

# Capstone Project - 1

## Telecom Churn Analysis

By - Karan Malhotra

## Churn Analysis

1. Defining Problem Statement
2. EDA(Exploring and Analysing Data)
3. Feature Selection
4. Visualising Data
5. Driving Conclusions

# Points for Discussion

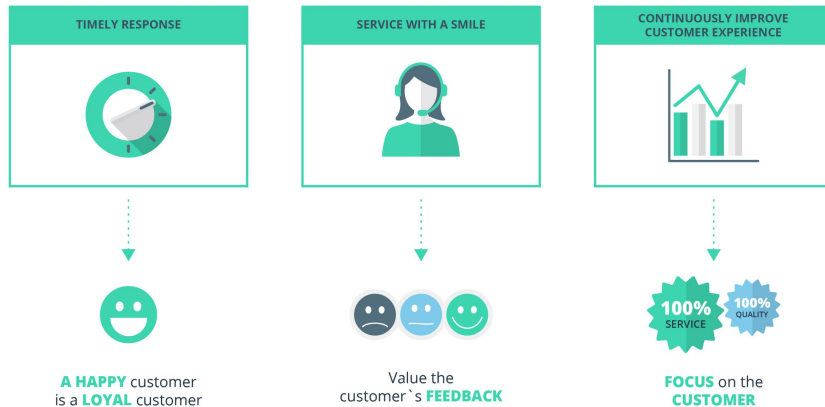
1. What is Churn Analysis
2. Data Summary
3. Analysing Churn Rate
4. Specific Causes of Churn
5. Preventing Customer Churn by analysis
6. Challenges while analysis
7. Conclusion

# The Dilemma

## How Telecom Churn Analysis Works?

Telcos apply machine learning models to predict churn on an individual customer basis and take counter measures such as discounts, special offers or other gratifications to keep their customers. For Telco companies it is key to attract new customers and at the same time avoid contract terminations (=churn) to grow their revenue generating base. Looking at churn, different reasons trigger customers to terminate their contracts, for example better price offers, more interesting packages, bad service experiences or change of customers' personal situations. The idea of this project is to analyse the data and reach out to some conclusions that can prevent churning.

## CUSTOMER SATISFACTION



# Data Pipeline

- Data processing -: In this first part we've learnt about columns and removed unnecessary features and found missing values if any.
- Data processing -: In this part, we manually go through each features selected from part 1, and analyse Churn rate.
- EDA: In this part, we do some exploratory data analysis (EDA) on the features selected in part-1 and 2 to see the Churn due to those features.

# Data Summary

1. Independent Variables - Account length, Number vmail messages, Total day minutes, Total day calls, Total day charge, Total eve minutes, Total eve calls, Total eve charge, Total night minutes, Total night calls , Total night charge , Total intl minutes, Total intl calls, Total intl charge, Customer service calls
2. Dependent Variables - Churn
3. Categorical Variables - International Plan, Voice mail Plan, State, Area code

# Data Summary

- 1.State: the US state in which the customer resides, indicated by a two-letter abbreviation.
- 2. Account Length: the length of the account.
- 3. Area Code: the three-digit area code of the corresponding customer's phone number
- 4. Int'l Plan: whether the customer has an international calling plan: yes/no
- 5. VMail Plan: whether the customer has a voice mail feature: yes/no

# Data Summary

- 6. VMail Message: presumably the average number of voice mail messages per month.
- 7. Day Mins: the total number of calling minutes used during the day.
- 8. Day Calls: the total number of calls placed during the day.
- 9. Day Charge: the billed cost of daytime calls.
- 10. Eve Mins: the total number of calling minutes used during the evening.
- 11. Eve Calls: the total number of calls placed during the evening.



# Data Summary

- 12. Eve Charge: the billed cost of evening time calls
- 13. Night Mins: the total number of calling minutes used during the night
- 14. Night Calls: the total number of calls placed during the night
- 15. Night Charge: the billed cost of nighttime calls
- 16. Intl Mins: the total number of international minutes
- 17. Intl Calls: the total number of international calls

