SYRIA TELECOMMUNICATIONS LIMITED

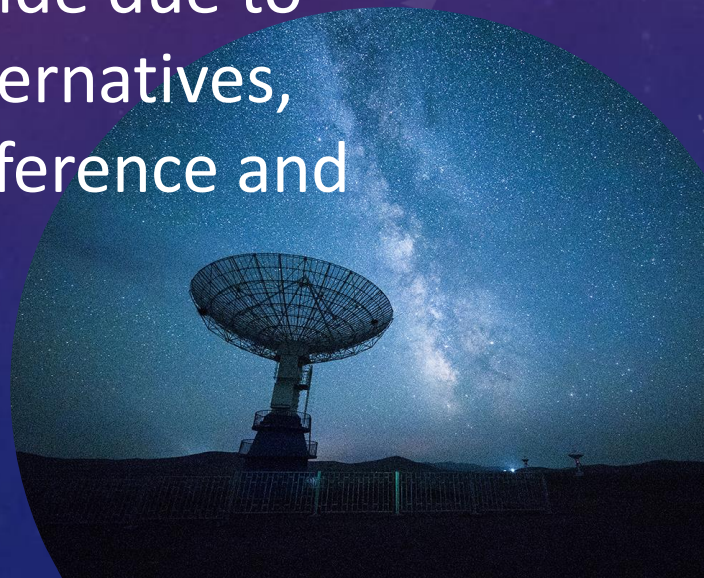
AN EXPLORATION OF THE HIGH CHURN RATE

A PRESENTATION BY GEORGE KARIUKI

BUSINESS UNDERSTANDING

OVERVIEW

- The telecommunications industry worldwide brags as being one of the few industries with the most stiff competition and unstable business conditions for any soft hearted telecommunications companies.
- Customer retention rates rock violently from side to side due to economic downturns, increased costs, competitive alternatives, increased technology, globalization, government interference and restrictions, among many other factors.



PROBLEM STATEMENT

- Management of Syria Telecommunications Limited realize the potential losses arising from this rampant churning of its customers from, what it thought to be, it’s firm grasp.
- The Data Science Department has been tasked with obtaining, scrubbing, exploring, and understanding the patterns with it’s customers and recommending any steps to be followed.

CHALLENGE AND BUSINESS PROBLEM

- With the booming profits and market share interests that could potentially be gained, Syria Telecommunications faces a downward trajectory in terms of customer retention which would inevitably lead to a deviation from the business goals of the organisation.

OBJECTIVES

- To identify factors leading to increased churn rates
- To create a classification model that predicts whether a customer will churn with a recall of over 80%
- To give customer retention recommendations

DATA UNDERSTANDING

- The data was sourced from Kaggle and was stored in a CSV file format.
- The data contained 3,333 rows and 21 records.
- Key Features include:

❖ Minutes and Calls – Consumer's purchase of minutes and calls made to the domestic and international regions

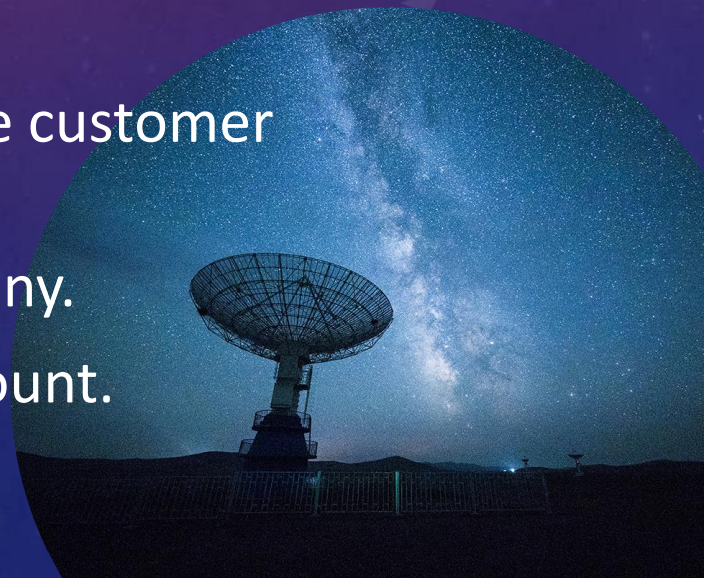
❖ Charges – Charges billed to consumers for making their domestic calls during the day, evening and night, and international calls.

❖ Subscriptions – A binary value indicating if the consumer has an international plan or a voice mail

❖ Customer Service Calls – Number of times a user has called the customer service

❖ Churn - This binary feature indicates if the user left the company.

❖ Account Length – period of time the user maintained their account.

