Capstone Project-1

EDA-Telecom churn Analysis

Team members

Nakul Pradeep

Kavya Sharma

Baliram Kumar

Aishani Barman Roy

Shubham Deodekar

Introduction

- Telecom industry is growing rapidly and also it faces major challenge with customer churn, as customers switch to alternate provider due to various reasons like lower cost, multi service offerings, marketing promotions by competitors etc.

- The cost of churn in the industry is large. Several studies have shown that attracting new customers is much more expensive than retaining existing ones.

- Proactive actions will prevent the loss of revenue for the company and will improve / retain the market share among the industry in terms of the number of active subscribers.

Project Objective

- In the telecom industry, customers are actively switching from one operator to another.
- Telecommunications industry experiences an average of (15 – 25)% annual churn rate
- The aim of this project is to investigate the main reasons for churning of customers in the telecommunication sector.

Overview

- The aim of the project is to identify the features or variables that make customer churn in order to prevent it and retain the customers.

- We will be working on the given telecom customer data and analyze to draw some insights by using the various charts and visualizations.

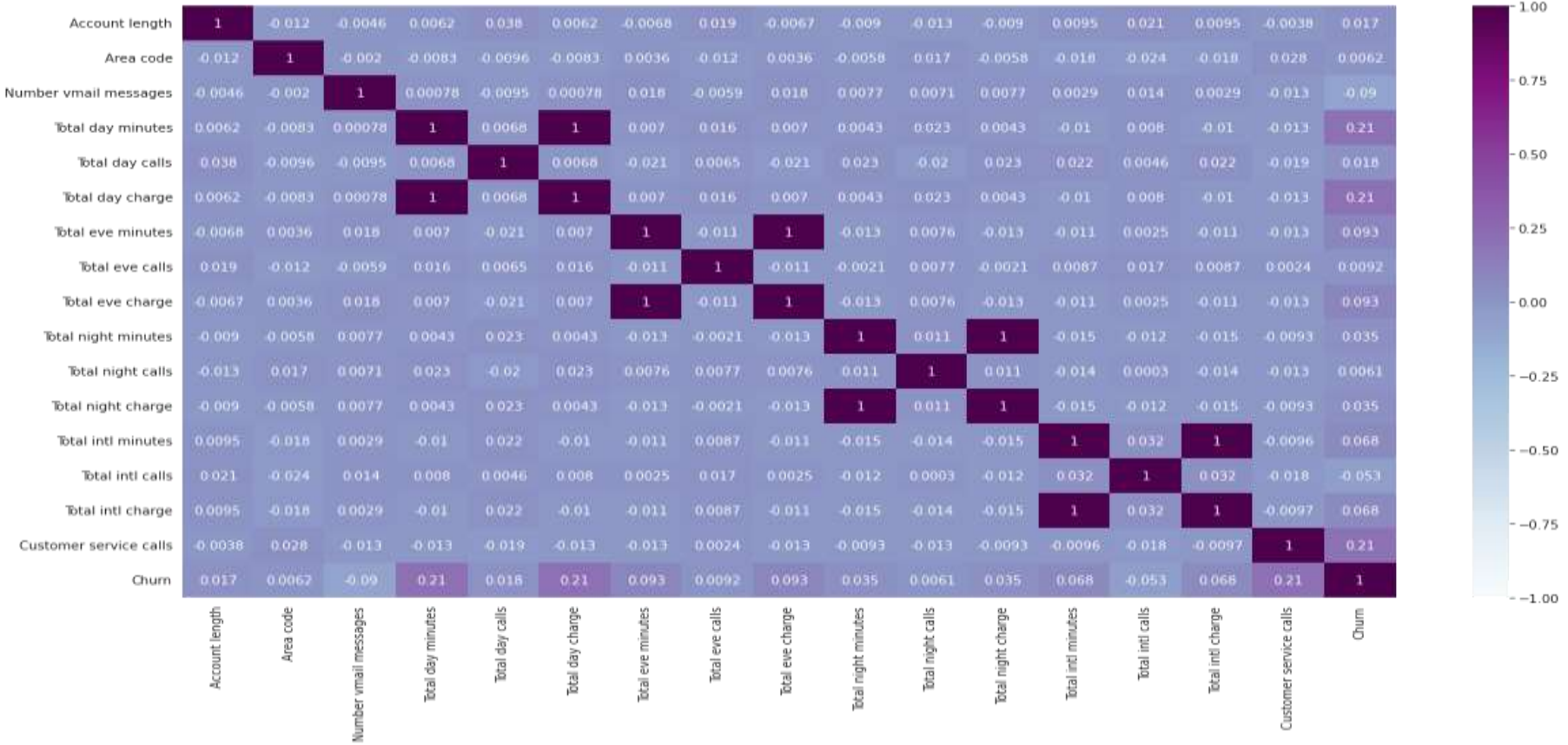
Steps of our Data Analysis

1. Study and overview the whole data .
2. Assess and clean the data.
3. Conduct exploratory data analysis by using various libraries.
4. Visualizations
5. Understanding the Factors
6. Summaries and valuable insights

Library used

- Numpy - provides high-performance multidimensional array objects and tools to work with the arrays.
- Pandas - provides high-performance, easy-to-use data structures and data analysis tools
- Matplotlib - You can create stories with the data visualized with Matplotlib
- Seaborn - seaborn is an extension of Matplotlib with advanced features.