# Data Summary

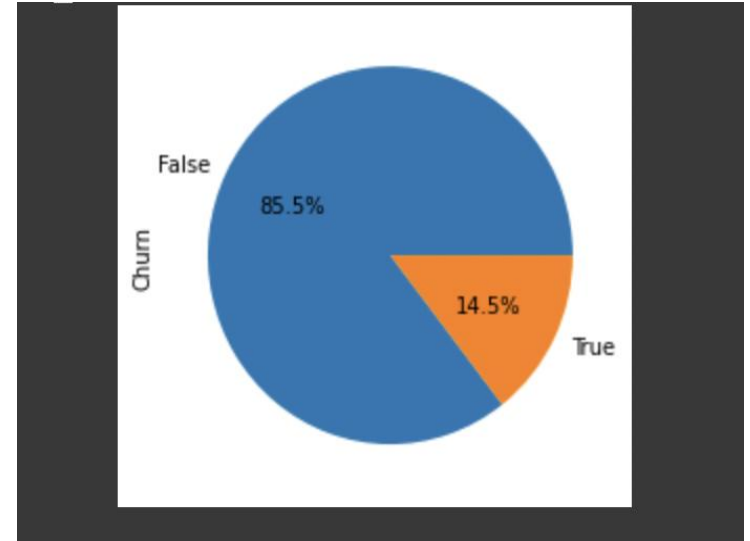
- 18. Intl Charge: the billed cost for international calls
- 19. CustServ Calls: the number of calls placed to Customer Service
- 20. Churn?: whether the customer left the service: true/false
- 21. Total Charge: the total amount charged by one customer

# Data-types of Column

```
#      Column      Non-Null Count  Dtype
---  -
0     State        3333 non-null    object
1     Account length 3333 non-null    int64
2     Area code     3333 non-null    int64
3     International plan 3333 non-null    object
4     Voice mail plan 3333 non-null    object
5     Number vmail messages 3333 non-null    int64
6     Total day minutes 3333 non-null    float64
7     Total day calls  3333 non-null    int64
8     Total day charge  3333 non-null    float64
9     Total eve minutes 3333 non-null    float64
10    Total eve calls   3333 non-null    int64
11    Total eve charge  3333 non-null    float64
12    Total night minutes 3333 non-null    float64
13    Total night calls  3333 non-null    int64
14    Total night charge 3333 non-null    float64
15    Total intl minutes 3333 non-null    float64
16    Total intl calls   3333 non-null    int64
17    Total intl charge  3333 non-null    float64
18    Customer service calls 3333 non-null    int64
19    Churn              3333 non-null    bool
dtypes: bool(1), float64(8), int64(8), object(3)
memory usage: 498.1+ KB
```

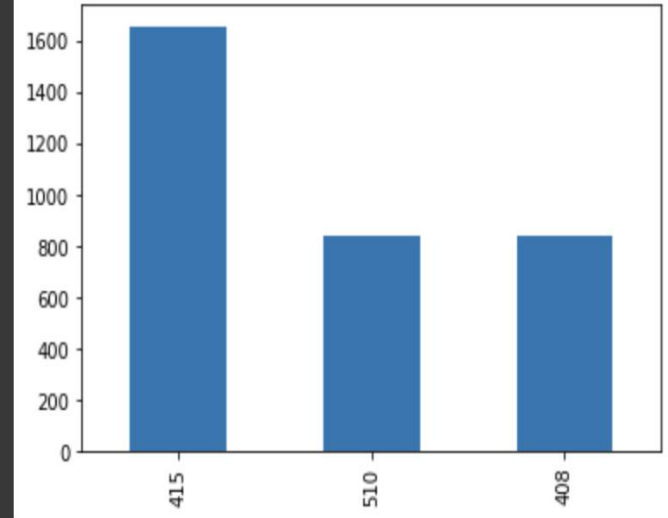
# Analysing 'Churn Rate'

Here we can see that the customers who chose to churn are roughly 14.5 percent whereas the loyal customers have a percentage of about 85.5 percent.



# Analysing Area Codes

We see here that there are 3 unique Area codes i.e 415,510,408 with area code 415 having the maximum population among the three.



