

PYTHON DATA ANALYSIS PROJECT



pwc

# Telecom Customer Churn Analysis

PwC Consulting Case Study

PREPARED BY

TAARA NGUYEN

CODE LINK

<https://github.com/taaranguyen/PwC-Customer-Churn-Analysis.git>



# BACKGROUND

PwC's Data & Analytics Consulting Division has been engaged by a prominent telecom company grappling with a high rate of customer churn. The company has enlisted PwC to analyze churn patterns, identify the key drivers of customer attrition, and propose data-driven strategies to reduce churn and improve customer retention. The goal of this analysis is to extract critical insights that will lead to actionable recommendations, ultimately reducing customer turnover and enhancing the company's retention efforts.

The telecom industry is highly competitive, with customers frequently switching providers in search of better pricing, superior service quality, or an improved overall experience. As a result, the company is experiencing significant monthly customer loss, which is negatively impacting its revenue and profitability. The CEO is seeking answers to several pivotal questions, including customer segments, underlying factors and strategies on predicting and preventing churn.

# EXECUTIVE SUMMARY

**Objective:** The task involves conducting an exploratory data analysis (EDA) on customer demographics and behaviors, followed by the development of a churn prediction model. This model will identify the key risk factors contributing to customer churn, providing valuable insights to inform retention strategies and reduce customer attrition.

## Key insights:

- Phase 1 - Customer Base Summary:
  - Churn percentage.
  - Churn Distribution in Numerical Variables (Num Tech Tickets, Total Charges and others).
  - Customer segments Analysis (Age groups, Gender, Tenure).
  - Breakdown of customers using Internet, Phone, TV services.
  - Billing Methods Investigation (Paperless Billing or Traditional methods).
  - Payment Method Analysis (Credit card, Bank transfer and others).
  - Contract Type Breakdown (Monthly, Annual and Two-year basis).
  - Feature Correlation
- Phase 2 - Customer Churn Prediction Model:
  - Model's Results (Logistic Regression, Decision Tree, Random Forest and Support Vector Machine)
  - Feature Importance Analysis: Top factors leading to churn
- Recommendation:
  - Improve technical support and service quality
  - Offer long-term contracts
  - Ensure quality and reliability of fiber optic internet service
  - Enhance customer engagement and experience for new customers issue

# CHURN ANALYSIS

## 1. CHURN PERCENTAGE

Based on the provided dataset, which includes service information for 7,043 customers, the overall churn rate has been calculated to be 26.58% (Figure 1).

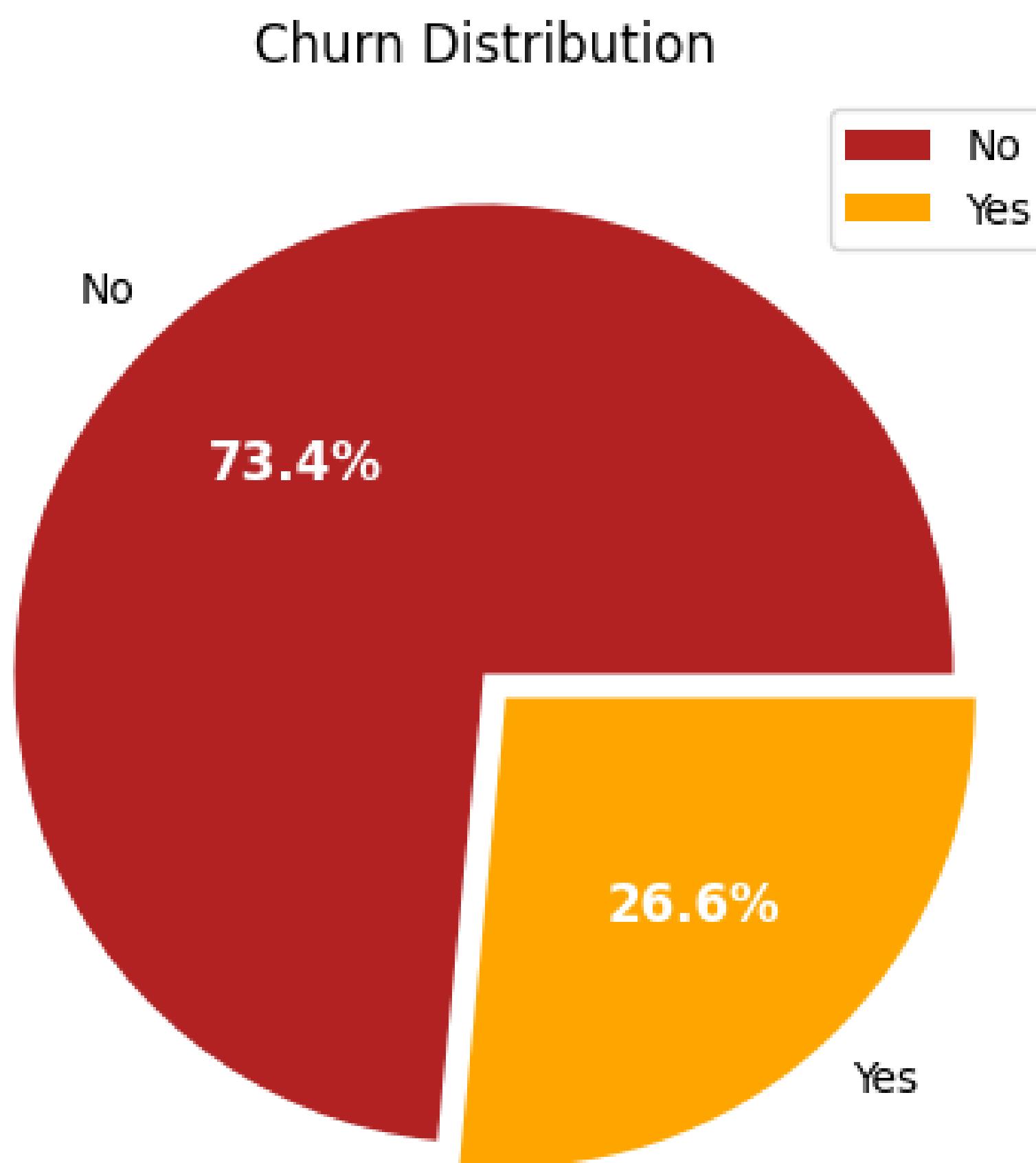


Figure 1. Churn Distribution

## 2. CHURN DISTRIBUTION IN NUMERICAL VARIABLES

(Monthly Charges, Total Charges, Num of Admin Tickets & Num of Tech Tickets)

The customer dataset contains a wide range of service information, including both numerical and categorical variables. Initially, the numerical variables were examined to gain an overview of the churn distribution across different customer segments.

