



**University of Information Technology**

## **Final Project CS313 - Data Mining**

**LECTURER: PhD. VO NGUYEN LE DUY**

# **TELCO CUSTOMER CHURN ANALYSIS**

**GROUP 2**

**MEMBERS:**

**Bui Ha Khanh - 21520282**

**Nguyen Quoc Hung - 21520253**

**Nguyen Tan Khang - 21520281**

**Nguyen Chanh Nghia - 21520362**



# Table of contents

I. INTRODUCTION

II. EXPLORATORY DATA  
ANALYSIS

III. CONCLUSION

IV. EXPERIMENTS

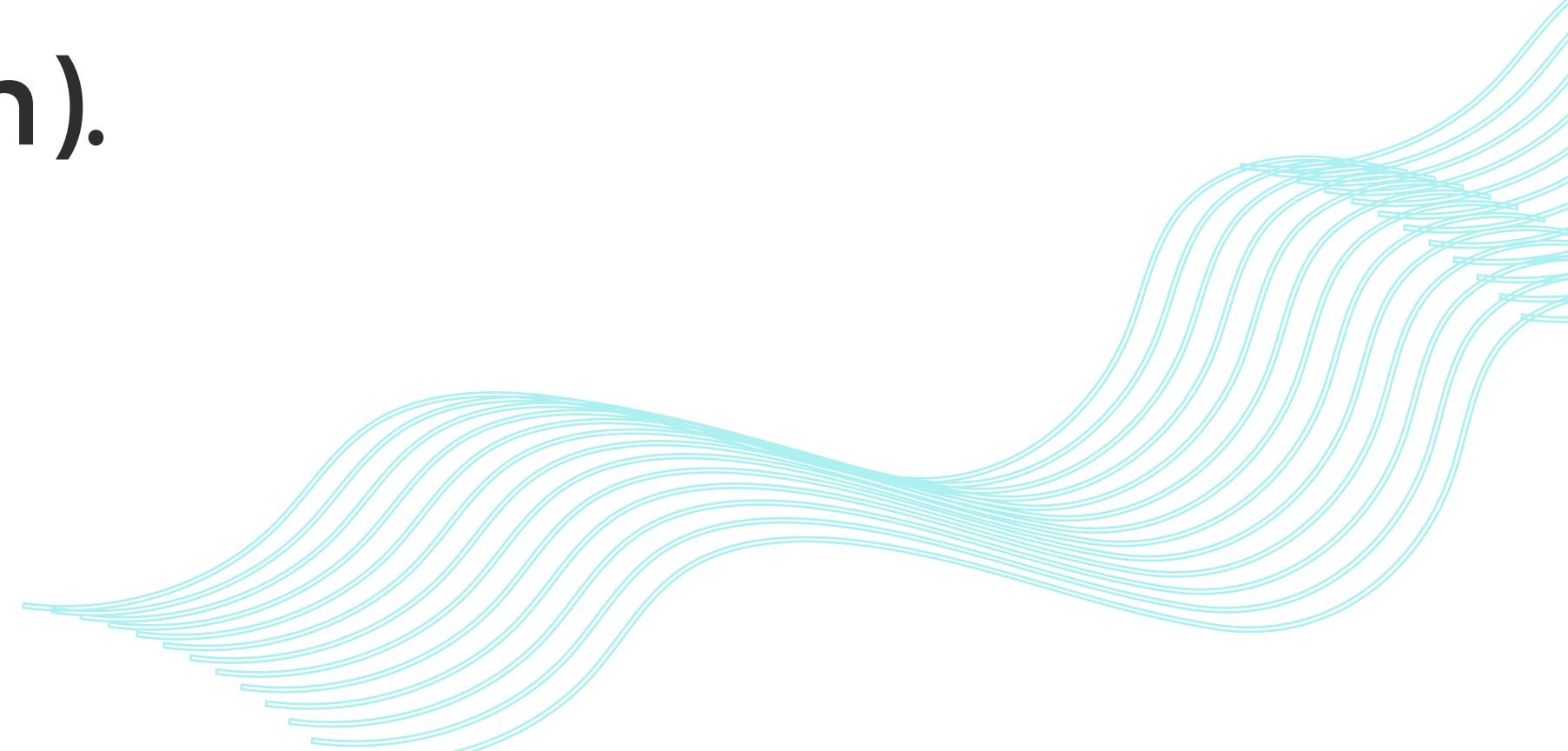
V. DEMO



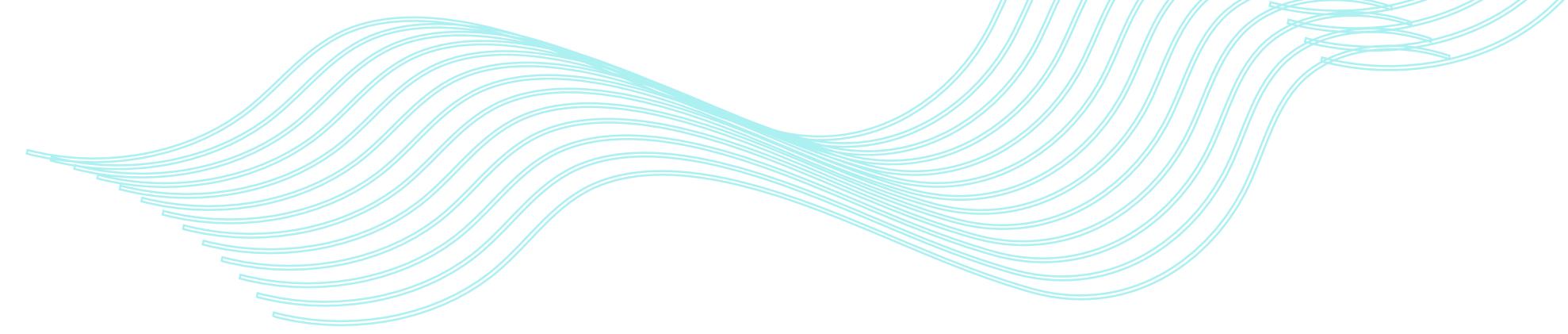
# INTRODUCTION

# PURPOSE

Analyze telecommunication company customer data to predict whether or not a customer is likely to leave the platform (churn).