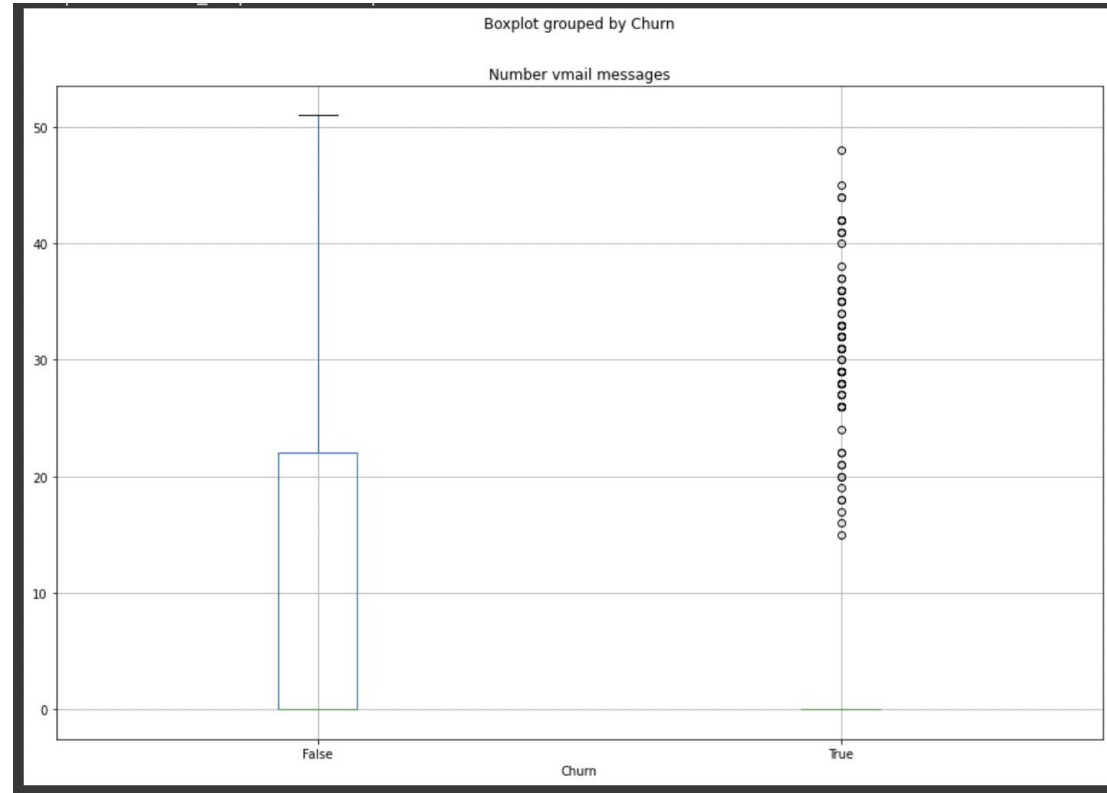
# Analysing Features or Columns

- Here we can see that the data is divided into
- 20 columns with each column having some
- number of unique fixed values while other values
- are repetition of those unique values.

```
Unique values :  
State                51  
Account length      212  
Area code           3  
International plan   2  
Voice mail plan      2  
Number vmail messages 46  
Total day minutes   1667  
Total day calls      119  
Total day charge     1667  
Total eve minutes    1611  
Total eve calls      123  
Total eve charge     1440  
Total night minutes  1591  
Total night calls    120  
Total night charge   933  
Total intl minutes   162  
Total intl calls     21  
Total intl charge    162  
Customer service calls 10  
Churn                2  
dtype: int64
```

# Voice Mail Feature

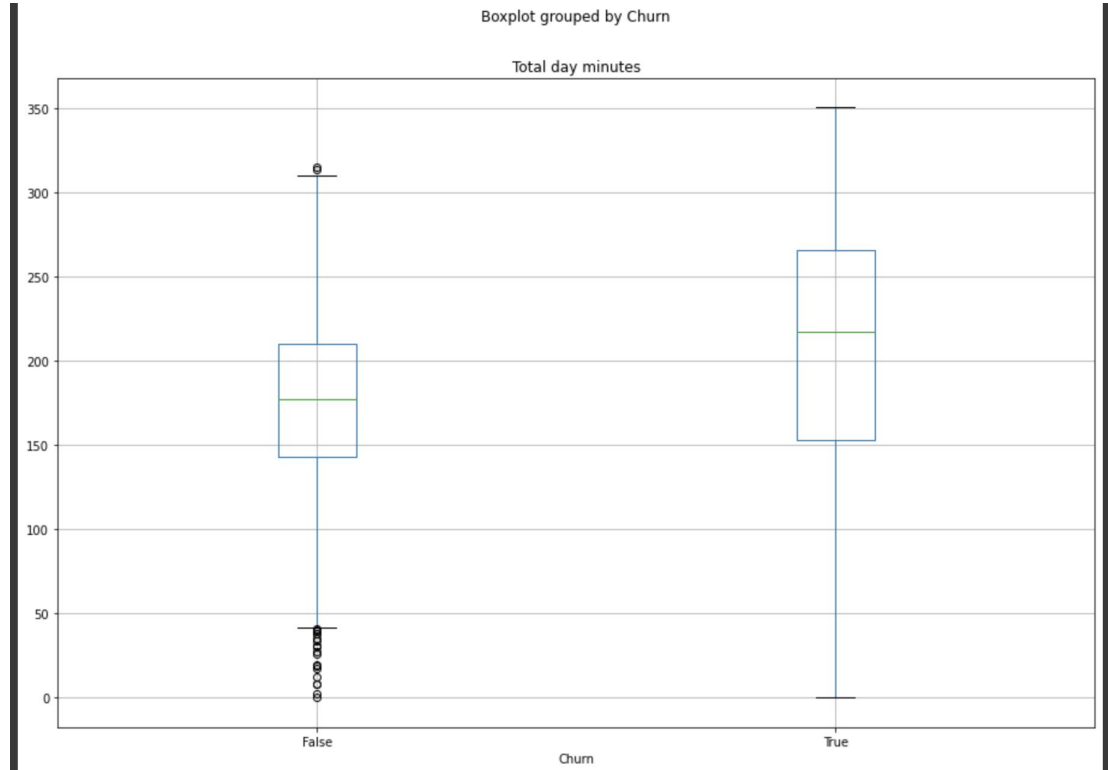
Here we check the voice Mail Feature of the company and realize that for Voice-Mail Feature when there are more than 20 voice-mail messages then certainly there is a churn indicating improving the voice-mail feature or setting a limit and check whether a customer is retained.



