



SYRIATEL

CUSTOMER CHURN PREDICTION MODEL

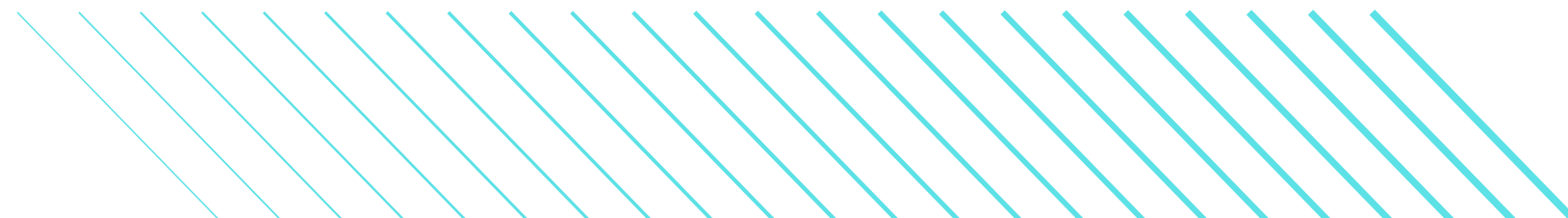


TABLE OF CONTENTS

- INTRODUCTION
- BUSINESS UNDERSTANDING
- METRIC FOR SUCCESS
- EXPLORATORY DATA ANALYSIS
- MODELLING
- LIMITATIONS
- RECOMMENDATIONS

INTRODUCTION

Customer retention refers to keeping existing customers loyal to a business by providing excellent service and experiences. It is important because retaining existing customers is more cost-effective than acquiring new ones, they can provide valuable feedback and referrals, and they are more likely to make repeat purchases.



BUSINESS UNDERSTANDING

PROBLEM STATEMENT

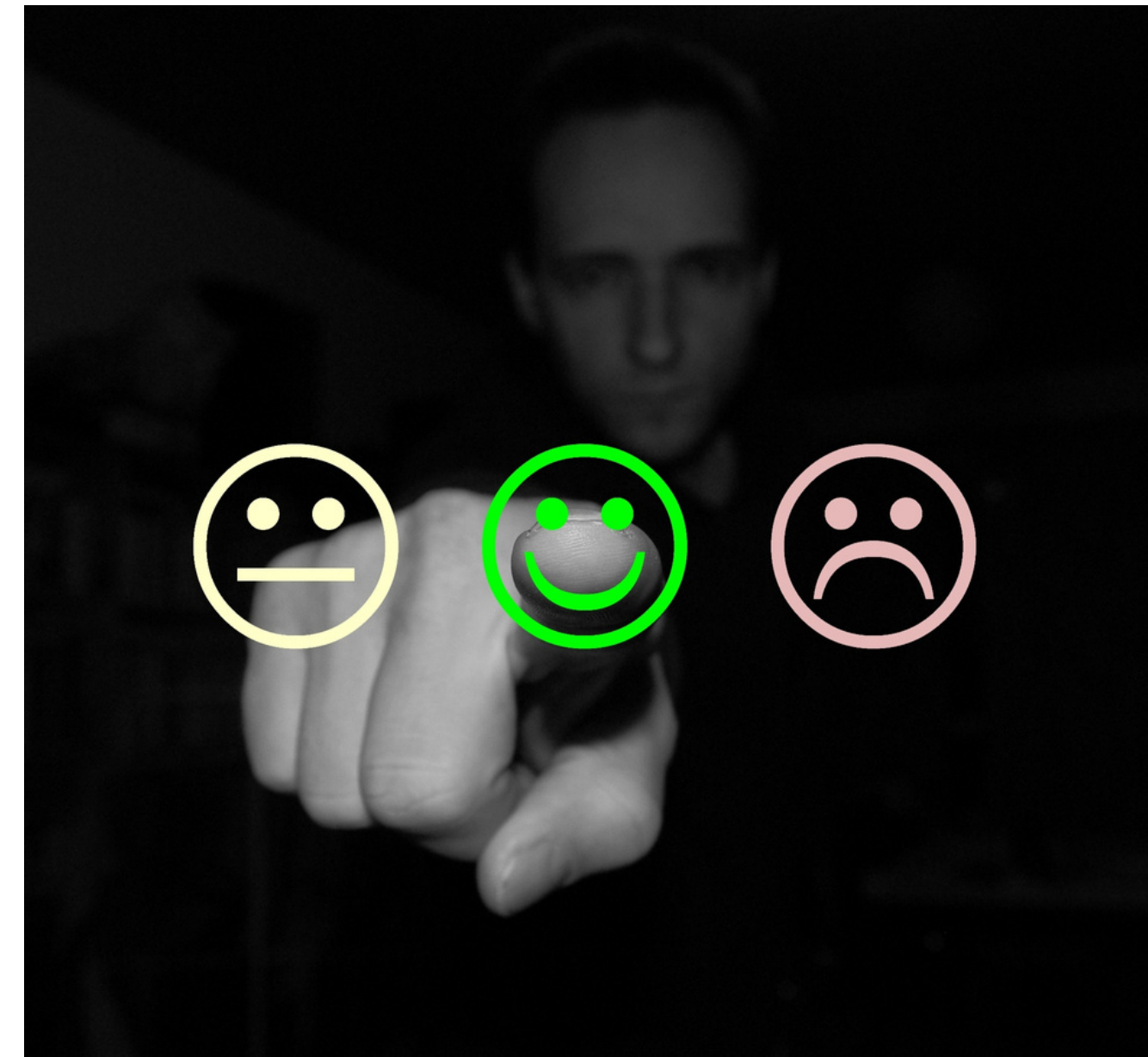
Customer churn is the loss of clients over time. Companies must pay close attention to it as it affects revenue directly. When customers depart, the company not only loses their current business but also future potential business. Also, gaining new clients is often costlier than keeping current ones, reducing churn saves costs. Monitoring and lowering churn is crucial for companies to retain and grow their customer base, ultimately increasing revenue. Improving customer satisfaction, creating customer loyalty, and growing the business can be achieved by identifying reasons for churn and retaining customers.

MAIN OBJECTIVE

The main objective of this project is to come up with a predictive model that will be able to flag customers who are likely to churn in order to have targeted mitigating actions focused on them in order to prevent them from churning

OTHER OBJECTIVES

- Determine the most important predictor of customer churn
- Look for potential areas of improvement for our model



