TELCOM CHURN ANALYSIS

IEC Portfolio Project 2

Sidra Nasir

**Introduction**

Customer churn is a significant issue for many businesses, especially in industries like telecommunications, where retaining customers is essential for profitability. This project aims to analyze customer churn using a dataset on Telcom Churn Analysis. The goal is to identify key factors contributing to customer churn and build a predictive model to determine which customers are likely to churn.

**Data Wrangling and EDA**

In the EDA phase, I perform data exploration and gain initial insights into the dataset.

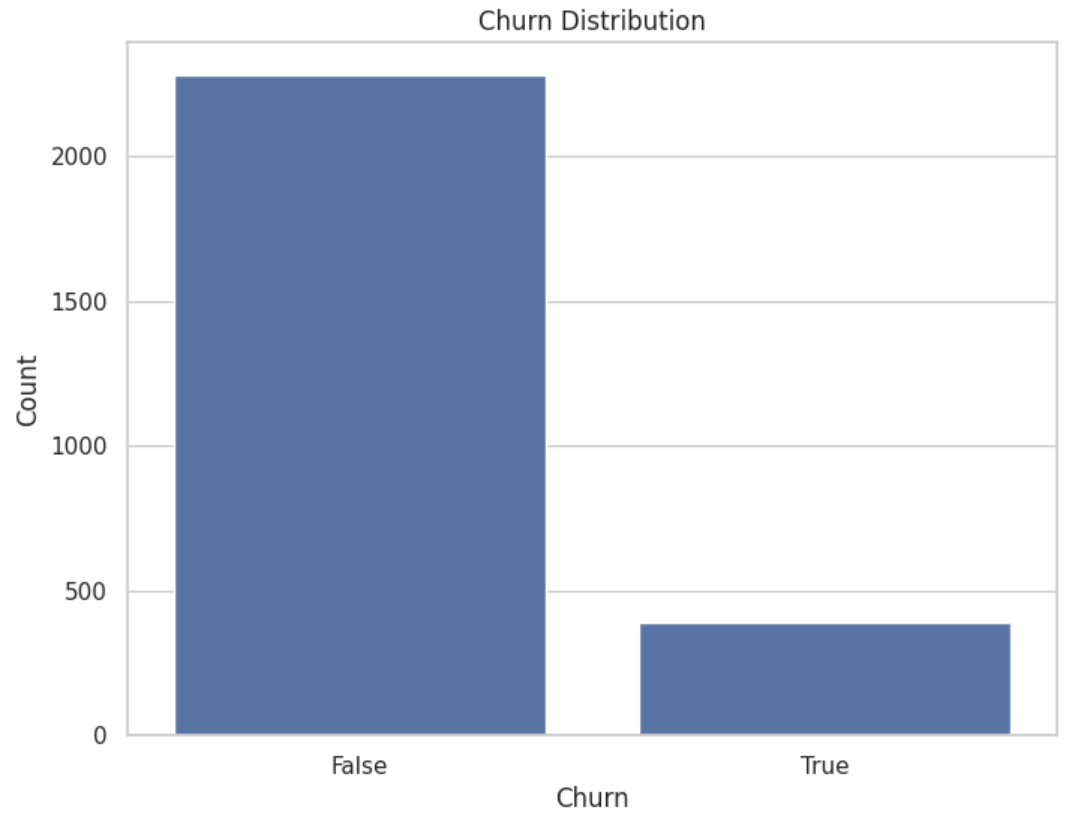
* **Importing Libraries**: Essential libraries like Pandas, NumPy, Matplotlib, and Seaborn were imported for data analysis and visualization.
* **Loading the Dataset**: The dataset was loaded into a Pandas Data Frame.
* **General Information**: The dataset contains 2666 rows and 20 columns, with a mix of numeric and categorical data.
* **Checking for Null Values**: The dataset was checked for null values, and none were found.
* **Data frame Shape and Column Types**: The dataset's shape and column types were examined to ensure consistency.
* **Summary Statistics**: Summary statistics were generated to understand the distribution of numeric variables.
* **Features:** I print the column names of the Data Frame.
* **Churn Counts:** I count the number of customers who churned and did not churn.
* **Churn Count Plot**: A count plot is created to visualize the churn distribution.
* **Box Plot:** A box plot is created to analyze the relationship between churn and account length.
* **Histogram distribution:** Histograms are created to visualize the distribution of various features in the dataset.
* **Checking for Correlation**: A correlation matrix is generated to identify highly correlated features.