# Total-Minutes in Morning Affecting the Churn Rate

We can infer from above box-plot that with users spending more 225 minutes or more i.e. approx 4hrs tend to switch to other operator.

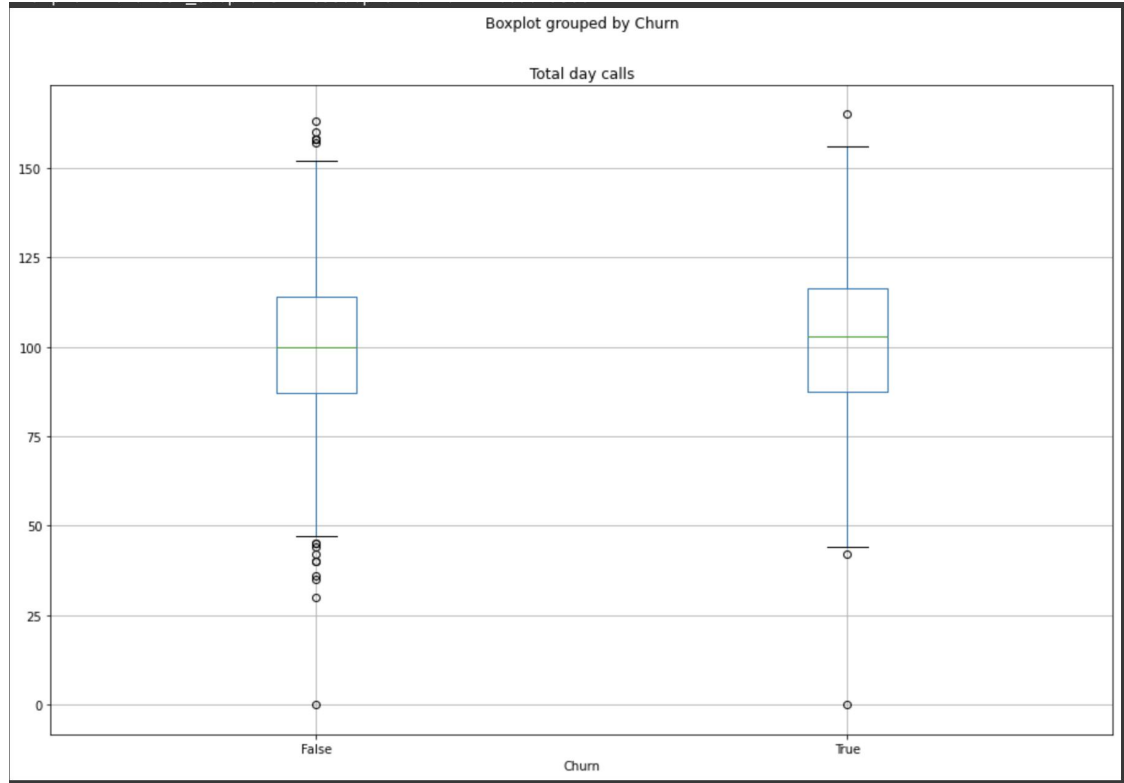
Need to Upgrade or make smarter use of technologies like VoLTE to improve Voice Quality.





# Total Day calls

We can infer here that on an average a 100 calls are made which is a good indication for the company. But we can also note that for the churn customer the median is slightly higher than 100 which indicates there are call drops which may lead to more calls in a morning.

