Customer Churn Analysis & Prediction Report

Introduction

This report summarizes the findings from the Customer Churn Analysis & Prediction study. The objective of this analysis is to understand factors contributing to customer churn and build a predictive model to identify at-risk customers.

Data Overview

The dataset contains 10,000 rows and 13 columns, including:

CustomerID: Unique identifier for each customer.

Age: Age of the customer.

Gender: Gender of the customer (Male/Female).

Subscription_Length_Months: Duration of the subscription in months.

Monthly_Charges: Monthly charges incurred by the customer.

Total_Charges: Total charges incurred by the customer.

Payment_Method: Method of payment (e.g., Bank Transfer, Debit Card).

Tenure: Duration of the customer's tenure.

Contract_Type: Type of contract (e.g., Month-to-Month, One-Year, Two-Year).

Support_Calls_Last_Year: Number of support calls made in the last year.

Streaming_Services_Used: Number of streaming services used.

Customer_Satisfaction_Score: Satisfaction score of the customer.

Churn: Whether the customer churned (Yes/No).

Data Cleaning & Preprocessing

The notebook begins with data cleaning and preprocessing steps:

Handling Missing Values: The dataset has no missing values, but the notebook includes code to fill missing values in numerical columns with the mean and categorical columns with the mode.

Outlier Detection: Outliers are detected using the Interquartile Range (IQR) method. The notebook identifies 104 outliers in the "Total_Charges" column.

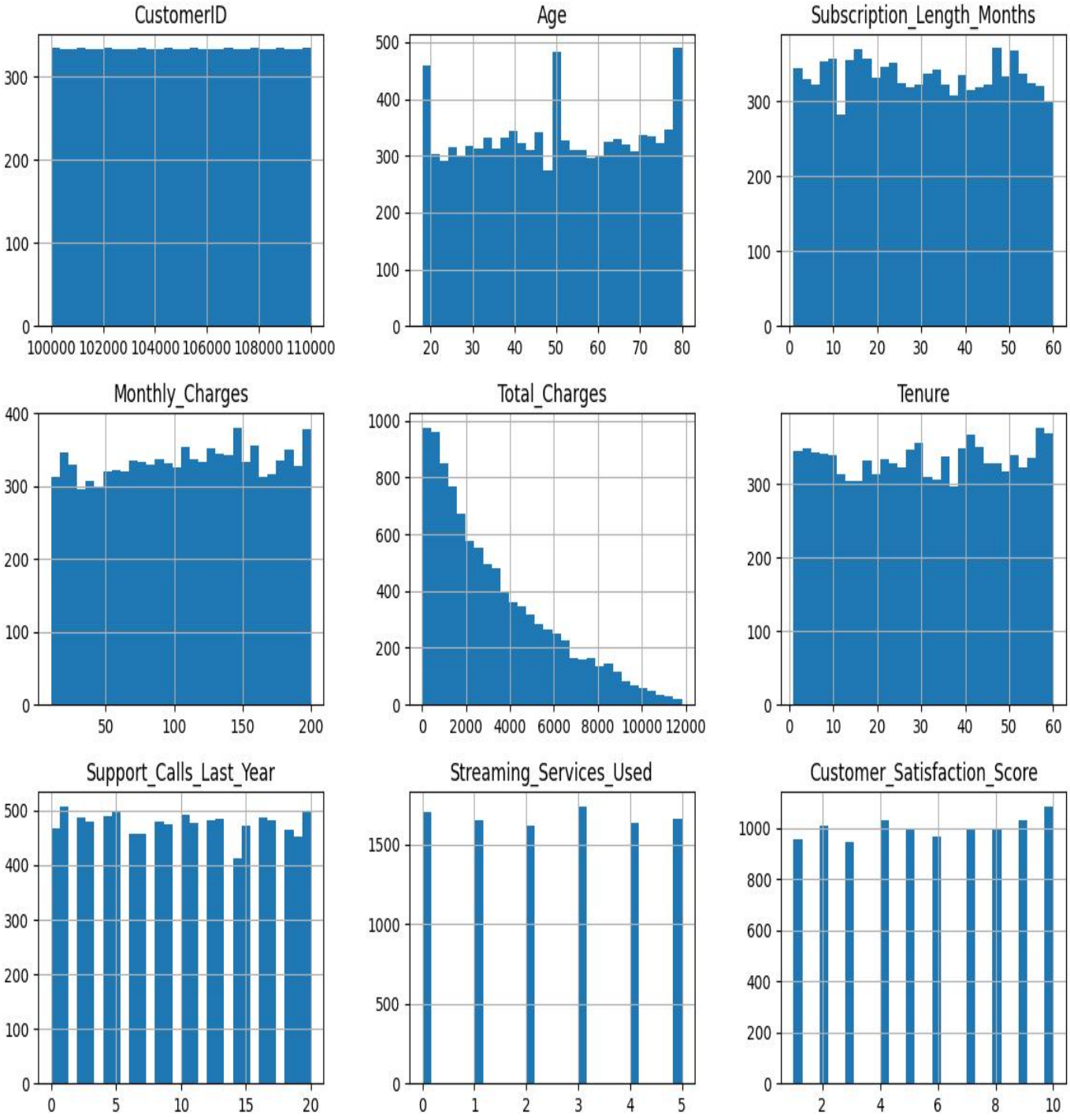
Standardization: Categorical data (e.g., "Gender") is standardized by converting to lowercase, and numerical data is checked to ensure it falls within expected ranges.

