

TELCONNECT'S BATTLE AGAINST CHURN



Table of CONTENTS

01

Introduction

02

EDA

03

Data Vizz

04

Summary EDA

05

Data
Preprocessing

06

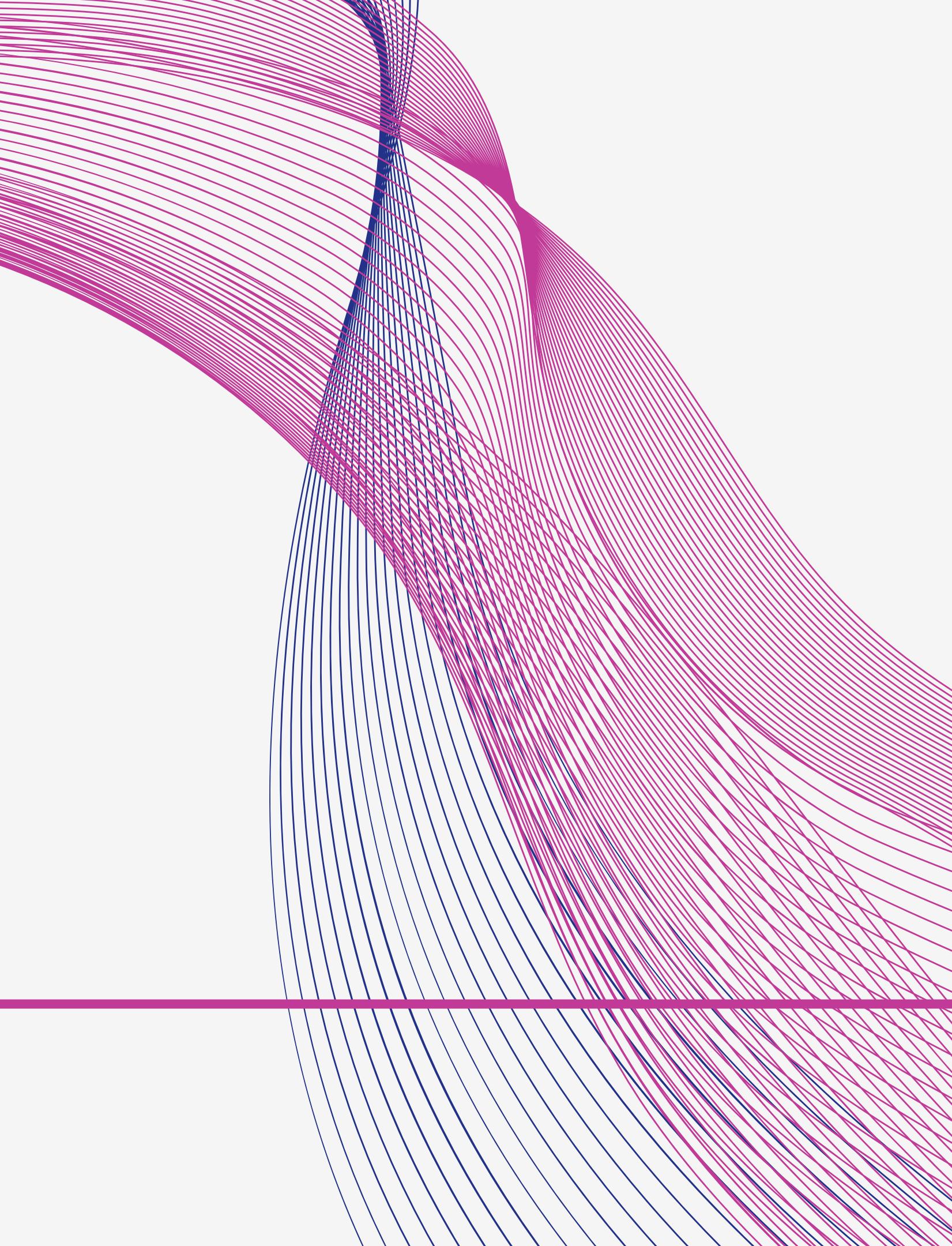
Data Modeling

07

Accuracy Compare

08

Thank You



INTRODUCTION

- TelConnect has been experiencing increasing customer churn rates.
 - Our goal is to analyze churn patterns and develop predictive models to improve retention.
 - We will use machine learning techniques to identify key churn drivers and optimize retention strategies
-

EDA

- The dataset consists of 7043 customer records with 21 features.
- Key features include demographics, service subscriptions, tenure, billing details, and churn status.
- Objective: Identify high-risk customers and predict churn likelihood.