METRIC FOR SUCCESS



1

Prediction accuracy of the model is above 90%



2

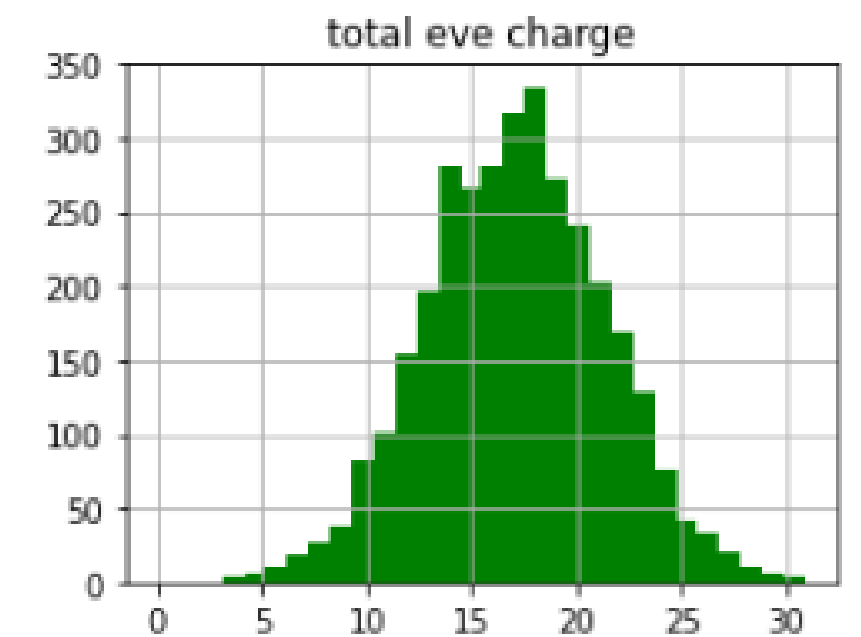
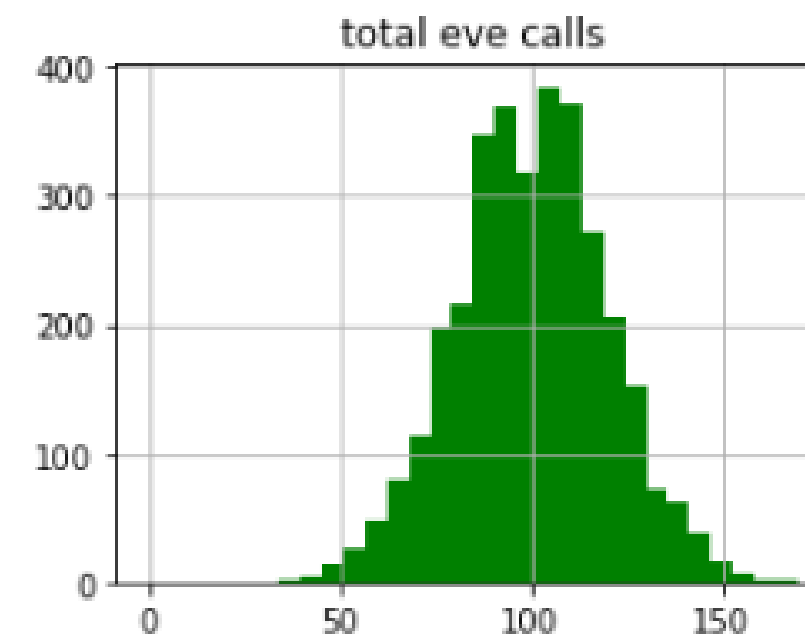
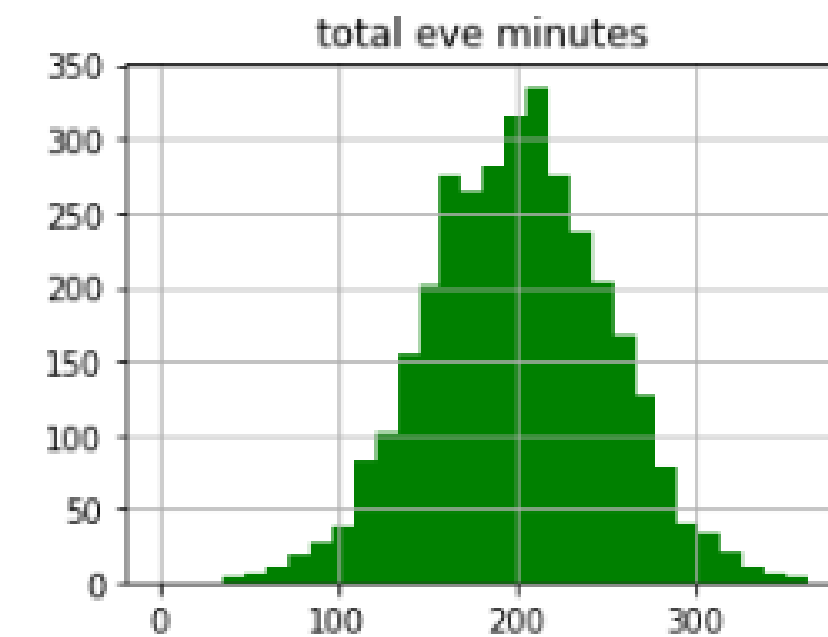
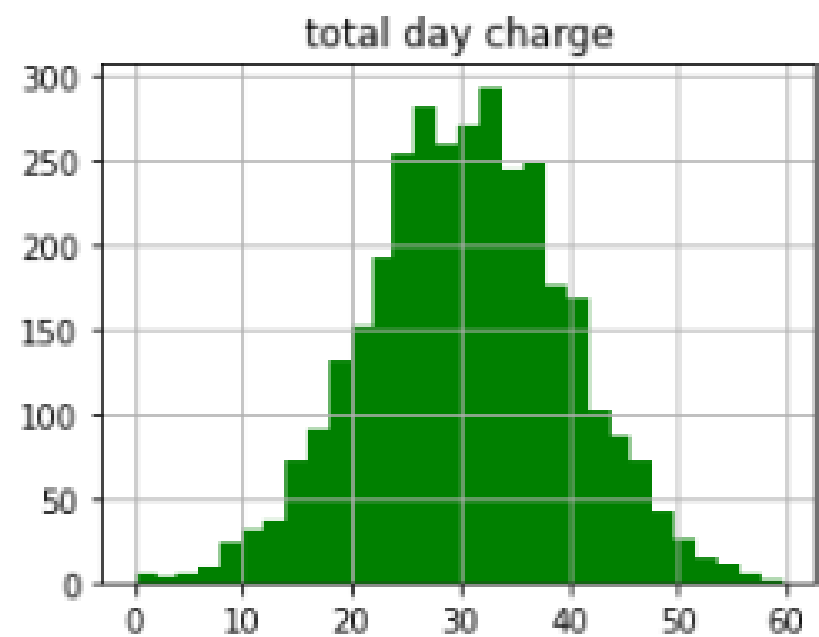
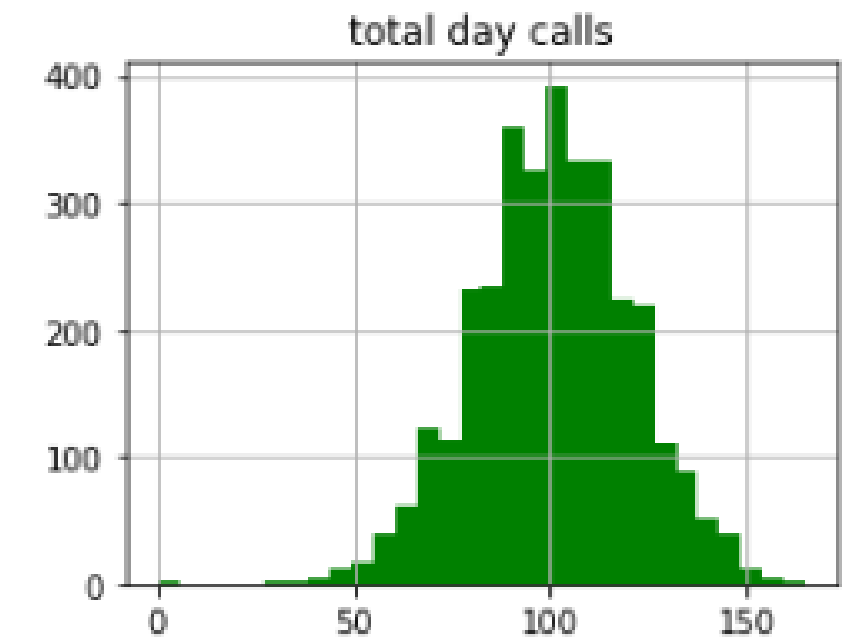
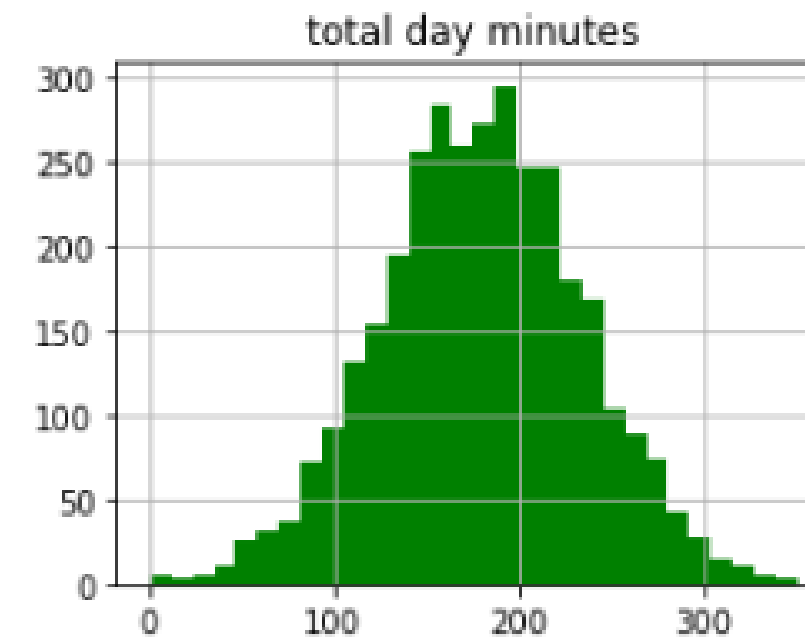
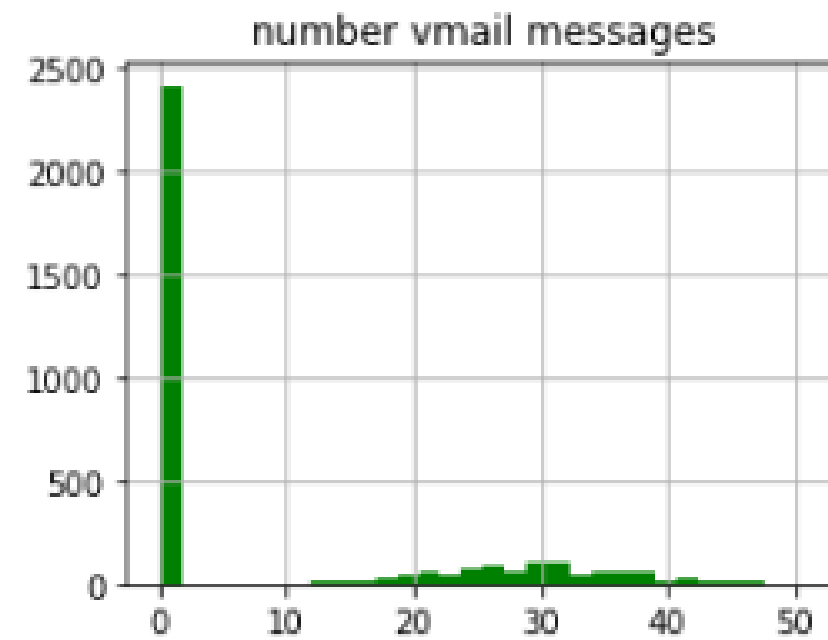
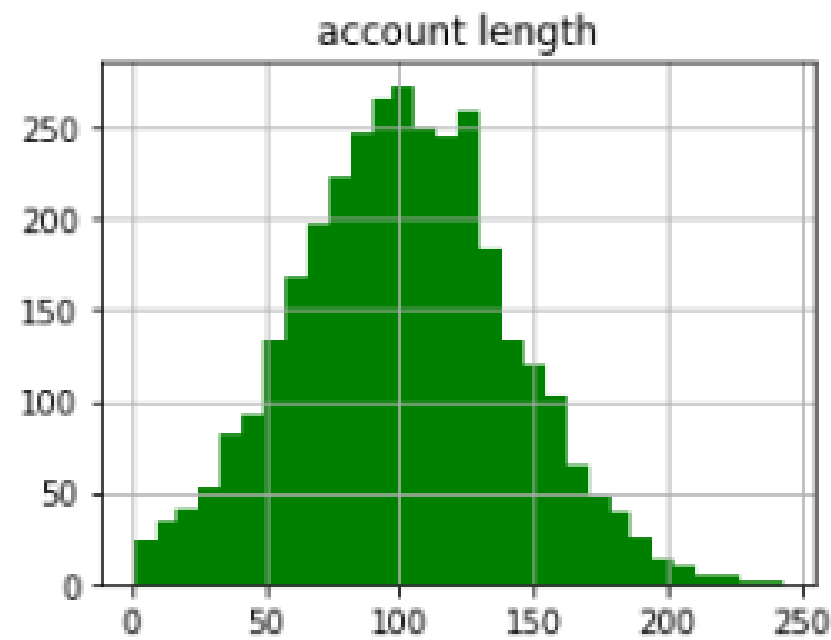
Recall score is above 75%