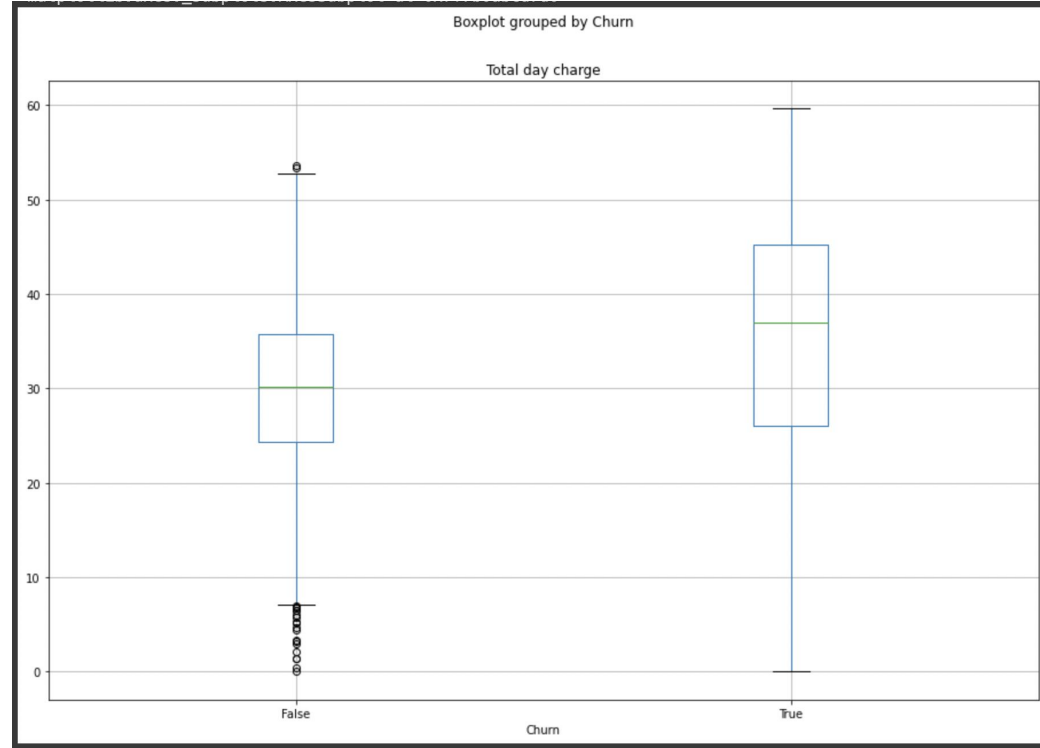


# Total day Charge

Here we can clearly indicate a good strategy to be implemented. As from above inferred box-plots we can conclude one thing i.e Customers having more minutes spent on the network tend to leave the it's subscription and from the above box-plot it clearly indicates that there is defect in the pricing startegy of the company.

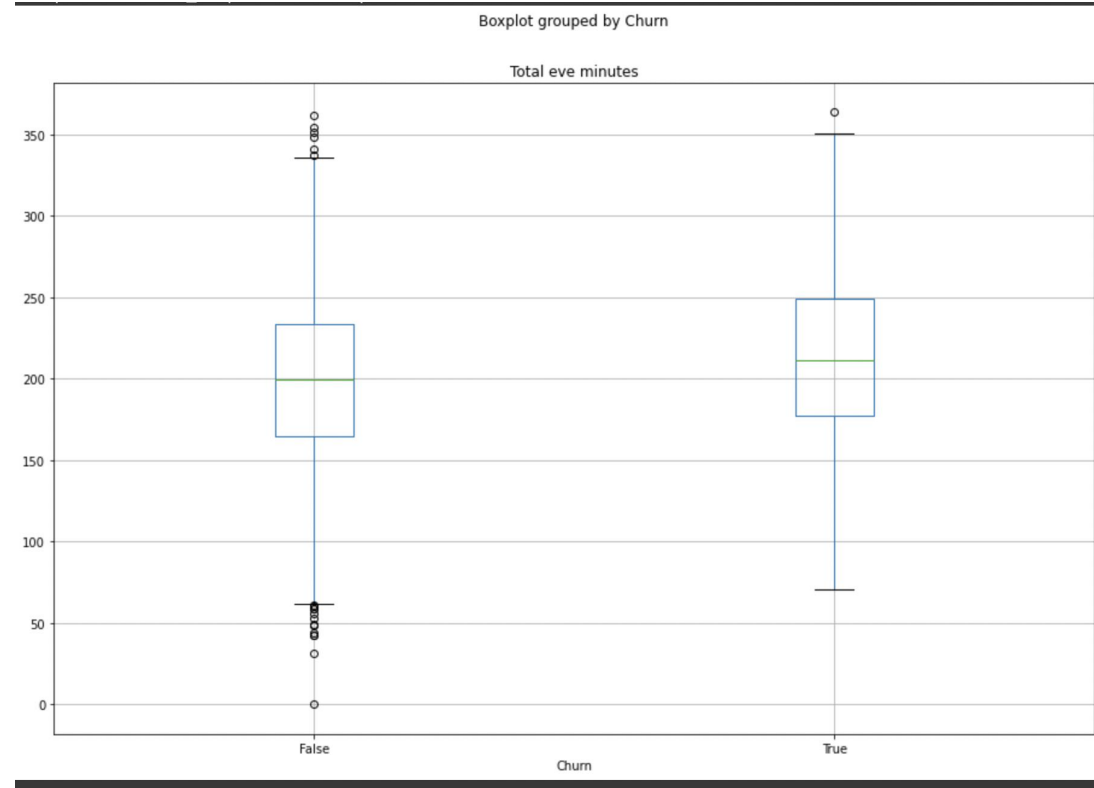
According to my Hypothesis:

- 1.Strategy of pricing needs to be re-evaluated.
- 2.The Clients who have high call minutes and calls need a discount in the end.



