

CAPSTONE PROJECT - EDA

Title: Telecom Churn Analysis

Team members

Rahat Bhambri

Manjari Lahariya

Manoj Patil

Krishna Das

Likhith Ram

Index

1. Problem Statement
2. Descriptive Statistics
3. Data Analysis
4. Data Visualization
5. Conclusion

Topic :- Telecom Churn Analysis

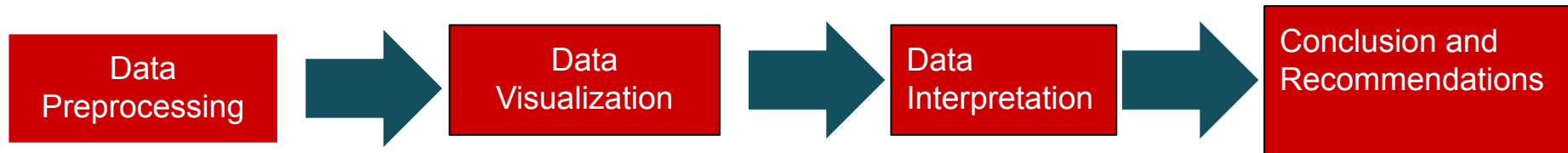
Customer churn is a major problem and one of the most important concerns for large companies. Due to the direct effect on the revenues of the companies, especially in the telecom field, companies are seeking to develop means to predict potential customer to churn. To meet the need of surviving in the competitive environment, the retention of existing customers has become a huge challenge. In the survey done in the Telecom industry, it is stated that the cost of acquiring a new customer is far more than retaining the existing one. Therefore, by collecting knowledge from the telecom industries can help in predicting the association of the customers as whether or not they will leave the company.



Data Set Overview

- Source dataset is in 'txt' format with '.csv'.
- Dataset contains 3333 rows and 20 columns.
- Out of the total, 19 Columns are independent variables and 1 is dependent variable.
- There is no missing values for the provided input dataset.
- Churn status is the variable which notifies whether a particular customer is churned or not.

Data Pipeline



Data Set Summary

1. **State :(String)**- It represents Letter codes for the different states where customers resides.
2. **Account Length :(Numerical)**- It shows how long the account has been active.
3. **Area Code :(String)**- These are the codes of the particular areas by which areas are identified.
4. **International Plan :(Yes/No)**- Yes - tells us customer had International Plan and No - tells us customer hadn't subscribed for International Plan.
5. **Voicemail Plan :(Yes/No)**- Yes- tells us customer had Voicemail Plan and No - tells us customer hadn't subscribed for Voicemail Plan.
6. **Number vmail messages :(Numerical)**- Messages done by customers.
7. **Total day minutes :(Numerical)**- Total number of minutes spent by customers in the morning.