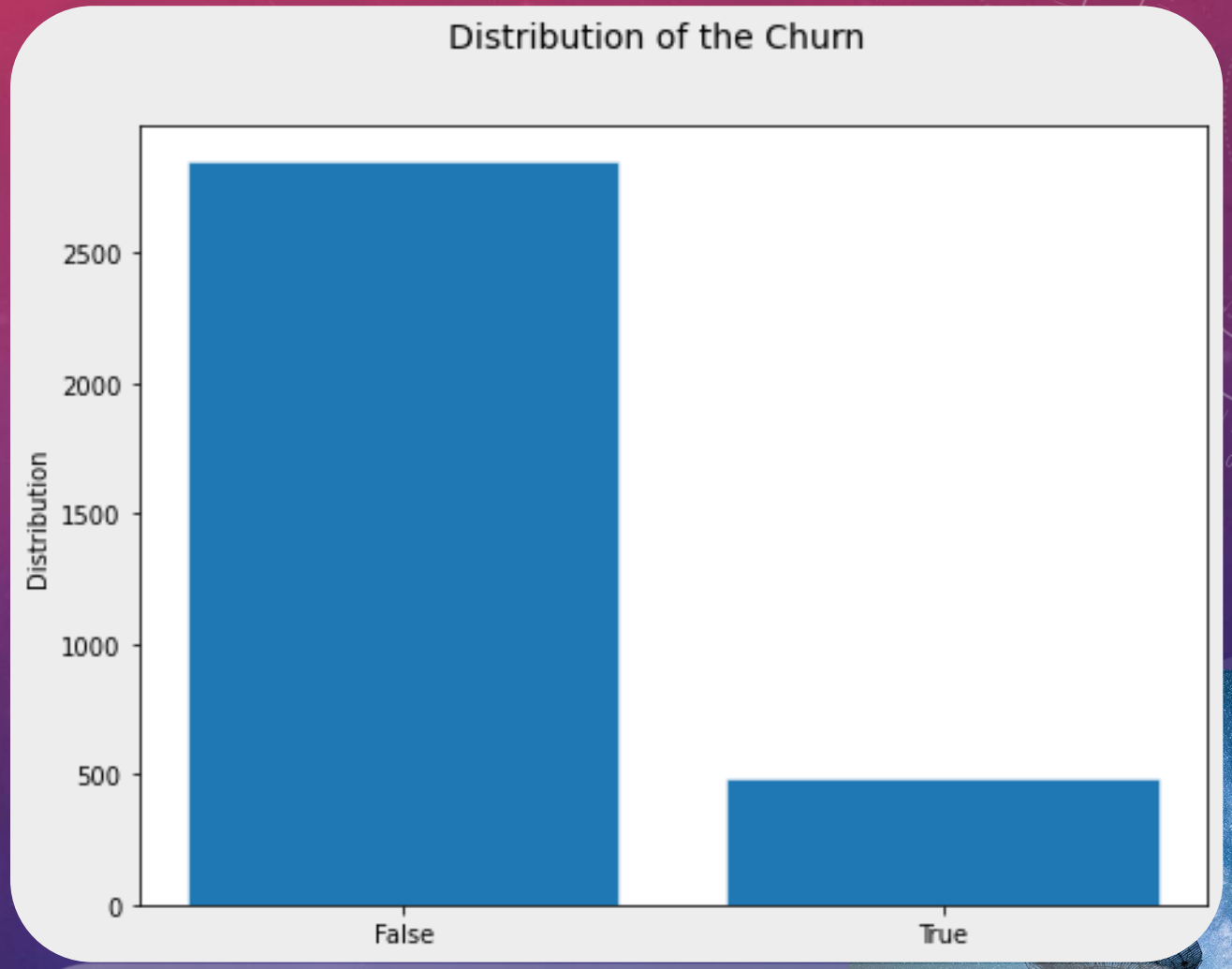


FINDINGS

UNIVARIATE ANALYSIS

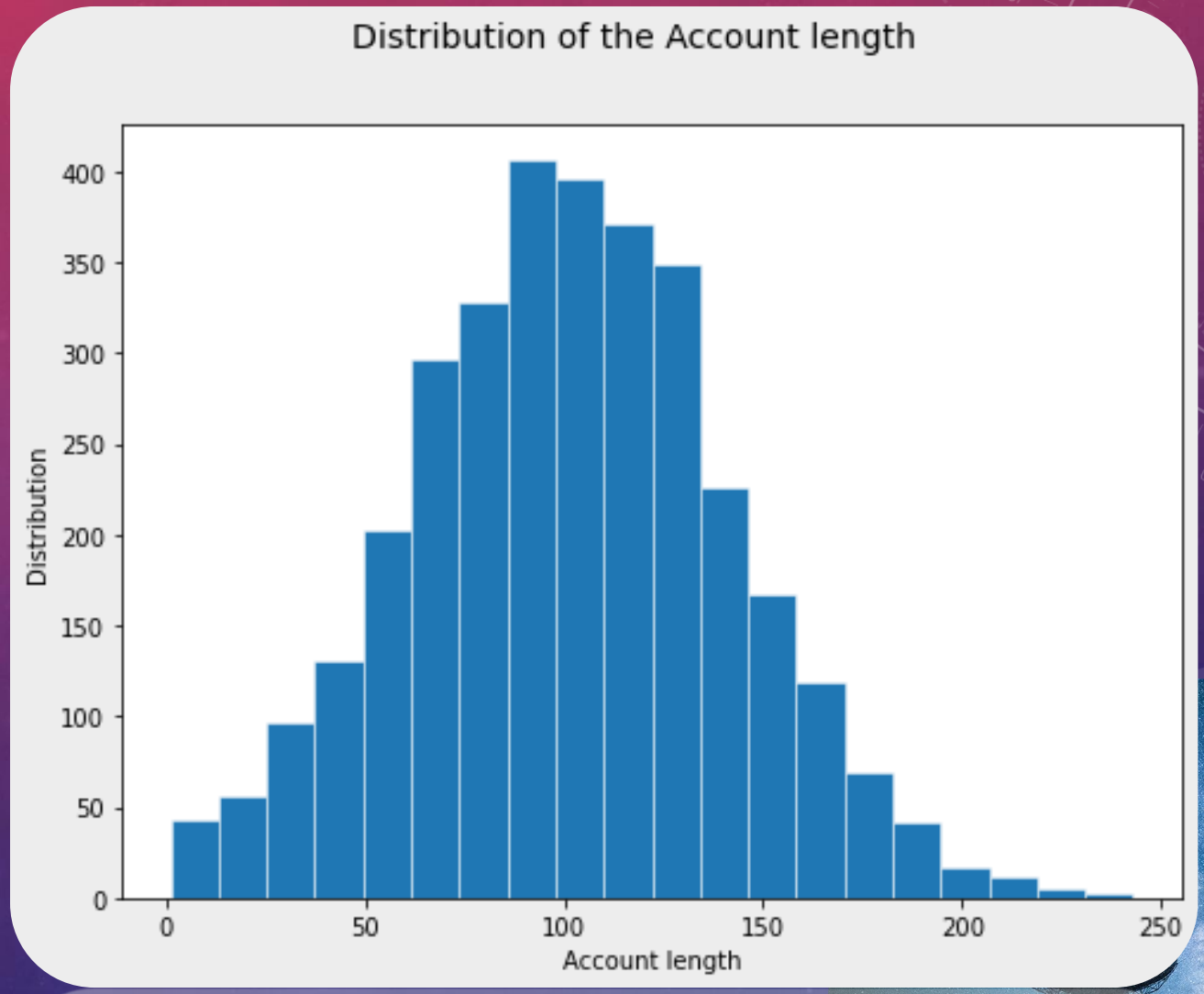
CHURN RATE

Near 85% of consumers stayed while 15% of the consumers churned.