Data correlation



Data correlation

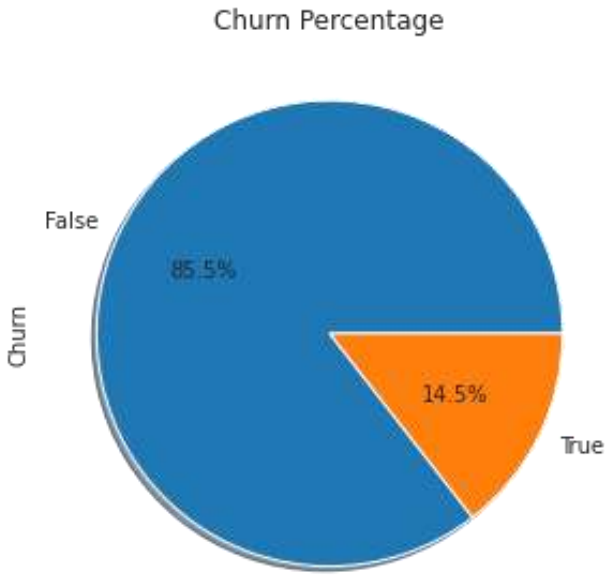
- From the above graph it appears that there exists a perfect linear relationship between ‘Minutes’ and its corresponding ‘Charges’.

- That is for international charge with international minutes, for day charges with day minutes, evening charges with evening minutes and finally night charges with night minutes.

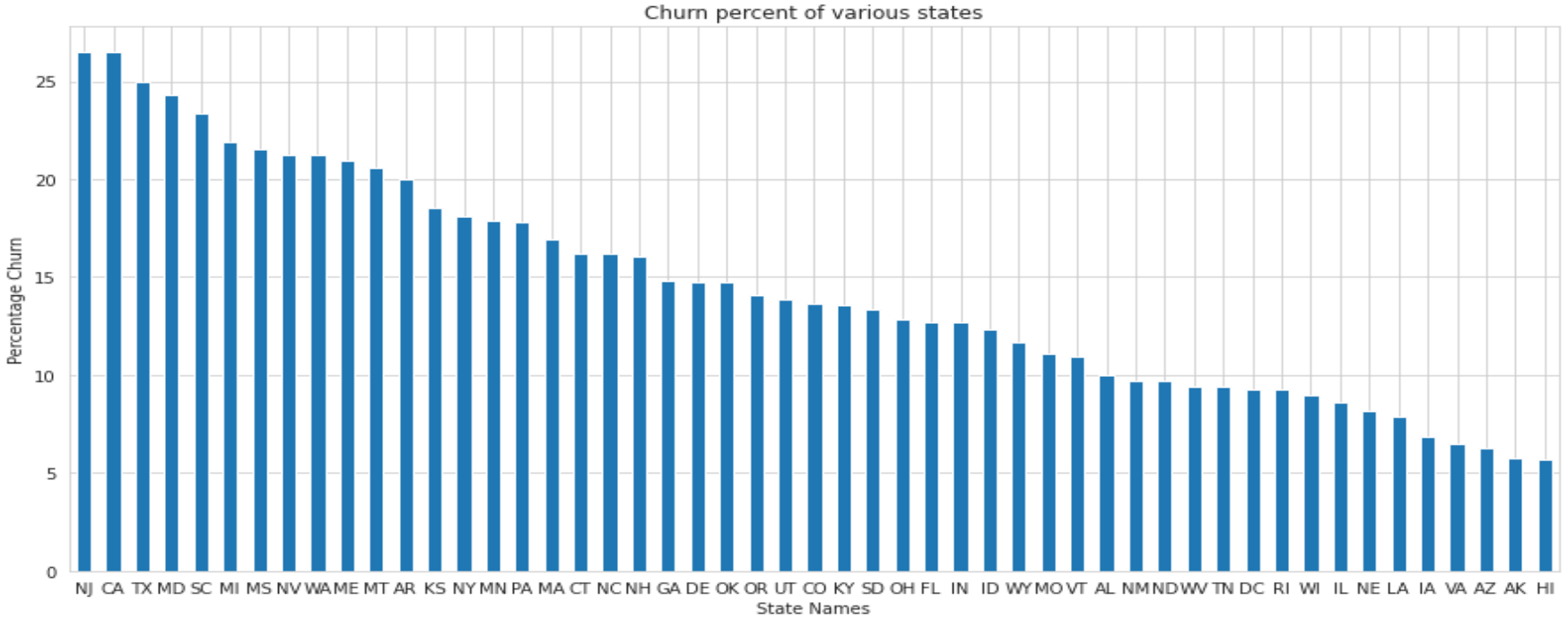
- The call minutes are highly correlated with call charges. So we consider only call minutes for further analysis.