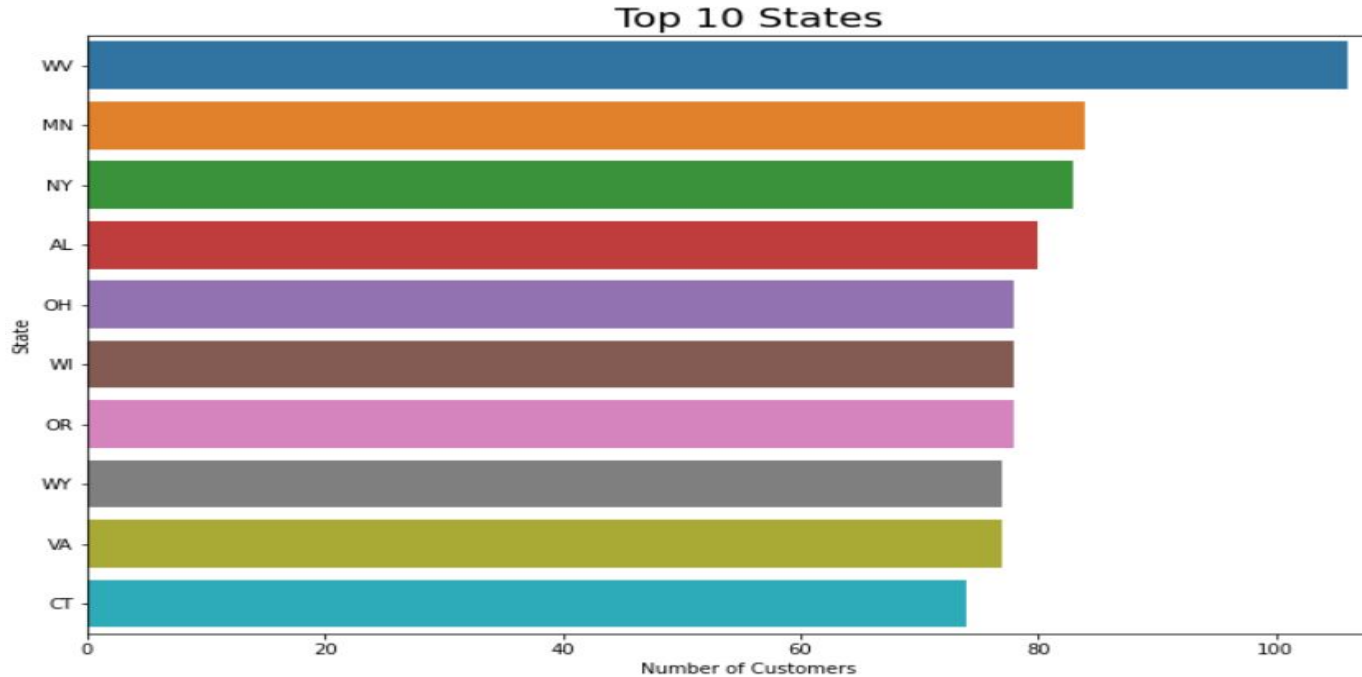
Data Set Summary

8. **Total day calls :(Numerical)-** Its Total of number of calls made by Customer in the morning.
9. **Total day charge :(Numerical)-** Its Total charge of services those are taken by the customers in the Morning.
10. **Total eve minutes :(Numerical)-** Total number of minutes spent by customers in the Evening.
11. **Total eve calls :(Numerical)-** Its Total of number of calls made by Customer in the Evening.
12. **Total eve charge :(Numerical)-** Its Total charge of services those are taken by the customers in the evening.
13. **Total night minutes :(Numerical)-** Total number of minutes Spent By Customers in the Night.
14. **Total night calls :(Numerical)-** Its Total of number of calls made by Customer in Night.

Data Set Summary

15. **Total night charge :(Numerical)-** Its Total charge of services those are taken by the customers in the Night.
16. **Total intl minutes :(Numerical)-** Total international minutes Spent By Customers.
17. **Total intl calls :(Numerical)-** Its Total of number of International calls made by Customers.
18. **Total intl charge :(Numerical)-** Its Total charge of International calls those are taken by the customers.
19. **Customer Service Calls :(Numerical)-** Number of calls to customer Service
20. **Churn :(Yes/No)-**This is our Dependent variable which describe that customers churned or not churned where, True represents customer churned and False represents customer not churned

Top 10 states from where more customers are associated with company

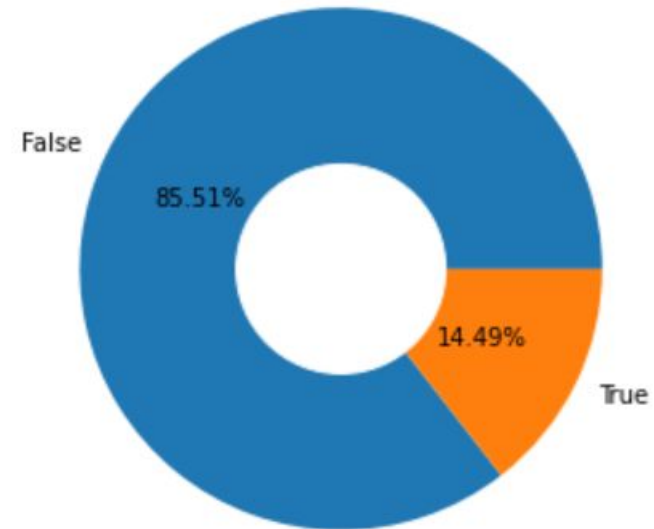


- WV has highest number of customers associated with the company

What percent of the churn is company suffering?