EXPLORATORY DATA ANALYSIS

What does the data tell us?

What patterns from the data can we observe

DISTRIBUTION OF DATA FOR CONTINUOUS DATA



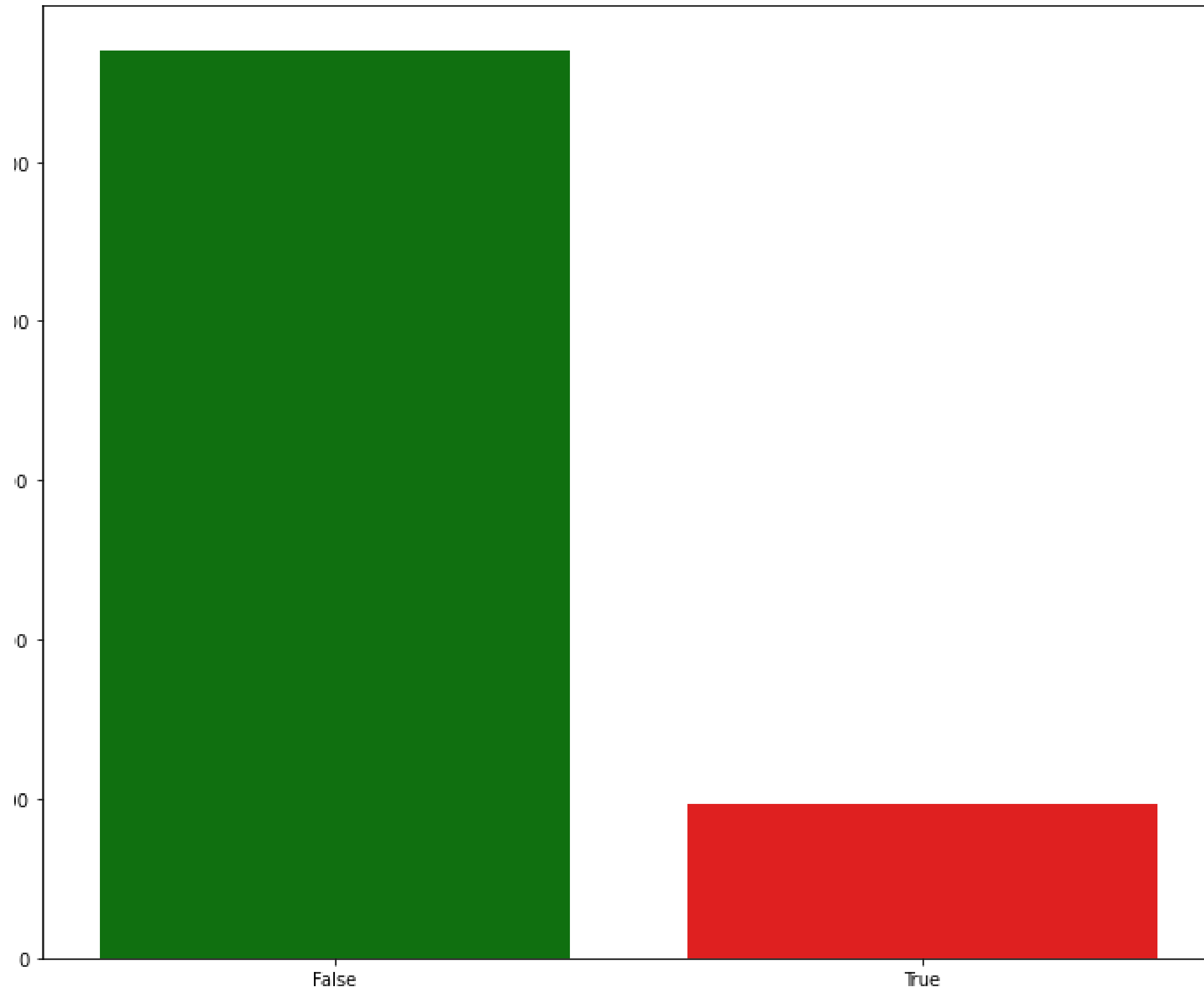
Data is normally distributed (there are few small and large values and many average values)



Data is normally distributed (there are few small and large values and many average values)