# **DATASET**

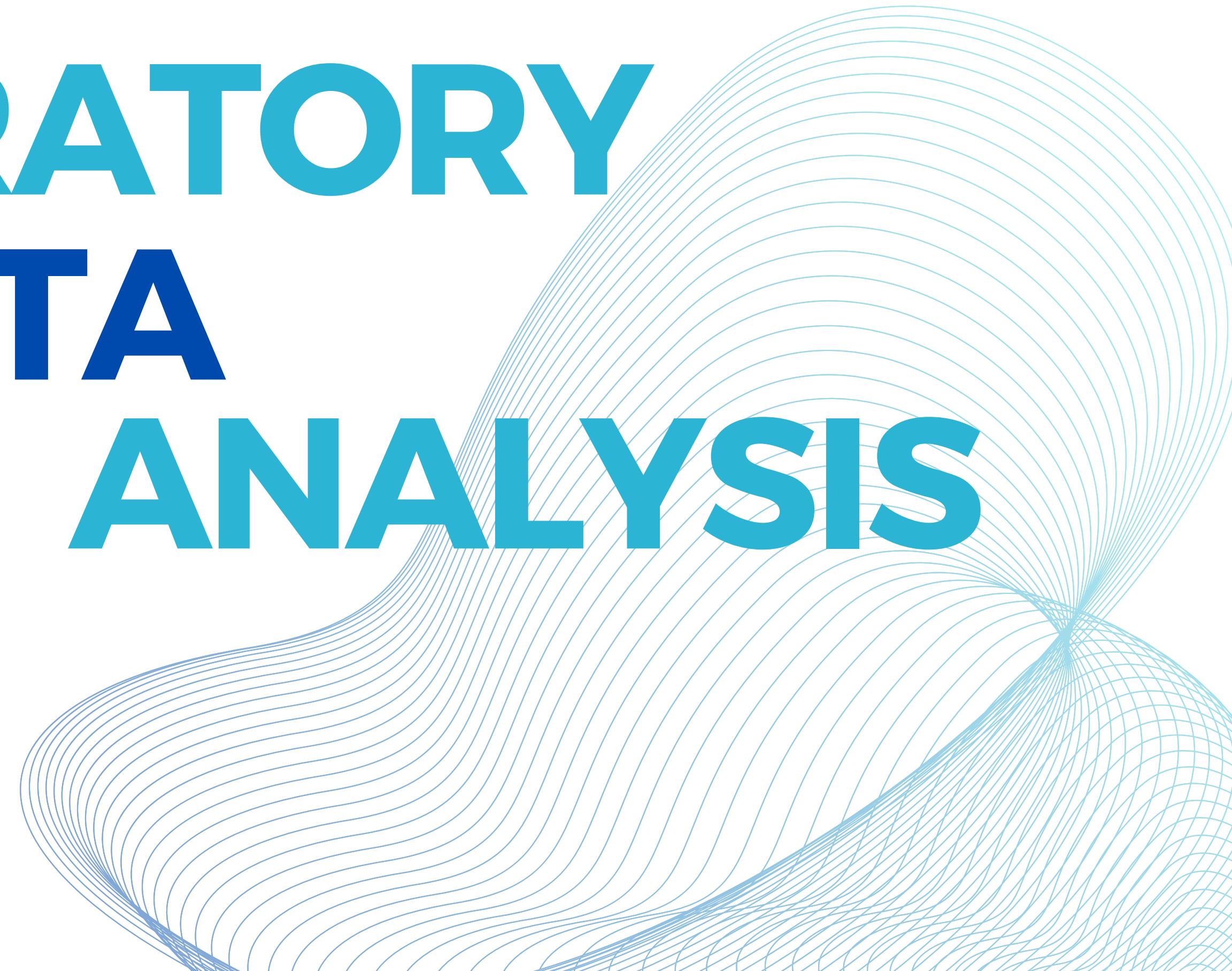


**Data from 7043 customers (33 features):**

- Churn (Yes or No).
- Customer account information (tenure, contract, payments, etc.).
- Demographic Information (Partner, Gender, Age, etc.).
- Add-on services provided by the platform.