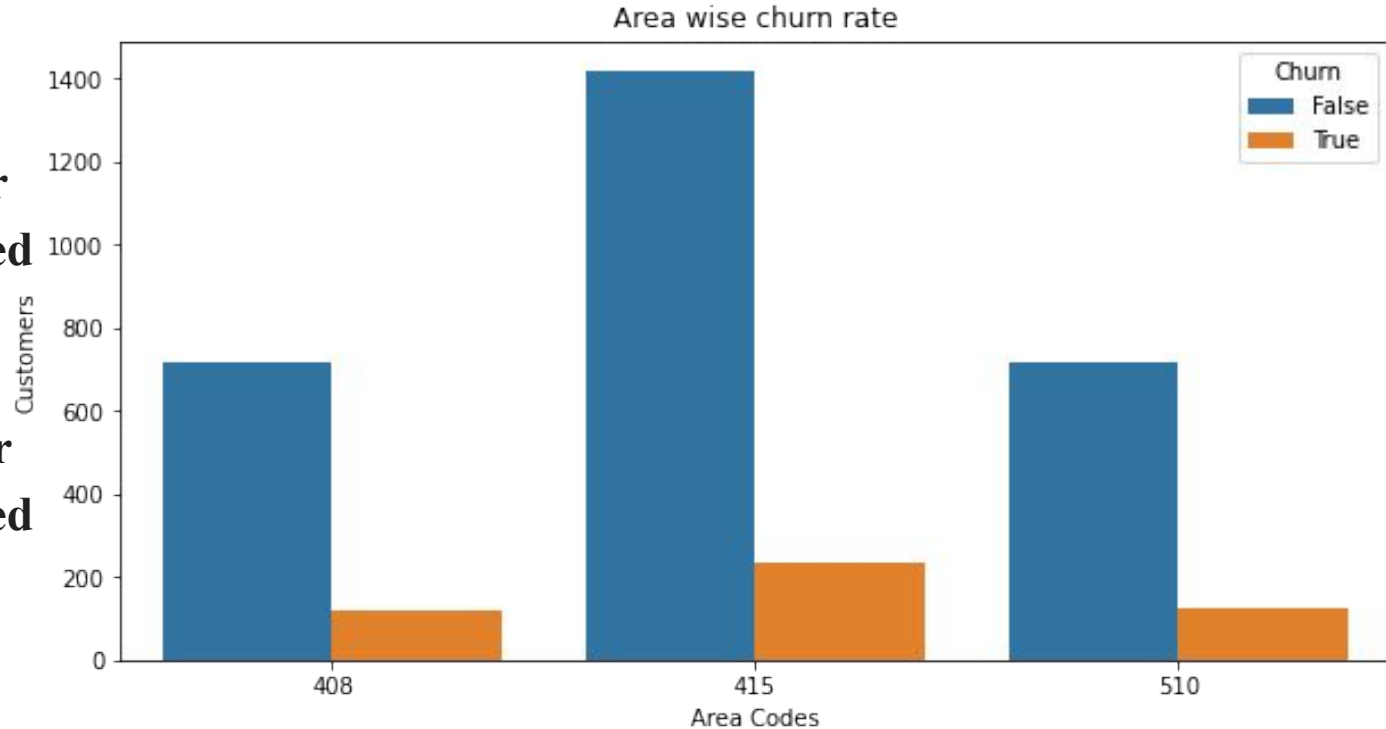
- **Overall Customer churn is 14.49%.**
- **85.51% customers have not churned.**