Total Churn Rate

- With the help of Data Visualisation. It can be seen that there is low customer churn rate which is 14.5% as shown in graph.
- It is necessary to find which features or variables are responsible for customer to switch to another service provider.
- Better strategies are needed to further reduce the churn rate.



Churn rate per State



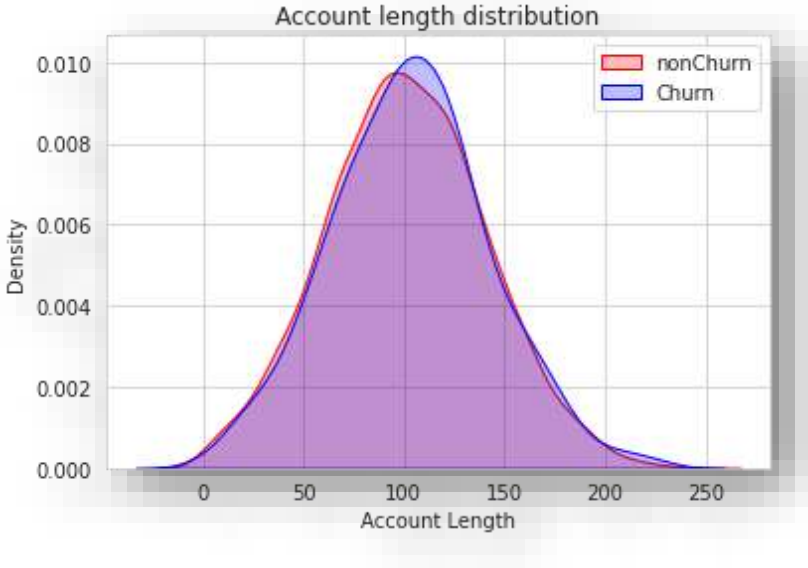
Churn rate per State

- From above graph ,It is clear that there are 51 states and every state have different churn rate.
- Graph shows out of 51 states, 5 states New Jersey(NJ), California(CA), TEXAS(TX), Mary Land(MD), South Carolina(SC) have higher customer churn rate.
- Company should focus there more and provide better facilities to reduce the churn rate.
- States like Hawaii(HI), Alaska(AK), Arizona(AZ),(Virginia)VA, (Iowa)IA has very low churn rate.
- Company can find what they are doing right in these states and try to implement it in states where there are higher churn rate.

Account Length

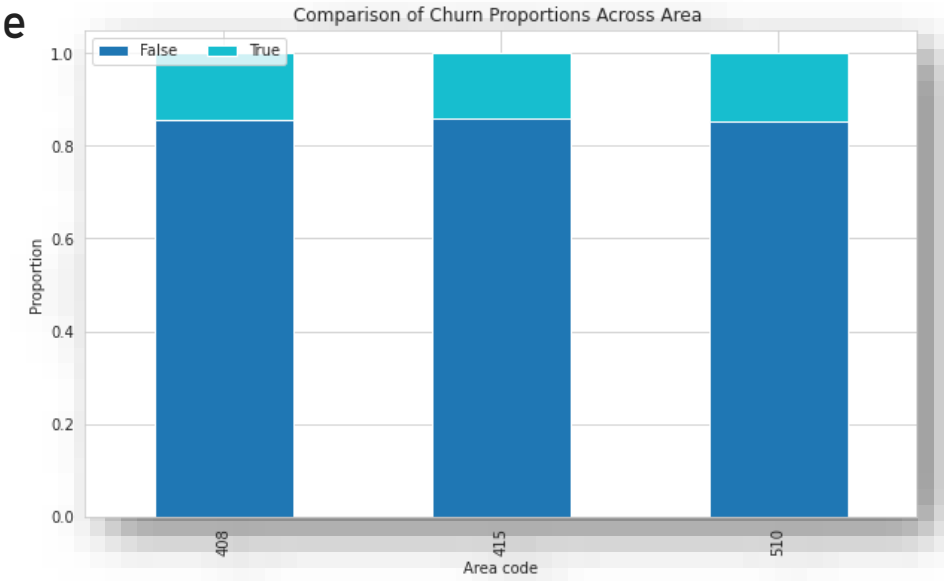
- It is the no of days the Customer account has been active.