Objective 1

- Check the function of all the columns and change the datatype if necessary

Objective 2

- Remove the duplicates and check the null values

Objective 3

- drop the column which is not necessary

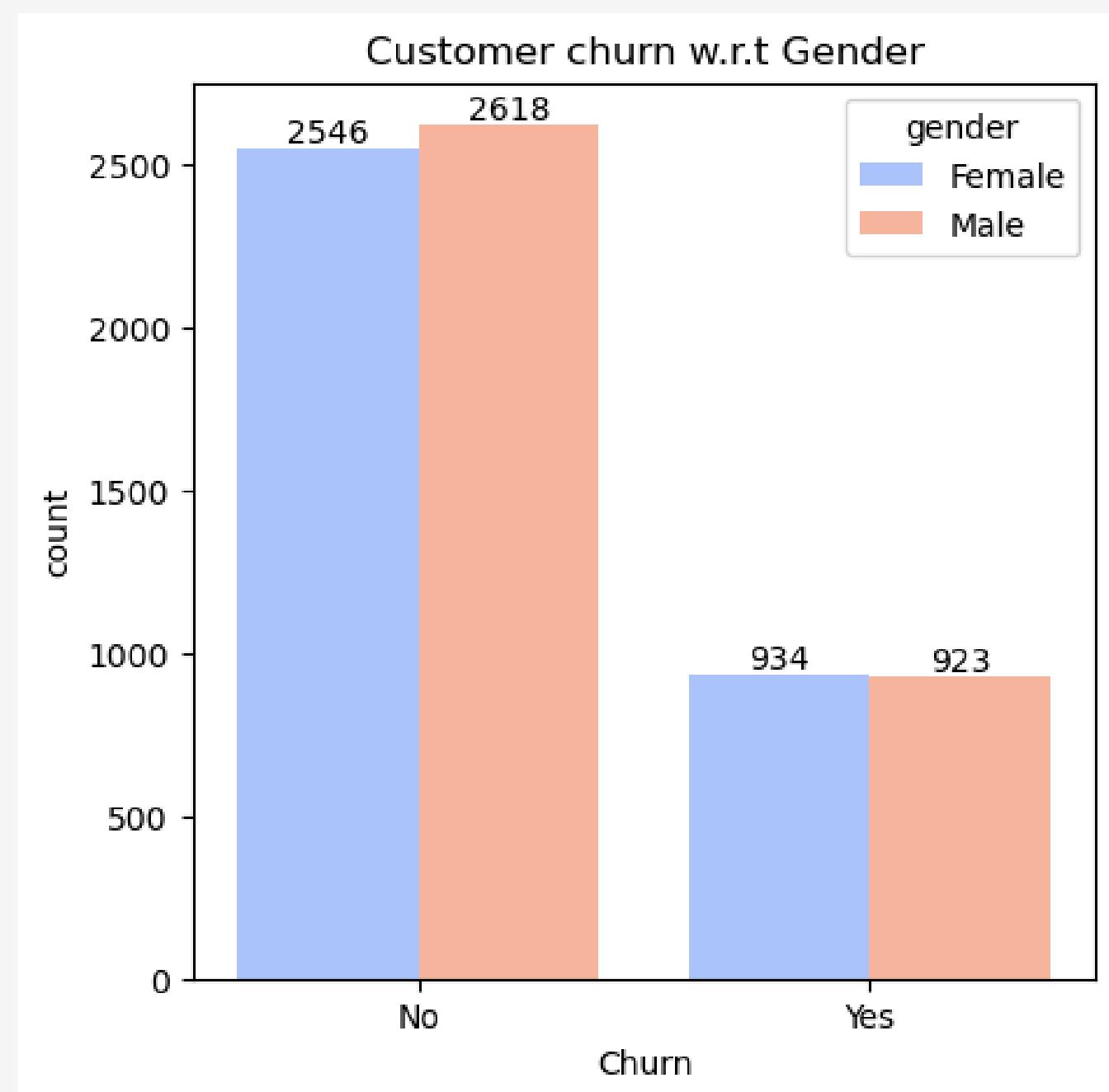


Data VISUALIZATION

Let's present the key visual insights derived from the dataset. Exploratory Data Analysis (EDA) revealed significant patterns influencing customer churn. Heatmaps highlighted correlations between tenure, monthly charges, and service subscriptions. Bar charts showed that customers on monthly contracts and higher billing plans had a higher churn rate. The confusion matrix demonstrated our model's effectiveness in classifying churned and retained customers, while the ROC-AUC curve validated its predictive power. These visualizations helped identify key churn drivers, allowing TelConnect to develop targeted retention strategies.

WHICH OF THE CUSTOMERS ARE MORE LIKELY TO CHURN(MALE,FEMALE)

CONCLUSION --> THIS SHOWS THAT BOTH THE MALE AND FEMALE CUSTOMERS EQUALLY STOPS USING THE SERVICE WITH A NEGLIGIBLE DIFFERENCE BETWEEN THE NUMBERS.

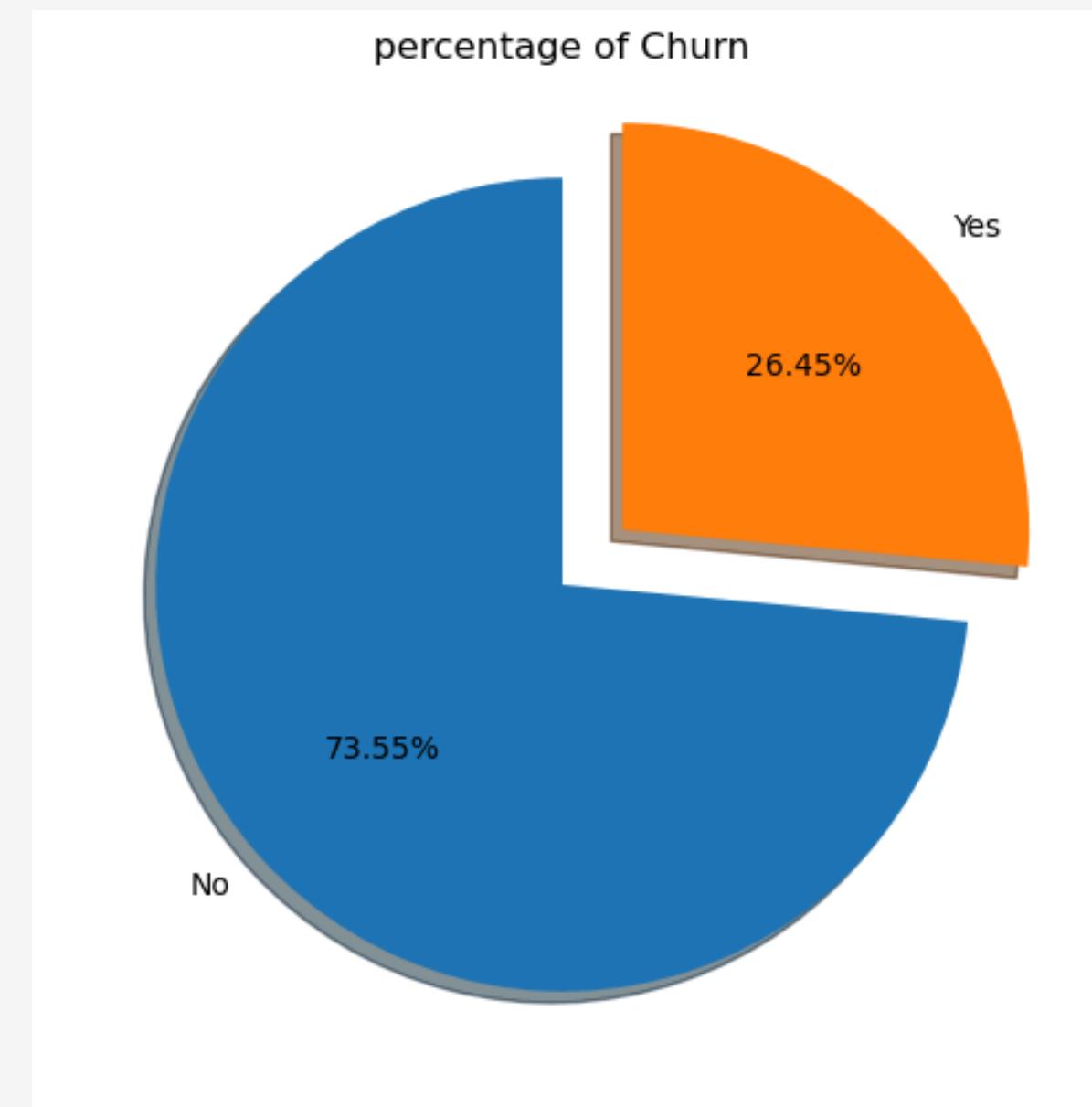


WHAT IS THE PERCENTAGE OF CHURN CUSTOMERS AND ACTIVE CUSTOMERS ?

CONCLUSION:

26.45% ARE CHURNED CUSTOMERS.

73.55% ARE ACTIVE CUTOMERS.