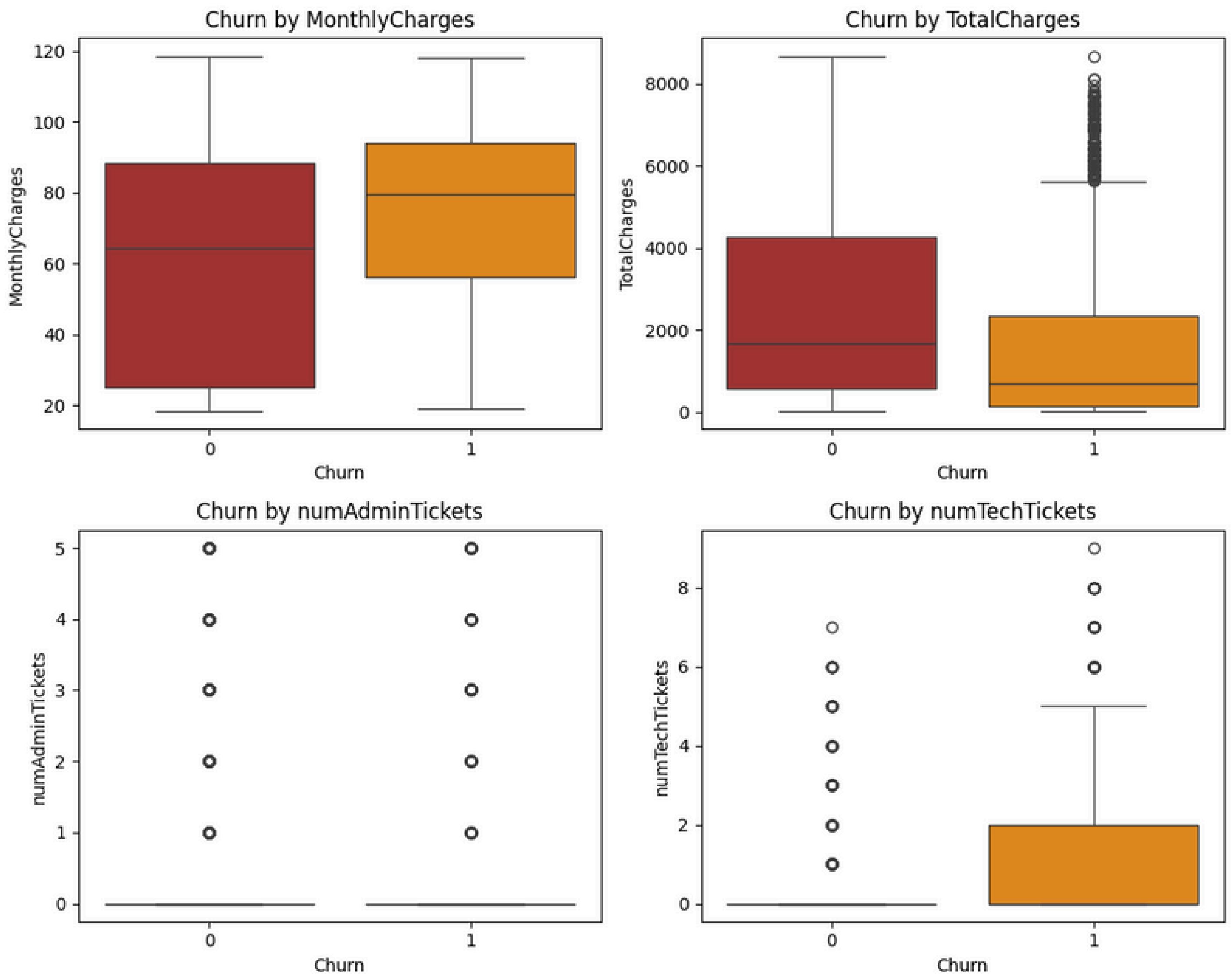


Figure 2. Churn By Monthly Charges, Total Charges, Number of Admin Tickets and Number of Tech Tickets

The charts suggest that customers who churn generally exhibit higher monthly charges, lower total charges, and a greater number of technical support tickets (figure 2).

Further analysis of the churn rate in relation to the number of technical support tickets reveals that customers with a higher frequency of support interactions are more likely to churn. This indicates that repeated service issues or dissatisfaction with the technical support experience may significantly contribute to customer attrition. The data shows a clear correlation between an increasing number of tech tickets and a higher probability of churn, highlighting this as a key area for targeted interventions aimed at enhancing customer retention (figure 3).

Num Tech Tickets	Count	Churn Rate (%)
9	1	100
8	11	100
7	29	96.55
6	72	81.94
5	116	75
4	133	69.17
3	151	66.89
2	201	62.69
1	256	65.63

Figure 3. Table of Churn Rate by Num Tech Tickets

### 3. CUSTOMER DEMOGRAPHIC ANALYSIS

#### 3.1. CHURN DISTRIBUTION IN CUSTOMER SEGMENTS

An analysis of customer demographics, including factors such as gender, age group, partner status, and dependents, was conducted to gain deeper insights into which customer segments are most susceptible to churn. This analysis aimed to identify patterns and characteristics that may correlate with higher churn rates, enabling the company to better understand the specific groups at greater risk of attrition.