PROPORTION OF CUSTOMER CHURN

Churn vs non-churn

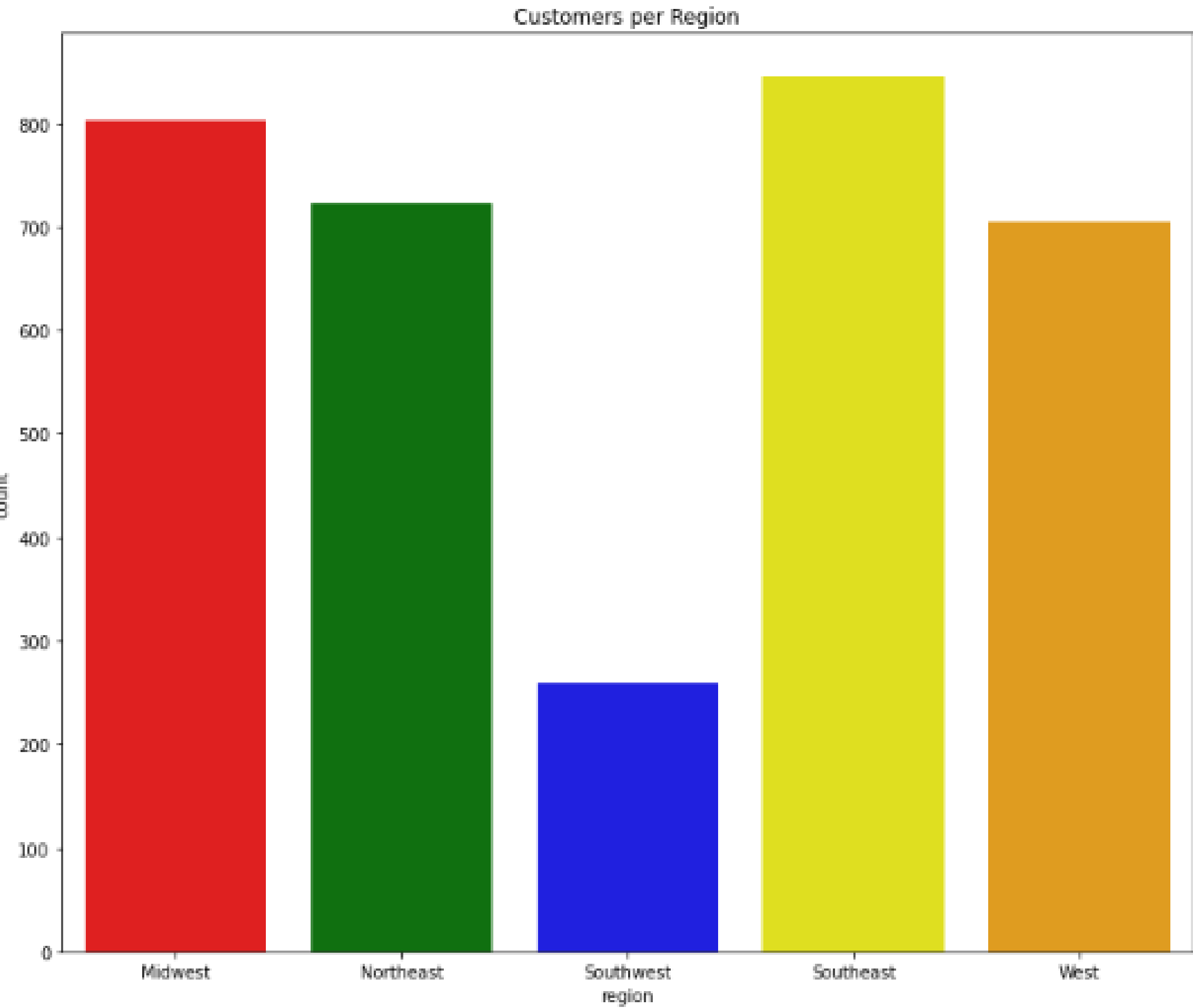


86%



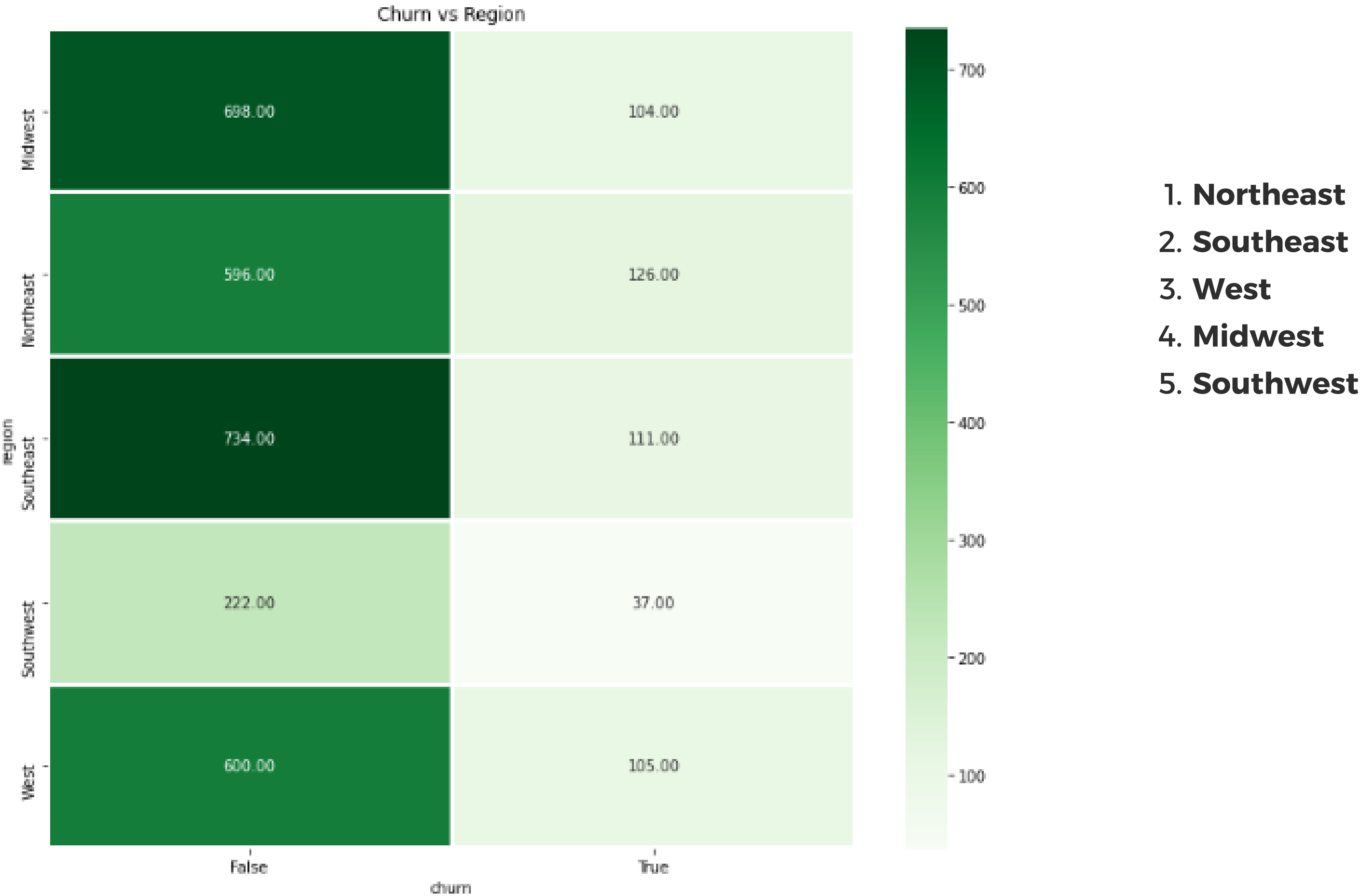
14%

CUSTOMERS PER REGION

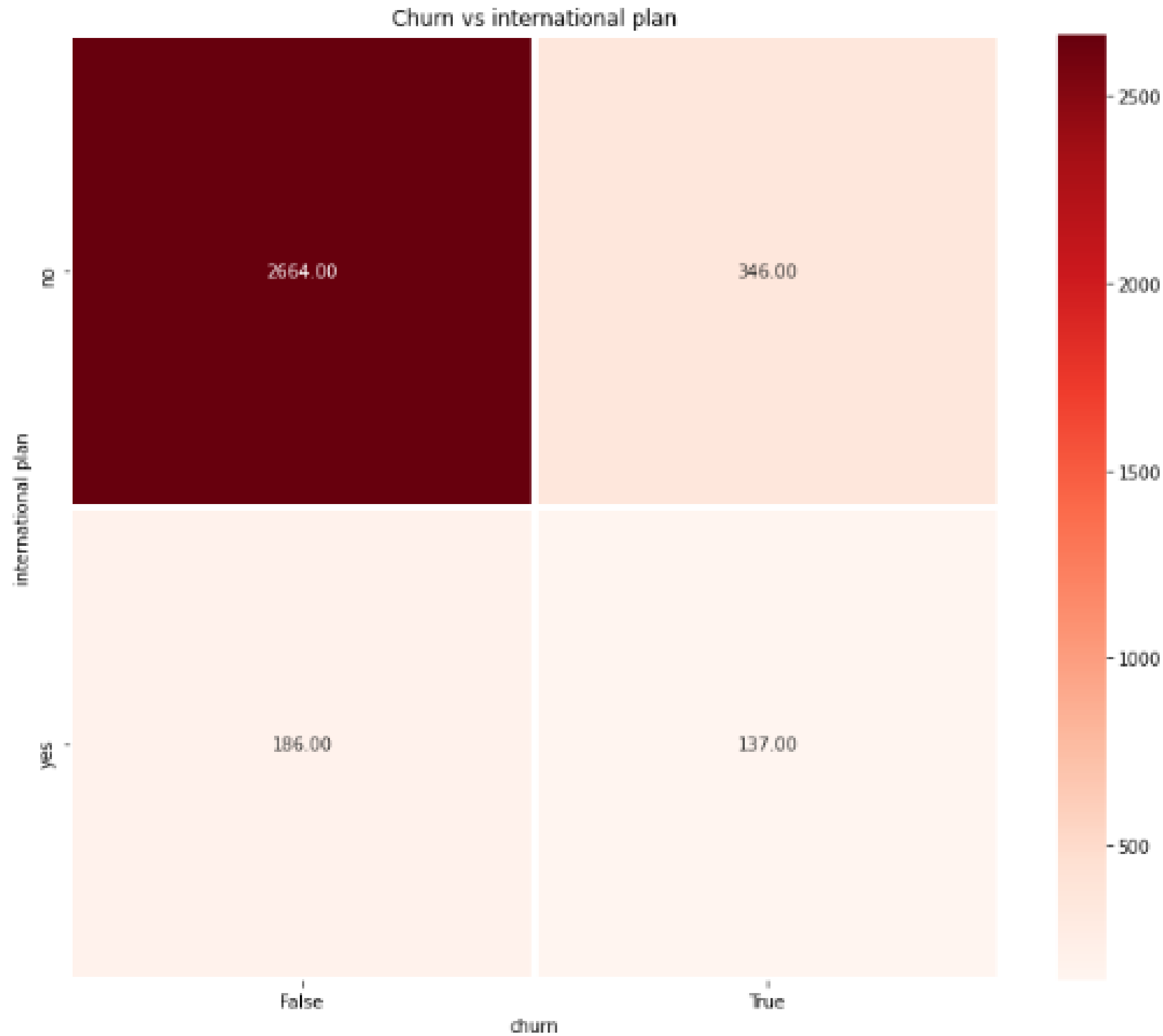


- 1. SouthEast
- 2. Midwest
- 3. Northeast