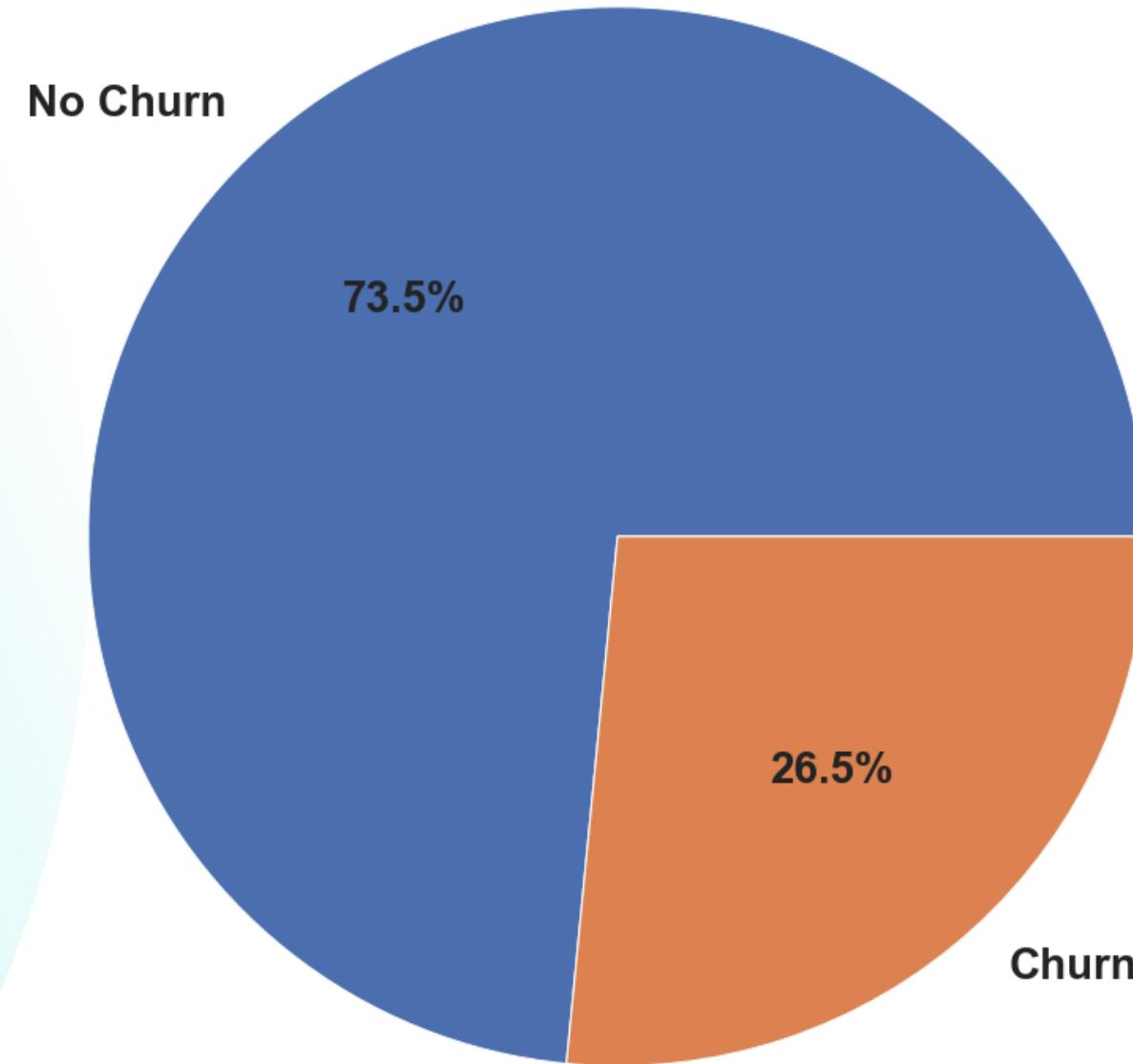
**Source: [Kaggle.com](#)**

# **EXPLORATORY DATA ANALYSIS**

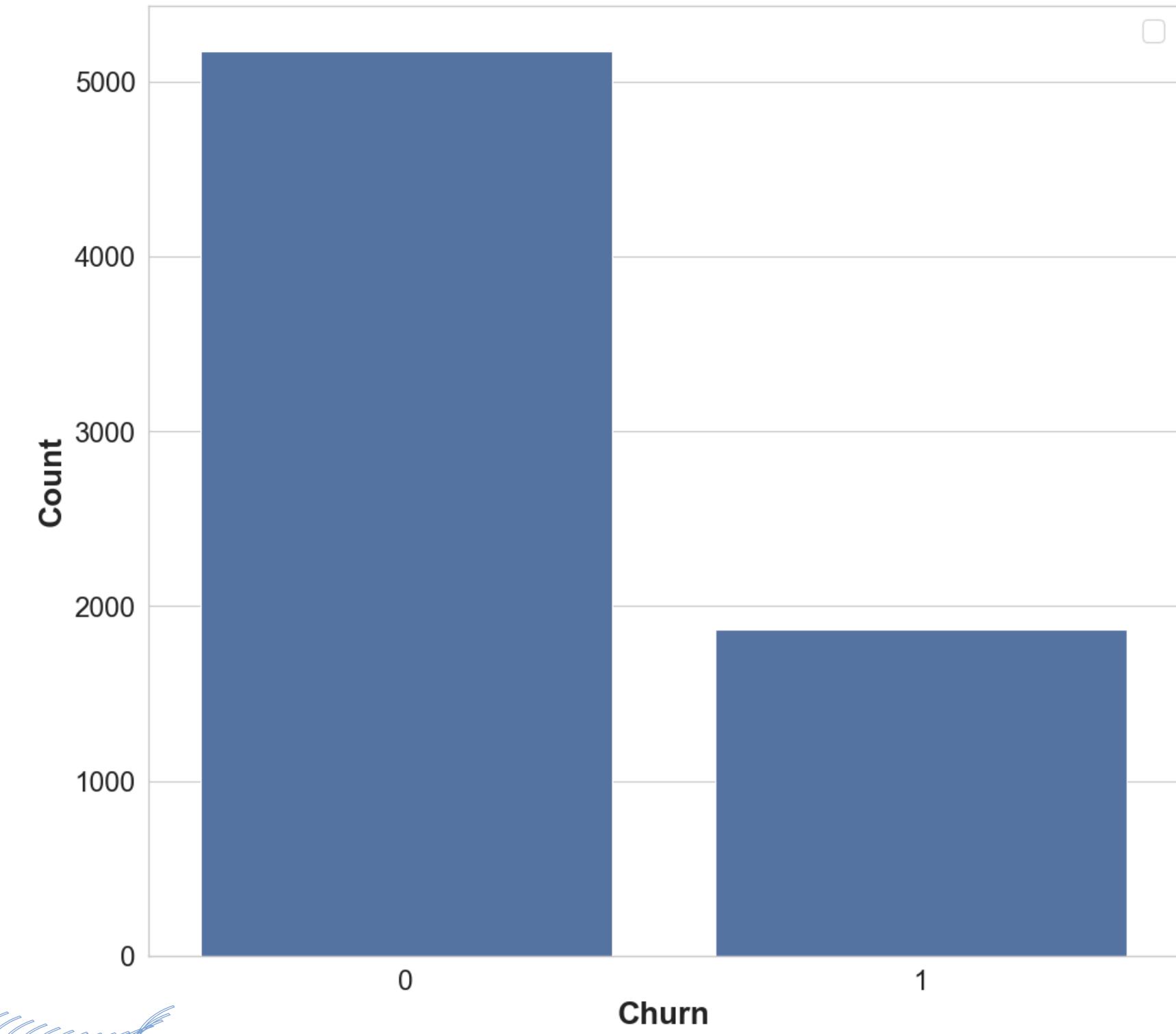


# Target (Churn)

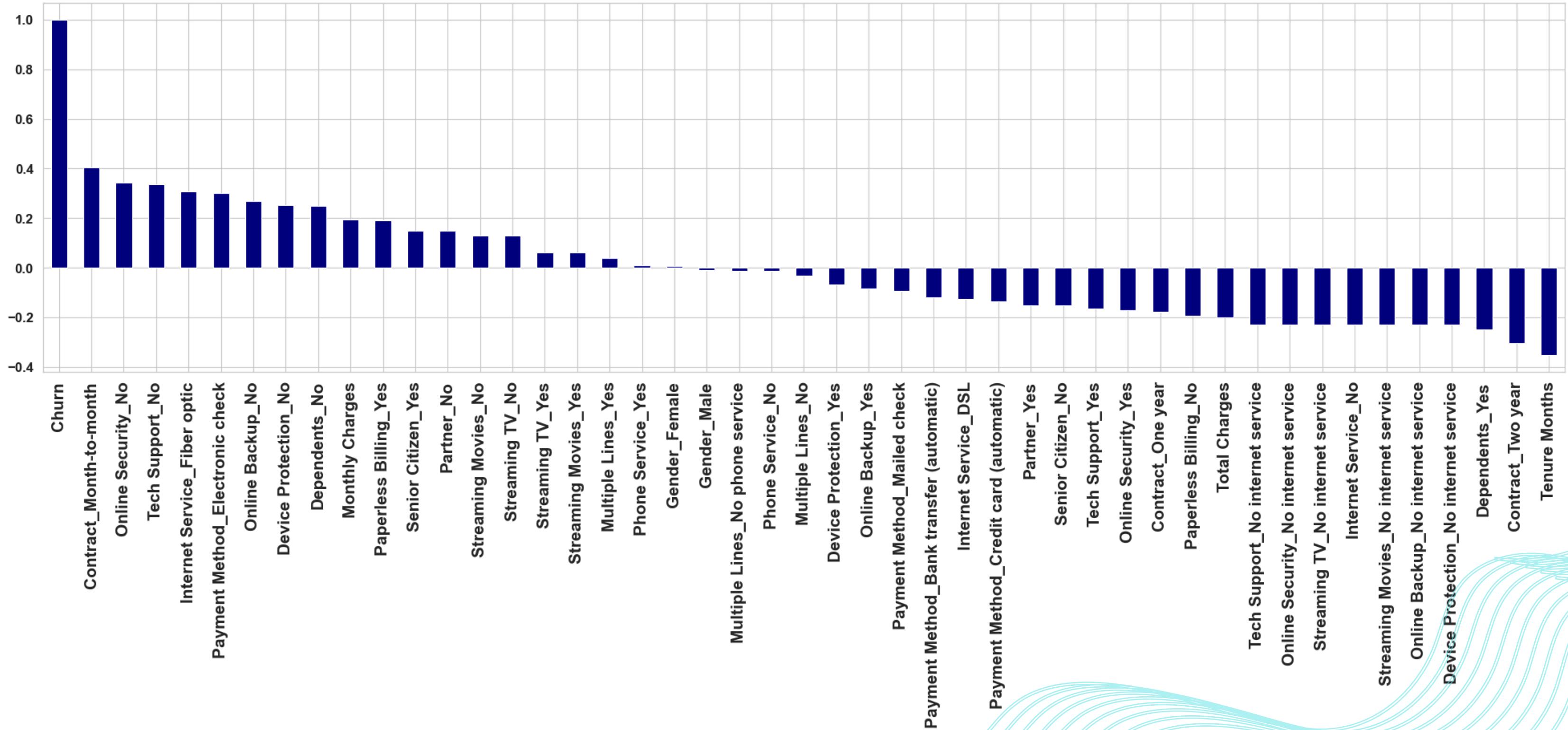
Churn Outcome Pie Chart



Churn Outcome Distributions



# Feature Correlation to Churn

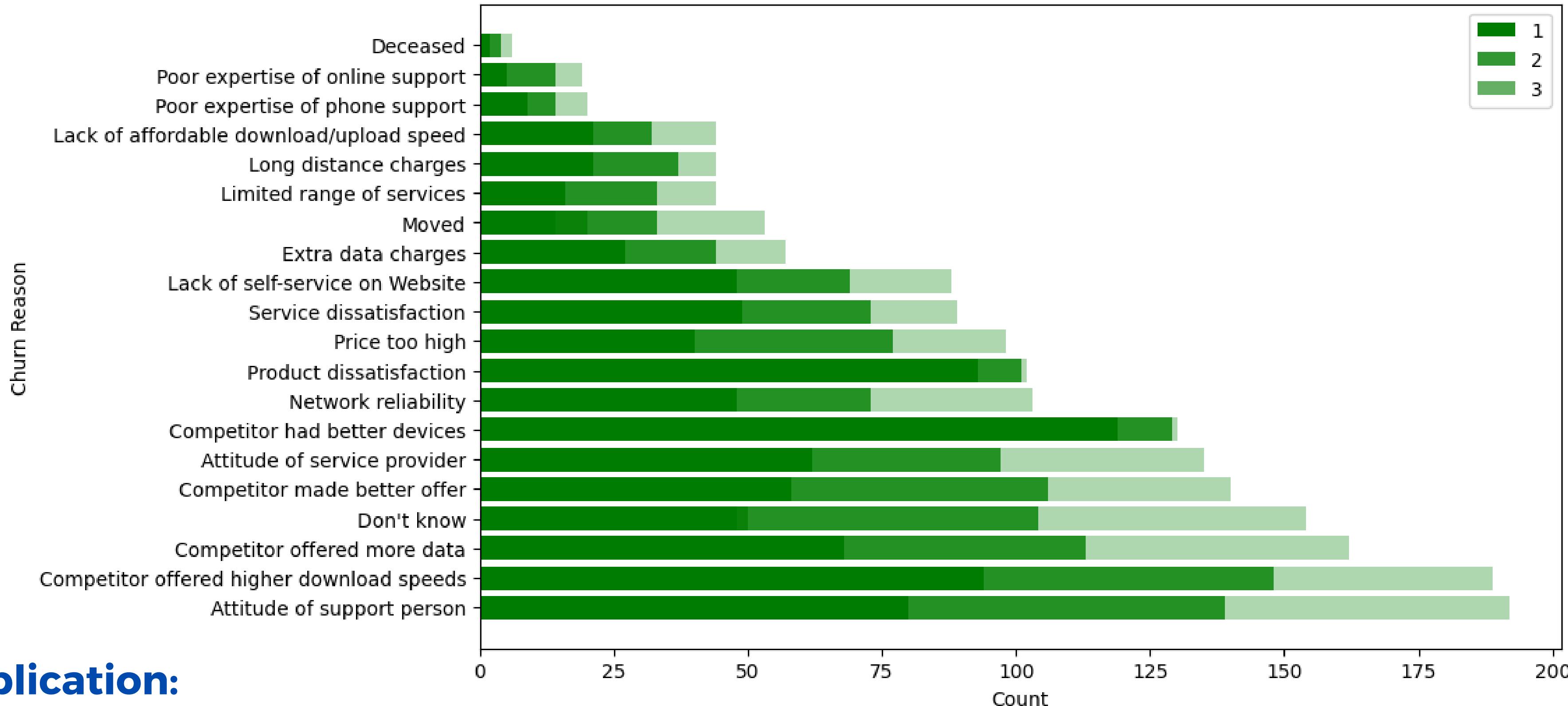