Exploratory Data Analysis (EDA)

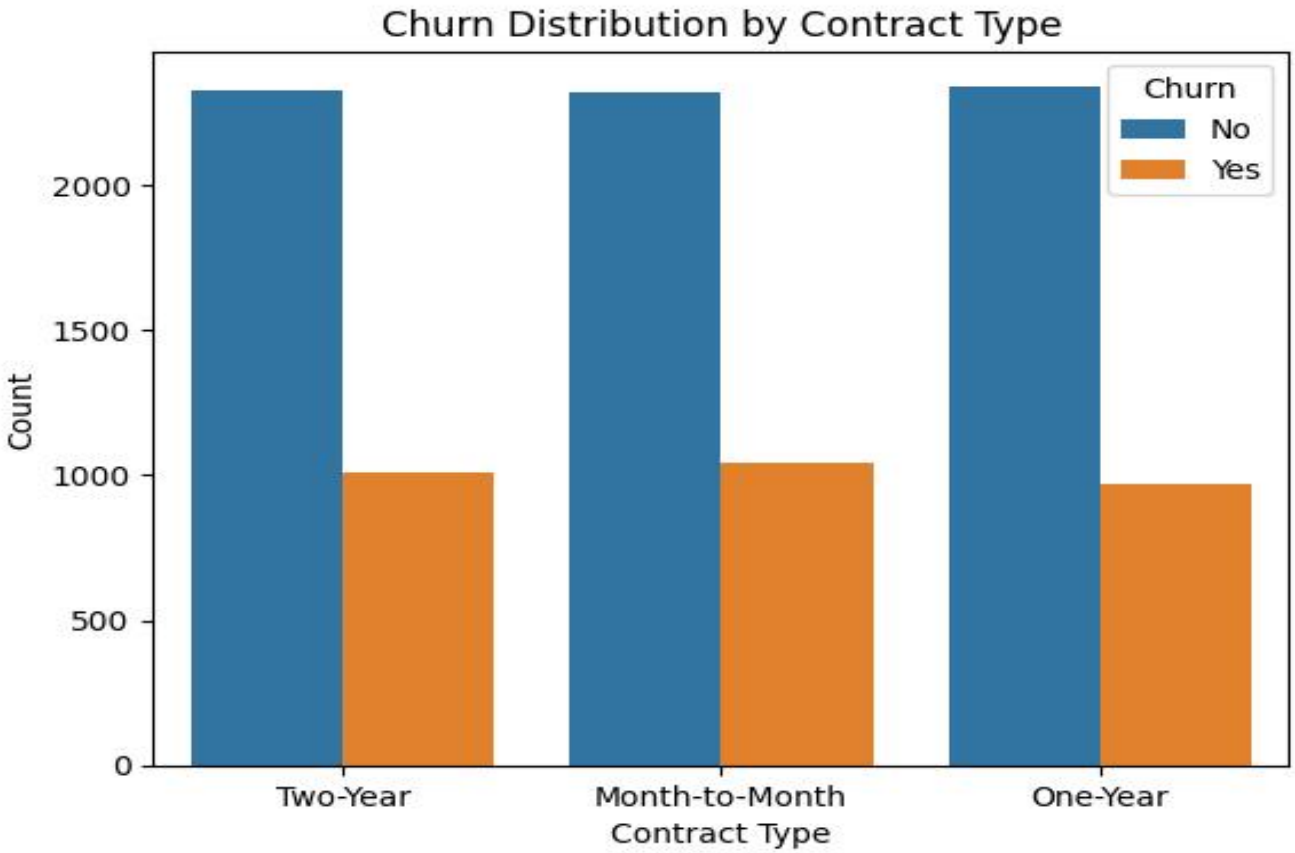
The EDA section provides insights into the dataset:

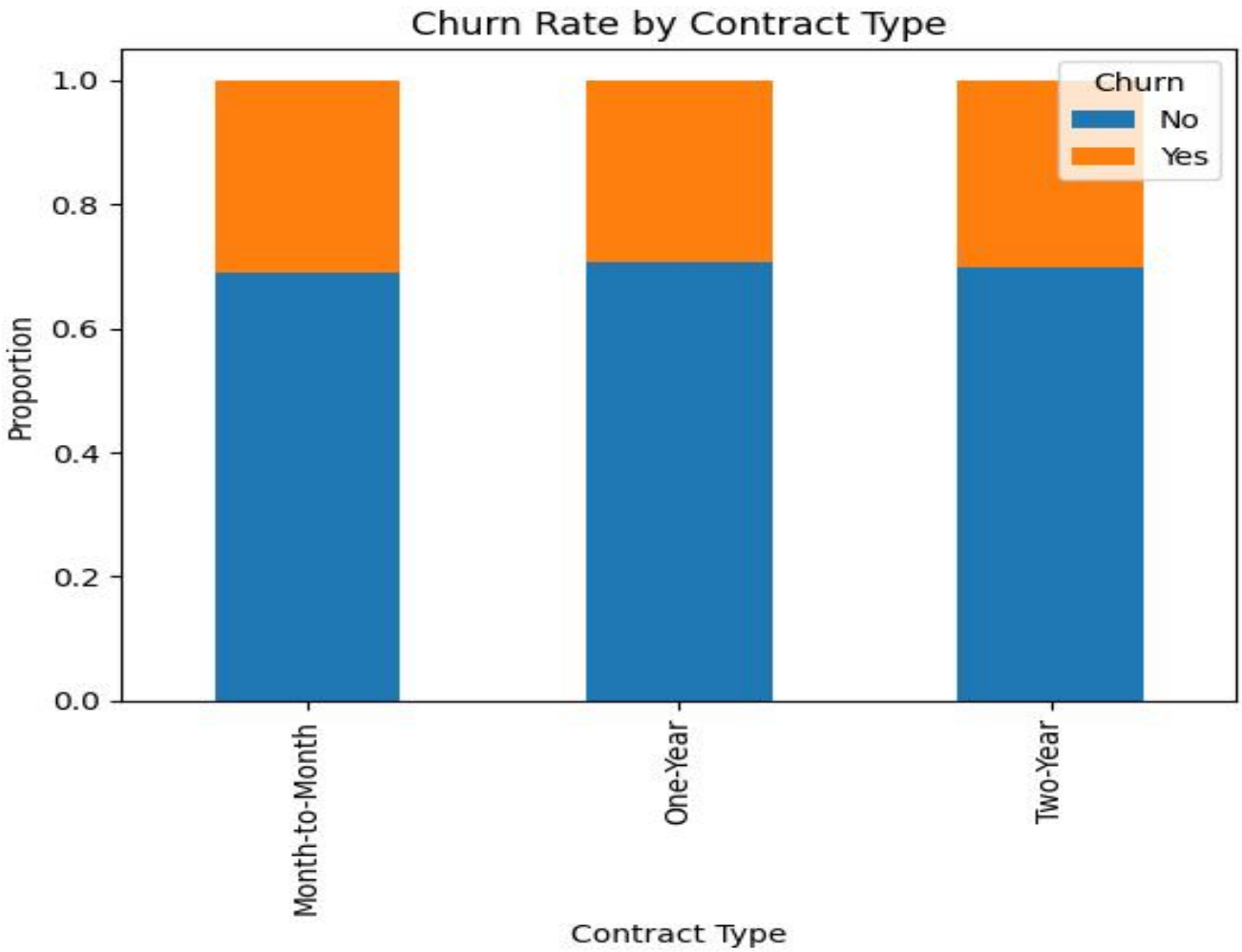
Summary Statistics: The notebook calculates summary statistics for numerical and categorical columns. For example, the average age of customers is 49.3 years, and the average monthly charge is $106.67.