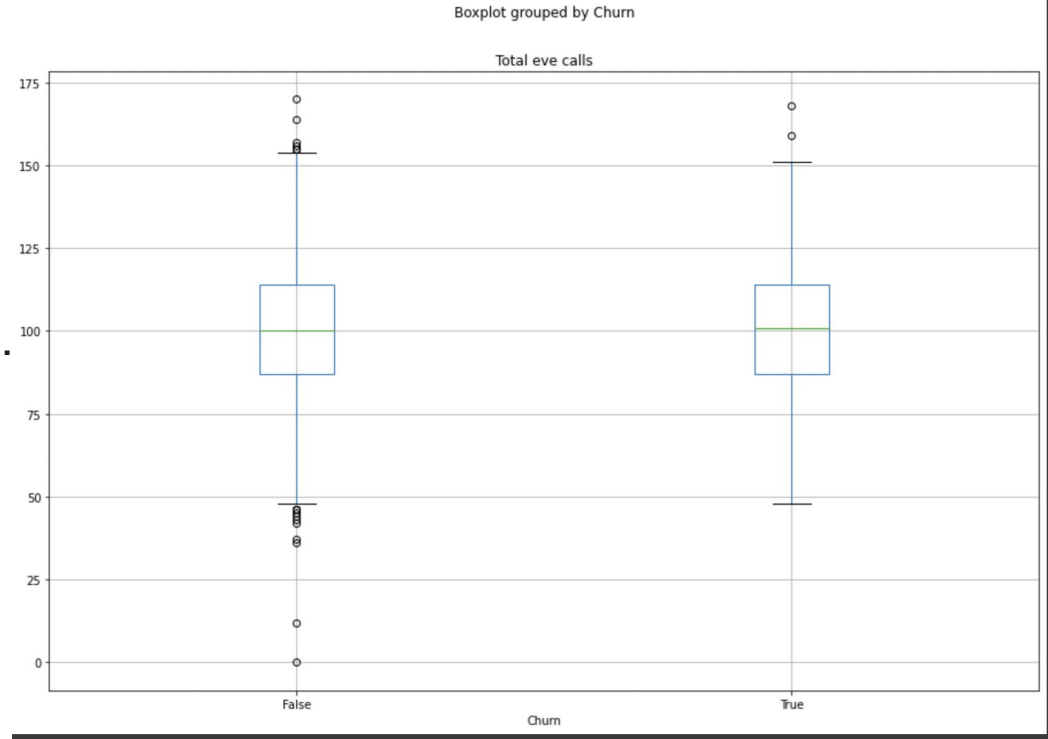
# Evening time Affecting the Churn Rate

The churned customers tend to spend more time in evening as compared to the loyal customers.