Churn rate Visualization



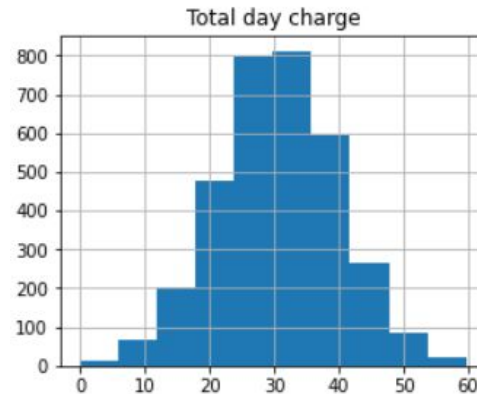
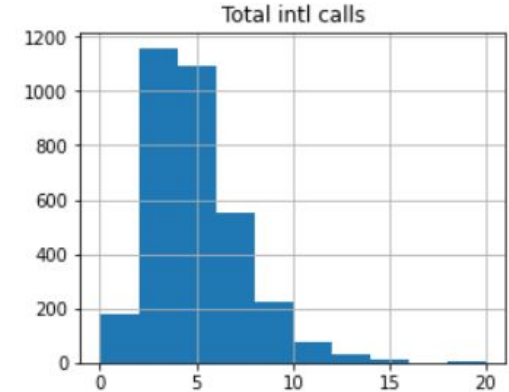
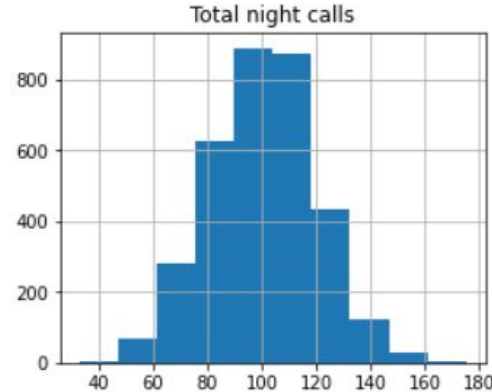
Area wise Customers Distribution

- For the Area Code '415' maximum number of customers are churned
- For the Area Code '408' minimum number of customers are churned



Distributions of the Features

- **Total night calls & Total day charge are roughly Symmetric.**
- **Customer Service Calls & Total intl Calls are Positively Skewed.**