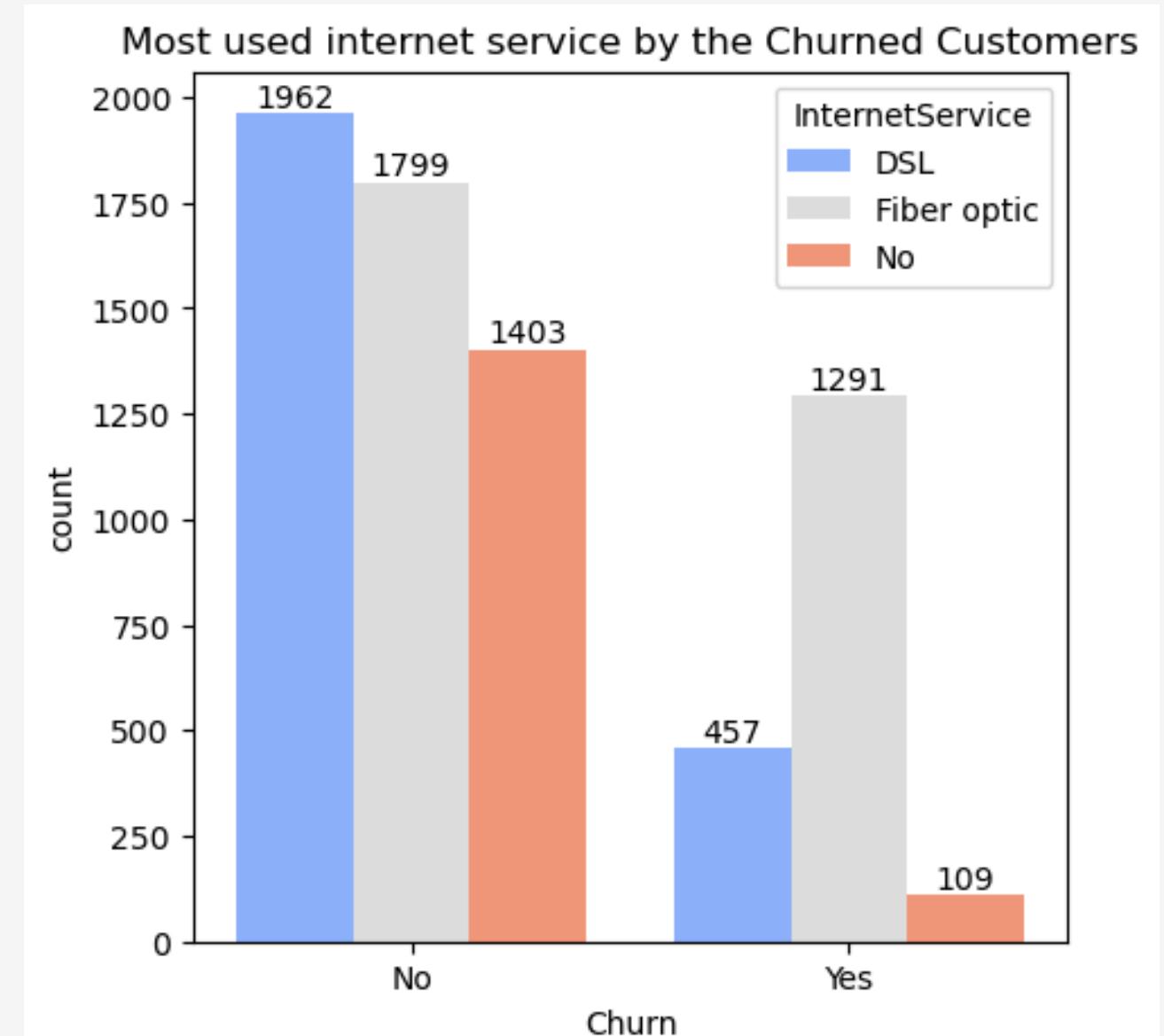


MOST USED INTERNET SERVICE BY THE CHURNED CUSTOMERS

CONCLUSION :

A LOT OF CUSTOMERS CHOOSE THE FIBER OPTIC SERVICE AND IT'S ALSO EVIDENT THAT THE CUSTOMERS WHO USE FIBER OPTIC HAVE HIGH CHURN RATE, THIS MIGHT SUGGEST A DISSATISFACTION WITH THIS TYPE OF INTERNET SERVICE.

CUSTOMERS HAVING DSL SERVICE ARE MAJORITY IN NUMBER AND HAVE LESS CHURN RATE COMPARED TO FIBRE OPTIC SERVICE.