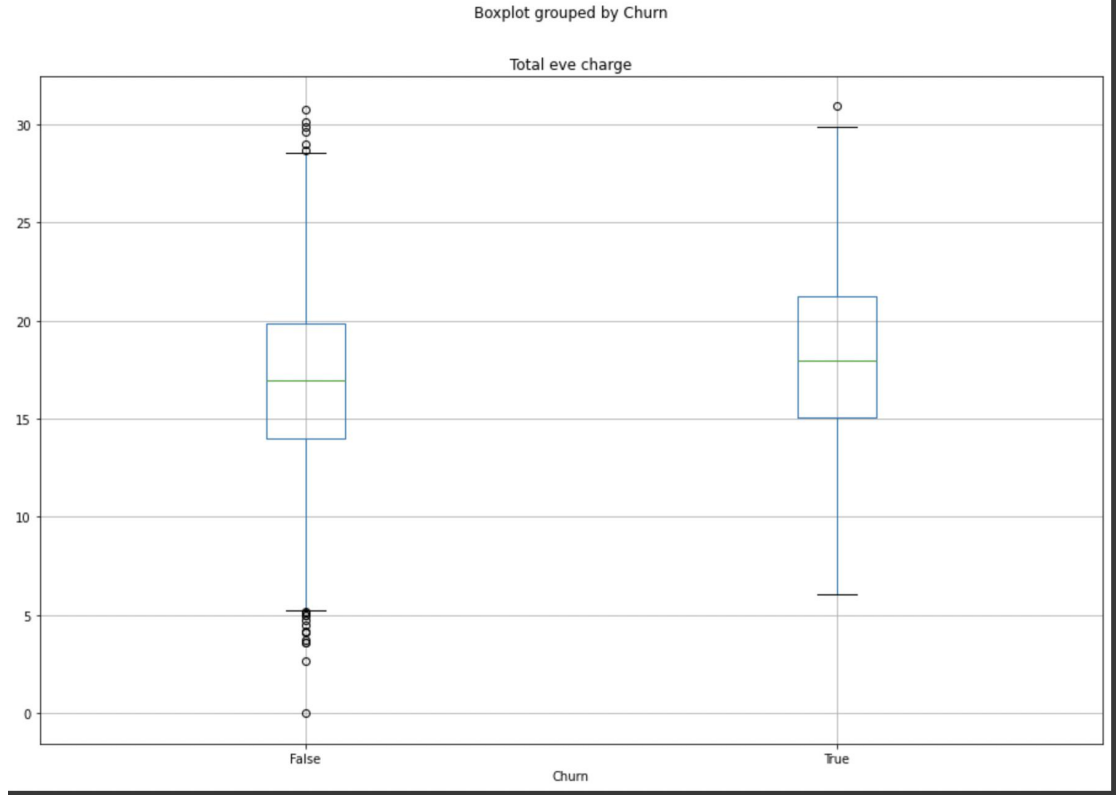
# Evening Calls

Total number of calls in the evening of churned customers is slightly more than the total no of calls of loyal customers.



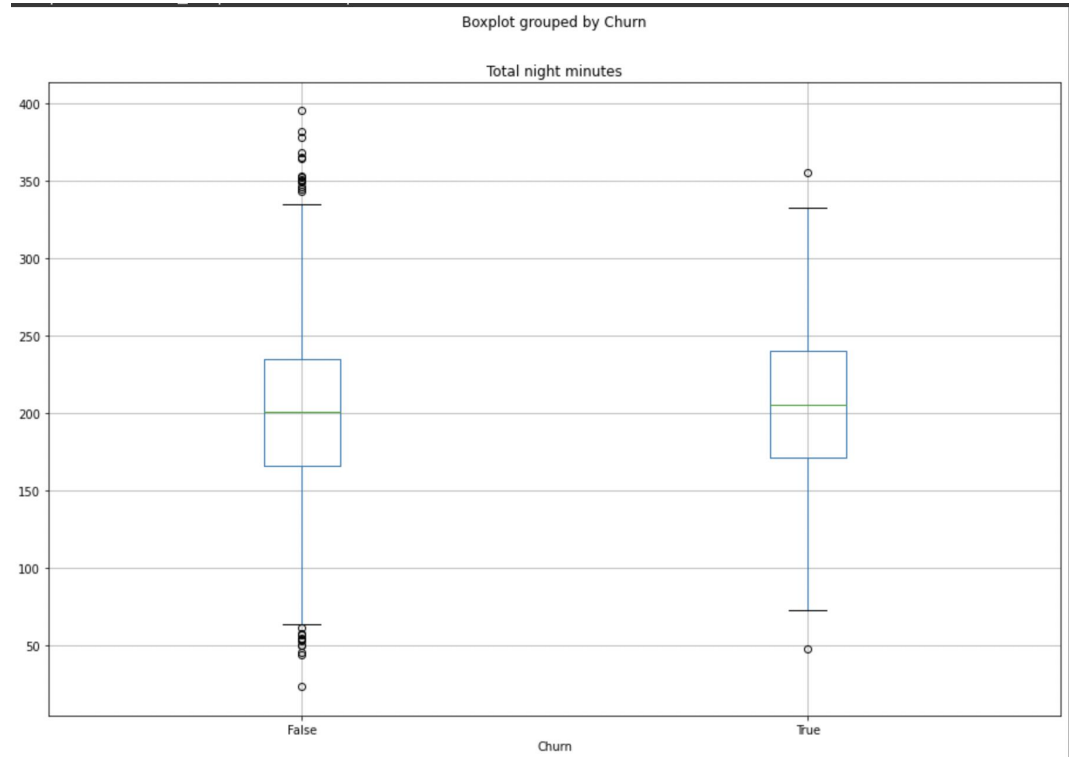
# Evening Charge

- The customers who had
- higher calls in the evening
- are the churned customers.