- The Account length for churned and non churned customers did not show any specific behaviour implying length of service has no bearing in the churn.



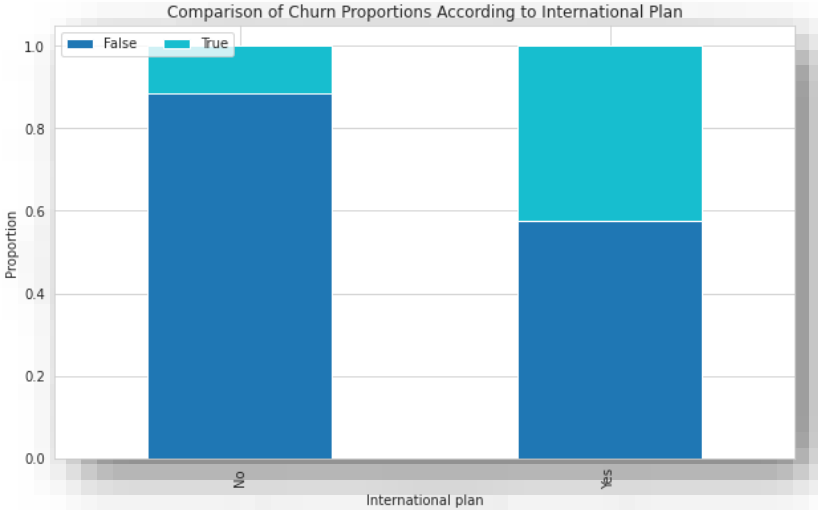
Area code Comparison

- Area code denotes the zone in which the customer is active or present.
- There are 3 unique area codes which are 408, 415 & 510.
- The graph shows the comparison between area codes and its clear that there is no major difference in churn rate across different area code.



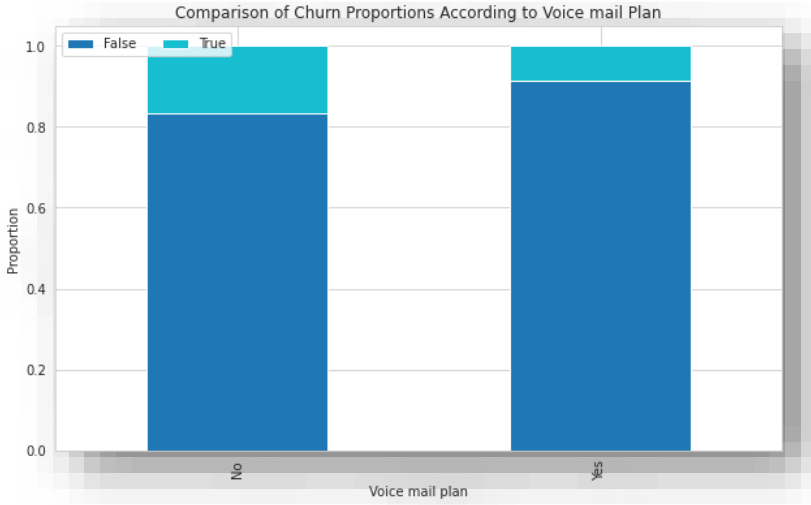
International plan

- International plans helps customer to receive calls and send when they are traveling outside the geographical location of the mobile service area and use another network.
- There is a 42.4% churn of customers with international plan while only a 11.4% churn of customers without international plan. It could be because many of the plan might have been taken by tourists who churned once their tour is over