**Most Positively Correlated:** Monthly Contracts, No Online Security Add-On Service.

**Negatively Correlated:** Tenure, Two-Year Contracts, No Internet Service.

# Churn Reason

Churn Reason and Satisfaction Score

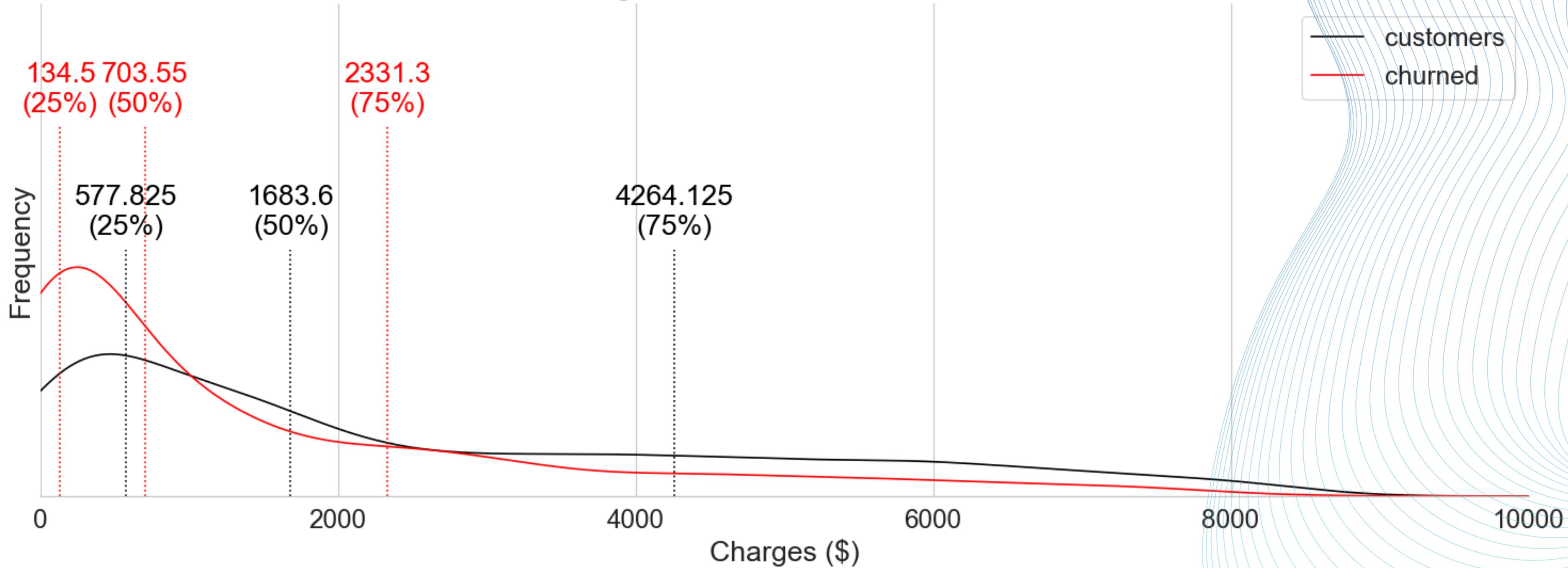


**Implication:**

2 major reason: attitude of support staff + better services from competitor

# Total Charges

Total charges of current vs churned customers

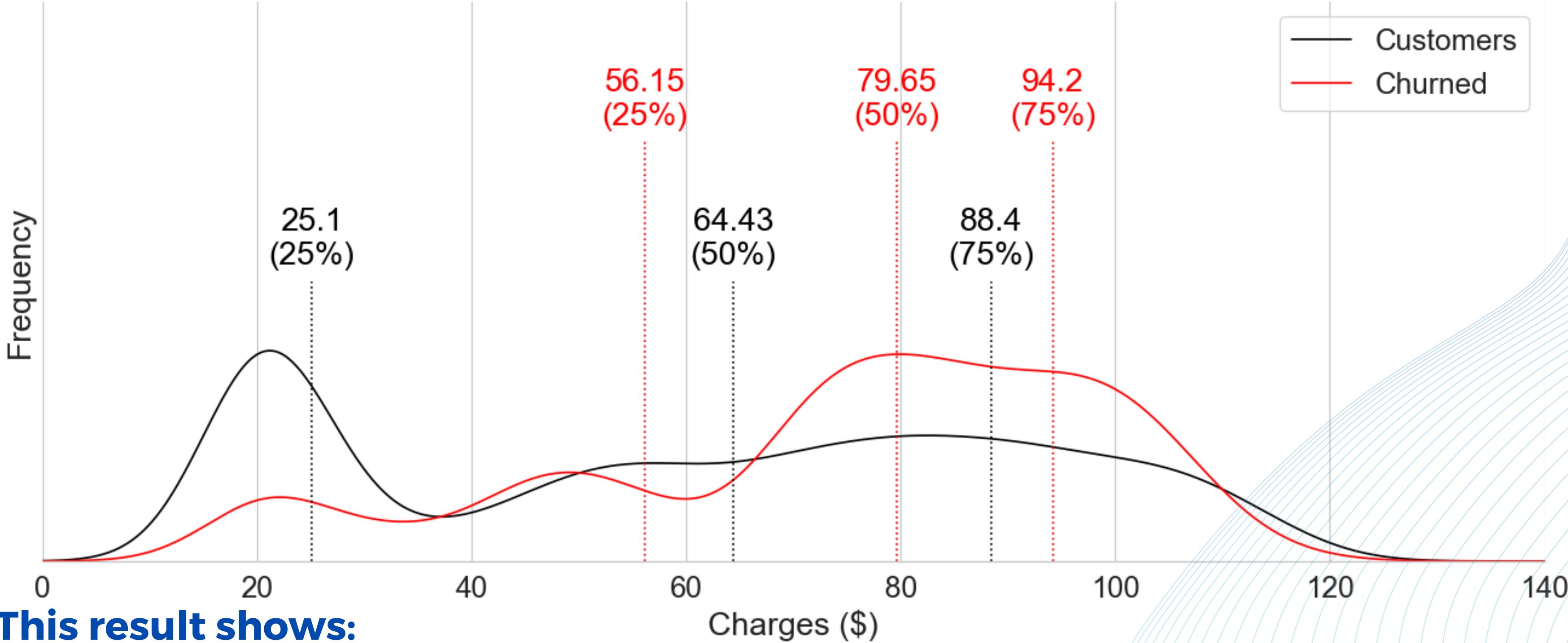