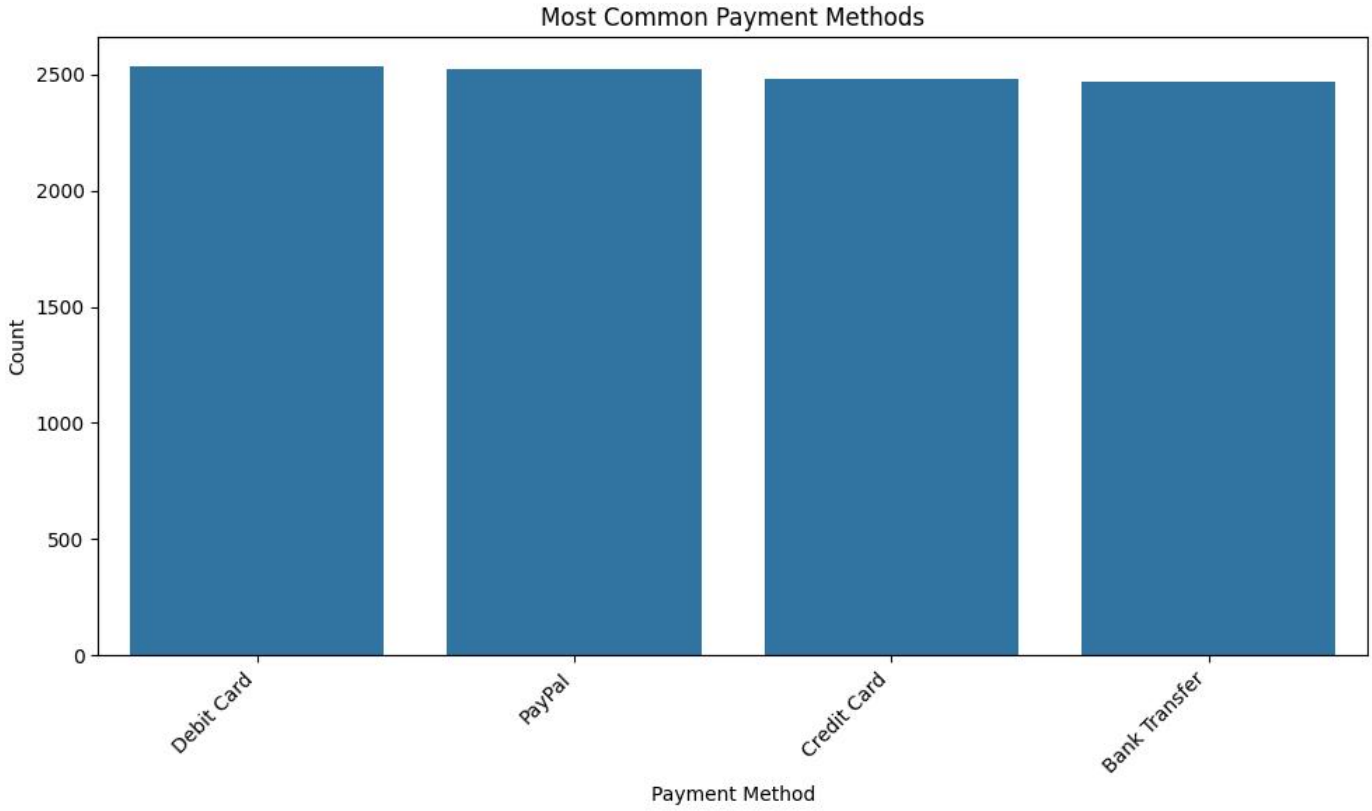
Data Distributions: Histograms are used to visualize the distributions of numerical columns, such as age, subscription length, and monthly charges.