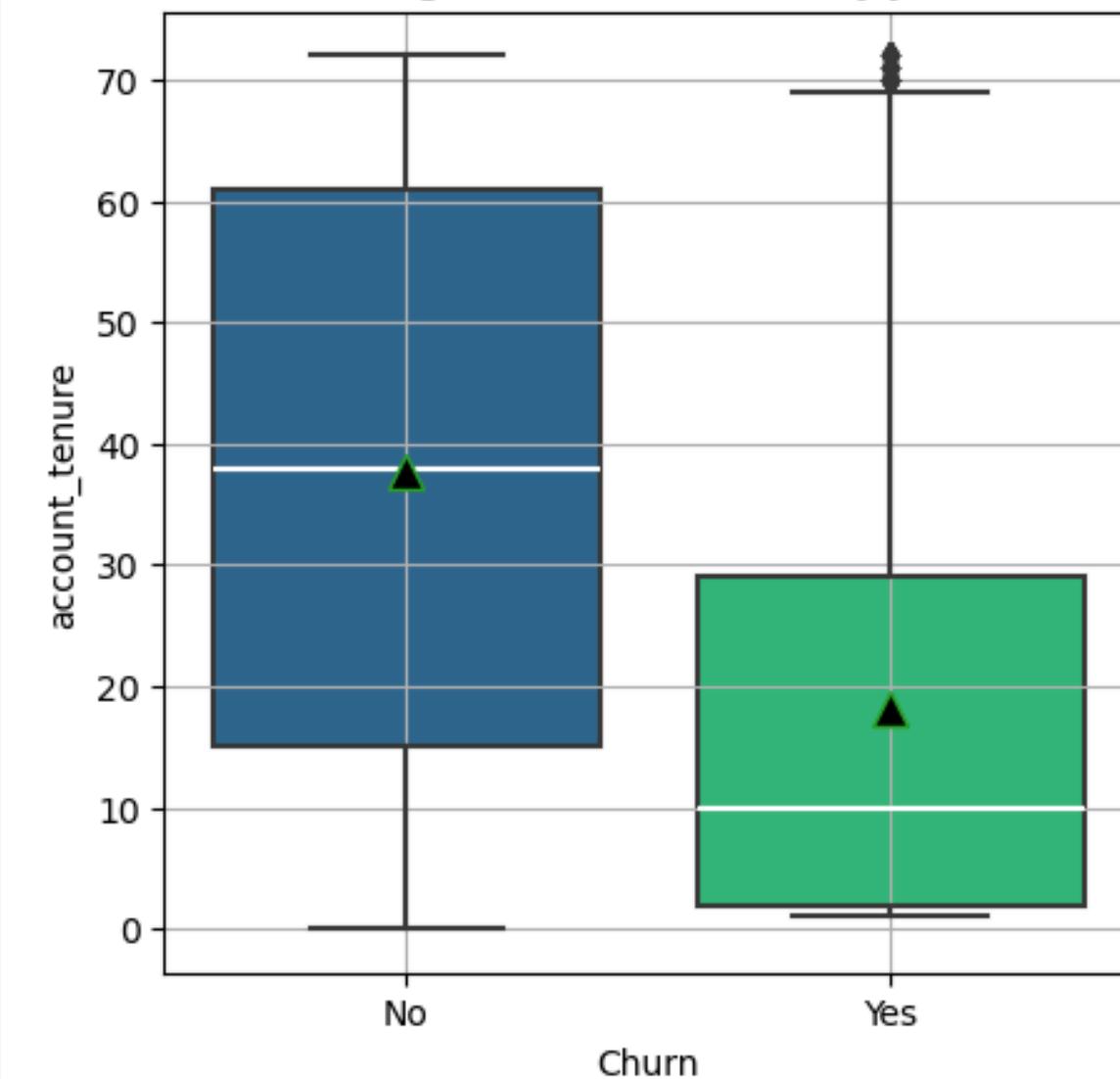


WHICH OF THE CUSTOMERS ARE MORE LIKELY TO CHURN EXISTING OR NEWLY JOINED

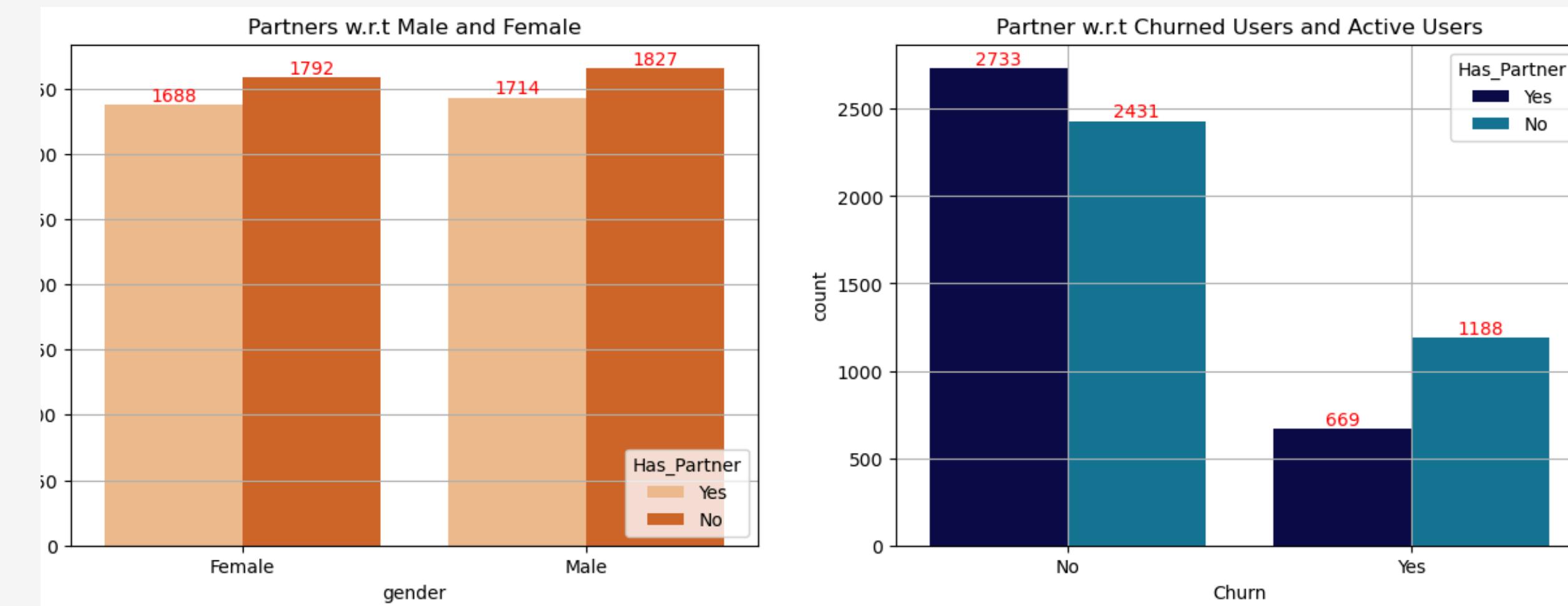
CONCLUSION:

THE LOWEST MEAN(SMALL TRIANGLE) IN CHURNED CUSTOMERS PLOT TELLS THAT THE NEW CUSTOMERS ARE MORE LIKELY TO UNSUBSCRIBE THE FIRM/SERVICE.

Tenure w.r.t Existing Customers vs Newly Joined Customers.



WHO HAVE MORE PARTNERS(MALE/FEMALE) AND IS USERS WITH PARTNERS ARE MORE LIKELY TO CHURN OR NOT HAVING ANY PARTNER ?

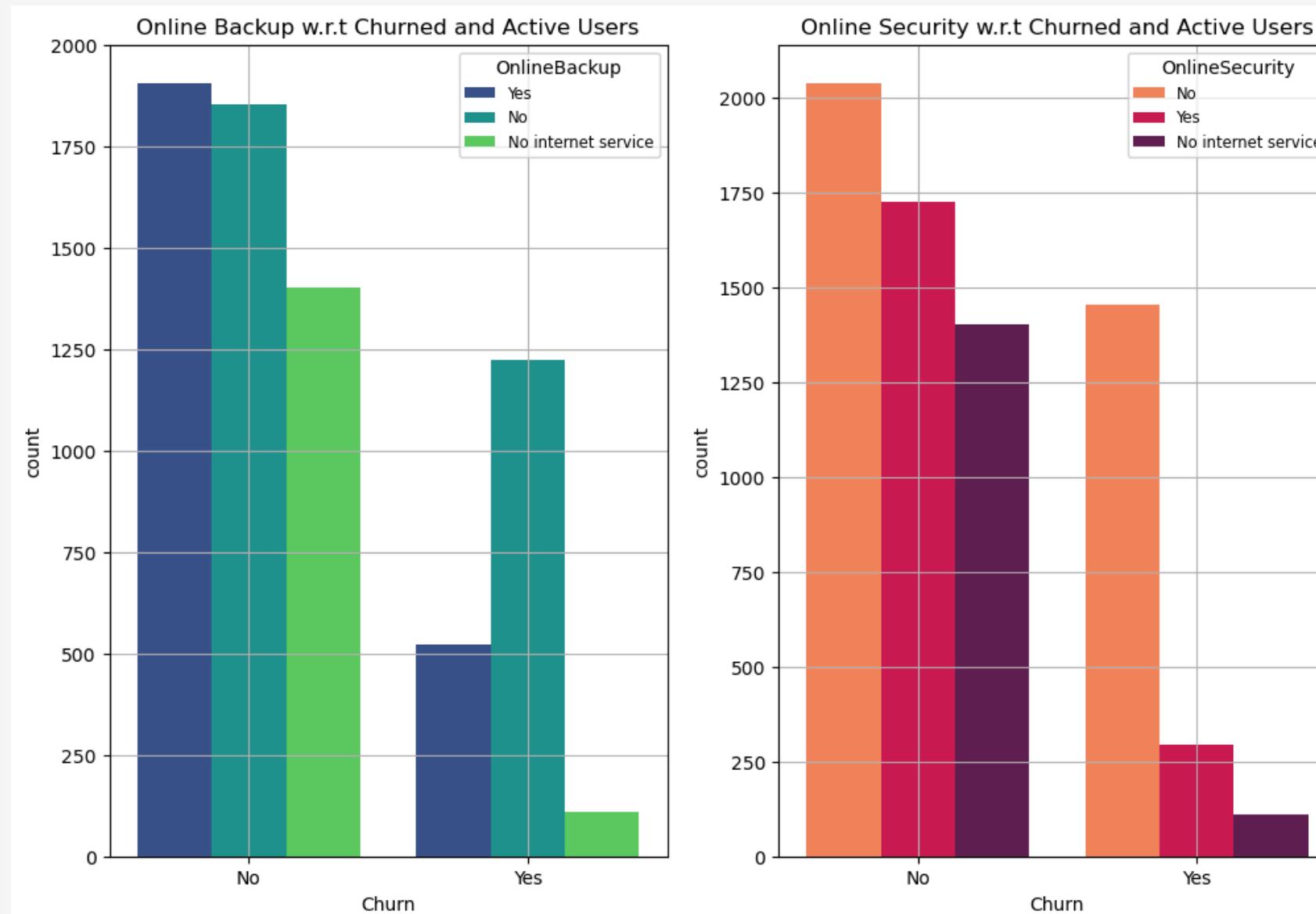


CONCLUSION :

THE 1ST PLOT SUGGEST THAT, MALE HAVE 1714 PARTNERS AND FEMALE HAVE 1688 PARTNERS.

THE 2ND PLOT SUGGEST THAT, USERS HAVING NO PARTNERS ARE MORE LIKELY TO CHURN.