**Problems**

**Question 1**

**Count the number of customers who churned and did not churn.**

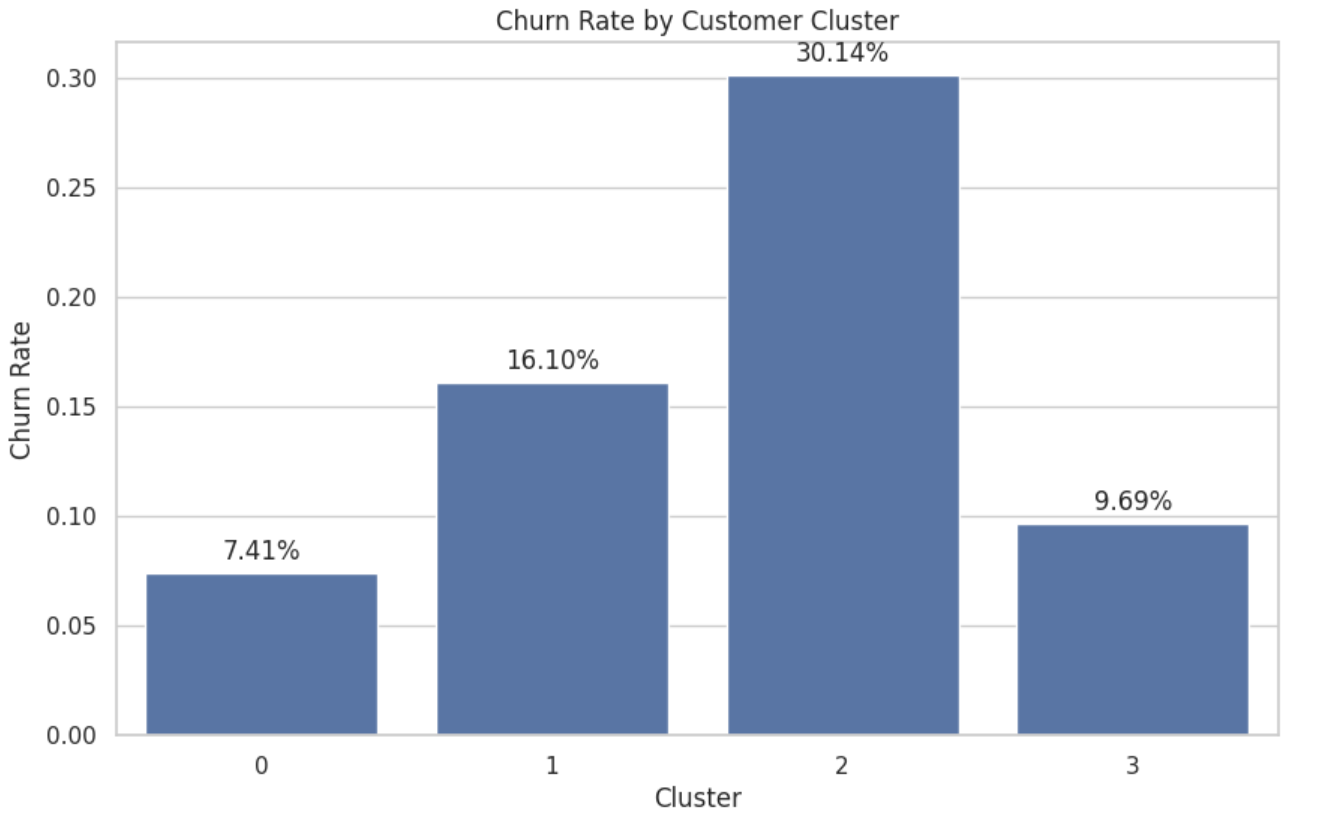
The churn count shows that 388 customers have churned, while 2,278 have stayed.



**Question 2**

**Identify distinct customer segments based on usage patterns to understand which segments are most likely to churn.**

**Interpretation**: The clustering analysis revealed four customer segments with varying churn rates, with Cluster 2 showing the highest risk of churn at 30.14%.