**This result shows:**

The current customers have higher total charges than the churned customers.

# Monthly Charges

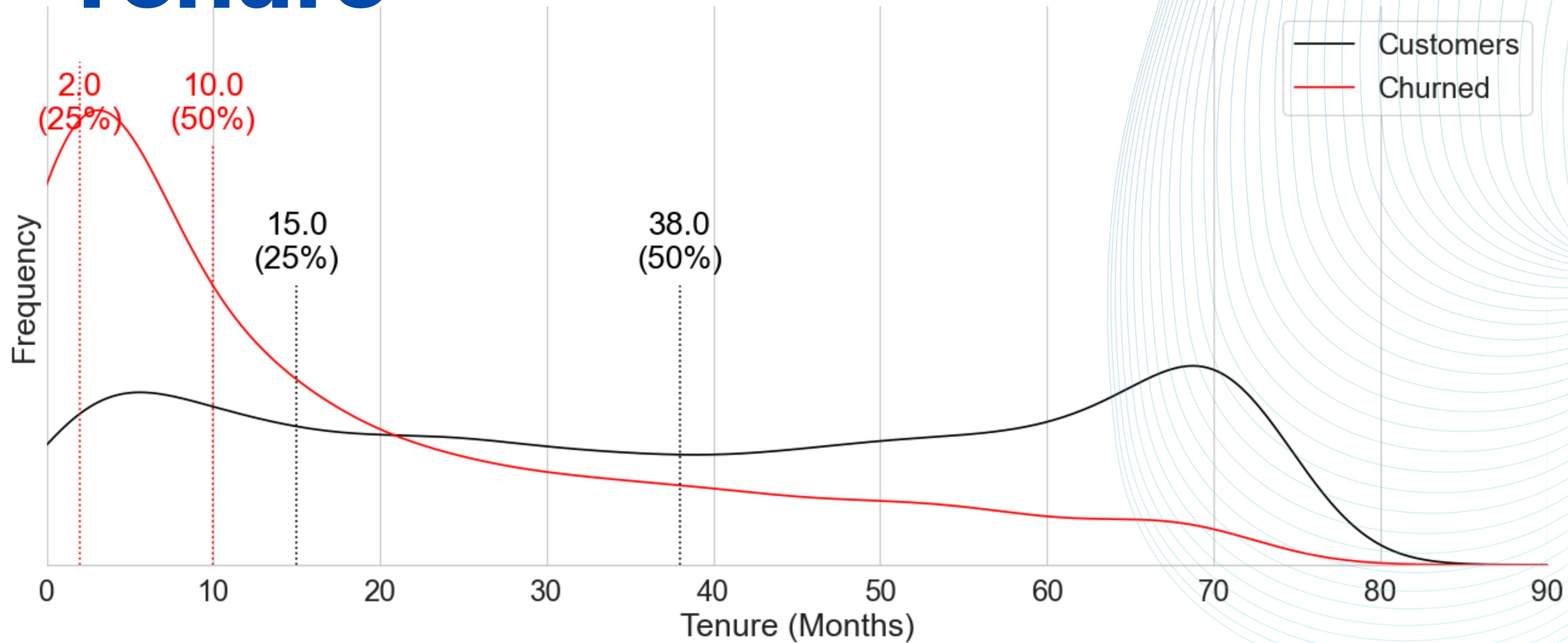
Monthly charges of current vs churned customers



**This result shows:**

The current customers have lower monthly charges than the churned ones.  
As monthly charges increase, the probability of customer churn increases.

