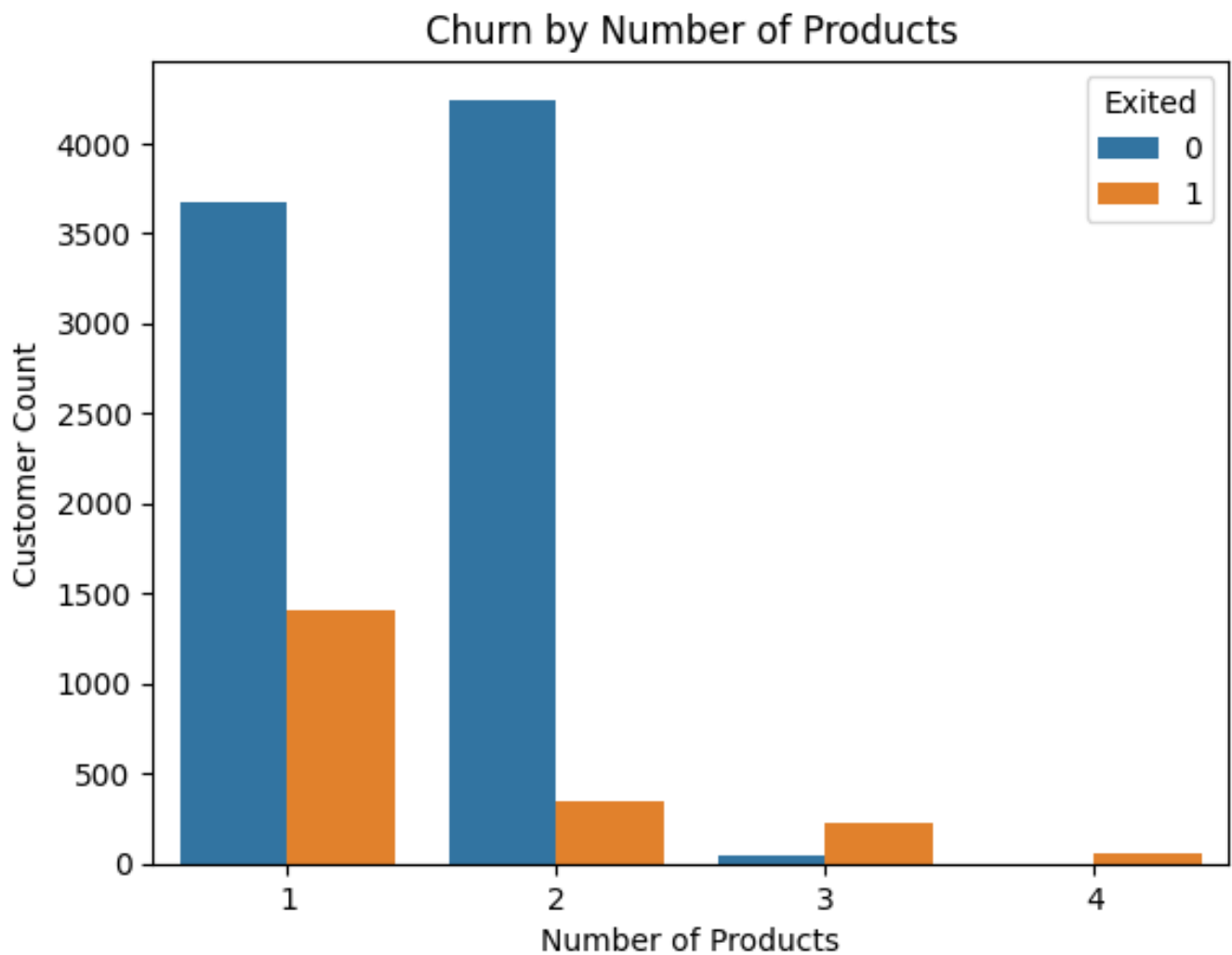
5. Credit Score vs Churn

Customers with lower credit scores were found to churn more frequently than those with higher credit scores.

Credit Score Distribution by Churn

Insight: Credit score is an important indicator of financial stability and trust. Customers with lower scores may require customized financial products or better support.