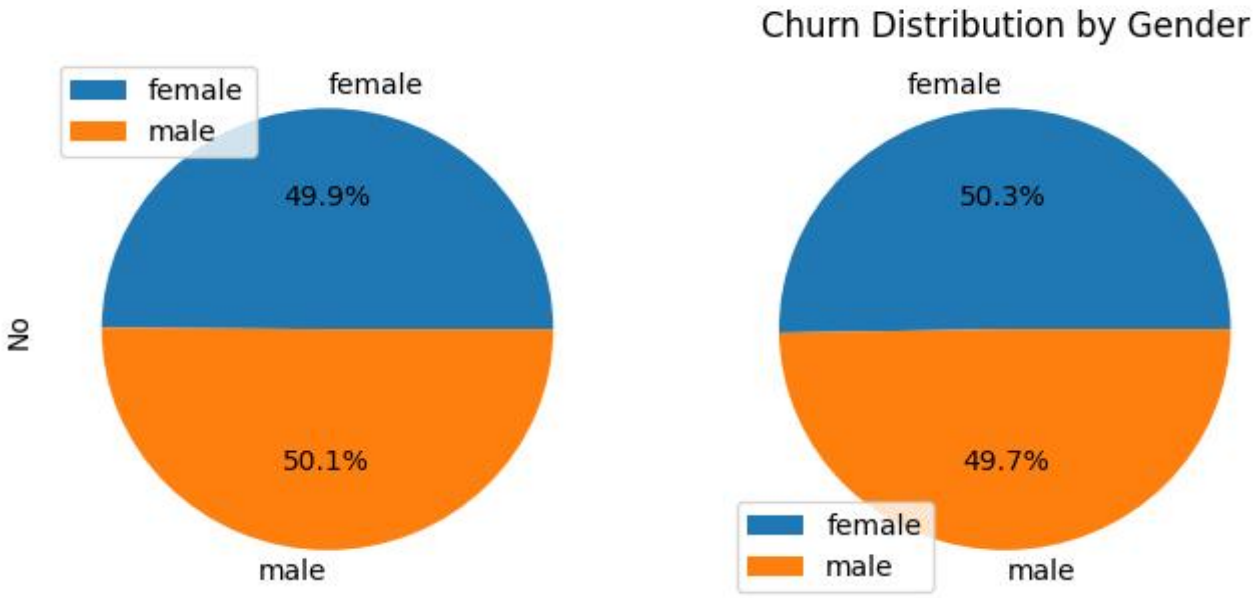


Categorical Data Analysis: The notebook summarizes categorical data, revealing that the most common payment method is "Debit Card," and the most common contract type is "Month-to- Month."









Predictive Modeling

Model Used: A classification model (e.g., Logistic Regression, Decision Tree and Random Forest, etc.) was trained.

Best Model - Random Forest Accuracy: 0.712

Performance Metrics:

Accuracy: 0.71**%** Precision: 0.51**%**

Recall: 0.71**%** F1 Score: 0.59**%**

The model successfully identified high-risk customers with reasonable accuracy.

Key Insights

Churn Rate: The dataset indicates that 30.16% of customers have churned (marked as "Yes" in the "Churn" column).

Customer Satisfaction: The average customer satisfaction score is 5.57, suggesting moderate satisfaction levels.

Support Calls: Customers who churned made an average of 9.94 support calls in the last year, compared to 7.23 for those who did not churn.

Monthly Charges: Customers who churned had slightly higher monthly charges (106.67)compared to those who did not churn(106.67)compared to those who did not churn(104.56).

Recommendations

Based on the analysis, the following recommendations are proposed to reduce customer churn:

Improve Customer Support: Since customers who churned made more support calls, improving customer support services could help retain customers.

Target High-Risk Customers: Customers with higher monthly charges and longer subscription lengths may be at higher risk of churn. Targeted retention strategies, such as personalized offers or discounts, could be effective.

Enhance Customer Satisfaction: Increasing customer satisfaction scores could reduce churn. This could be achieved through better service quality, more engaging content, or improved user experience.

Analyze Payment Methods: Since "Debit Card" is the most common payment method, offering incentives for using other payment methods might reduce churn.

Further Analysis: The notebook could be extended to include predictive modeling (e.g., logistic regression, random forest) to identify key factors influencing churn and predict future churn.

Conclusion

The analysis provides valuable insights into customer churn, highlighting key factors such as support calls, monthly charges, and customer satisfaction. By addressing these factors, the company can implement targeted strategies to reduce churn and improve customer retention. Further predictive modeling could enhance the understanding of churn and support more effective decision-making.