- 4. West
- 5. Southwest

CHURN RATE PER REGION

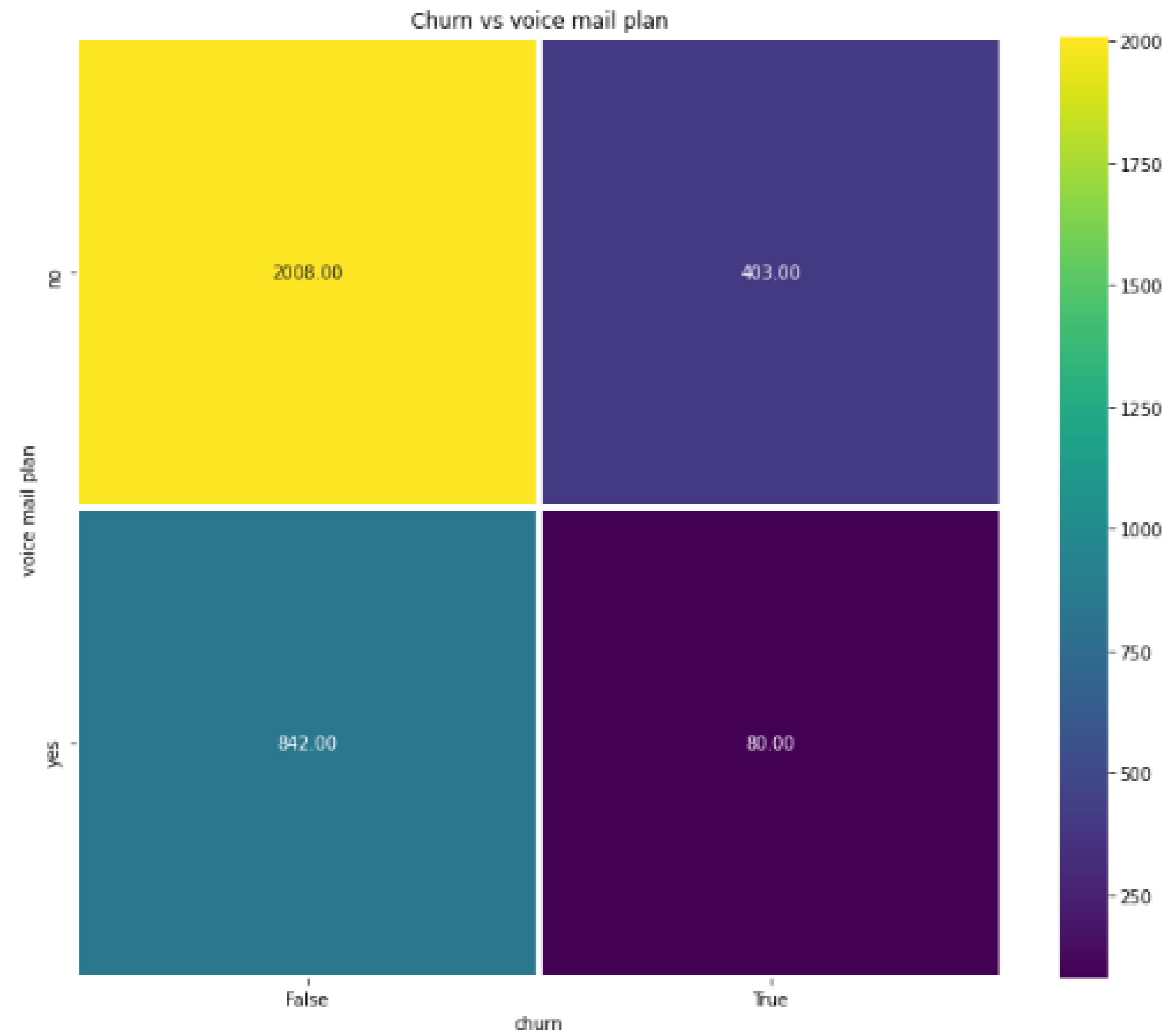


CHURN VS INTERNATIONAL PLAN



Customers who do not have an international plan are twice likely to churn that people who have an international plan