6. Number of Products vs Churn

Customers owning only one product show the highest churn rate, while customers with multiple products are more loyal.
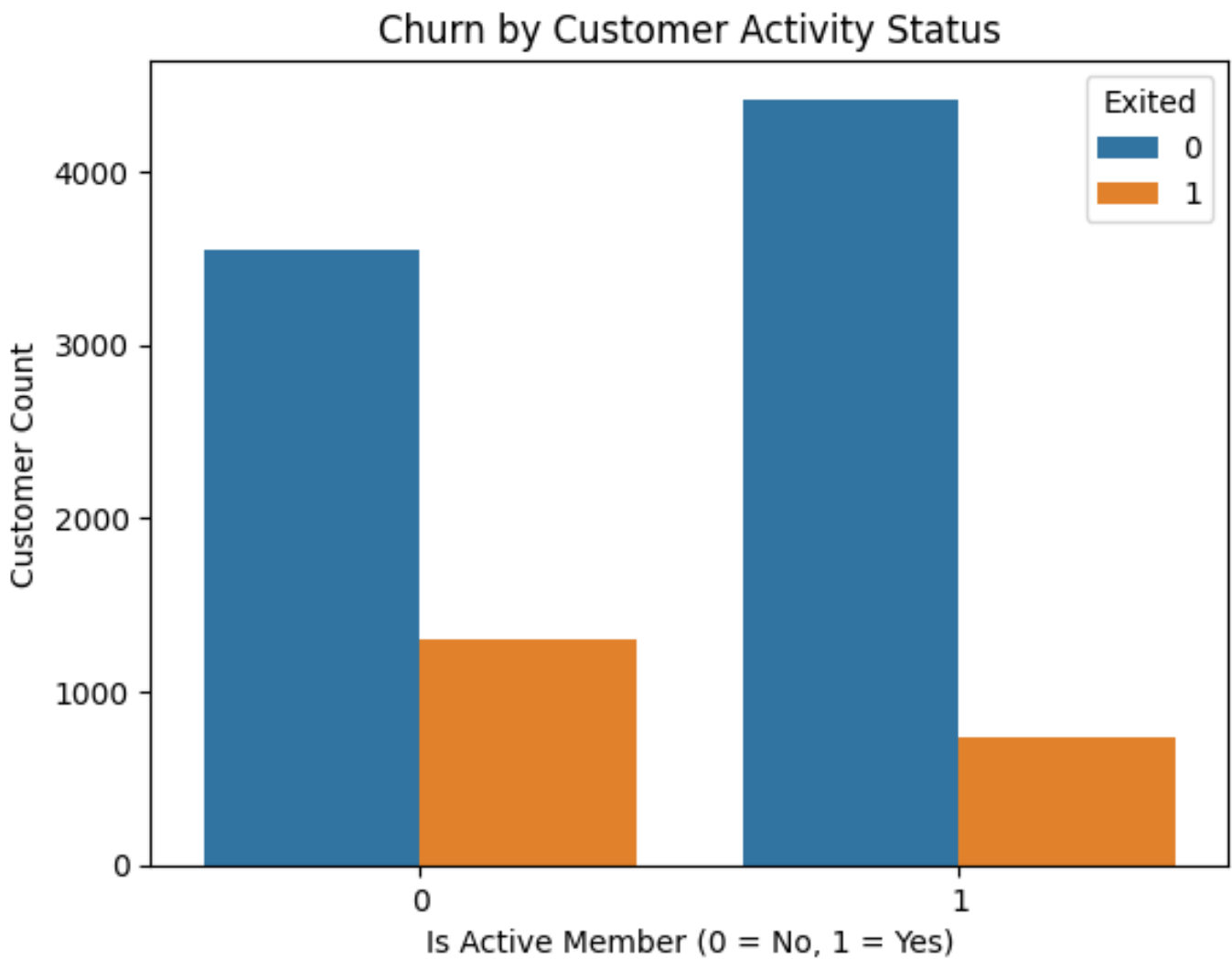

Churn by Number of Products

Insight: Cross-selling and encouraging customers to adopt multiple products can significantly reduce churn.

7. Activity Status vs Churn

Inactive members have a much higher churn rate compared to active members.

Churn by Customer Activity Status

Insight: Customer engagement is one of the strongest predictors of churn. Increasing customer activity through loyalty programs, notifications, and personalized communication can improve retention.

Key Insights Summary

- Inactive customers are at the highest risk of churn

- Customers with fewer products are more likely to leave

- Geography and age significantly influence churn behavior

- Low credit score customers show higher churn probability

- Engagement and product diversification reduce churn

Conclusion

This project demonstrates how exploratory data analysis using Python can uncover valuable insights into customer churn behavior. The analysis highlights the importance of customer engagement, product usage, and demographic factors in determining churn.

By identifying high-risk customer segments, businesses can take proactive measures to improve retention, enhance customer satisfaction, and reduce revenue loss.

Recommendations

- Launch engagement campaigns targeting inactive customers

- Implement cross-selling strategies to increase product adoption

- Develop region-specific retention plans

- Offer personalized financial products for low credit score customers

- Continuously monitor churn metrics using dashboards

Tools & Technologies Used

- Python

- Pandas

- Matplotlib

- Seaborn

- Jupyter Notebook

Final Note

This project can be extended further by building a machine learning model to predict customer churn and integrating the insights into a real-time business intelligence dashboard.