**Question 3**

**Solution**: **Segment the customer lifetime value on the basis of churn.**

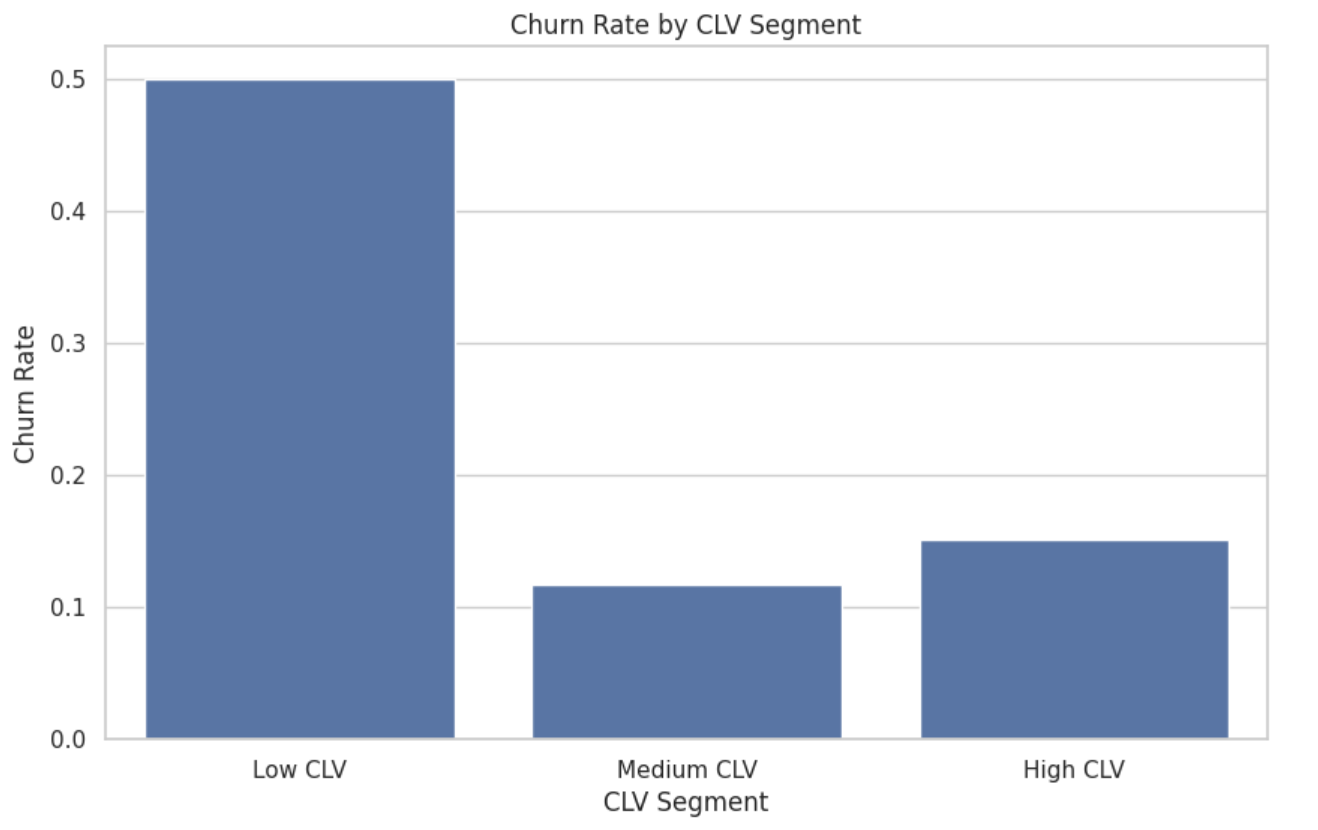
CLV Segment

Low CLV 0.500000

Medium CLV 0.116803

High CLV 0.151654

**Interpretation**: This indicates that customers with **low CLV** have a significantly higher churn rate compared to those with medium or high CLV.