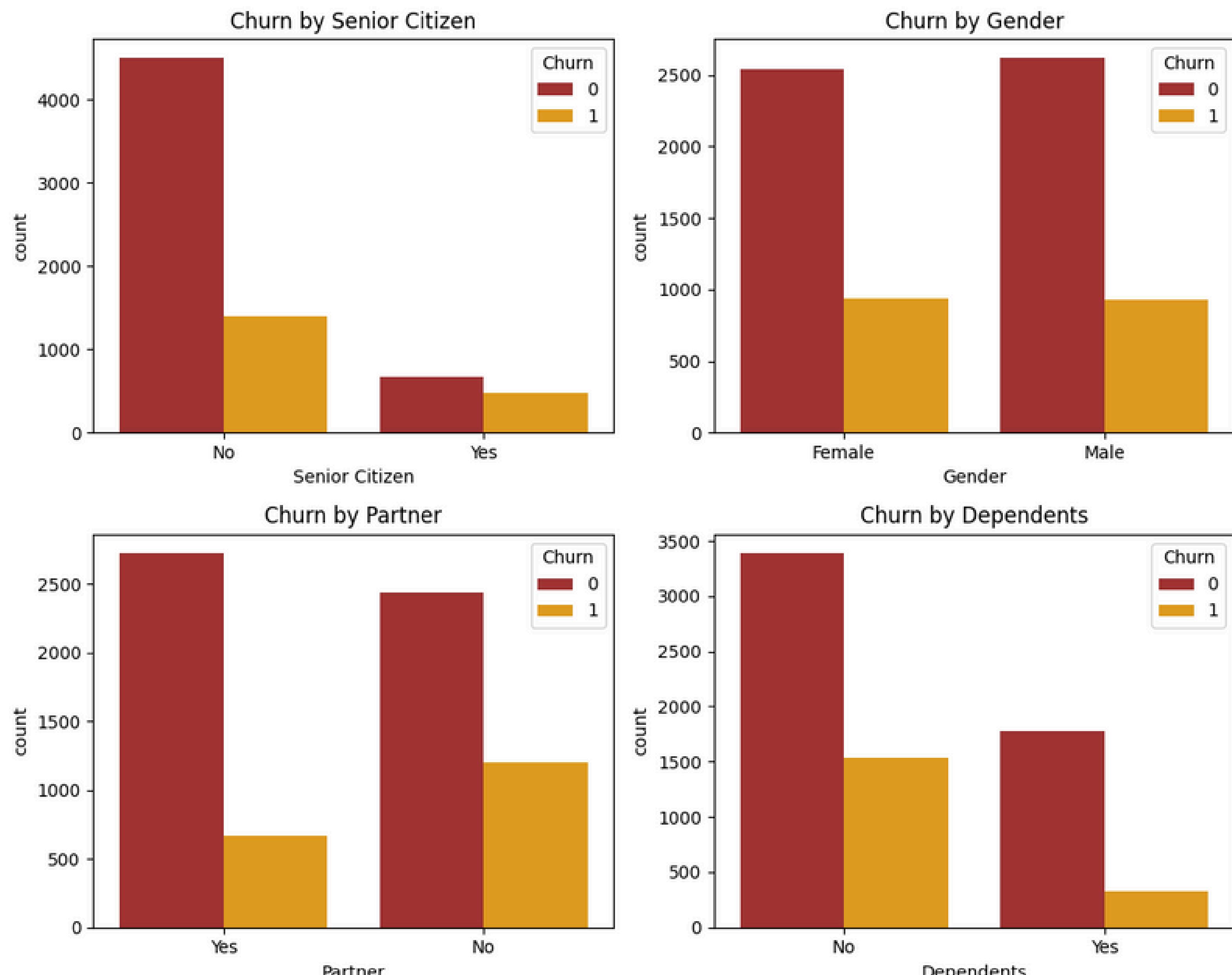


Figure 4. Churn By Age Groups, Gender, Partner Status and Dependents

Senior citizens represent a small portion of the total customer base (16.24%), yet nearly 50% of them cancel their services. The churn rate shows no significant variation between male and female customers, indicating that gender does not have a substantial influence on the decision to churn. Additionally, customers who do not have partners or dependents are more likely to discontinue their services (figure 4 and firgure 5).

Variable	Category	Count	Churn Rate (%)
Senior Citizen	Yes	1142	41.68
	No	5890	23.65
Gender	Female	3483	26.96
	Male	3549	26.20
Partner	No	3639	32.98
	Yes	3393	19.72
Dependents	No	4933	31.28
	Yes	2099	15.53

Figure 5. Table of Churn Rate by Age Groups, Gender, Partner and Dependents

## 3.2.CHURN DISTRIBUTION IN CUSTOMER TENURE

Customer tenure is also a significant factor in determining churn probability, with tenure ranging from 1 month to 72 months (figure 6).

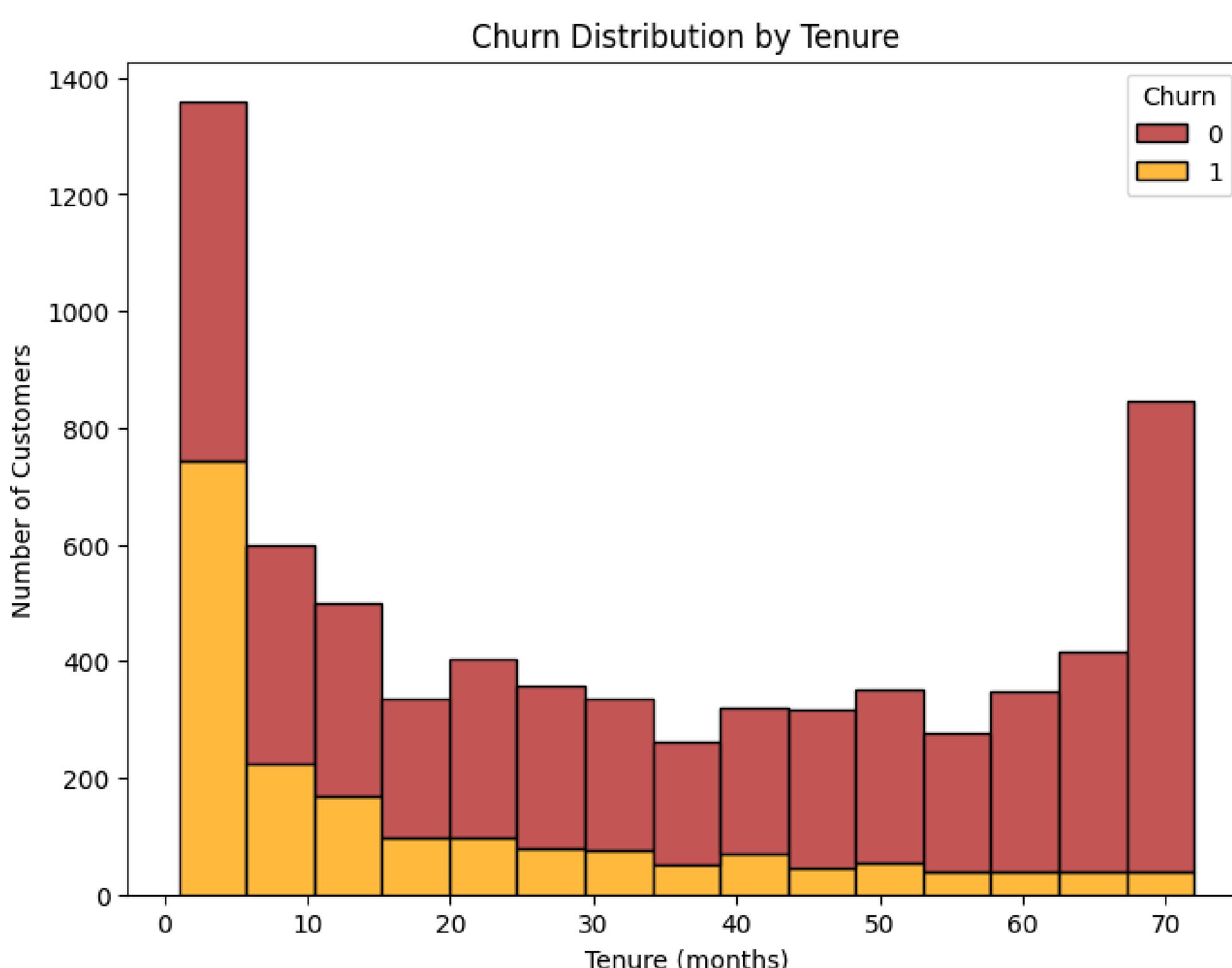


Figure 6. Churn Distribution by Tenure

Upon analyzing the churn rate across different tenure lengths, it is clear that customers with shorter tenures are more prone to churn. This suggests that as customer relationships lengthen, the likelihood of retention improves, with longer-tenured customers showing a lower risk of attrition (figure 7).