DOES ONLINE SECURITY AND ONLINE BACKUP IS A CONCERN FOR THE CHURNED USERS



CONCLUSION :

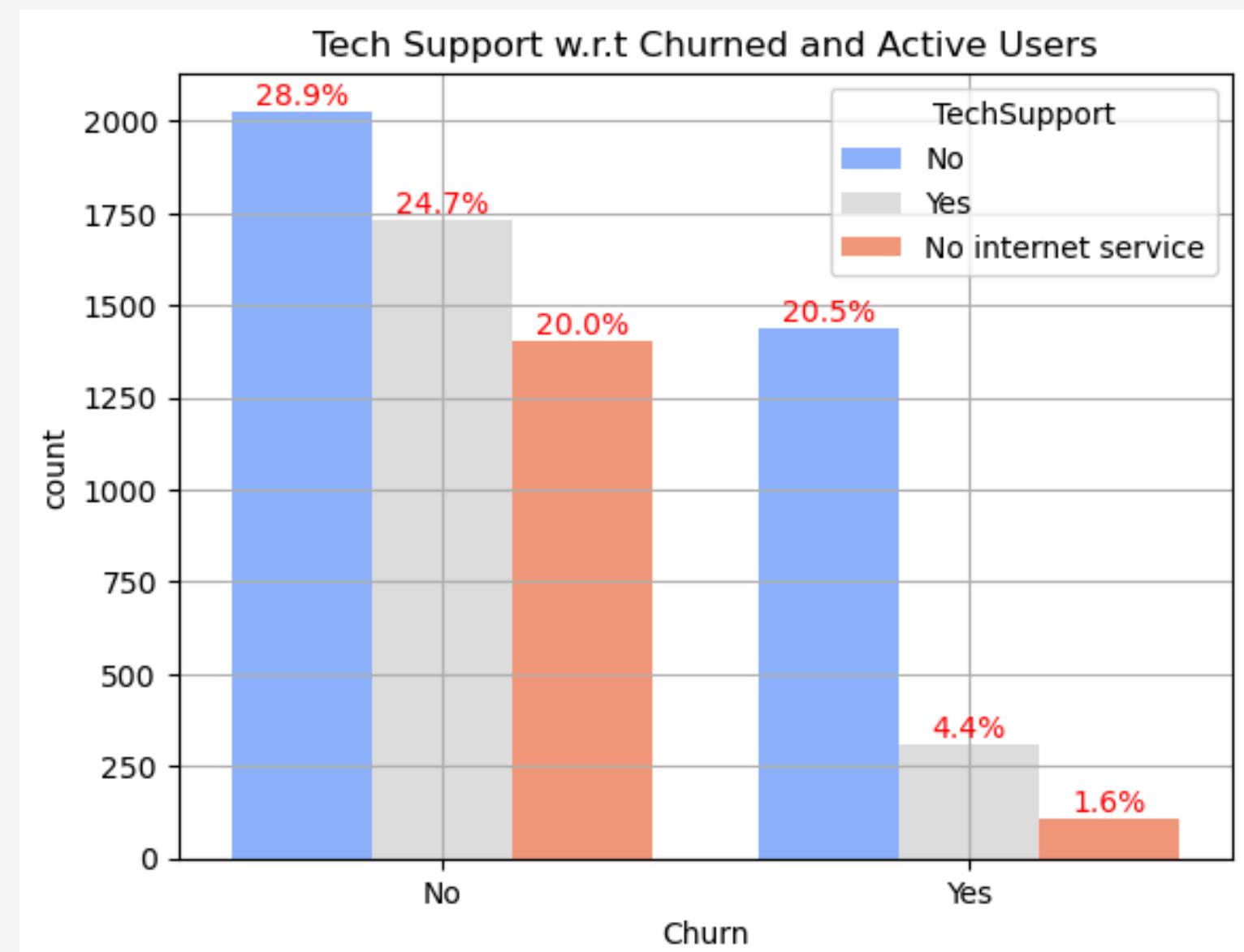
USERS WITH LESS ONLINE SECURITY ARE MORE LIKELY TO CHURN

USERS WITH LESS ONLINE BACKUP ARE MORE LIKELY TO CHURN.

DOES TECH SUPPORT FEATURE, A CONCERN FOR THE CHURNED CUSTOMERS

CONCLUSION:

USERS WITH NO TECH SUPPORT ARE MORE LIKELY TO CHURN.

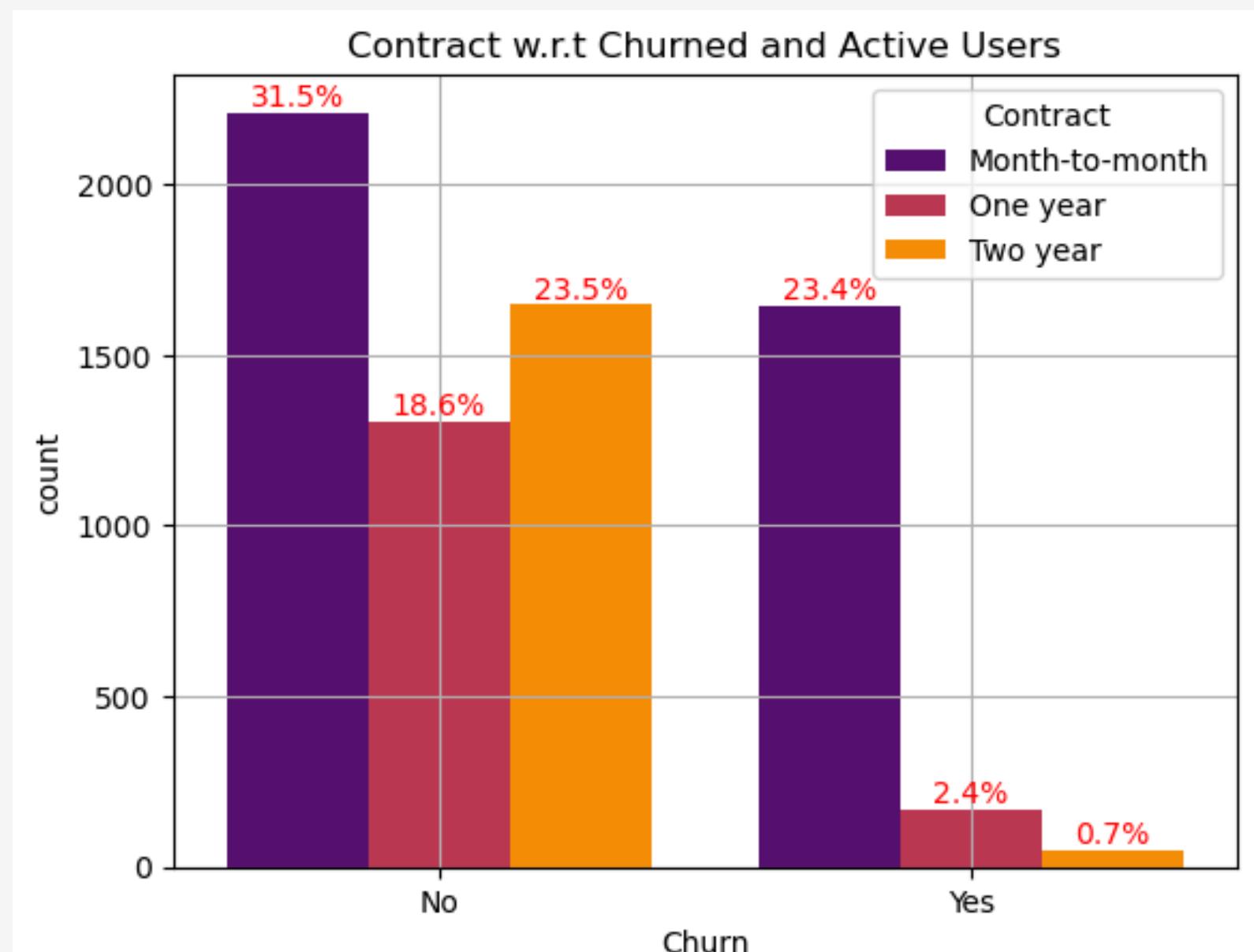


DOES SERVICE CONTRACTS REDUCE ACTIVE CUSTOMERS ?