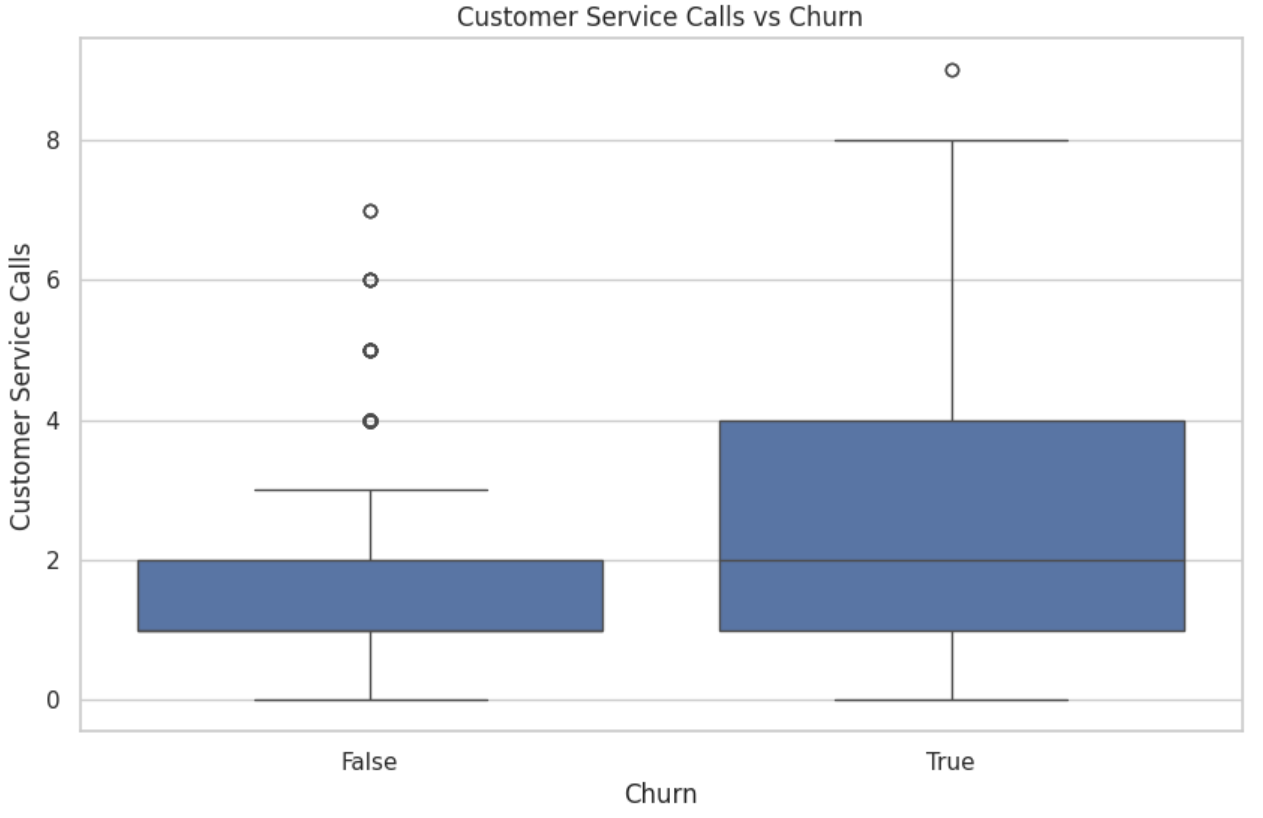


**Question 4**

**Analyze the impact of customer support interactions on the likelihood of churn.**

**Solution**: The correlation between customer service calls and churn is 0.20, indicating a positive but moderate relationship.

**Interpretation**: This suggests that customers who make more service calls are somewhat more likely to churn.



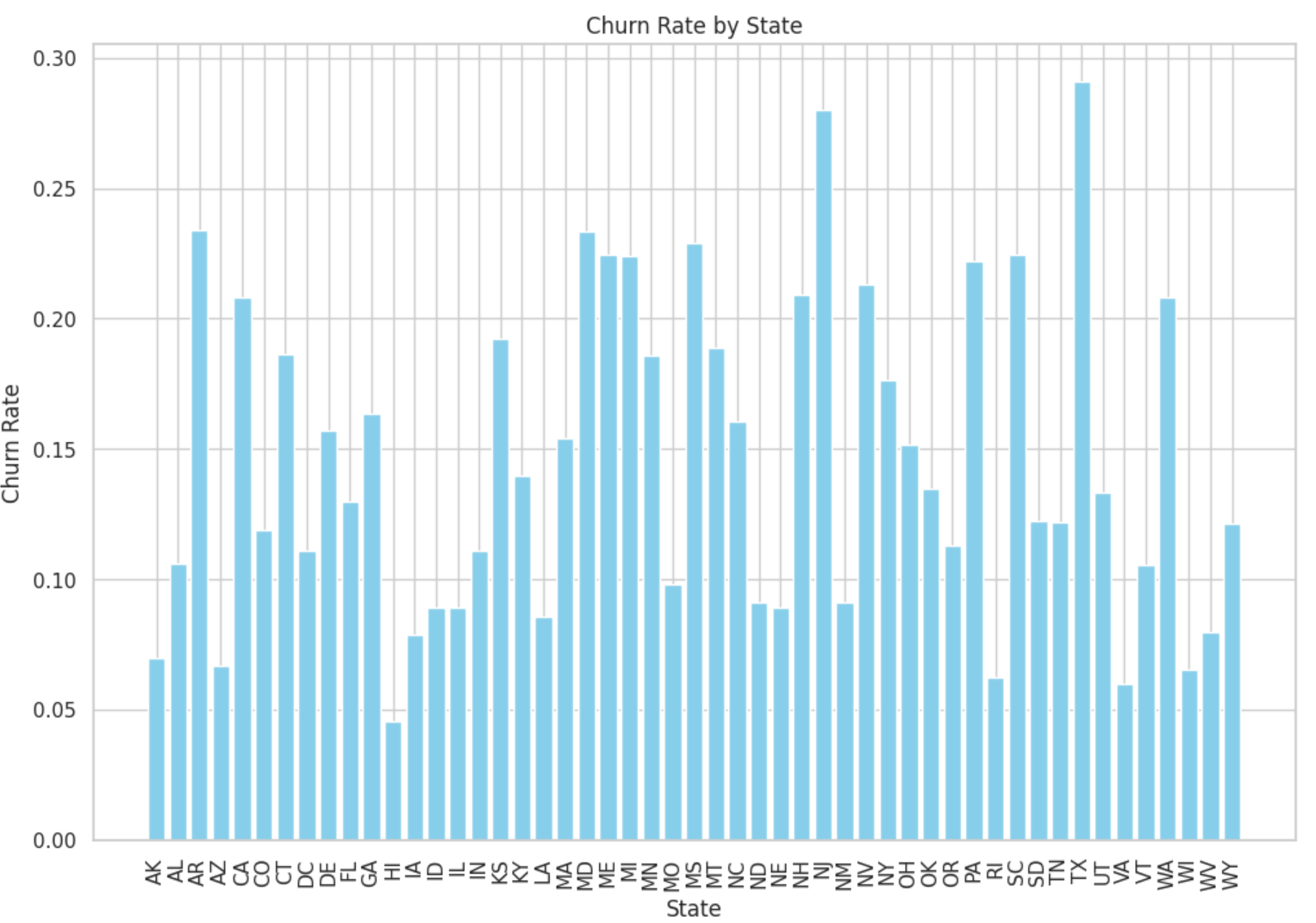
**Question 5**

**Investigate whether customers in specific geographical regions are more likely to churn.**

**Solution**: State with the highest churn rate is TX (29.09%)

State with the lowest churn rate is HI (4.55%)

**Interpretation**: The analysis reveals higher churn rates in certain states like TX, NJ, MN, and NV, indicating a potential geographical impact on customer retention.