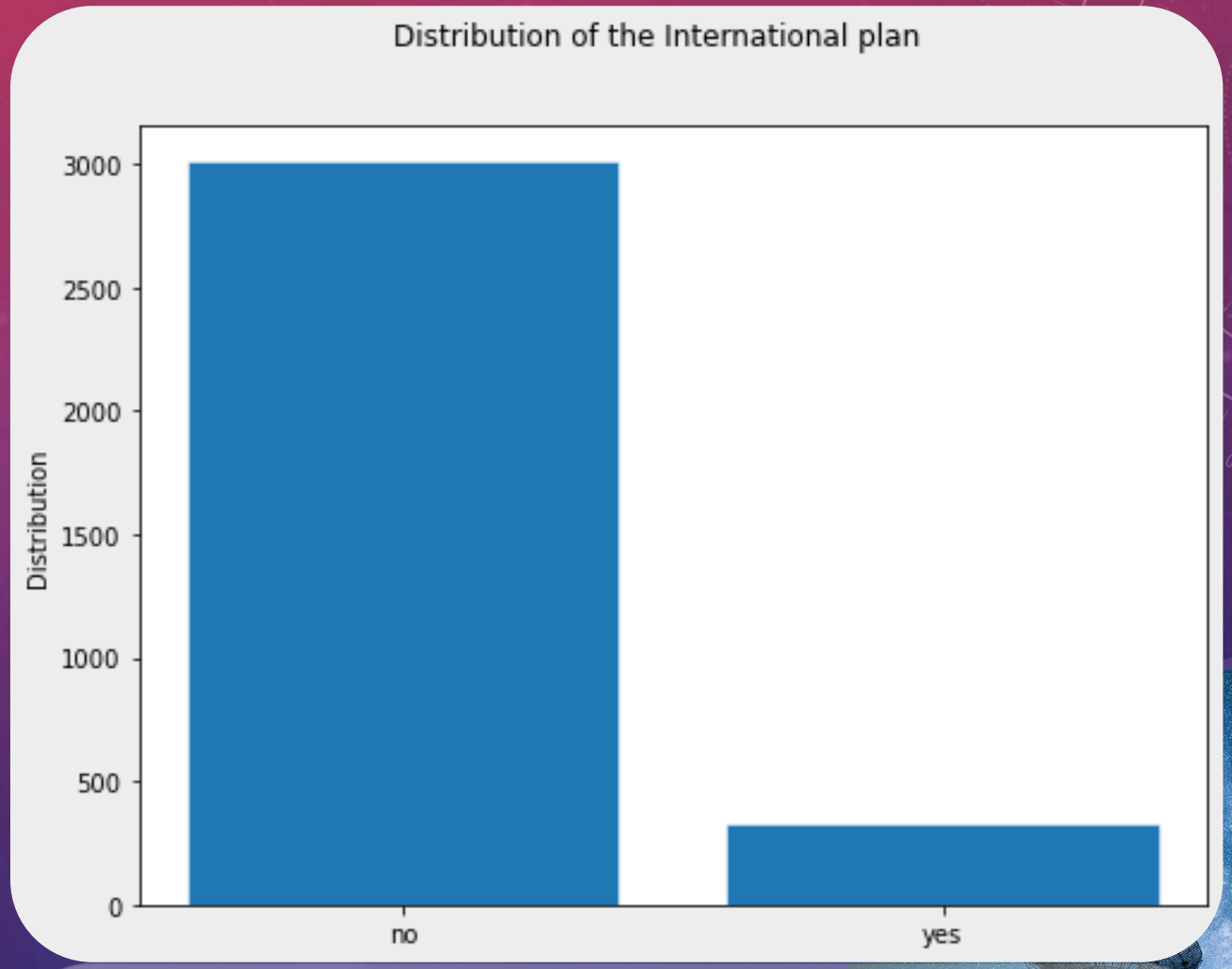
ACCOUNT LENGTH

The mean and median of users’ account length was around 101 days.