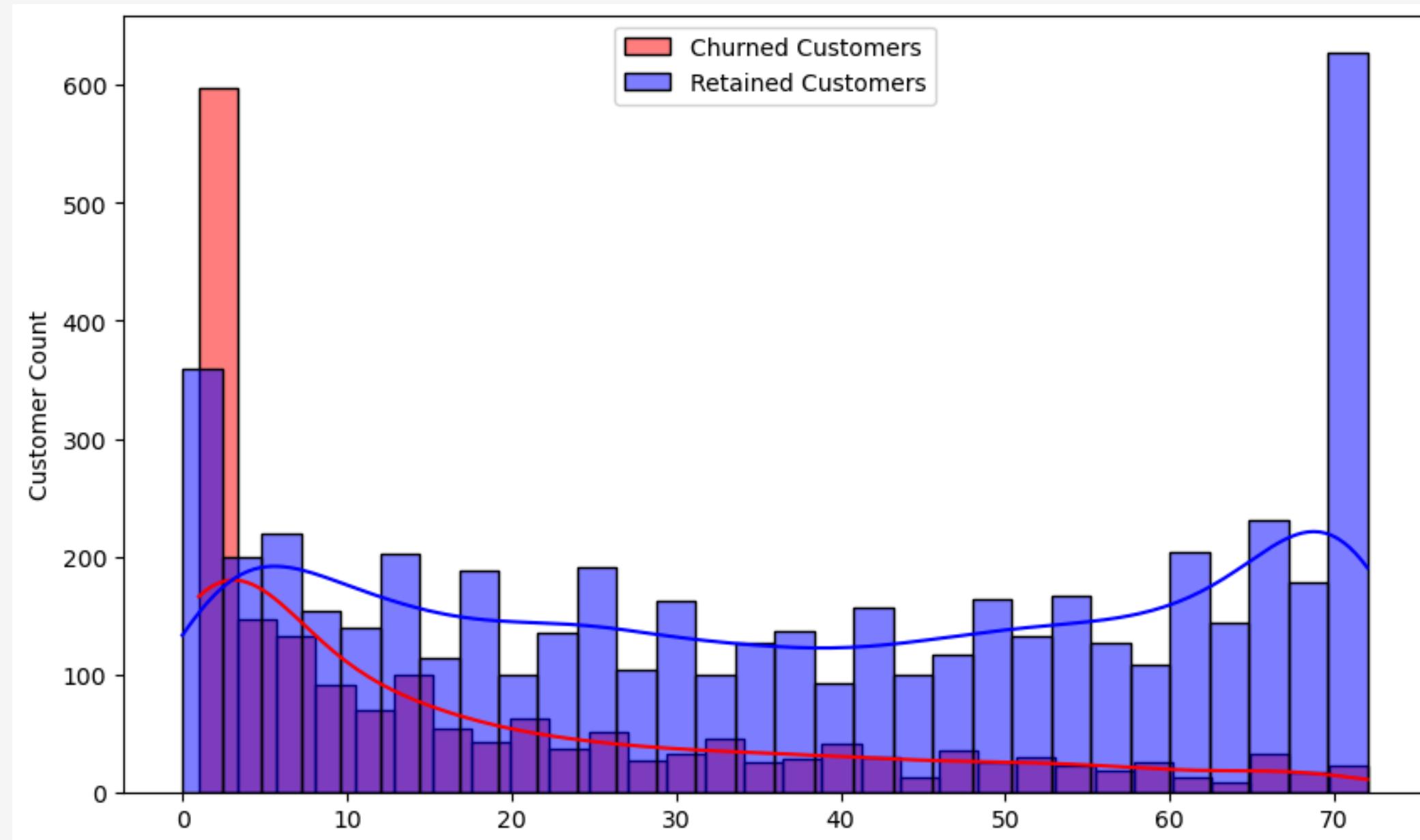
CONCLUSION :