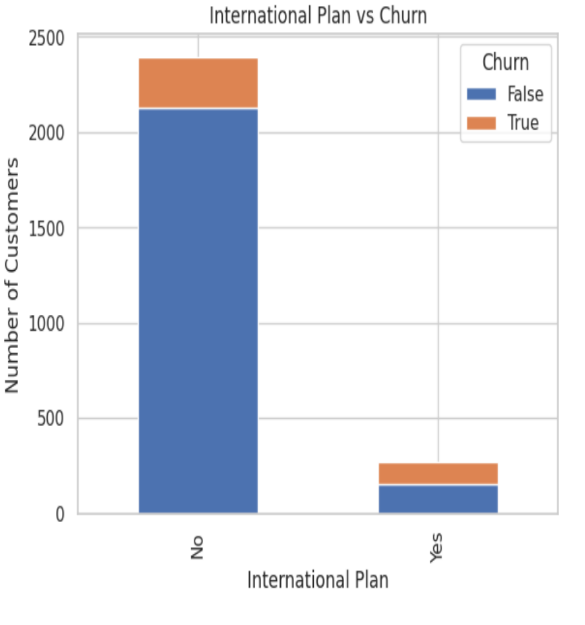


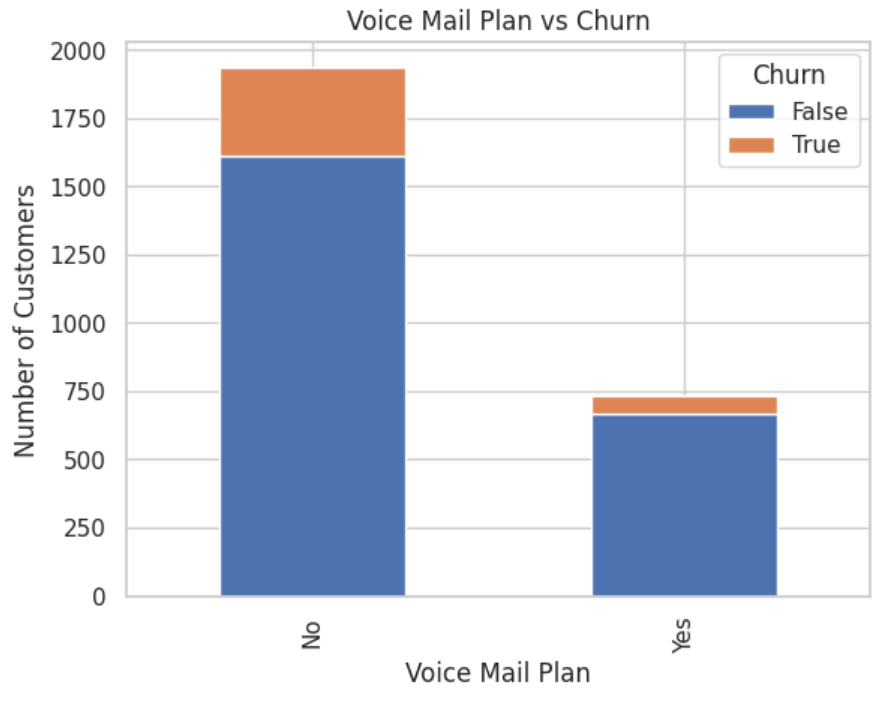
**Question 6**

**Determine if certain plans are associated with higher churn rates.**

**Interpretation**:

* **International Plan**: Customers with an international plan show a higher proportion of churn (118 out of 270) compared to those without (270 out of 2126). This indicates that having an international plan is associated with a higher likelihood of churn.
* **Voice Mail Plan**: Customers with a voice mail plan have a lower churn rate (65 out of 668) compared to those without (323 out of 1610), suggesting that a voice mail plan is associated with lower churn.