# Tenure



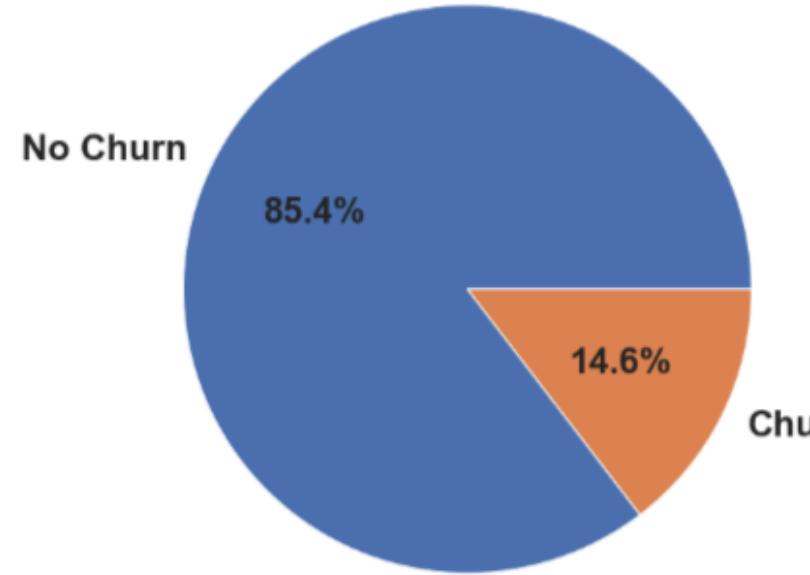
**This result shows:**

Current customers have longer tenure.

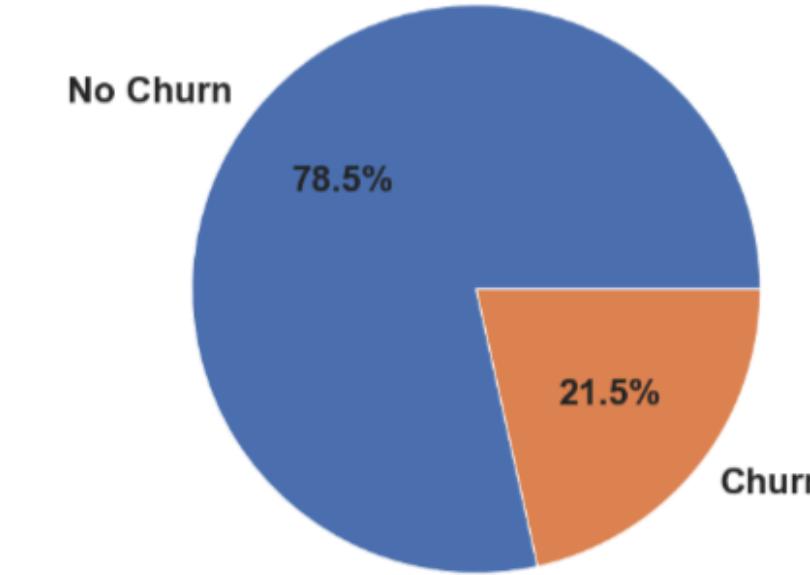
As tenure increases, probability of churn decreases.