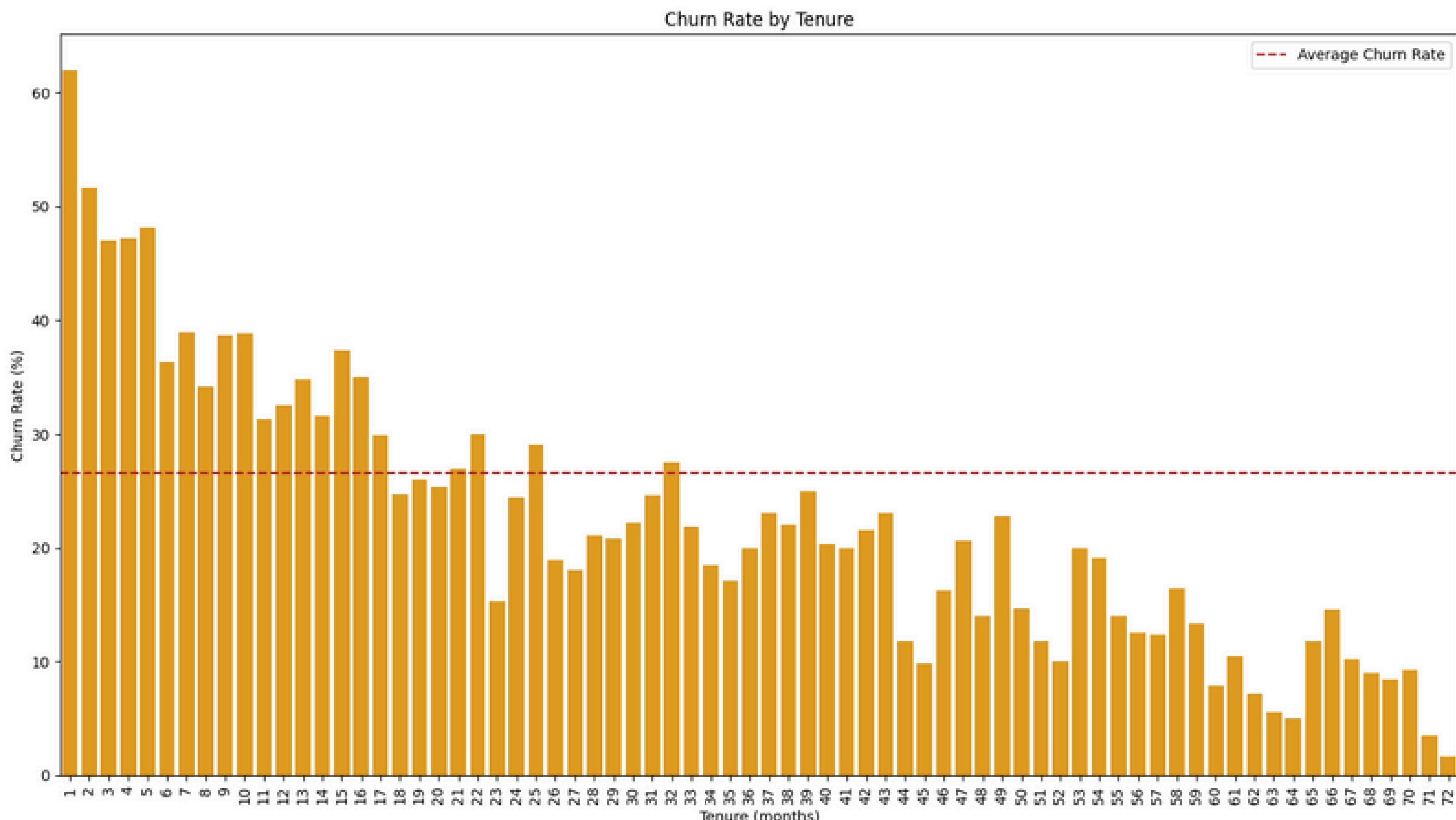


Figure 7. Churn Rate by Tenure

Particularly within the first 16 months of tenure, churn rates are notably higher, ranging from 31.31% to 61.99%, indicating that the probability of discontinuing their services is significantly elevated during this initial period of customer engagement (figure 8).

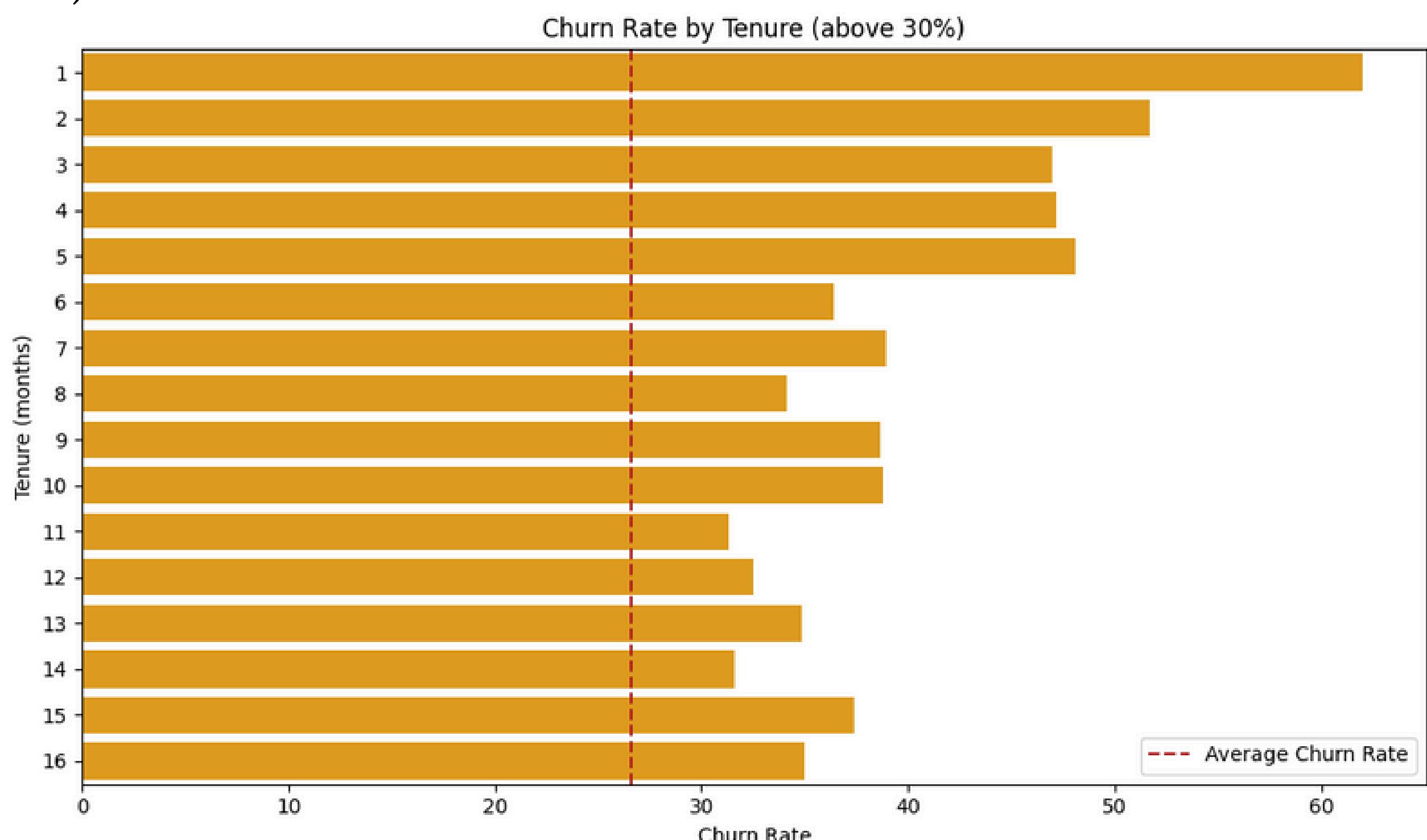


Figure 8. Churn Rate by Tenure (above 30%)

### 3.3.CHURN DISTRIBUTION IN SERVICES USED

A breakdown of the services used was conducted to assess whether any critical factors influence the likelihood of customer churn. The analysis focused on key services, including Phone Service, Internet Service, and TV Service (Streaming TV), to determine their impact on customer retention.