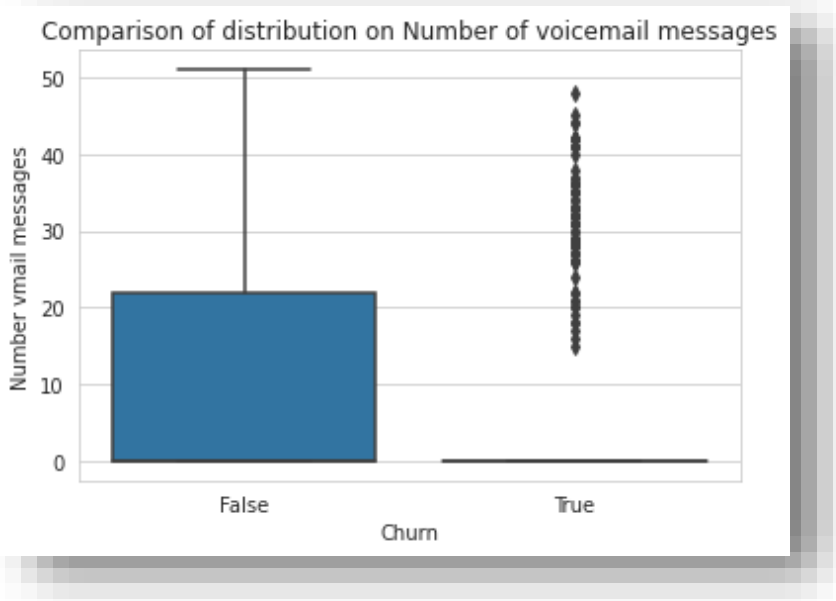
Voice mail plan

- A voicemail is an electronically stored voice message that is left by a caller to be retrieved later by the intended recipient.
- The customers who don't have a voice mail plan are more churned.ie(With voicemail plan there is a (16.7%) churn rate compared to people with no voice mail plan (8.67%).



Number of voicemail messages

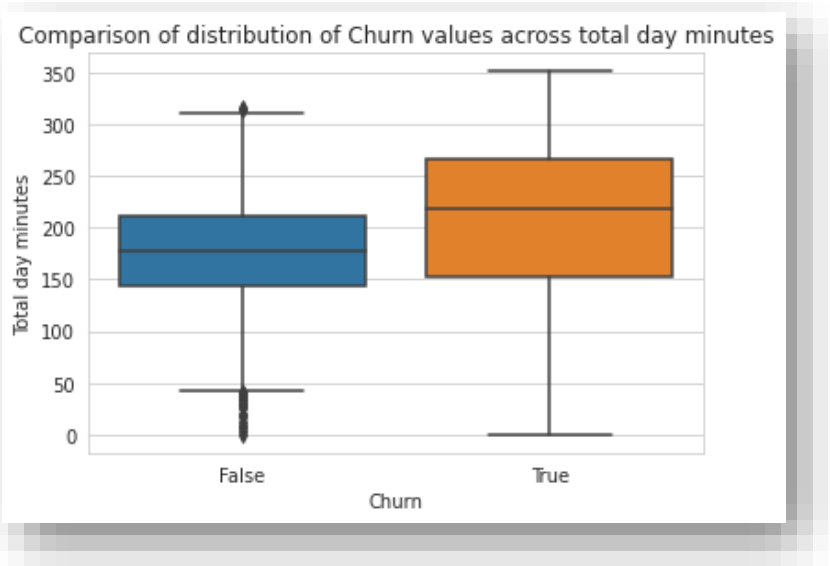
- The customers who retained service most of them did about 0-20 voicemail messages after that the distribution of retained customers is less indicating some underlying issue.
- So setting a limit of 20 voice calls might help reduce churn rate.



Total day minutes of Call

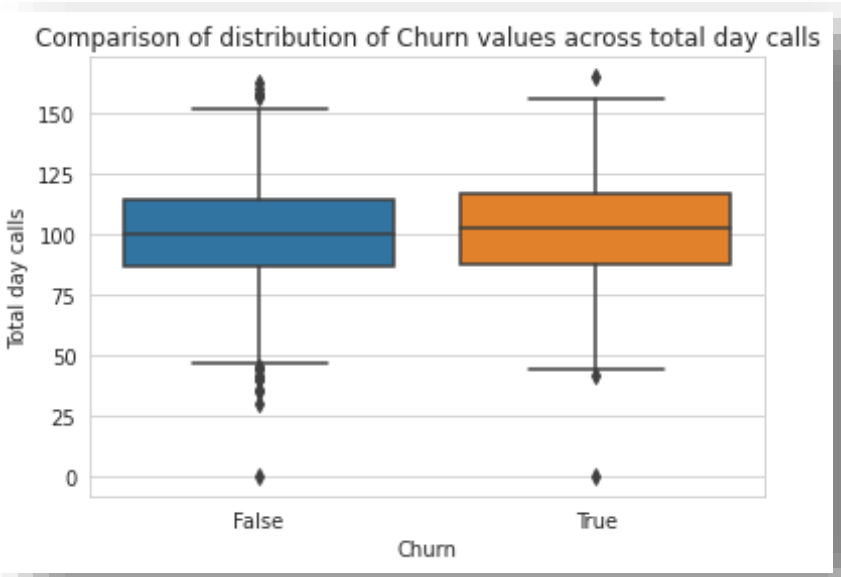
- It can be observed from box-plot that with users spending more 220 minutes or more i.e. approx. 4hrs tend to switch to other operator.ie when day minutes are more then more people are leaving.