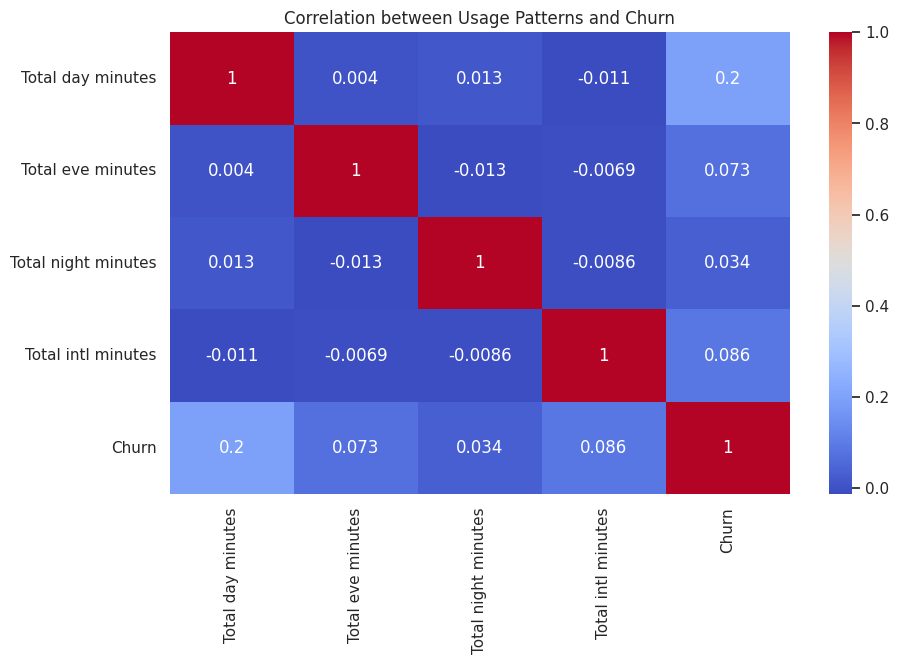


**Question 7**

**Explore whether specific usage patterns (call minutes) are correlated with higher churn.**

The correlation matrix shows:

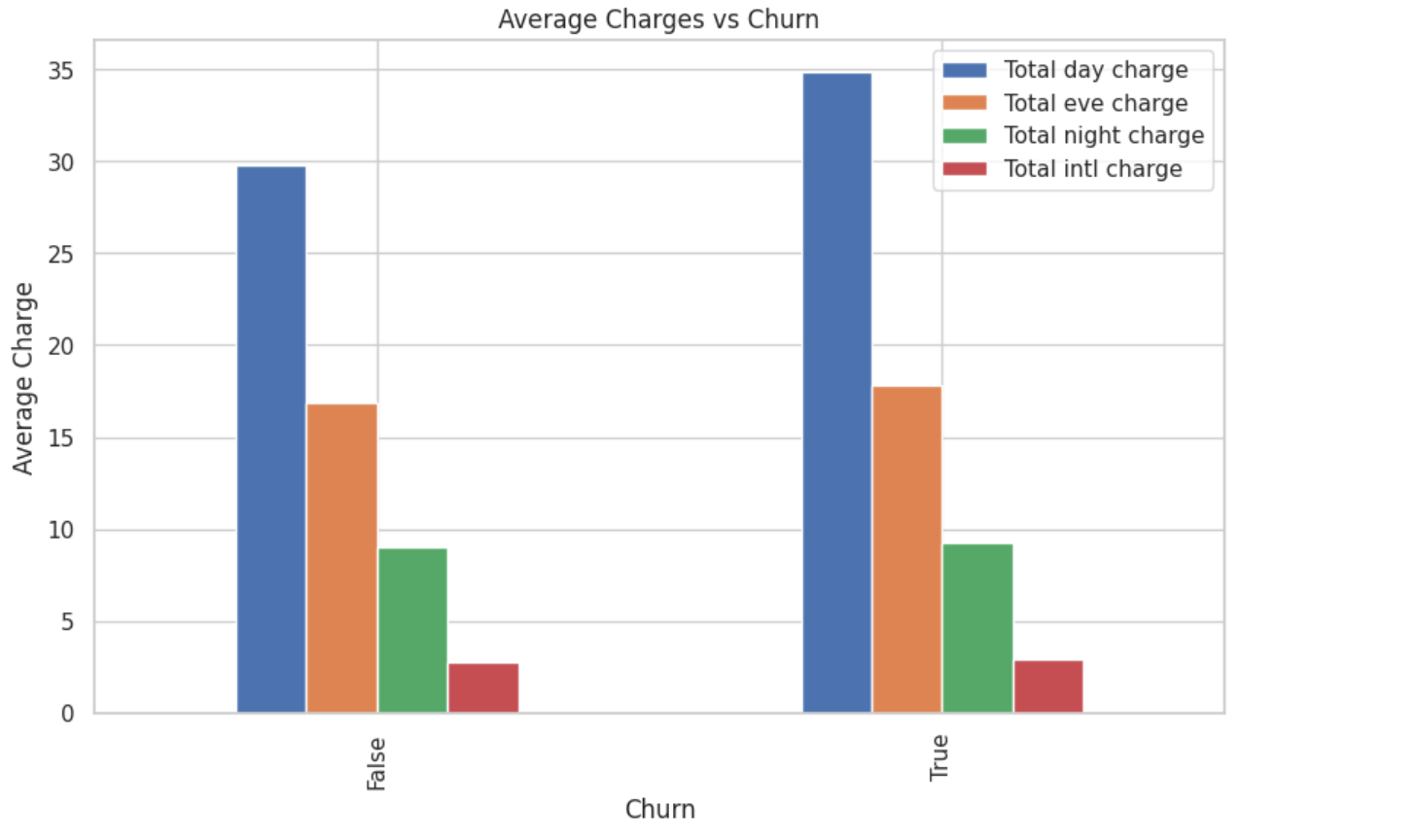
* **Total Day Minutes**: Has the highest correlation with churn at 0.196, suggesting a modest positive relationship—higher usage of daytime minutes is associated with slightly higher churn.
* **Total Eve Minutes**: Shows a minimal correlation with churn (0.073), indicating a weak relationship.
* **Total Night Minutes and Total Intl Minutes**: Have very low correlations with churn (0.034 and 0.086, respectively), suggesting minimal impact on churn.