Approximately 1,700 customers with phone services have terminated their subscriptions, accounting for nearly a quarter of the total customer base. Around 44% of customers use fiber optic internet service, but nearly half of them eventually discontinue their service. The churn rate shows little difference between customers with TV services and those without, indicating that TV service usage does not significantly impact the risk of churn (figure 9 & 10).

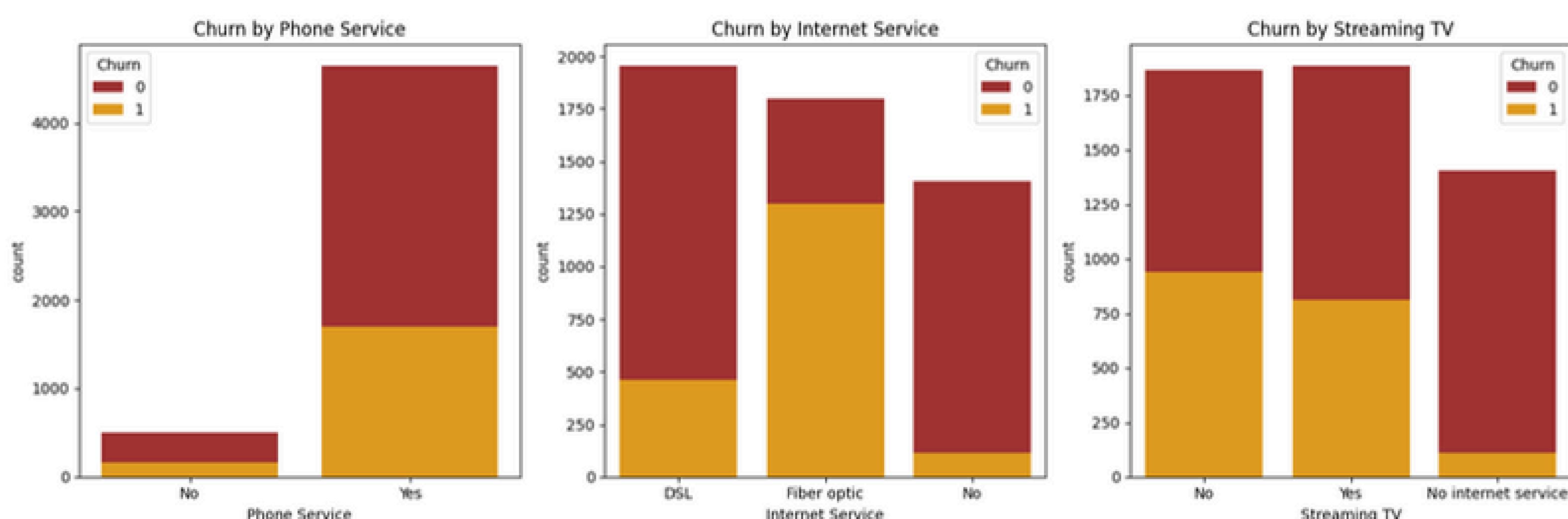


Figure 9. Churn Distribution by Phone Service, Internet Service and Streaming TV

Service	Category	Count	Churn Rate (%)
Phone Service	Yes	6352	26.75
	No	680	25.00
Internet Service	Fiber optic	3096	41.89
	DSL	2416	19.00
	No	1520	7.43
Streaming TV	No	2809	33.54
	Yes	2703	30.11
	No internet service	1520	7.43

Figure 10. Table of Churn Rate by Phone Service, Internet Service and Steaming TV

### 3.4.CHURN DISTRIBUTION IN BILLING AND PAYMENT METHODS

An investigation into billing and payment methods was carried out to uncover any patterns associated with the risk of customer churn. This analysis considered factors such as the use of paperless billing and the preferred payment methods, including electronic check, mailed check, bank transfer, or credit card, to identify potential correlations with service discontinuation.

Around 60% of customers opt for paperless billing, and about one-third of these customers end up canceling their services. Furthermore, approximately one-third of customers prefer electronic checks, with nearly 50% of them subsequently opting to churn (figure 11 & 12).