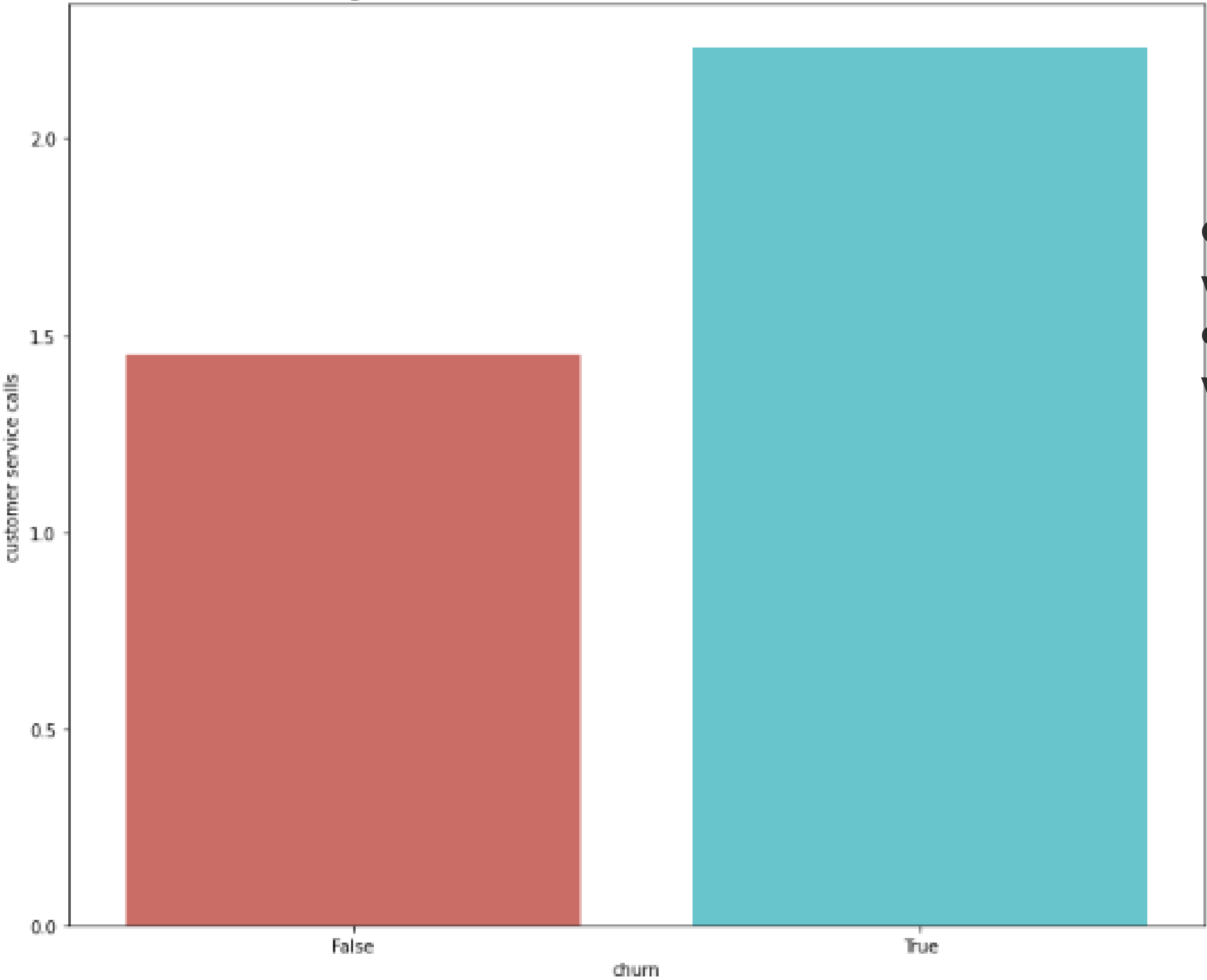
CHURN VS VOICEMAIL PLAN



Customers who do not have a voicemail are five times likely to churn than people who have a voicemail plan

HOW DOES CUSTOMER SERVICE CALLS AFFECT CHURN

Average customer service calls for Churn and Non-churn customers



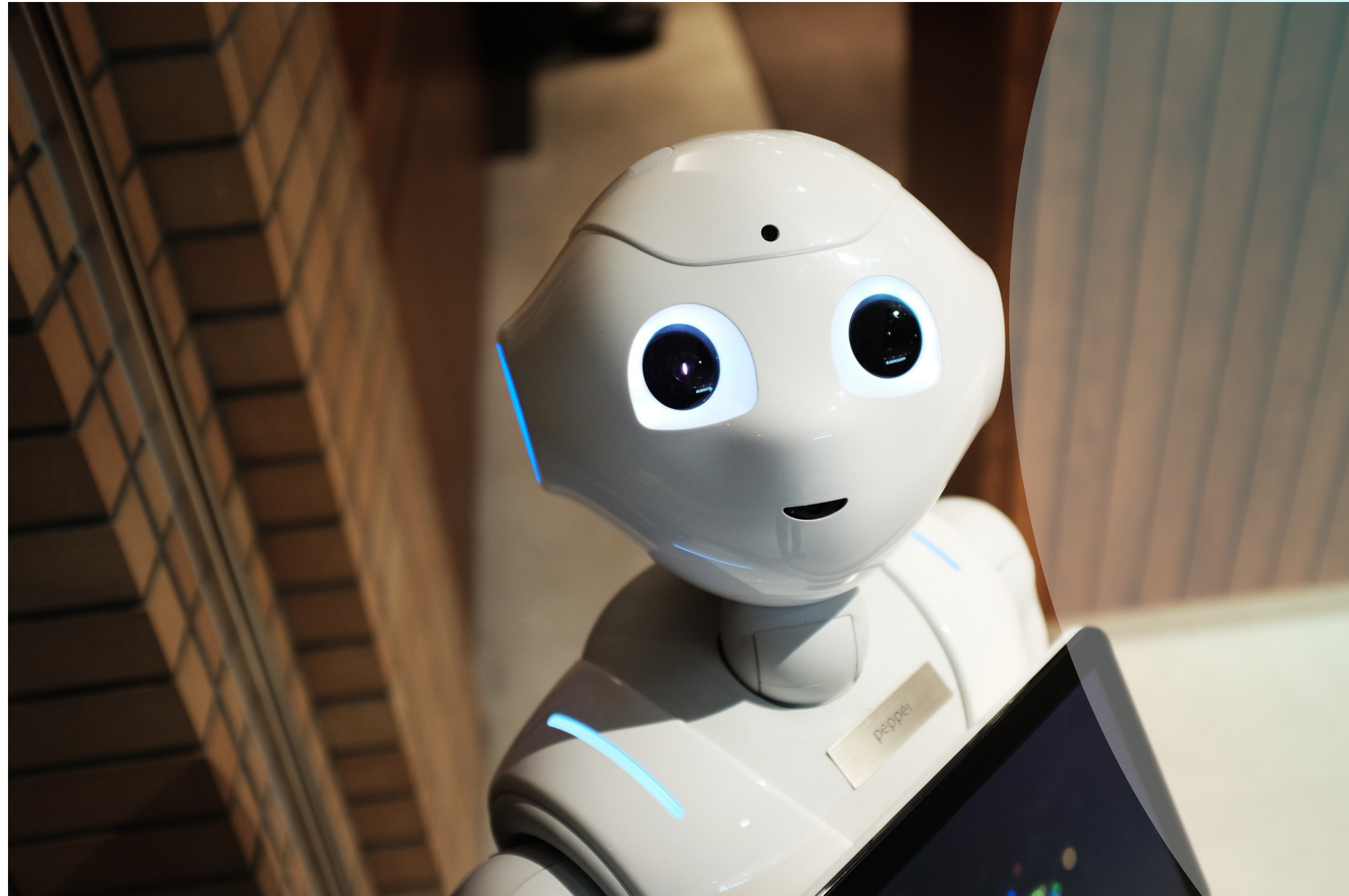
Customers who do not have a voicemail are five times likely to churn that people who have a voicemail plan