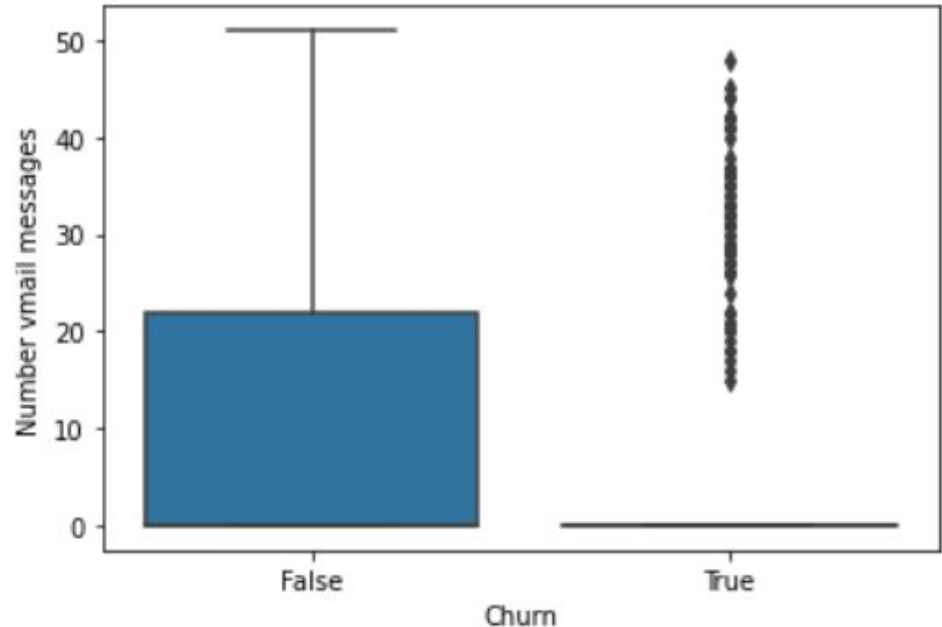


Voice mail messages affecting the Churn Rate

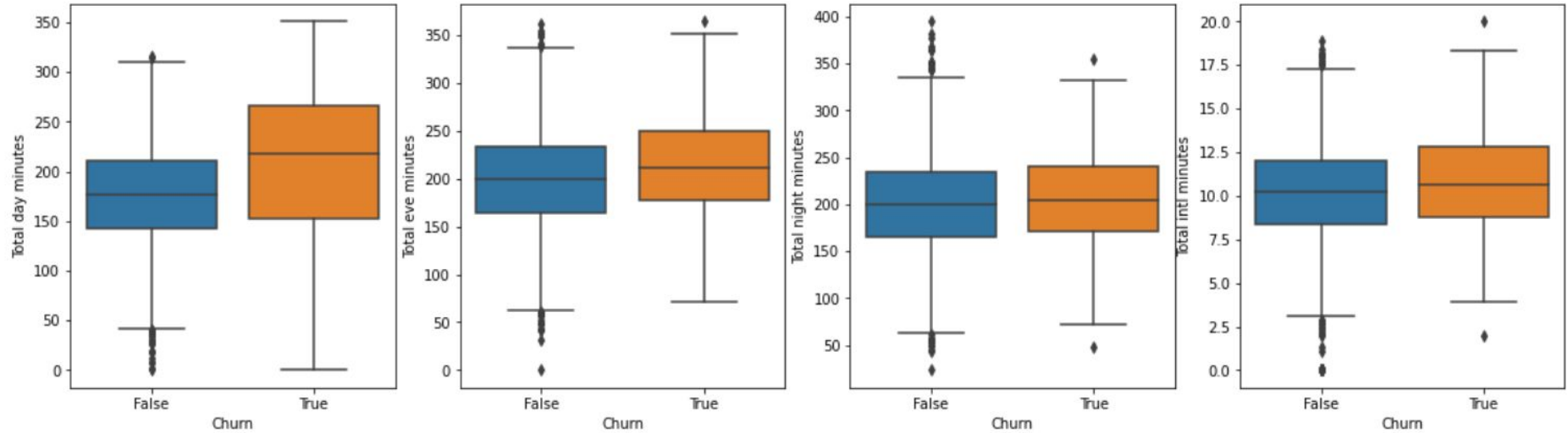
We can notice here for Vmail messages Feature, when there are more than 15 voice-mail messages then certainly there is a churn

According to Analysis following are the recommendations :

1. Voice-Mail Service Upgradation required
2. Provide better rate plans for Voice-Mail Service



Total-Minutes Affecting the Churn Rate

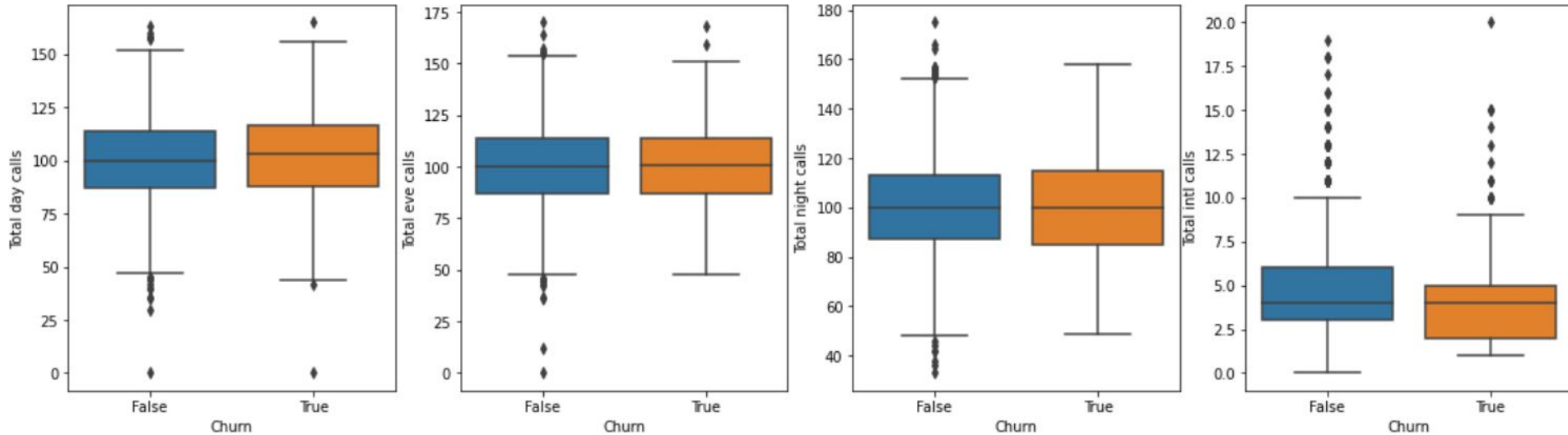


We can infer here that the customer who left the service provider spent more minutes in all types(day,night,eve,intl) of calls.