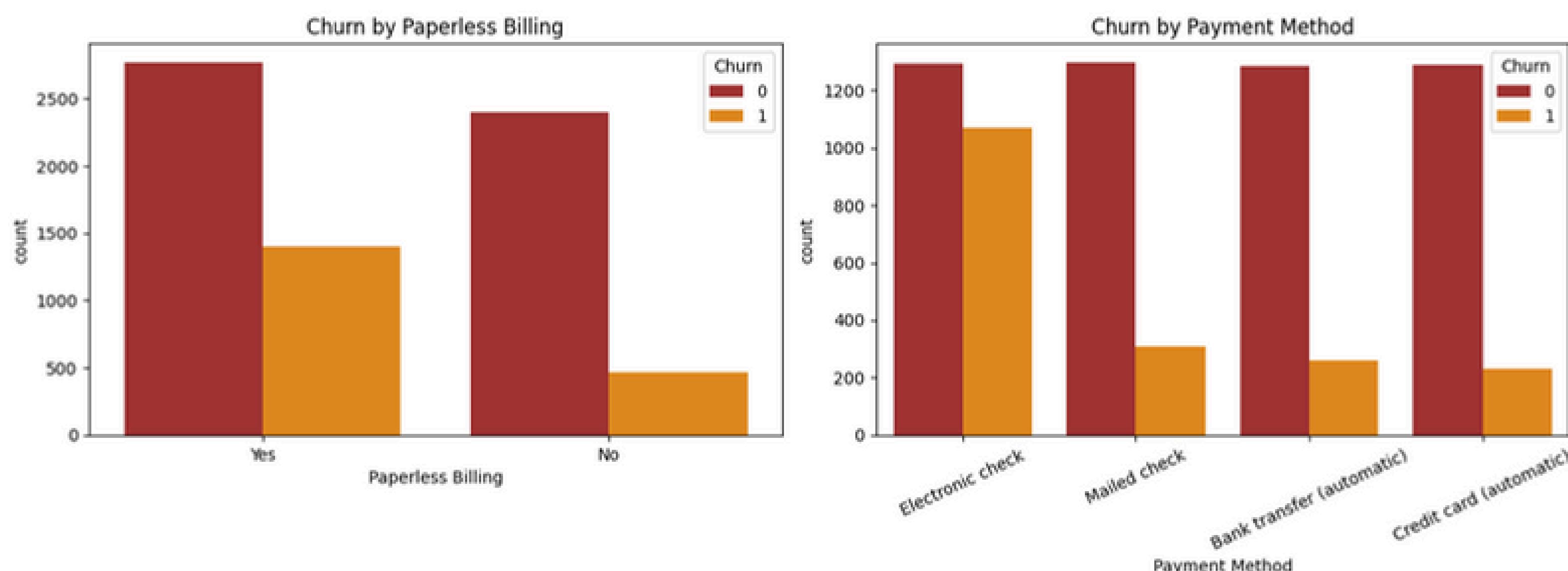


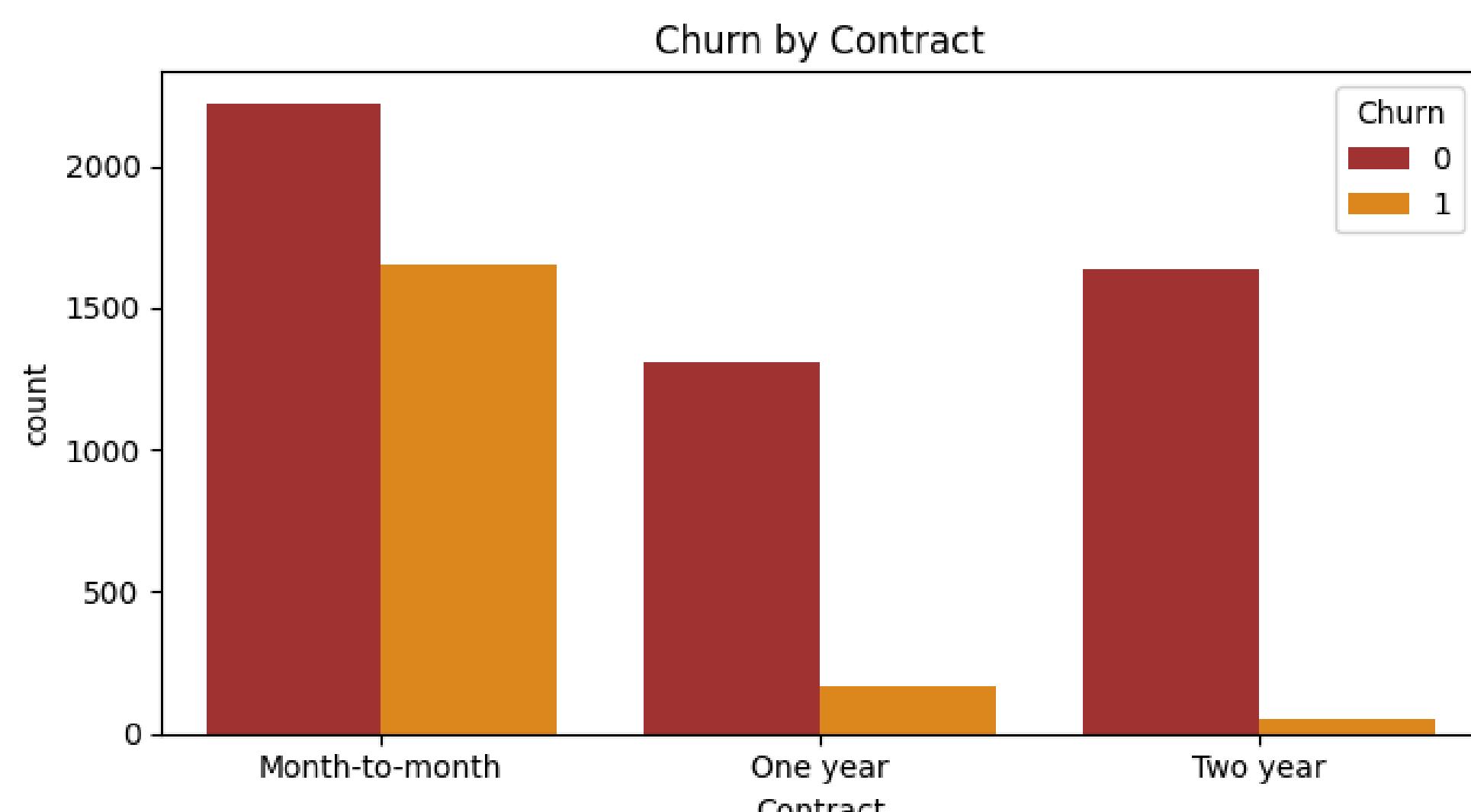
Figure 11. Churn Distribution by Billing and Payment Methods

Service	Category	Count	Churn Rate (%)
Paperless Billing	Yes	4168	33.59
	No	2864	16.38
Payment Method	Electronic check	2365	45.29
	Mailed check	1604	19.20
	Bank transfer	1542	16.73
	Credit card	1521	15.25

Figure 12. Table of Churn Rate by Billing and Payment Methods

## 3.5.CHURN DISTRIBUTION IN CONTRACT TYPES

Given the correlation between short customer tenure and an increased risk of churn, an analysis of contract types was conducted. The contract types examined include month-to-month, one-year, and two-year contracts, to assess their impact on customer retention.



Nearly half of customers with month-to-month contracts tend to churn (42.71%). As the contract duration increases, customer churn rates decrease, suggesting that longer contracts contribute to better customer retention (figure 13 & 14).

Figure 13. Churn Distribution by Contract

Variable	Category	Count	Churn Rate (%)
Contract	Month-to-month	3875	42.71
	One year	1472	11.28
	Two year	1685	2.85

Figure 14. Table of Churn Rate by Contract

## 3.6. FEATURE CORRELATIONS

To accurately assess the relationships between these variables, a correlation matrix was created to visually represent their interconnections (figure 15).

In the correlation matrix, a value of 1.00 signifies a perfect positive linear correlation, meaning an increase in one variable leads to a rise in another variable. Conversely, -1.00 indicates a perfect negative linear correlation, where an increase in one variable results in a decrease in another variable.