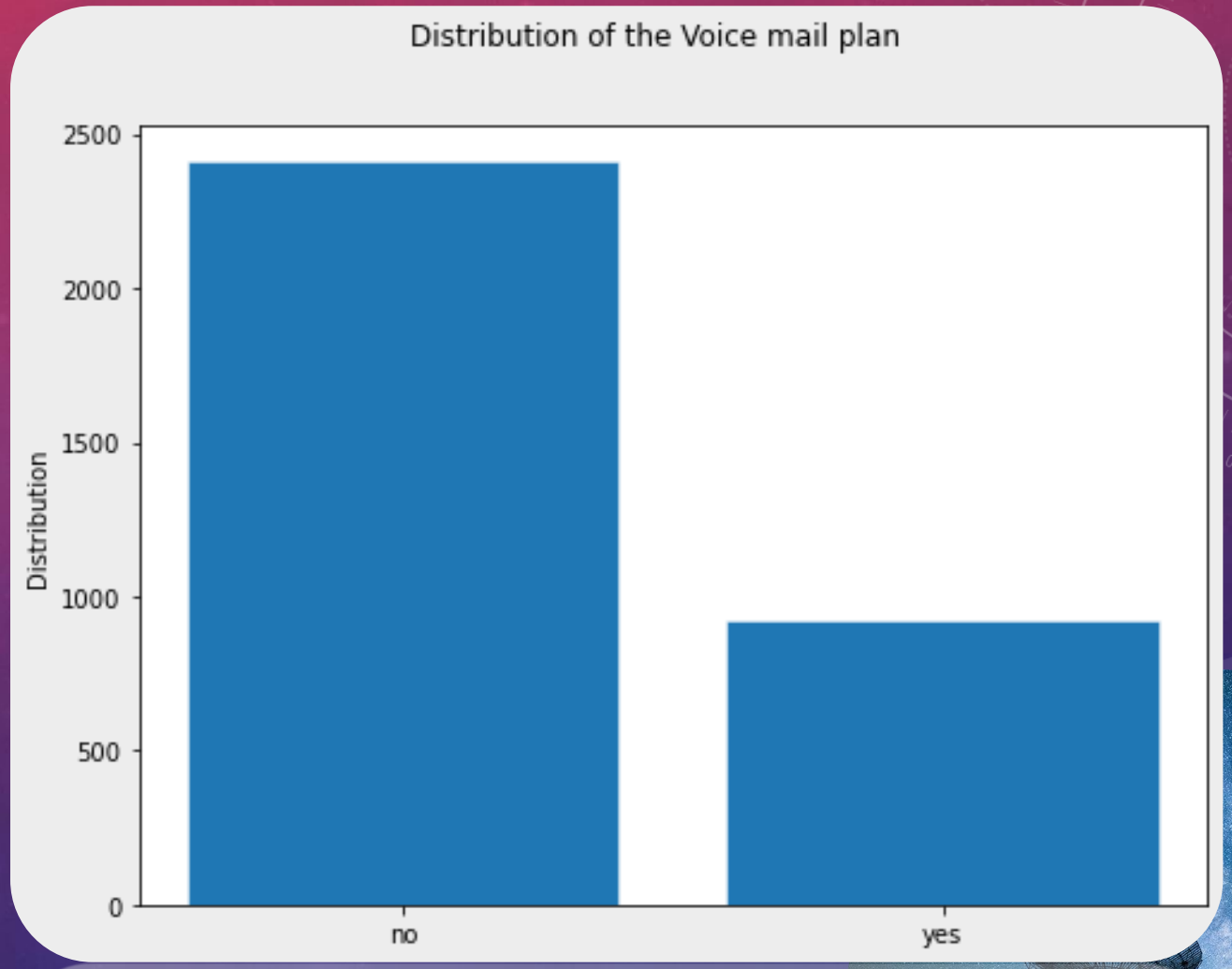
INTERNATIONAL PLAN

Near 90% of consumers do not have an international plan while 10% of the consumers have an international plan.