USERS WITH MONTH_TO_MONTH CONTRACTS ARE MORE LIKELY TO CHURN WHILE

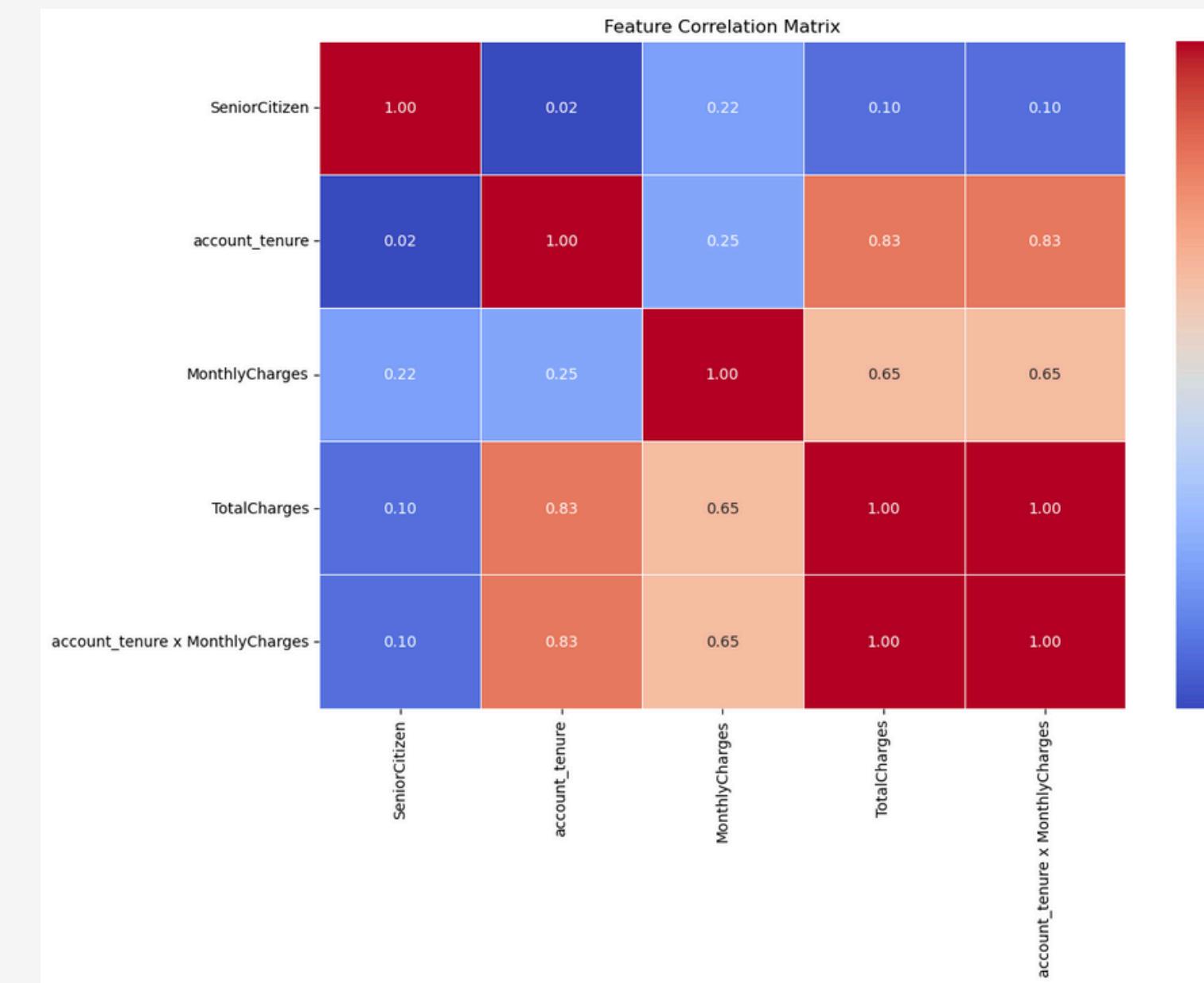
USERS WITH ONE YEAR AND TWO YEAR CONTRACTS ARE LESS LIKELY TO CHURN



CHURN DISTRIBUTION & CUSTOMER BEHAVIOR ANALYSIS



CORRELATION MATRIX TO UNDERSTAND FEATURE RELATIONSHIPS



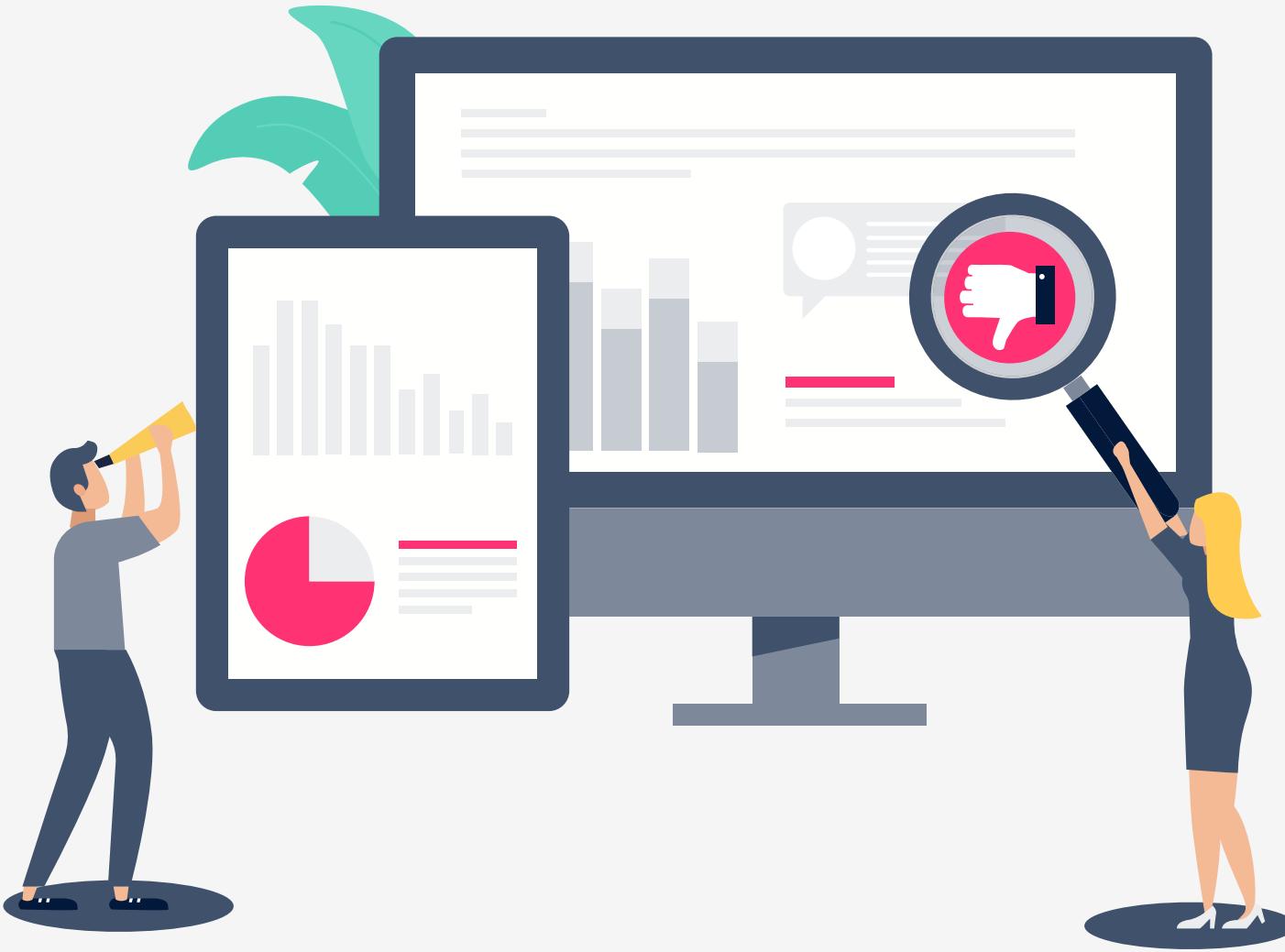
Summary EDA

Order / Values of features for customer churn cases :

- Categorical Features (Order) :
 - gender : Male = Female
 - SeniorCitizen : No SeniorCitizen > SeniorCitizen
 - Partner : No Partner > Partner
 - Dependents : No Dependent > Dependent
 - PhoneService : PhoneService > No PhoneService
 - MultipleLines : MultipleLines > No MultipleLines > No PhoneService
 - InternetService : Fiber Optic > DSL > No InternetService
 - OnlineSecurity : Absent > Present > No InternetService
 - OnlineBackup : Absent > Present > No InternetService
 - DeviceProtection : Absent > Present > No InternetService
 - TechSupport : Absent > Present > No InternetService
 - StreamingTV : Absent > Present > No InternetService
 - StreamingMovies : Absent > Present > No InternetService
 - Contract : Month-to-Month > One year > Two year
 - PaperlessBilling : Present > Absent
 - PaymentMethod : Electronic check > Mailed check > Bank Transfer (automatic) > Credit Card (automatic)!
- Numerical Features (Range) :
 - tenure : 1 - 5 months
 - MonthlyCharges : 65 - 105
 - TotalCharges : 0 - 1000



Data PREPROCESSING



1 Handled Missing Values

- Checked for missing data and replaced TotalCharges missing values with the median.

2 Dropped Irrelevant Columns

- Removed customerID since it is not a predictive feature.

3 Encoding Categorical Variables

- Applied Label Encoding for binary categorical columns (e.g., gender, Has_Partner).
- Used One-Hot Encoding for multi-category variables (e.g., InternetService, Contract).

4 Feature Scaling

- Standardized numerical columns (MonthlyCharges, TotalCharges, account_tenure) using StandardScaler to improve model convergence.

5 Balanced the Dataset

- Used stratified train-test split to maintain class balance.

6 Correlation Analysis & Feature Selection

- Dropped highly correlated or redundant features to avoid multicollinearity.
- Selected the most relevant features for model training.

Data MODELING

Model	Accuracy (%)	Precision (%)	F1-Score (%)	Recall (%)
Stacking Classifier	84.4%	82.5%	83.8%	85.2%
XGBoost	83.6%	80.9%	83.1%	85.5%
AdaBoost	83.2%	80.4%	82.9%	85.5%
Gradient Boosting	83.1%	80.9%	82.4%	84.1%
Random Forest	81.4%	80.1%	80.5%	80.9%
Logistic Regression	80.8%	78.3%	80.1%	82.0%
Decision Tree	81.1%	79.4%	80.3%	81.2%