MODELLING

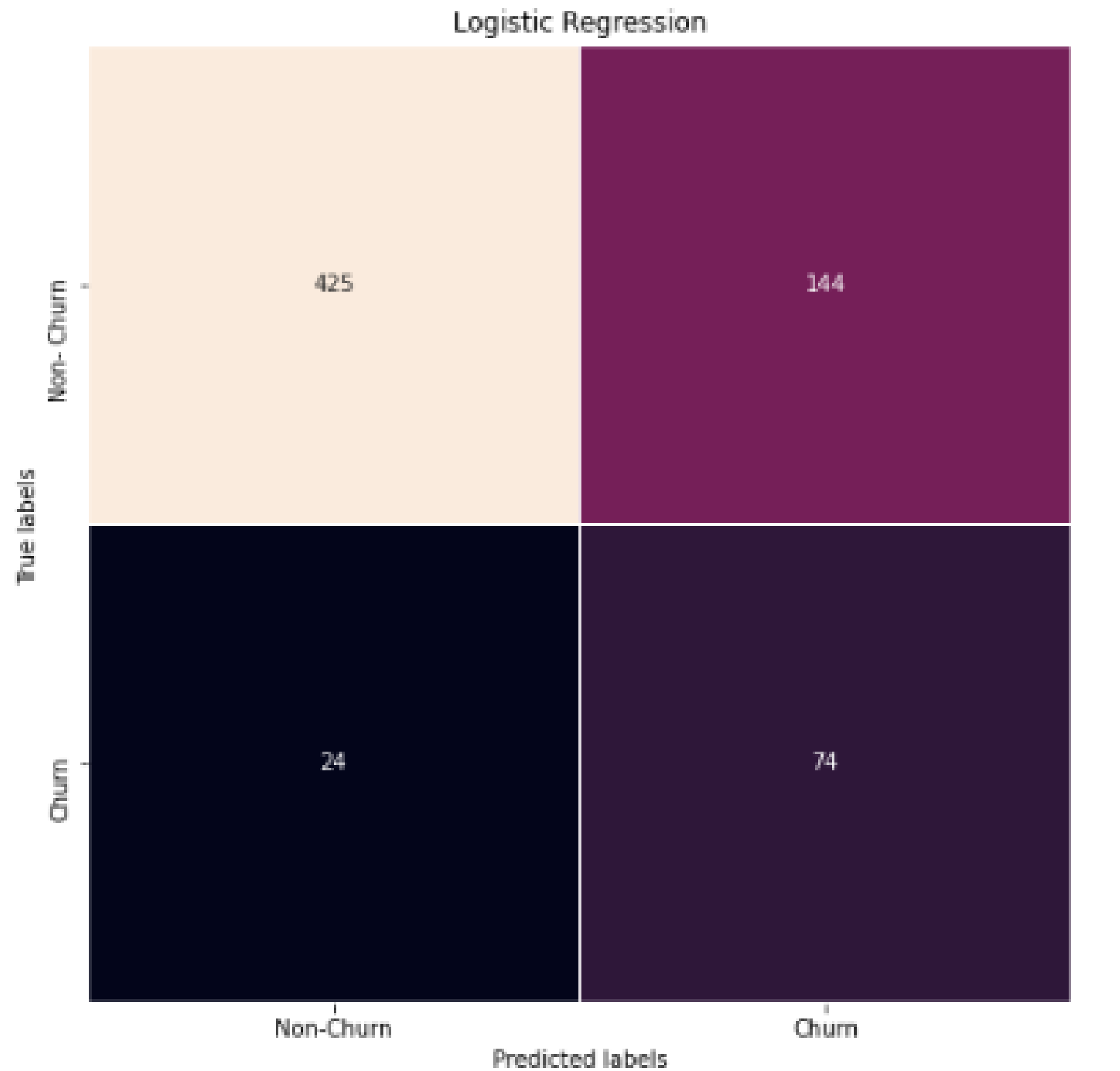


80% Accuracy

75% Recall



BASELINE MODEL (LOGISTIC REGRESSION MODEL)



Accuracy; 75%

Recall: 76%

BENCHMARK

OTHER POTENTIAL MODELS FITTED

Random Forest

Support Vector Machine

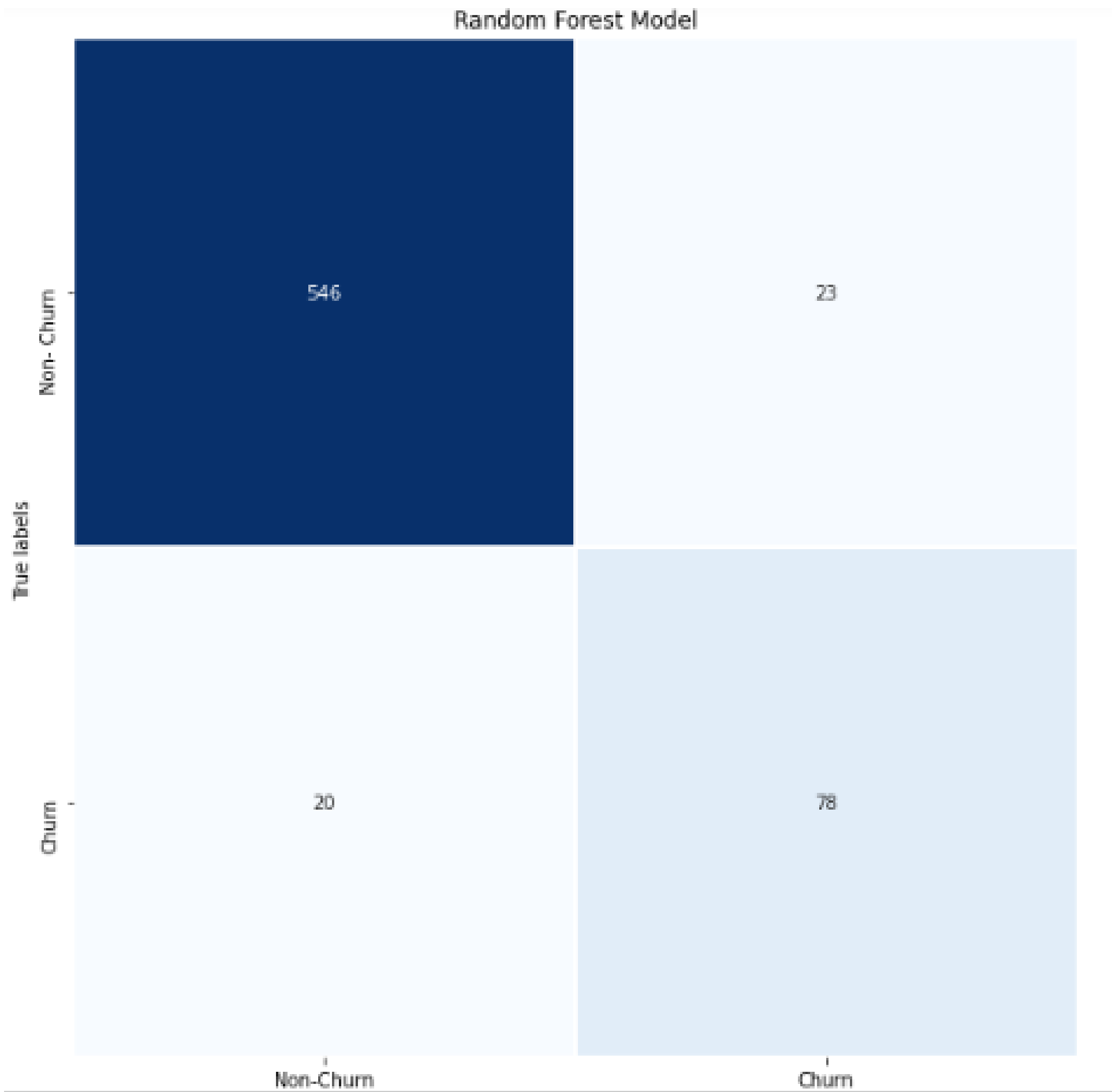
Adaboost

XgBoost



Random Forest and Support Vector Machine selected as the two best models for model tuning