VOICE MAIL PLAN

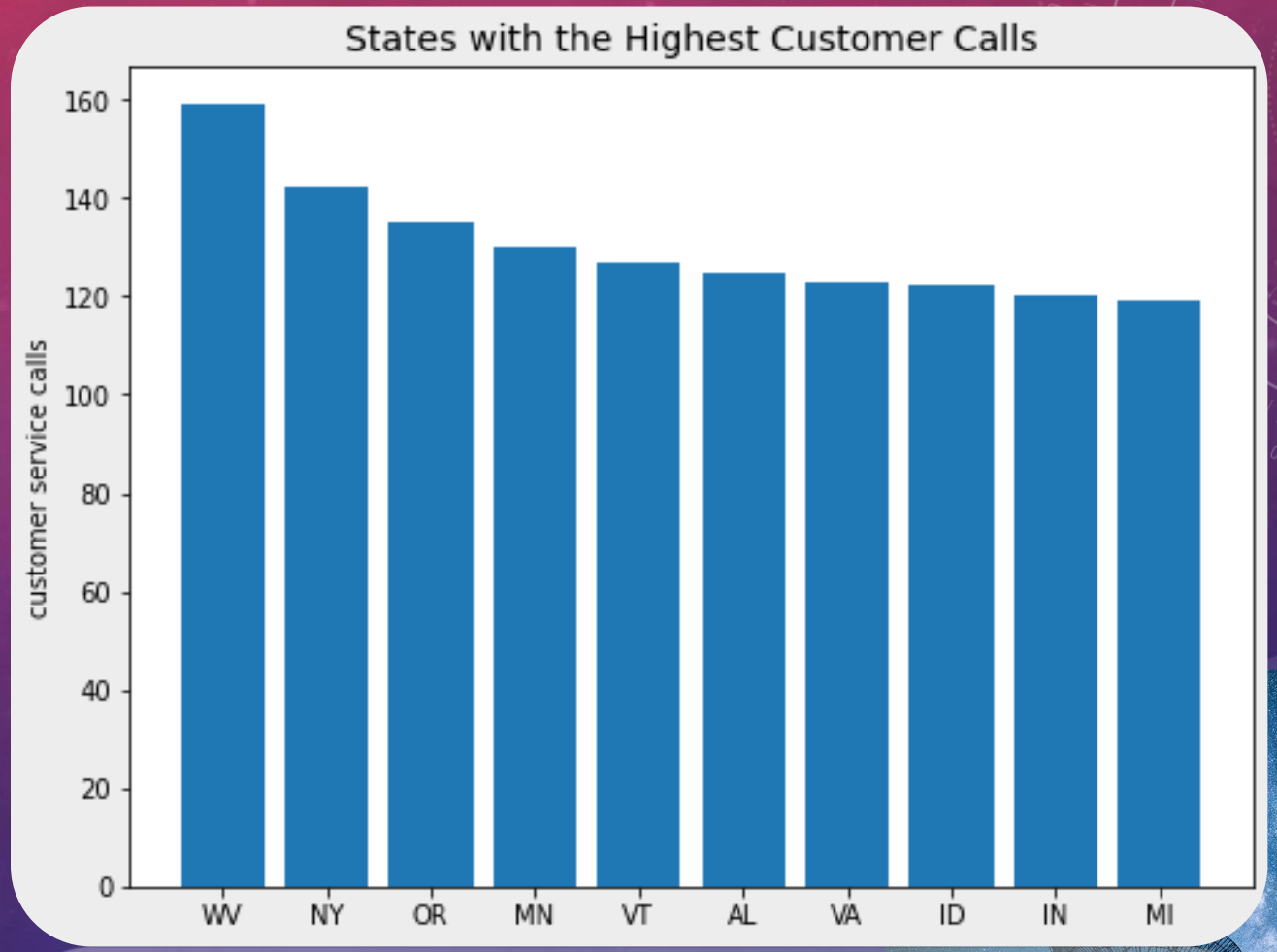
Near 86% of consumers do not have a voice mail plan while 14% of the consumers have a voice mail plan.