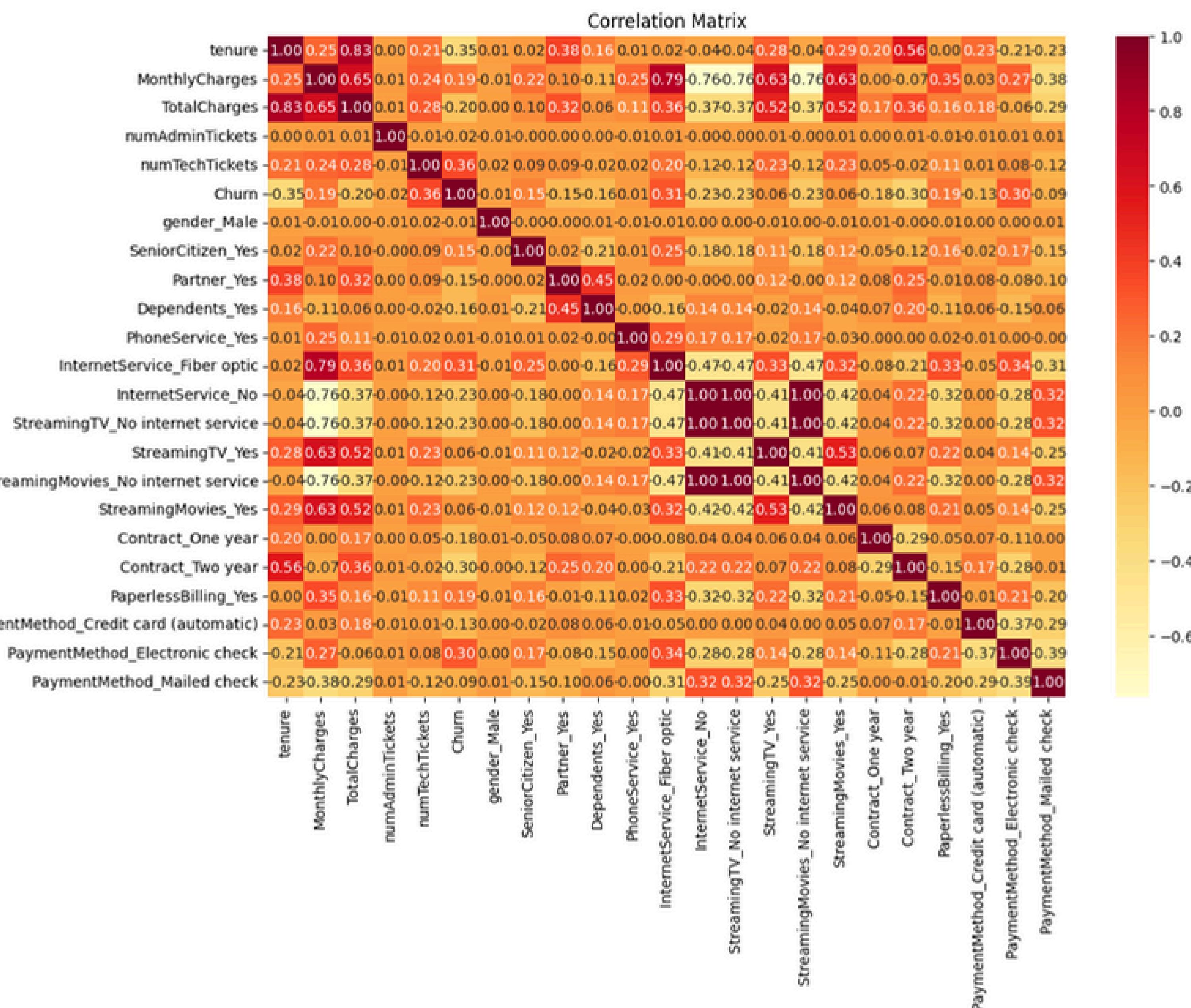


Figure 15. Correlation Matrix

# CUSTOMER CHURN PREDICTIVE MODEL

## 1. MODEL'S RESULTS

To develop a customer churn prediction model, multiple algorithms were utilized for testing and training, including Logistic Regression, Decision Tree, Random Forest, and Support Vector Machine (SVM).

Among these models, Logistic Regression emerged as the most effective, achieving the highest accuracy of 85.99% and accurately predicting the largest number of churns, totaling 378 (figure 16 & 17).

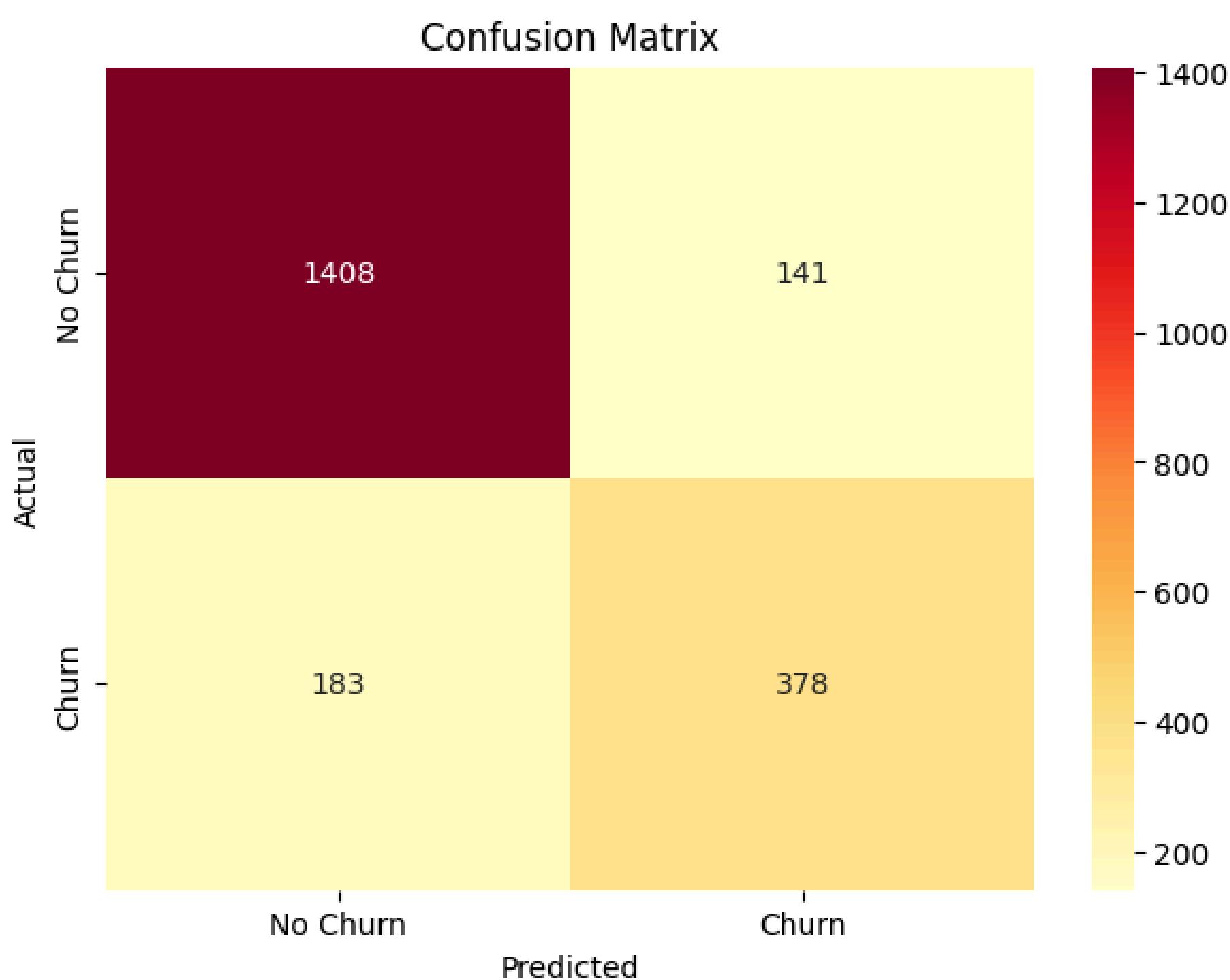


Figure 16. Confusion Matrix

	precision	recall	f1-score	support
0	0.88	0.91	0.90	1549
1	0.73	0.67	0.70	561
accuracy			0.85	2110
macro avg	0.81	0.79	0.80	2110
weighted avg	0.84	0.85	0.84	2110

Figure 17. Classification Report

## 2. FEATURE IMPORTANCE ANALYSIS

A feature importance analysis was conducted using the aforementioned models, revealing that certain features consistently ranked at the top in most analyses. These key features include number of tech tickets, month-to-month contract, fiber optic internet service and tenure.