According to Analysis following are the recommendations :

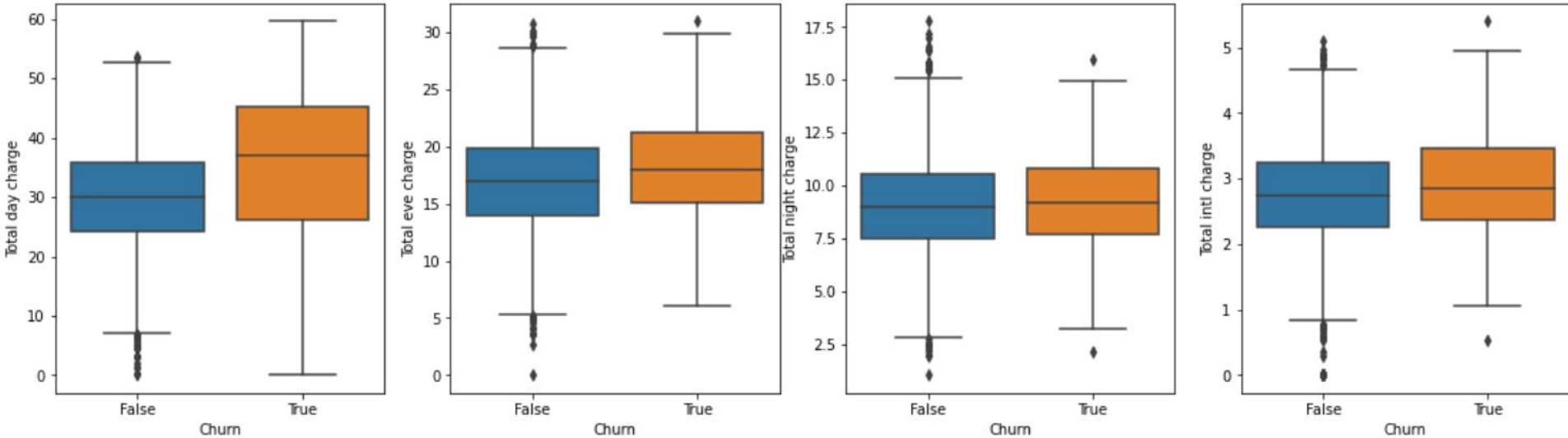
1. Network Upgradation
2. Removing Network Disturbance during a Call

Total-Calls Affecting the Churn Rate



- For the churned customer the median is slightly higher than 100 for Total day calls and Total eve calls.
- For the churned customer the median is almost equal to 100 for Total night calls
- For the churned customer the median is less than unchurned customers for Total intl calls

Total Call Charges Affecting the Churn Rate

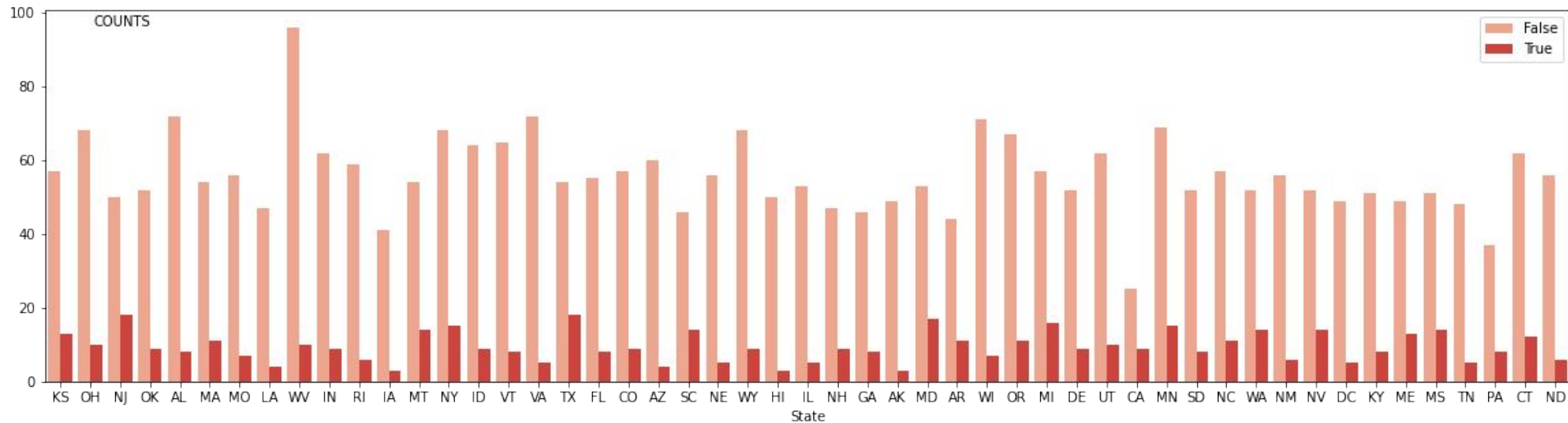


We can infer from above box-plots that Customers spending more on the service tend to leave the subscription and it clearly indicates that there is defect in the pricing strategy of the company.