BIVARIATE ANALYSIS

STATES AND CUSTOMER SERVICE CALLS

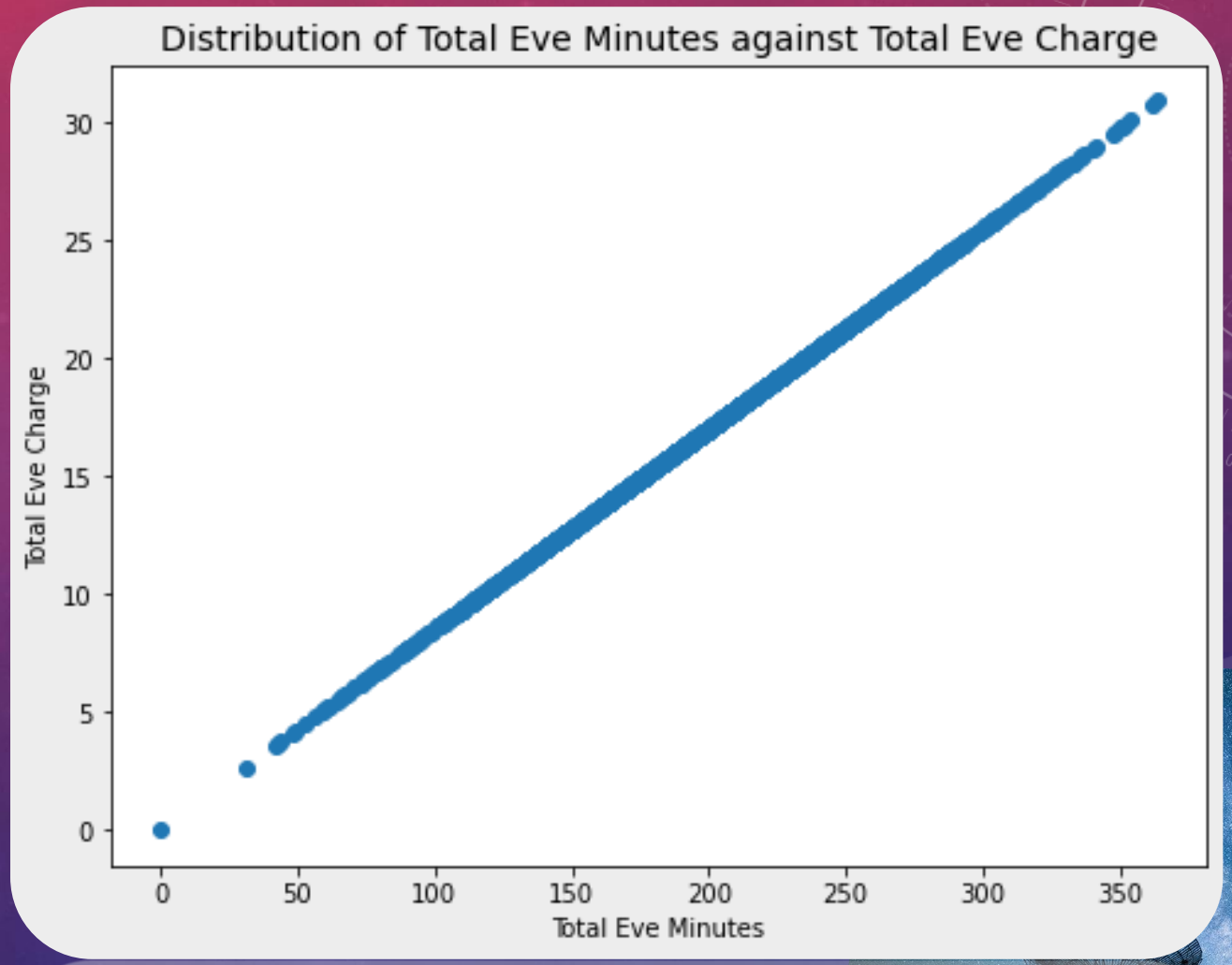
Customers who frequently call the service are from these top 10 locations – showing an issue in these locations.