# Add-On Services

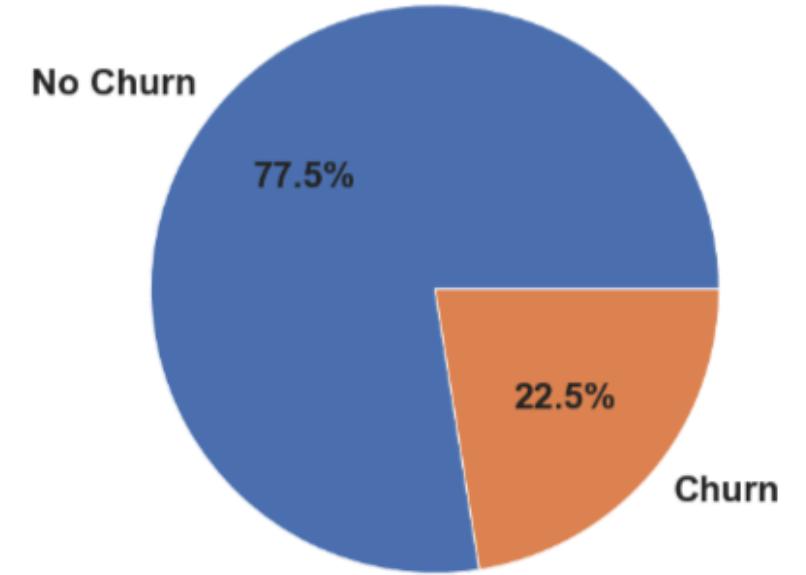
Online Security - Churn %



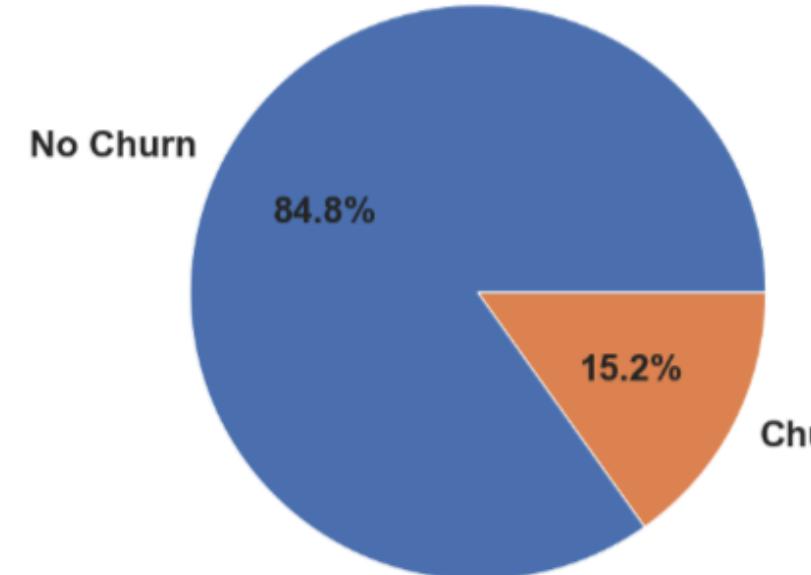
Online Backup - Churn %



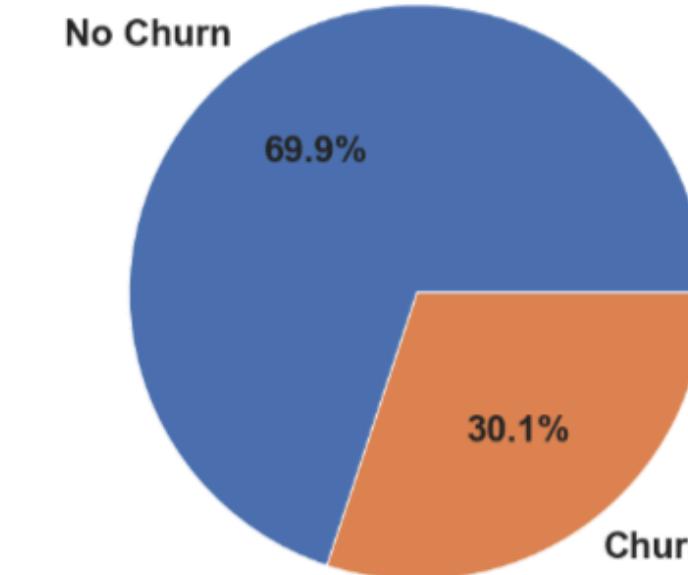
Device Protection - Churn %



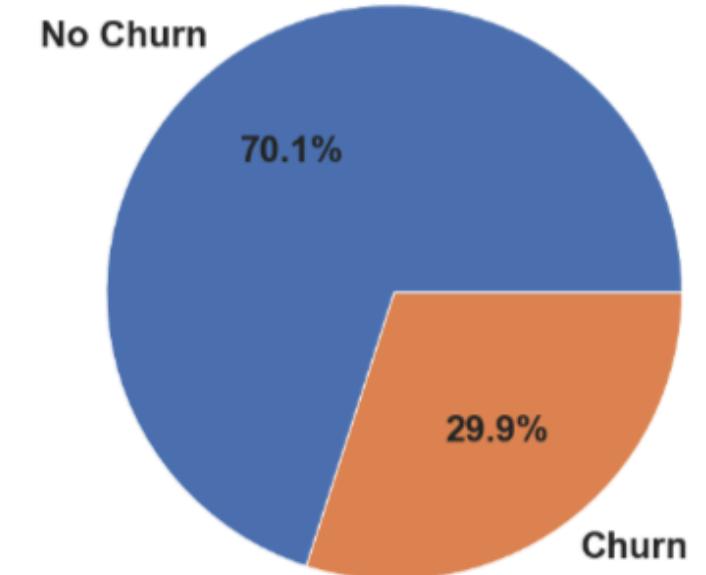
Tech Support - Churn %



Streaming TV - Churn %



Streaming Movies - Churn %



**This result shows:**

Customers with online security and/or tech support add-ons will churn the least

Customers with Streaming services (TV/Movies) will churn the most

# Association Rules

antecedents	consequents	confidence	lift
(Streaming TV_Yes, Device Protection_Yes)	(Streaming Movies_Yes)	0.800255	1.616035
(Streaming TV_Yes, Device Protection_Yes, Phon...)	(Streaming Movies_Yes)	0.799427	1.614363
(Streaming TV_Yes, Device Protection_Yes)	(Streaming Movies_Yes, Phone Service_Yes)	0.712189	1.614275
(Device Protection_Yes, Streaming Movies_Yes, ...)	(Streaming TV_Yes)	0.792051	1.614239
(Device Protection_Yes, Streaming Movies_Yes)	(Streaming TV_Yes)	0.786207	1.602329

**This result shows:**

Phone services, device protection, streaming movies and streaming TV services are usually opened together.

# Discounting Strategy

Average monthly charges for customers with both Movies and TV = 93.2375

Average monthly charges for customers with only Movies = 76.81174242424242

Average monthly charges for customers with only TV = 77.35280312907432

Average monthly charges for customers with only 1 service = 77.08227277665837

Difference in monthly charges of both services and 1 service = 16.155227223341626

The marginal profit is: 281.2867387992129 when the discount is 0.03

The marginal profit is: 487.01547987485696 when the discount is 0.06

The marginal profit is: 617.1862232269323 when the discount is 0.09

**The marginal profit is: 671.7989688554385 when the discount is 0.12**

The marginal profit is: 650.8537167603763 when the discount is 0.15

The marginal profit is: 554.3504669417453 when the discount is 0.18

The marginal profit is: 382.2892193995449 when the discount is 0.21

The marginal profit is: 134.66997413377703 when the discount is 0.24

The marginal profit is: -188.50726885556287 when the discount is 0.27

The marginal profit is: -587.2425095684666 when the discount is 0.3

# Conclusions

## To minimize the rate of churn:

- Increase the friendliness and enthusiasm with customers, and specially improve the quality of services to compete with others.
- Maximize the tenure months of customer to lower the monthly charges

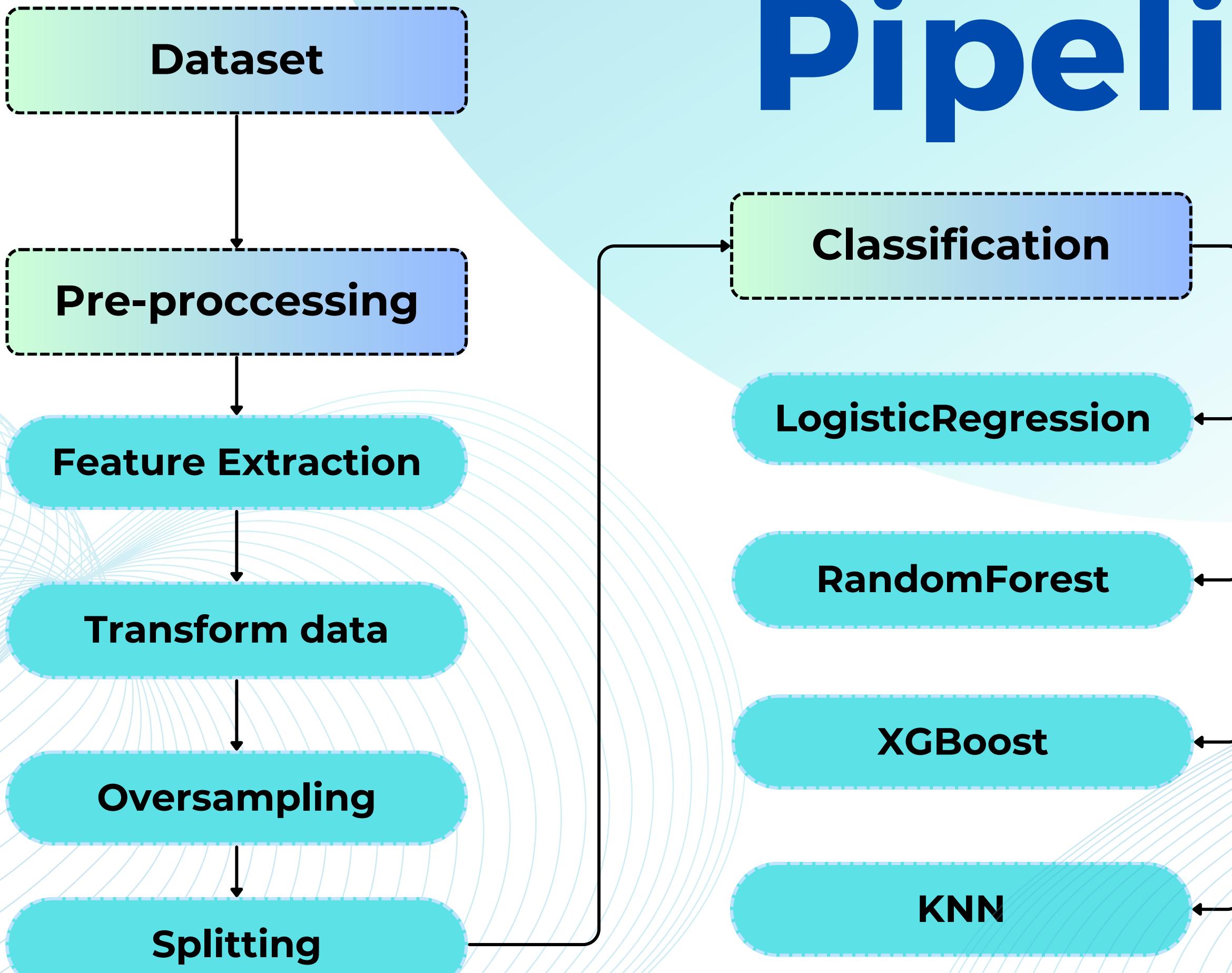
## Suggesting new business approach :

- Phone services, device protection, streaming movies and streaming TV services are the best combination.
- The Telco company should have discount of 12% for the existing customers with only one service to subscribe for more service.