- Since the minutes are highly correlated to charges implying when charges are high people are churned



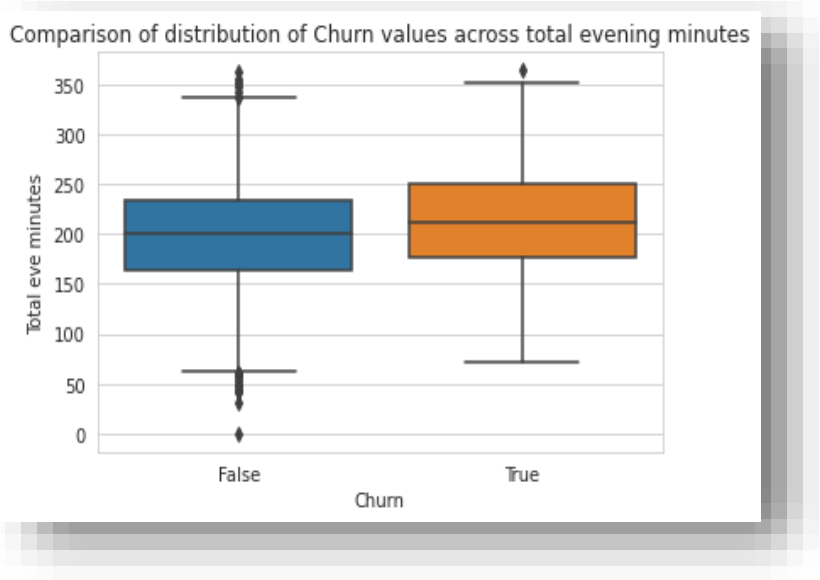
Comparison total day calls

- It can be inferred that on an average a 100 calls are made which is a good indication for the company.
- It can also be noted that for the churn customer the median is slightly higher than 100 which indicates there are call drops as there are more calls in the morning.



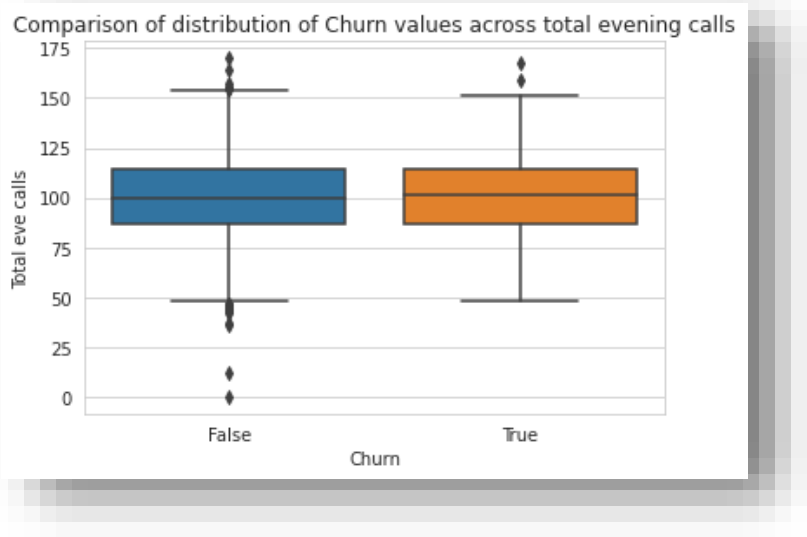
Total evening minutes of Call

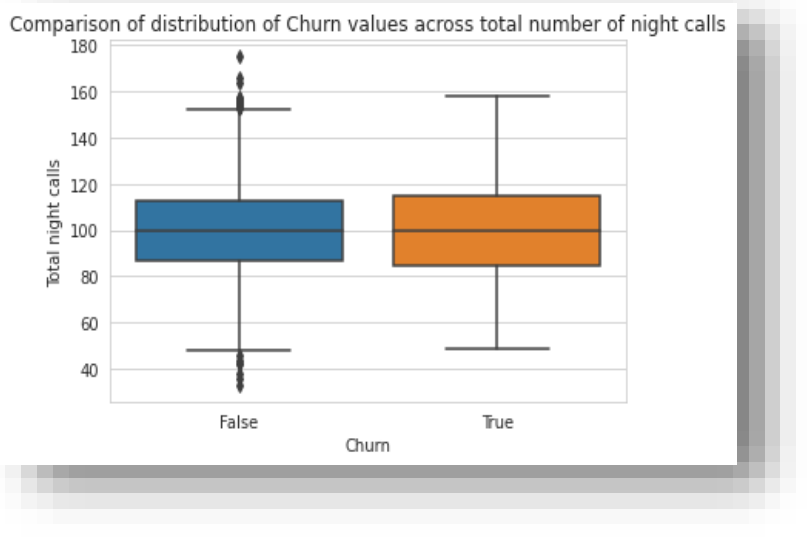
- It can be inferred from the box plot that the evening minutes across churned population is slightly more than that of the non churned population.
- The reason for this could be same like morning where more usage in minutes means more charges as minutes is highly corelated to charges which cause people to churn.