MINUTES AGAINST CHARGES

There is a perfect correlation between the minutes and the charges.

This relationship holds true for the day, night and international minutes and their charges.

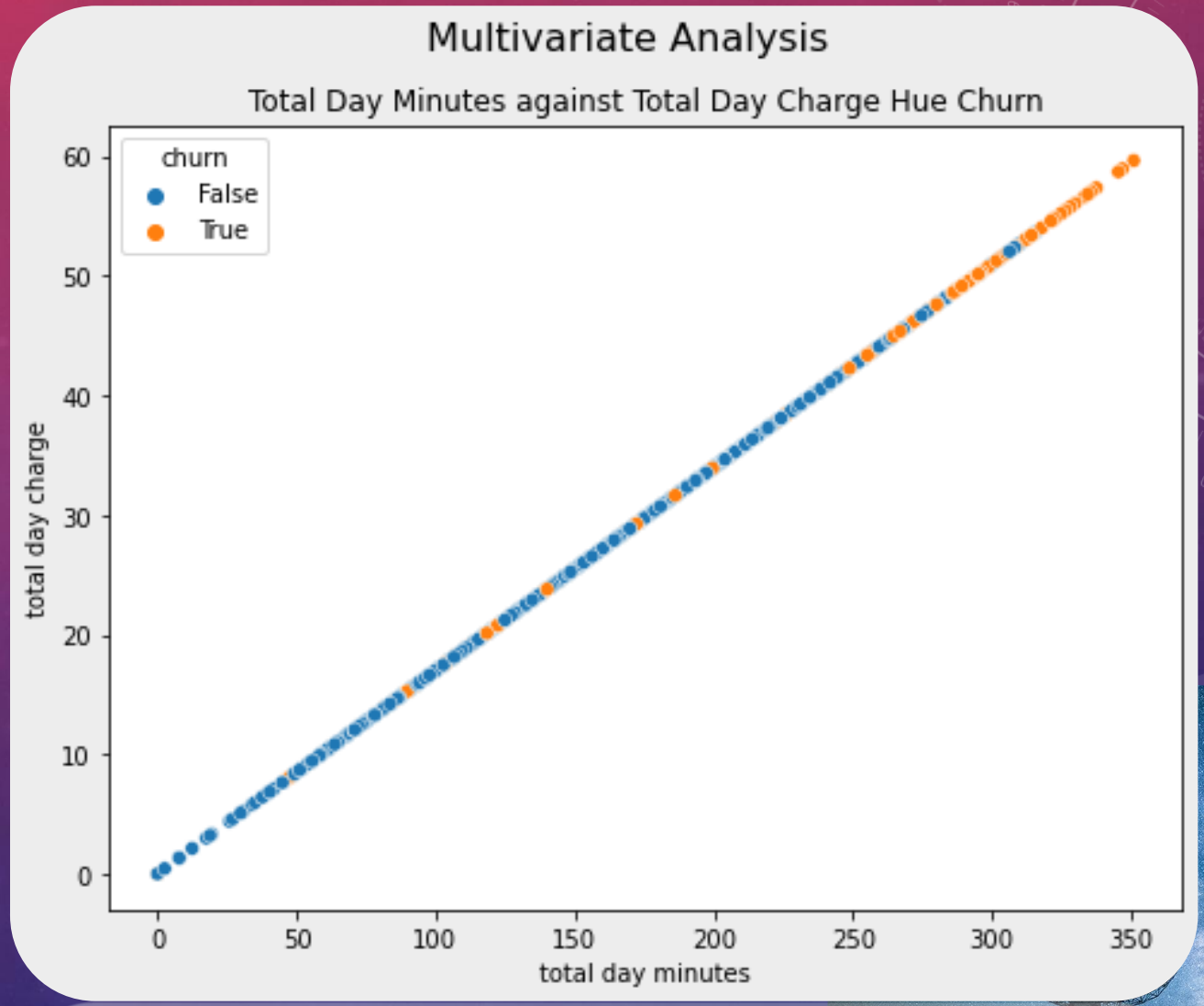


MULTIVARIATE ANALYSIS

MINUTES V CHARGE V HUE

Consumers churned when they were billed for charges above the average charge and were also purchasing a lot of talk-time minutes.