According to Analysis following are the recommendations :

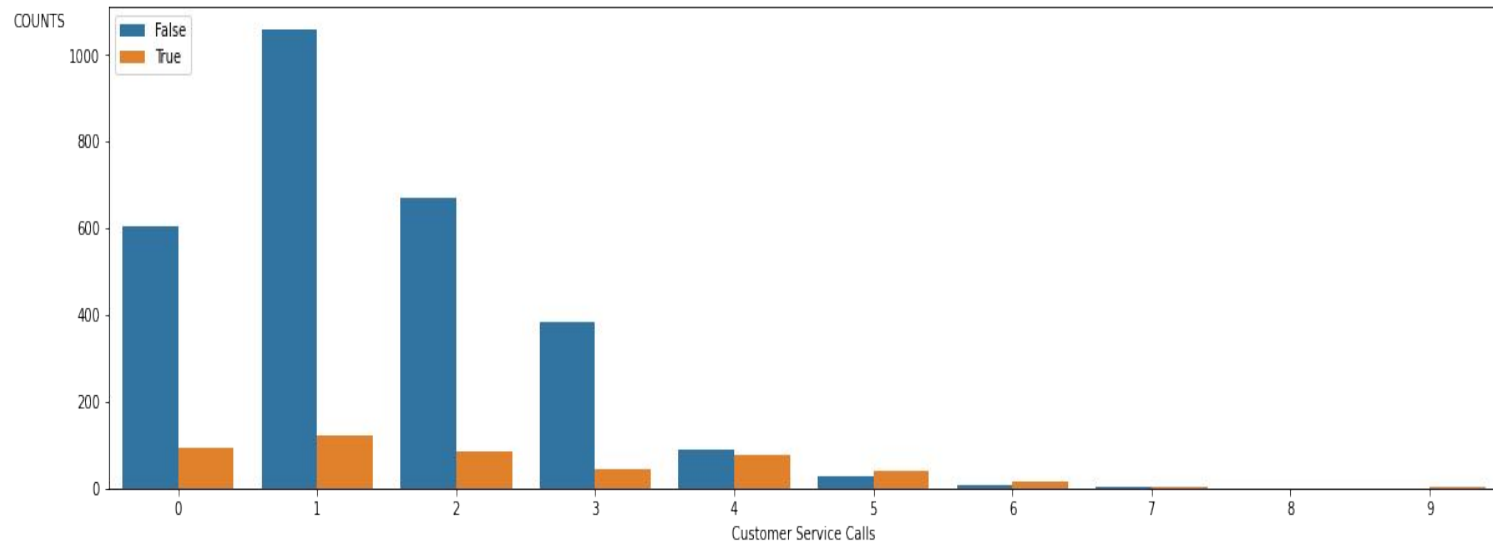
1. Strategy of pricing needs to be re-evaluated.
2. Premium customers should be offered prioritized services and good discounts.

State wise Churn Behavior



- Number of people Churned are higher in Texas(TX),Maryland (MD).
- It may be due to high competition from rival telecom companies.
- A Network Upgradation would be strongly suggested in these areas.