Total evening and night calls

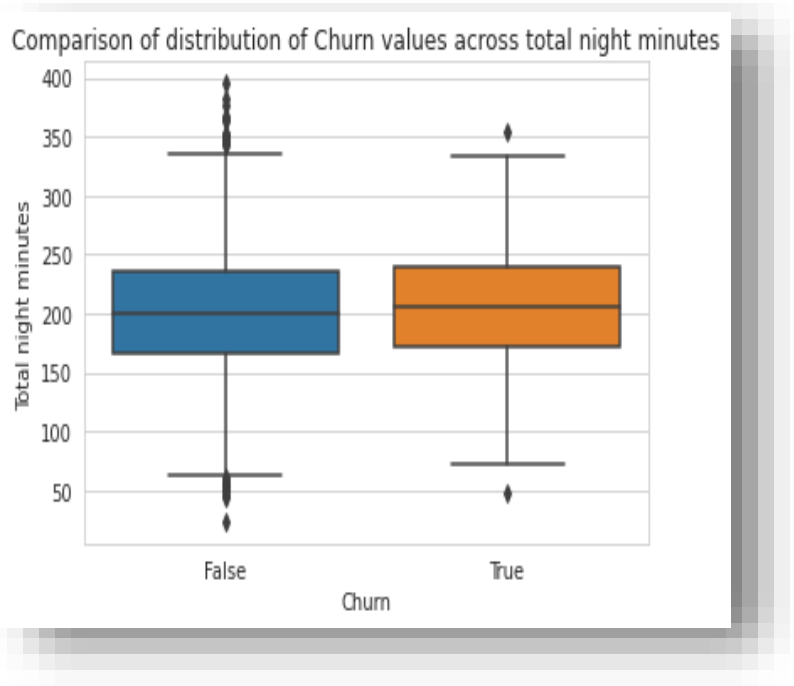
- The number of total evening and night call distribution is roughly same implying no significant relation to the churn rate.





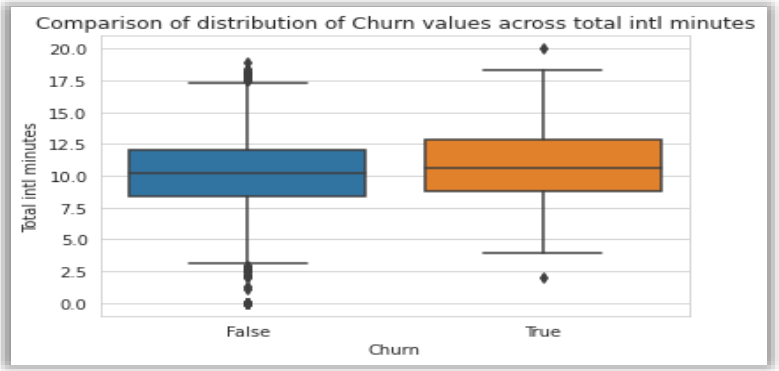
Night Minutes

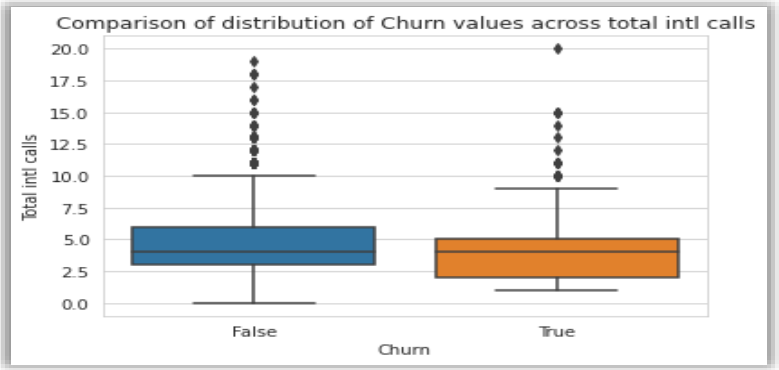
- It can be inferred from the box plot that the night minutes across churned population is slightly more than that of the non churned population.
- There is a correlation seen in both evening and morning minutes where churn is high in users who have more used as more minutes used is highly correlated to charge