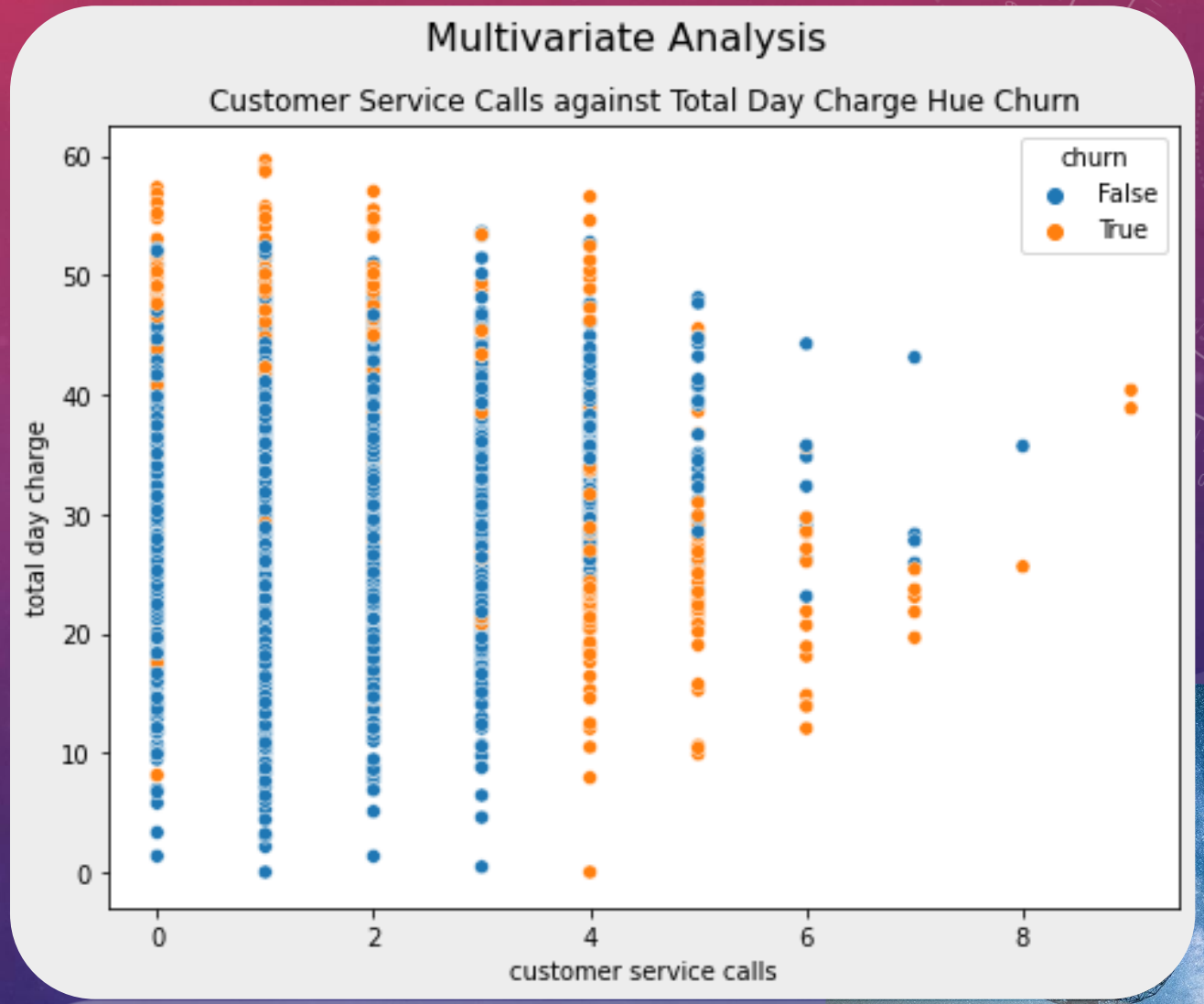
This relationship was evidenced in other features like evening, night and international.



SERVICE CALLS V CHARGE V HUE

Consumers churned when they were billed above average rates and called the center below 3 times.

Consumers also churned when they paid lower rates but had to call the center more than 3 times.



MODELLING

CLASSIFICATION PROBLEM

- The exploration of the data reveals that this is a classification problem where we attempt to predict whether a customer would churn or not.
- Minimizing misclassifying the status of a client as a non-churner would be our biggest objective.
- The appropriate metric to use based on this problem is recall which can defined simply as a measure of how well our model identifies the true positive cases from all the actual positive cases.
- Recall is often used when the cost of false negatives is high – in this case it would be detrimental to classify a customer as a non-churner when in reality the customer churns.

RANDOM FOREST MODEL

- Out of all tested models, Random Forest Classifier provided the highest metrics – a recall score of about 80% - to predict churning in consumers.
- Our model is very good at identifying the consumers who will churn, with very few misclassified as non-churners.
- Focusing on retaining the 80% will reduce the churn rate and increase the retention rate.

RECOMMENDATIONS

Bespoke packages, Incentives and Rewards – Minutes, Voice Mail and International Plans could be bundled together

Trained Customer Service Agents and Customer Feedback Recording

Upgrade Technology in problem areas where there are high customer service rates.

Re-evaluation of price points by using a mix of linear and constant prices.

QUESTIONS AND FEEDBACK

THANK YOU!

GEORGE KARIUKI