'VANILLA' RANDOM FOREST



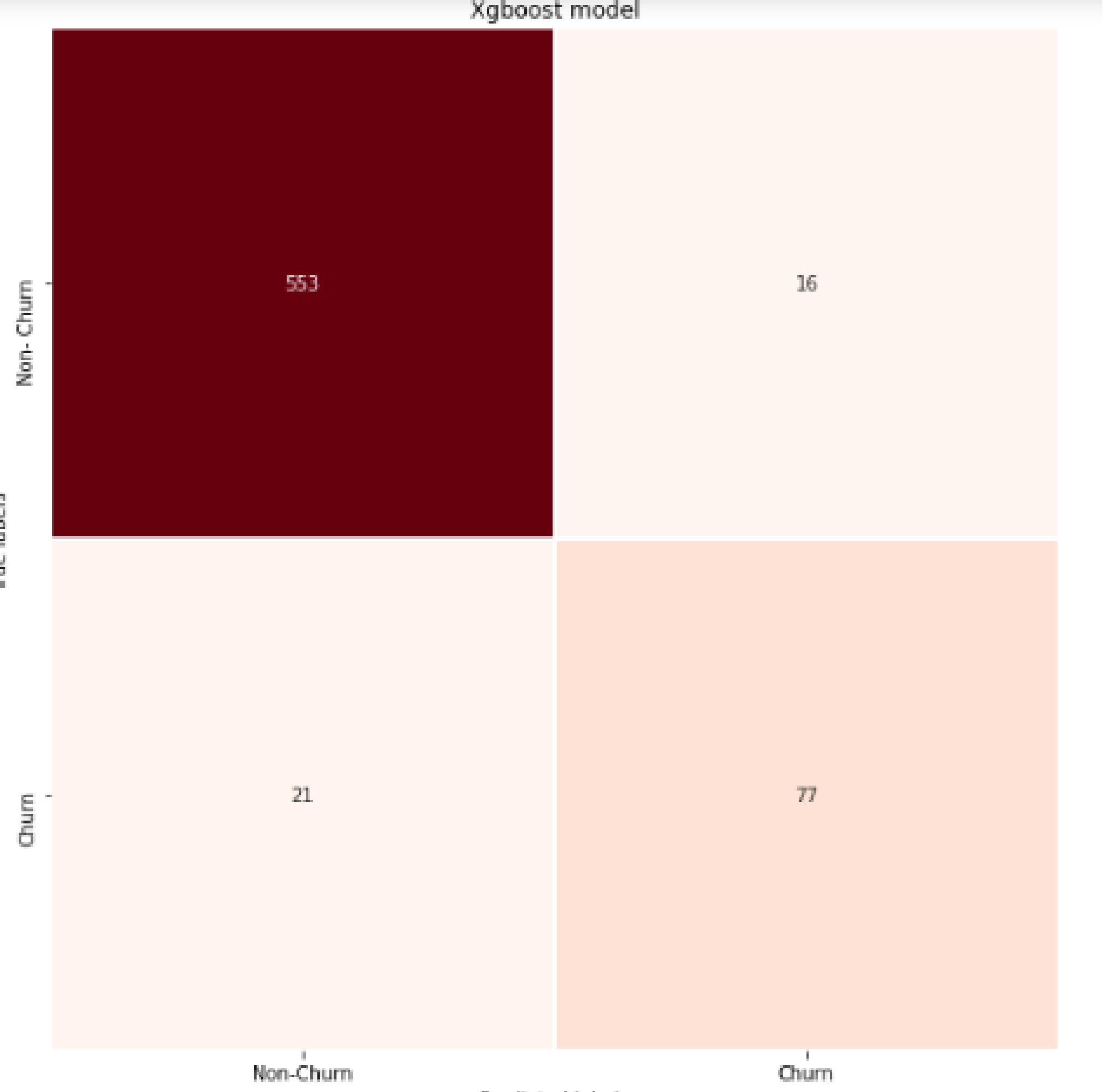
Accuracy; 93%

Recall: 80%

Required recall

Good overall accuracy

XGBOOST CLASSIFIER



Accuracy; 94%

Recall: 79%

Required recall

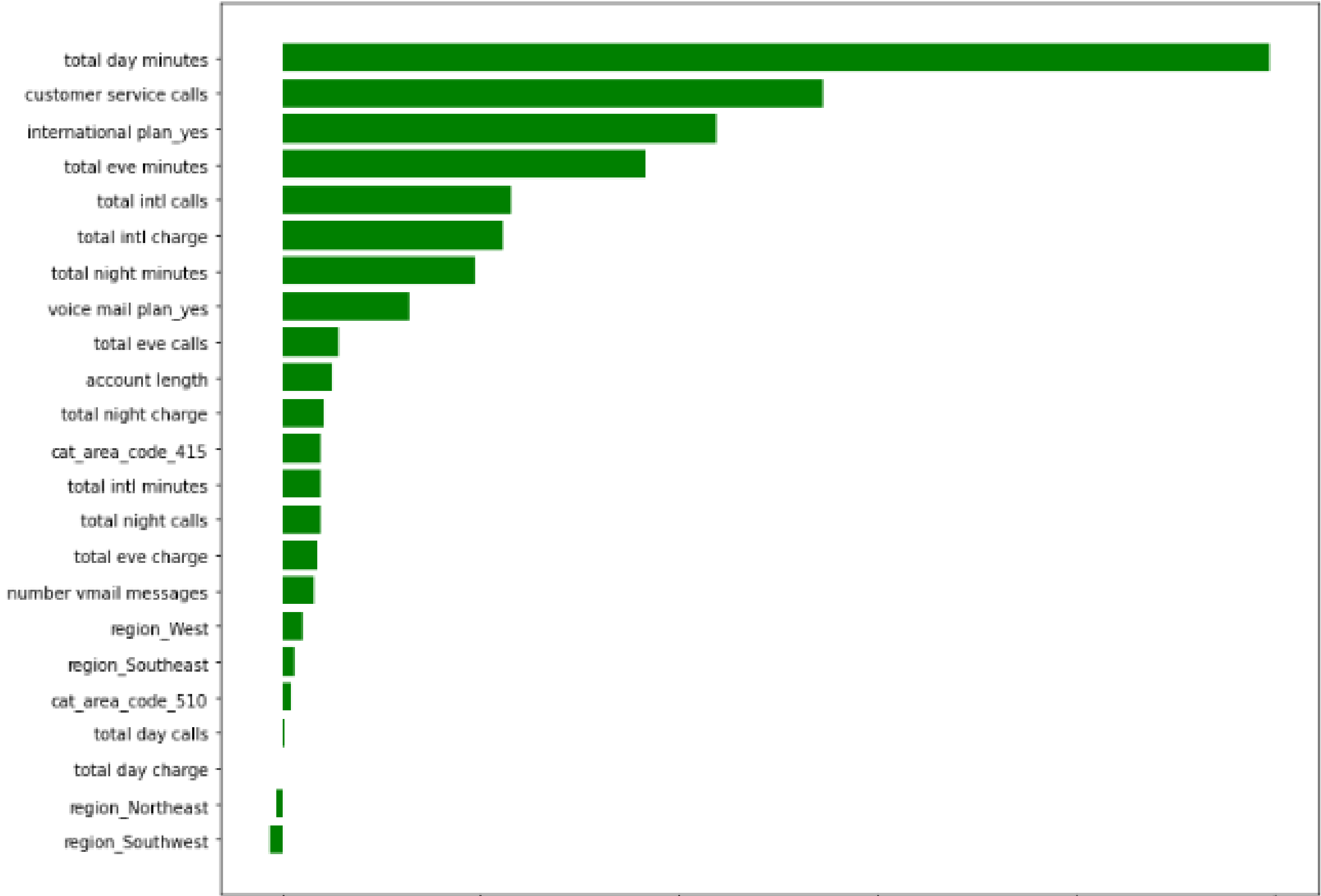
Very good overall accuracy as required

Chosen as the best model



VARIABLE IMPORTANCE

Variable importance



1. **Total day minutes**
2. **Customer service calls**
3. **International plan**
4. **Total eve minutes**
5. **Total int calls**
6. **Total intl charge**
7. **Voice mail plan**

LIMITATIONS

As a popular machine learning says '**All models are wrong , but some are useful**', the chosen model is not a silver bullet in correctly classifying churning and non churning customers. However, the model does a brilliant job in capturing well the churning customers. On the flip-side, making correct classifications on churning customers comes at a cost where more non-churning customers are misclassified as churning customers. This can often mean **that resources to retain may be channeled towards a non-churning customer when there is no need.**



CHALLENGING THE SOLUTION

Collecting demographic data on customers

- **age,**
- **gender**

RECOMMENDATIONS

- **Emphasis on customer service calls as they impact customer churn the most**
 - **Good customer service improves customer satisfaction and retention**
 - **Provides quick and easy problem resolution and feedback**
 - **Helps build and maintain a positive company reputation**
- **Strong emphasis on marketing and pricing for international plans to reduce churn**
 - **Evidence that customers with international plans are less likely to churn**
- **Review of marketing and pricing strategies for voicemail plans to reduce churn as customers who do not have a voicemail are five times more likely to churn than people who have a voicemail plan**
- **Review of marketing and pricing policies to ensure that customers enjoy favourable offers when making calls**