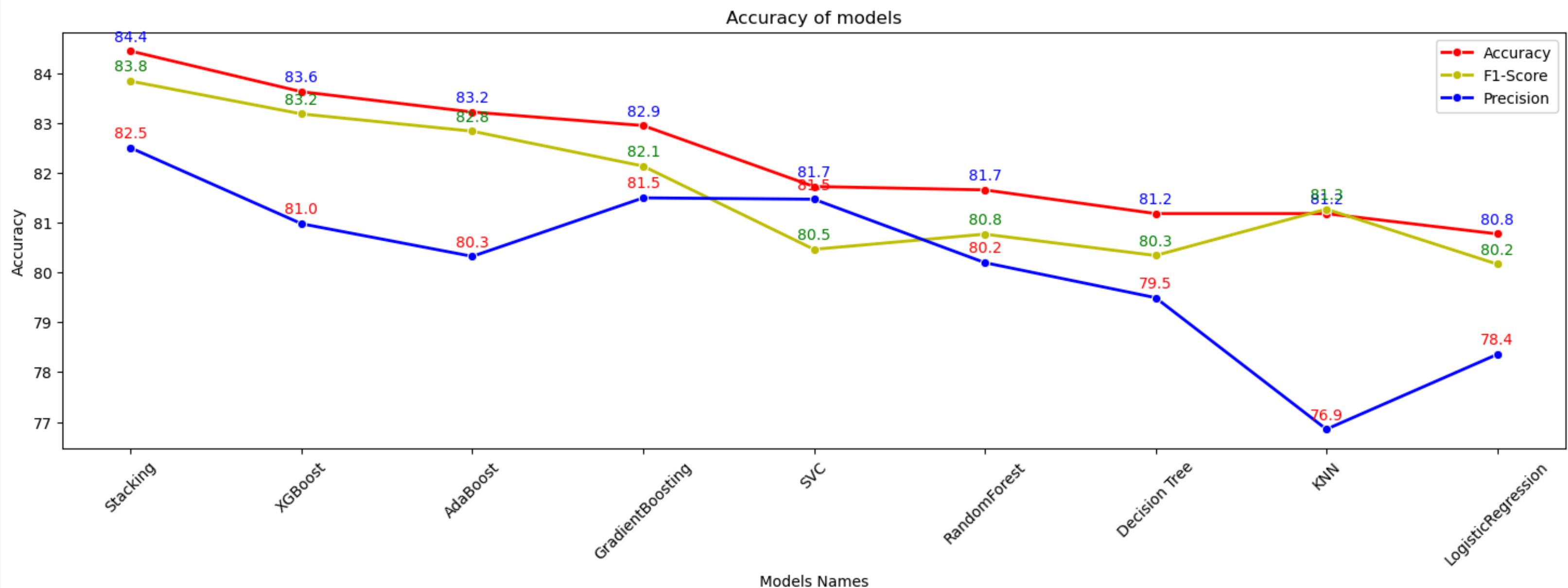
Data modeling is a critical step in the customer churn analysis for TelConnect because it helps predict which customers are likely to leave. The main reasons for data modeling in this assessment are:

- Predict Customer Churn Before It Happens
- Identify Key Drivers of Churn
- Improve Customer Retention Strategies
- Optimize Business Decisions & Revenue
- Automate & Scale Customer Insights

Evaluated multiple models using GridSearchCV:

- Logistic Regression
- Decision Tree
- Random Forest
- XGBoost
- AdaBoost
- Gradient Boosting
- Stacking Classifier (Best Model)

Accuracy of Models



ANALYSIS

TelConnect's current reactive approach to customer retention is costly and inefficient. By implementing machine learning-based churn prediction, the company can shift to a proactive strategy, reducing churn while optimizing retention costs.

1 Proactive Retention Strategy:

- Use predictive modeling to target at-risk customers before they churn.
- Offer loyalty discounts to customers with tenure < 6 months

2 Better Cost Allocation

- Reduce reliance on StreamFlix vouchers since they cost \$30 per customer.
- Instead, provide low-cost incentives like free service upgrades.

3 Customer Segmentation for Retention

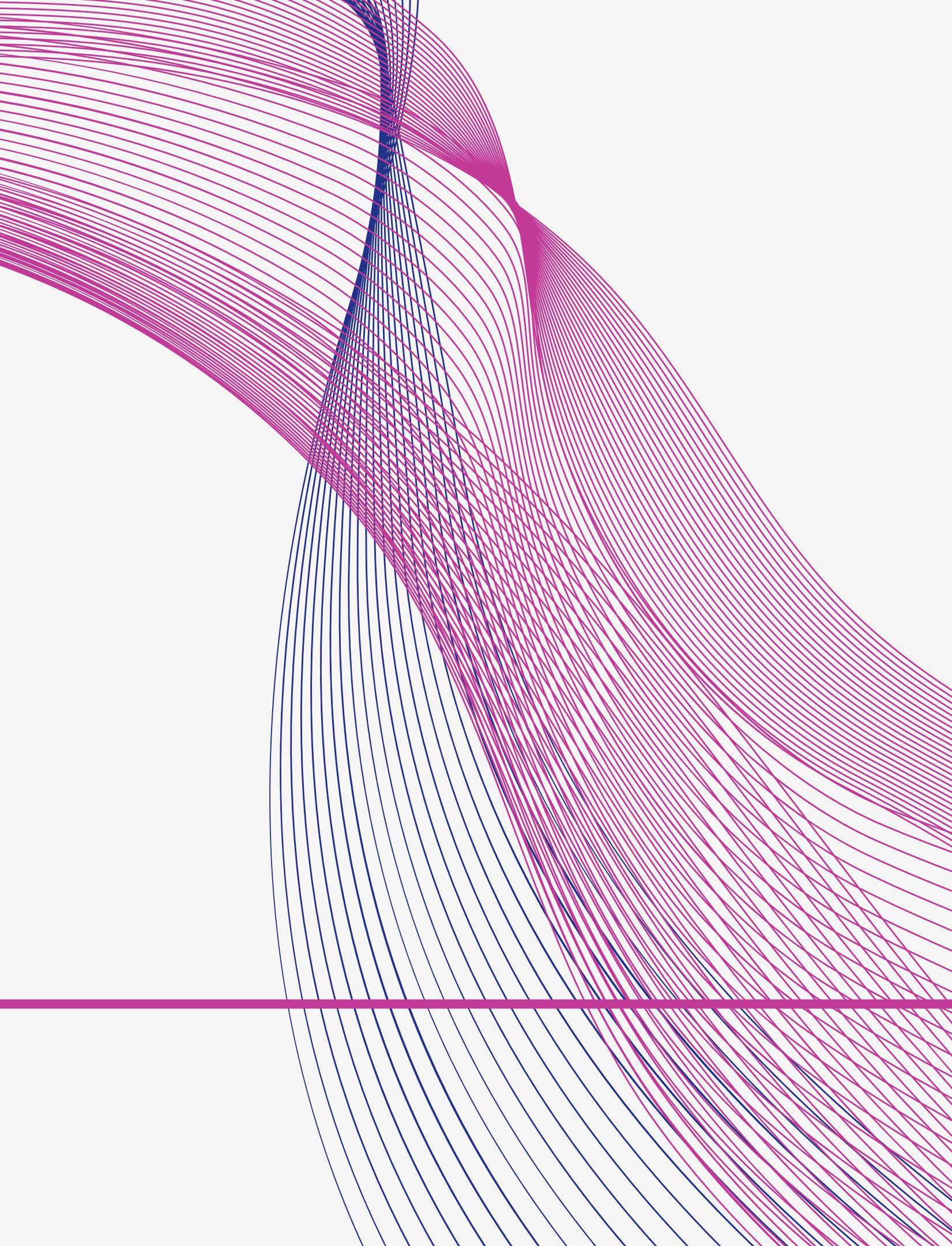
- Focus on customers with high monthly charges & low tenure
- Introduce personalized offers based on usage behavior

4 Improve Service Experience

- Increase customer support touchpoints for high-risk customers.
- Introduce a proactive engagement mode (e.g., satisfaction surveys).

5 Leverage Additional Data

- If possible, collect customer feedback data on service quality.
- Track support interaction history to enhance churn predictions.



THANK YOU