Below is the feature importance derived from the Logistic Regression model, which supports the findings mentioned above. The top factors identified are:

- Number of Tech Tickets (1.25): A high volume of technical support tickets is strongly correlated with churn, suggesting potential service issues or customer dissatisfaction.
- Month-to-month Contract (1.15): This is the most significant predictor of churn, indicating that customers with month-to-month contracts are at a higher risk of cancellation.
- Fiber Optic Internet Service (0.85): The presence of fiber optic internet service has a notable impact on churn, with customers using this service more likely to churn.
- Customer Tenure (-1.5): A shorter tenure is strongly associated with an increased probability of churn, suggesting that customers with less time on their contracts are more likely to leave.

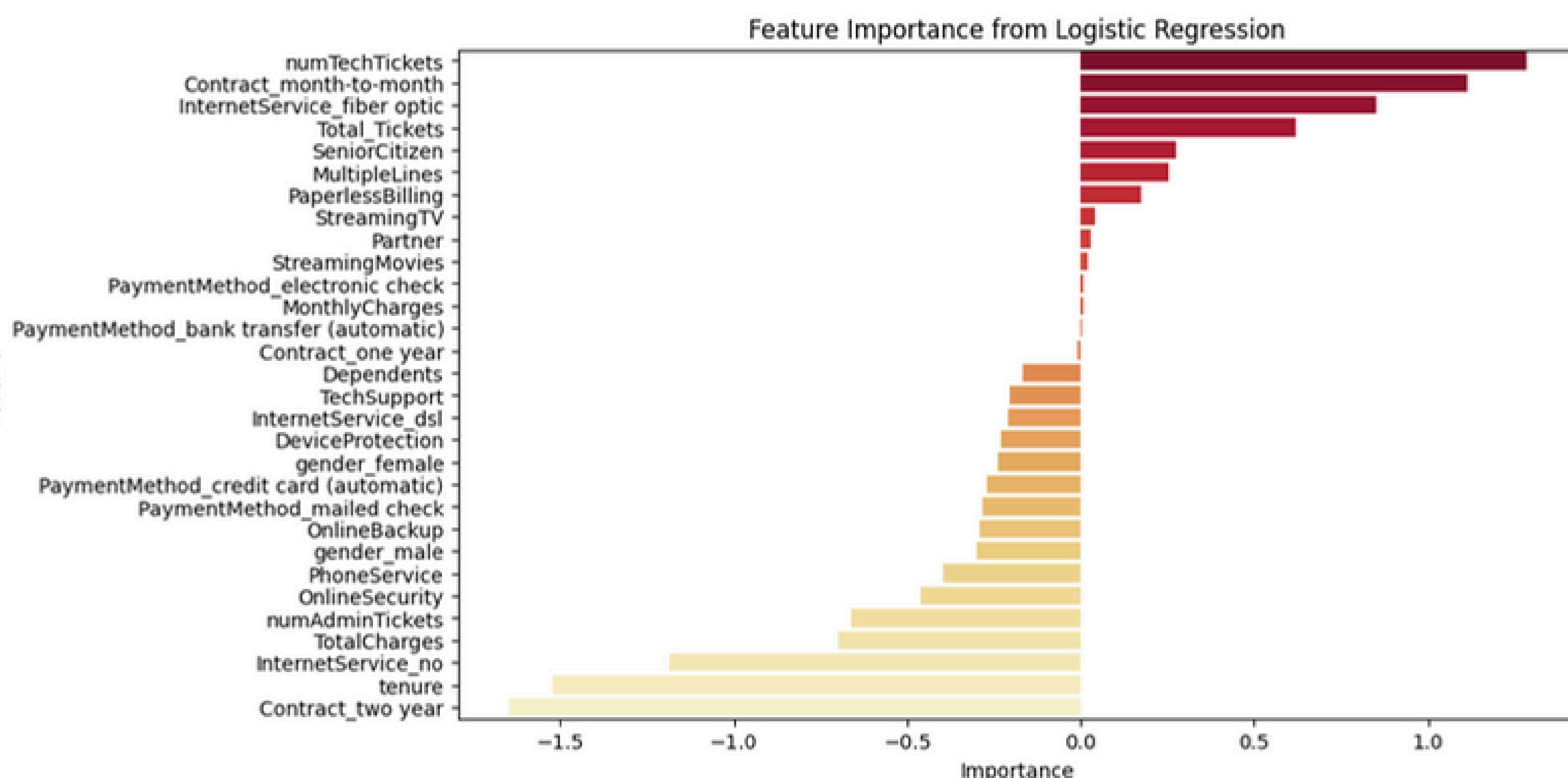


Figure 18. Feature Importance from Logistic Regression

# INSIGHTS AND RECOMMENDATIONS



## Improve Technical Support and Service Quality

**Issue:** A high number of technical support tickets is linked to churn, indicating service issues or dissatisfaction.

**Recommendation:** Invest in improving the quality of customer support and technical service. This could include:

- Training customer service representatives for faster resolution times.
- Implementing a more effective troubleshooting process to address common issues before they escalate.
- Proactively reaching out to customers who are experiencing frequent technical issues to resolve them before they decide to churn.
- Create a feedback system to gain insights into the causes of customer dissatisfaction. Conduct regular surveys to assess customer experiences, pinpoint areas of concern, and take action based on the feedback. Make it easy for customers to report problems and ensure their voices are heard.



## Offer Long-Term Contracts

**Issue:** Customers with month-to-month contracts are at a higher risk of churn.

**Recommendation:** Encourage customers to switch to longer-term contracts by offering attractive incentives, such as discounts or added benefits for committing to longer terms (e.g., 1-year or 2-year contracts). This would help increase customer retention and reduce churn rates



## Ensure Quality and Reliability of Fiber Optic Internet Service

**Issue:** The presence of fiber optic internet service affects churn likelihood.

**Recommendation:** Invest in ensuring the quality and reliability of fiber optic services, as it significantly influences customer retention. If fiber optic is a premium service, emphasize its advantages in marketing materials, and offer targeted promotions to retain fiber optic customers by addressing any network or service issues proactively.



## Enhance Customer Engagement and Experience for New Customers

**Issue:** Shorter tenure correlates with increased churn probability.

**Recommendation:** Focus on increasing engagement with new customers through:

- Personalized onboarding experiences.
- Regular follow-up to ensure customers are satisfied and fully utilizing their services.
- Offering incentives or rewards for customers who stay for a longer period (e.g., free upgrades or service bundles after a certain number of months).