# EXPERIMENTS



# Pipeline



# DATA BALANCING

```
## oversampling
from imblearn.over_sampling import SMOTE
## creating an instance of SMOTE
smote = SMOTE(k_neighbors=5, random_state=42, sampling_strategy=1)
## resampling the dataset using SMOTE oversampling
X_resampled, y_resampled = smote.fit_resample(X_new, y)
X_train, X_test, y_train, y_test = train_test_split(X_resampled, y_resampled,
random_state = 42, test_size = 0.2)
```

# Models Performance



**LogisticRegression**

79% Accuracy

**RandomForest**

87% Accuracy

**XGBoost**

87% Accuracy

**KNN**

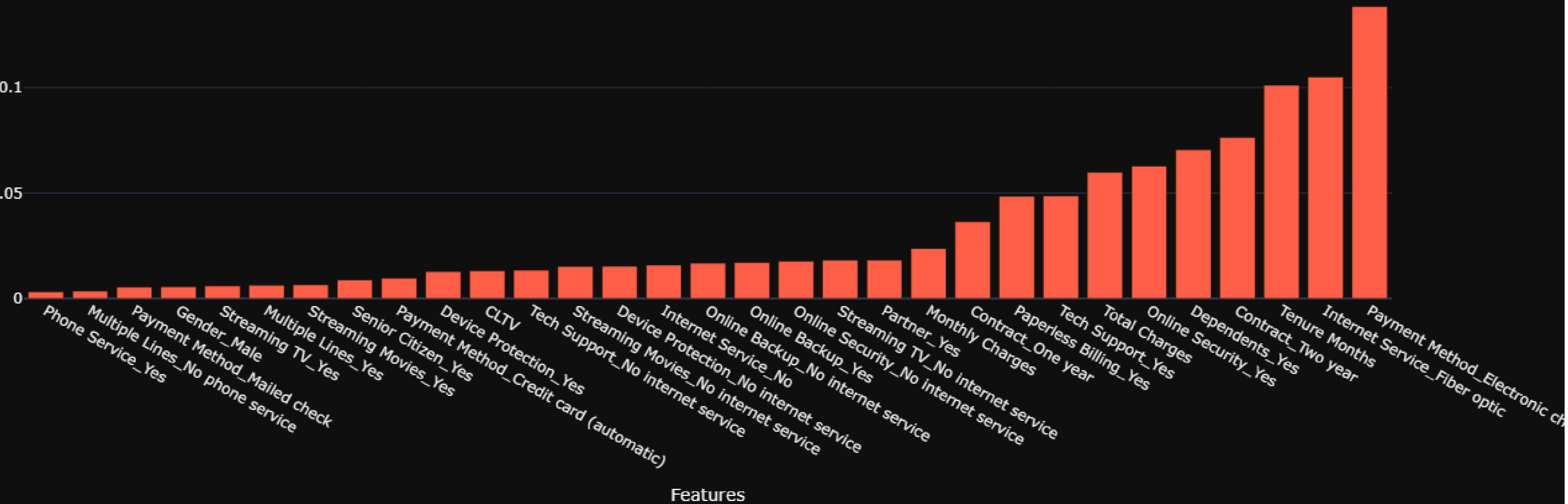
79% Accuracy

# Metrics

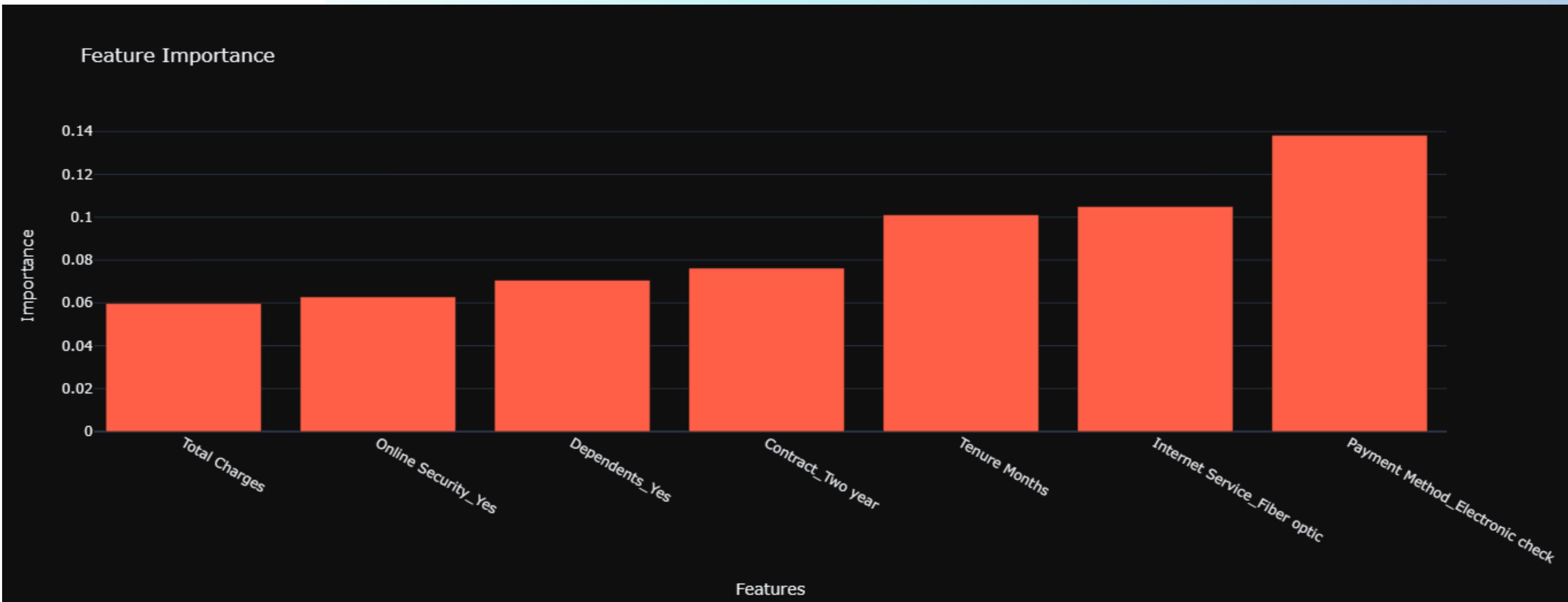
Method	Churn	Precision	Recall	F1-score	Accuracy
LogisticRegression	No	0.81	0.75	0.78	0.79
	Yes	0.77	0.83	0.80	
RandomForest	No	0.87	0.86	0.86	0.87
	Yes	0.86	0.87	0.87	
XGBoost	No	0.86	0.89	0.87	0.87
	Yes	0.89	0.86	0.87	
KNN	No	0.84	0.70	0.76	0.79
	Yes	0.75	0.87	0.81	

# Feature Importance

Feature Importance



# Feature Importance



# DEMO



# Thank for Watching

## Q & A