International minutes and calls

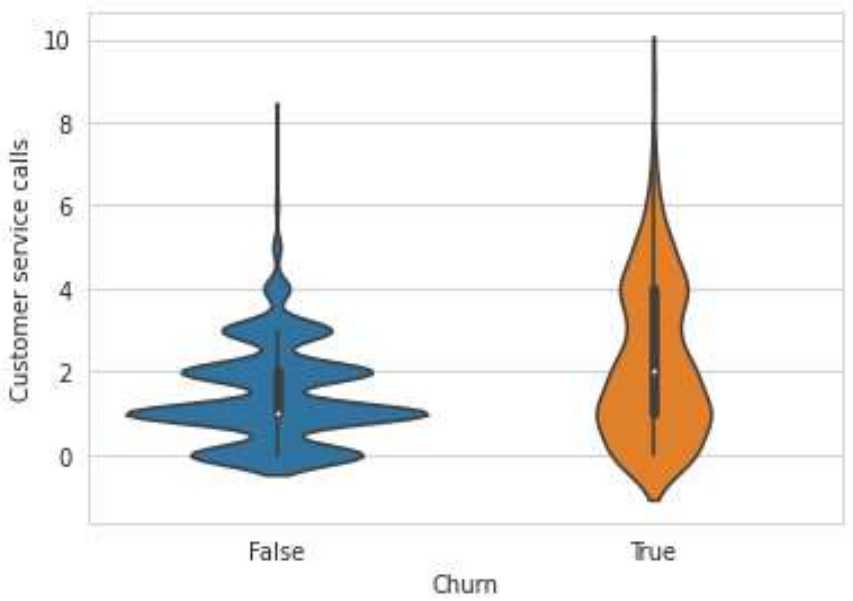
- From the box plot it is clearly seen that the international minutes across churned population is bit higher than that of the non churned population.
- It is also observed that the distribution of total international calls is negatively skewed. But most of values for churned is lower than the non churned





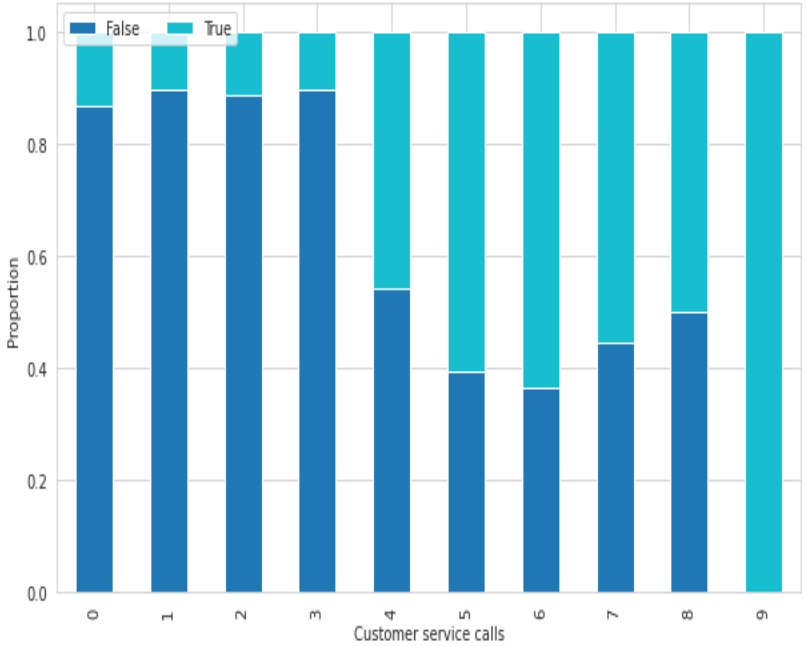
Customer service calls

- Majority of people are retained when the customer calls is just one which we can see from the violin plot after that the intensity of distribution of people retained reduces.



Customer service calls

- It can be observed that as number of customer service calls increase our churning rate is also increasing .
- In our bar chart it is very easy to see that as number of service calls are 3 or more then our churn rate increases drastically to more than 40 percent



Conclusion

- Company should introduce a longer validity prepaid plan with discounted prices so customers using more data get comparatively cheaper plan and at the same time retained, thus reducing churn rate.
- Company should focus on their quality of services as users start to churn with more number of customer service calls.
- Some states where churn rate is high like New Jersey, California, Texas, South Carolina should receive more focus and the company should implement new strategy here.
- Company should offer attractive benefits to gain new users and retain them.
- Company should offer better benefits to customers with international plans to retain them.
- **With voicemail plan there is a (16.7%) churn rate compared to people with no voice mail plan (8.67%). So introducing voicemail offers for customers with no voicemail plan can reduce churn by half.**

- The customers who retained service most of them did about 0-20 voicemail messages after that the distribution of retained customers is less indicating some underlying issue. So setting a limit of 20 voice calls might help reduce churn rate.
- The total night call distribution is similar to total evening call distribution, also the night minutes and evening minutes follow a similar pattern(which can help us reduce feature by making a feature including evening and night time values).Implementing this can reduce model complexity and can help reduce curse of dimensionality on applying dataset to an machine learning techniques.

THANK YOU