**Question 8**

**Analyze how pricing strategies (call charges) affect customer churn.**

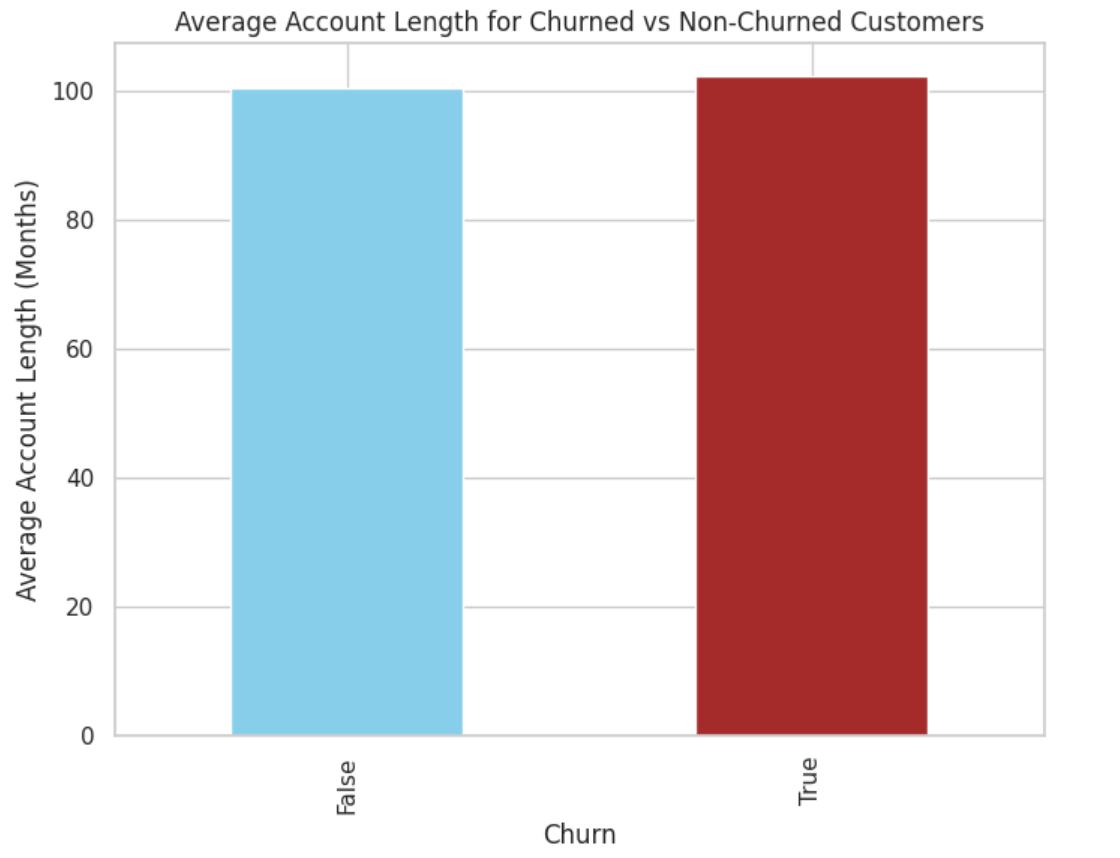
Customers who churn have higher average day charges ($34.88) compared to non-churners ($29.77), suggesting that higher daytime usage charges are linked to churn. The impact on evening, night, and international charges is minimal.