Customer service calls affecting churn



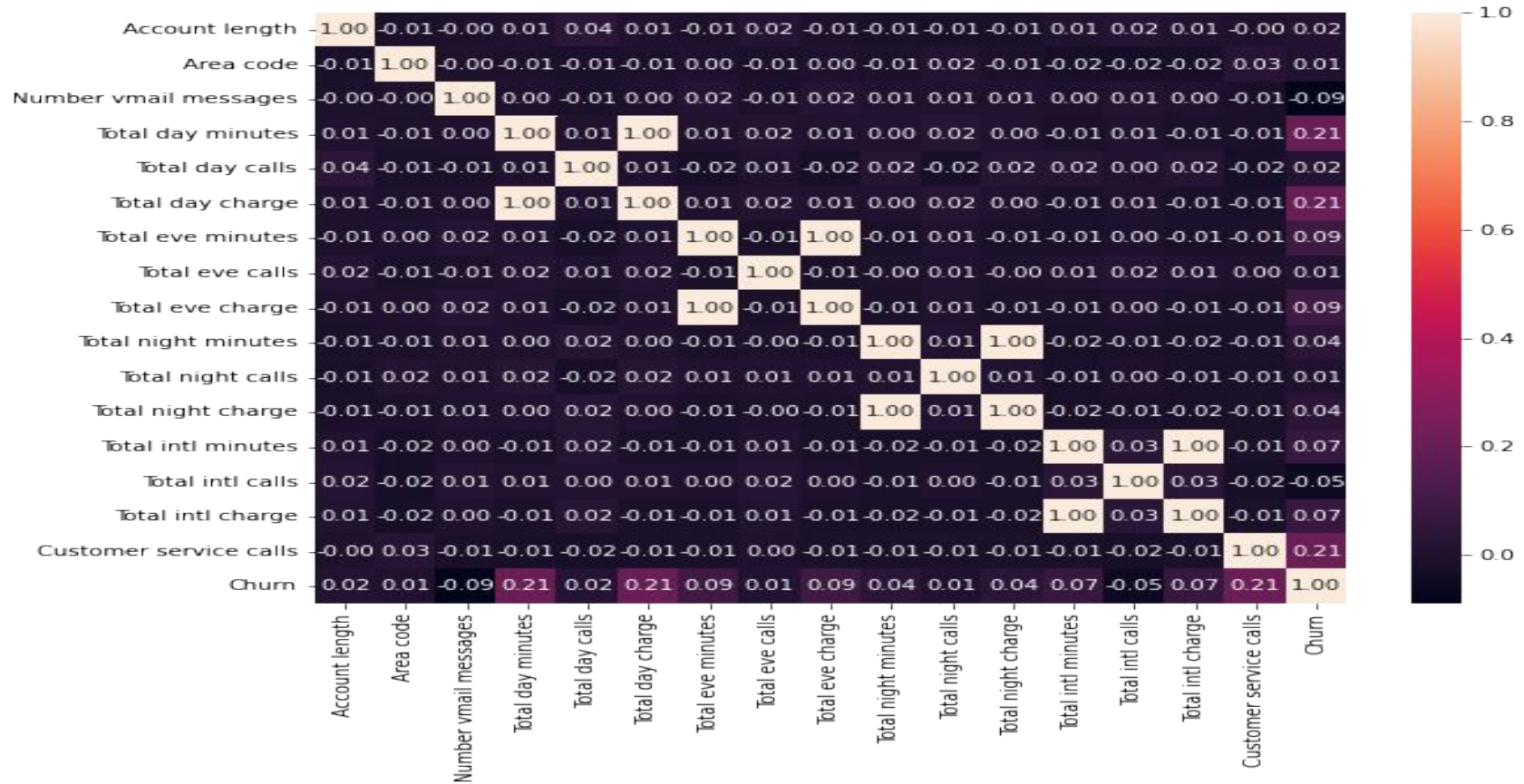
While some customers are lazy and hence without resolving the issue they have jumped to other telecom company,

While the customers who have called once also have high churn rate indicating their issue was not solved in first attempt.

According to Analysis following are the recommendations :

1. Company should take actions regarding to the customer feedbacks time to time.

Correlation in Data



Correlation in Data

- Total intl minutes and Total intl charges have very high positive correlation
- Total day charge and Total day minutes have very high positive correlation
- Total eve minutes and Total eve charges have very high positive correlation
- Total night minutes and Total night charges have very high positive correlation
- Total day charge, total day minute and customer service calls are weakly correlated with Churn variable.

Conclusion!

- Overall Customer churn is 14.49%.
- For the Area Code '415' maximum number of customers are churned.
- For Vmail messages Feature, when there are more than 15 voice-mail messages then certainly there is a churn.
- Number of people Churned are higher in Texas(TX),Maryland (MD).
- Customers who left the service provider spent more minutes in all types(day,night,eve,intl) of calls.
- Customers spending more on the service tend to leave the company.
- Total minutes(day,night,eve,intl) and Total charges(day,night,eve,intl) have very high positive correlation.
- Total day charge, total day minute and customer service calls are weakly correlated with Churn variable.

Thank You!