**Question 9**

**Evaluate if the length of a customer's contract influences their likelihood to churn.**

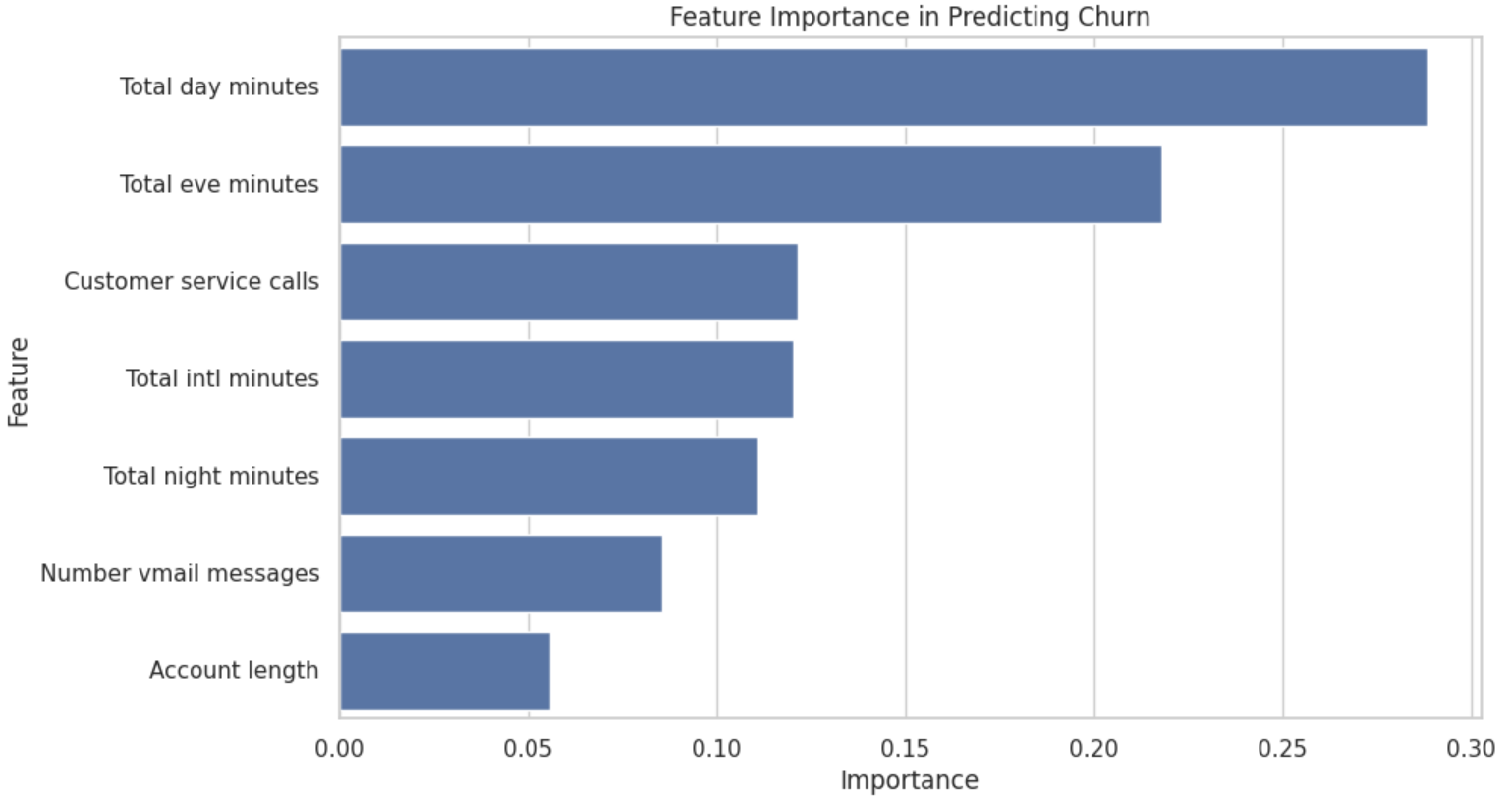
Customers who churned have a slightly higher average account length (102.32 months) compared to non-churned customers (100.33 months). This suggests that, on average, longer-term customers are slightly more likely to churn.



**Question 10**

**Identify key factors in the customer experience that lead to churn**.

The model identifies **Total Day Minutes** (28.81%) as the most important feature in predicting churn, followed by **Total Eve Minutes** (21.78%), and **Customer Service Calls** (12.14%). **Account Length** is the least important feature with only 5.57% importance.



**Recommendations:**

* Develop targeted retention strategies for high-risk customer segments, especially those in Cluster 2 with the highest churn rate.
* Refine CLV estimates with additional features and target high-CLV customers with personalized offers while focusing on retaining predicted churners.
* Enhance customer service quality to reduce churn, especially for customers who make frequent service calls.
* Implement region-specific retention strategies and promotions in states with higher churn rates, such as TX, NJ, and MN.
* Review and adjust international plan offerings to address higher churn rates, and promote voice mail plans as they correlate with lower churn.
* Focus on reducing churn by addressing issues related to high daytime usage, as it shows the strongest correlation with churn.
* Reevaluate pricing strategies, particularly for day charges, and consider offering discounts or incentives to mitigate churn among high-charge customers.
* Investigate why longer-term customers are churning and implement retention programs specifically for these customers.
* Prioritize features with high importance in predicting churn, such as day and evening minutes, and enhance customer service to improve overall retention.