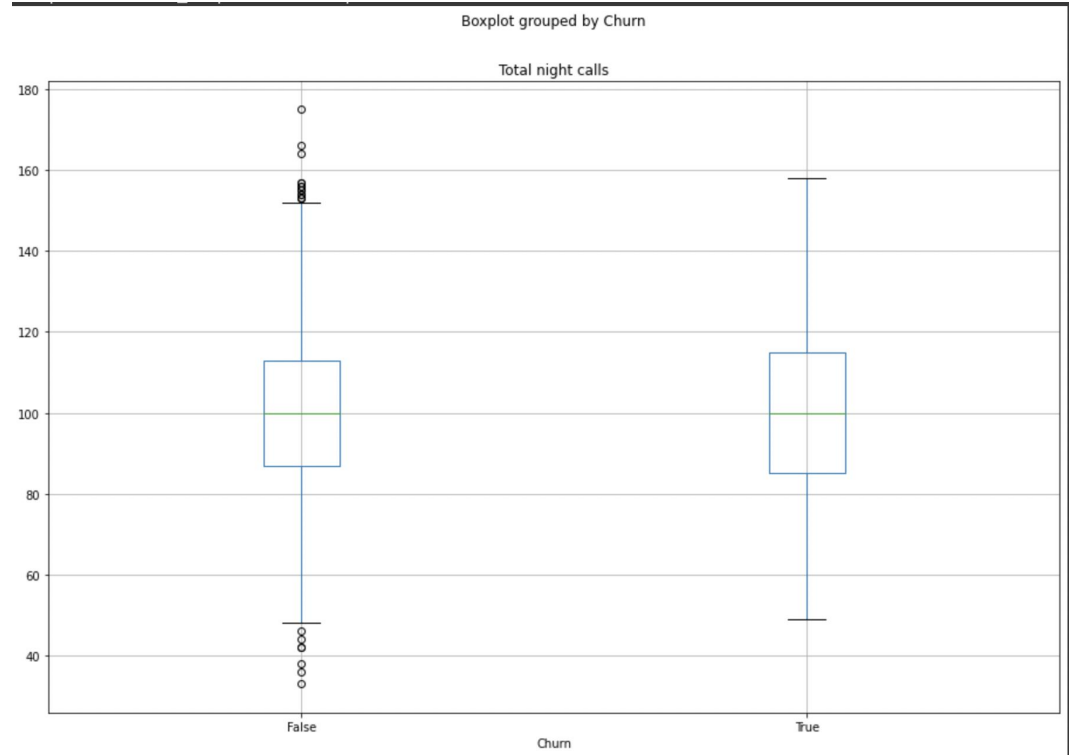
# Night Time Affecting Churn Rate

The Loyal Customers spend more time in night according to the box-plot.



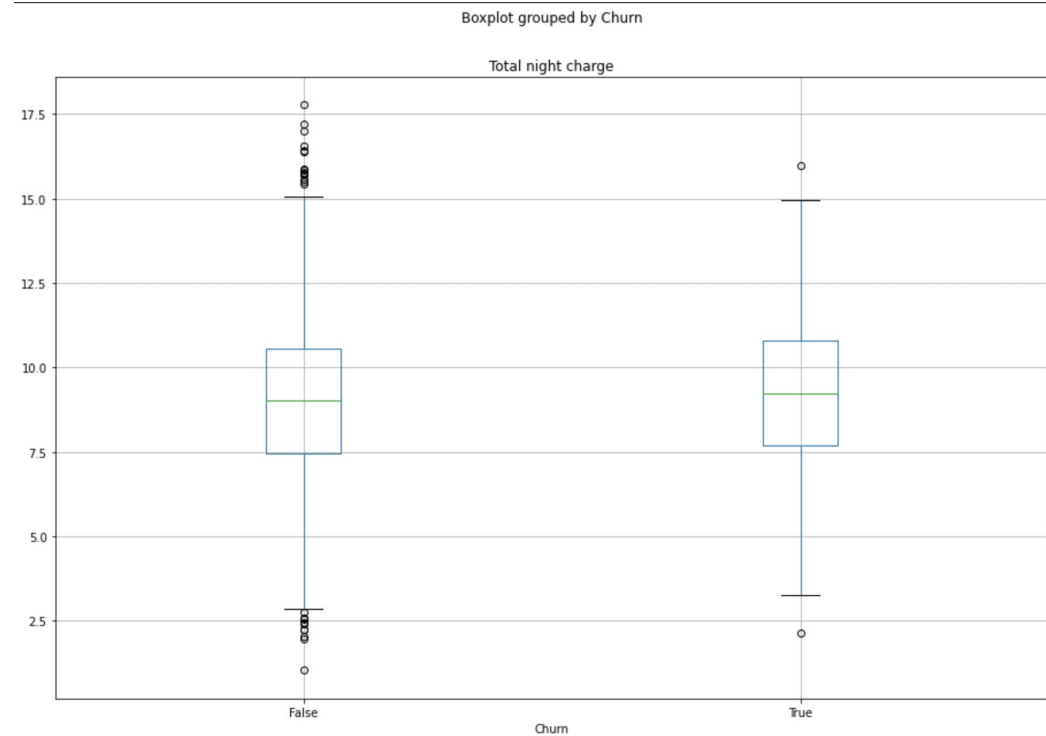
# Total Night Calls

Still Calls are made more by the churned customers.



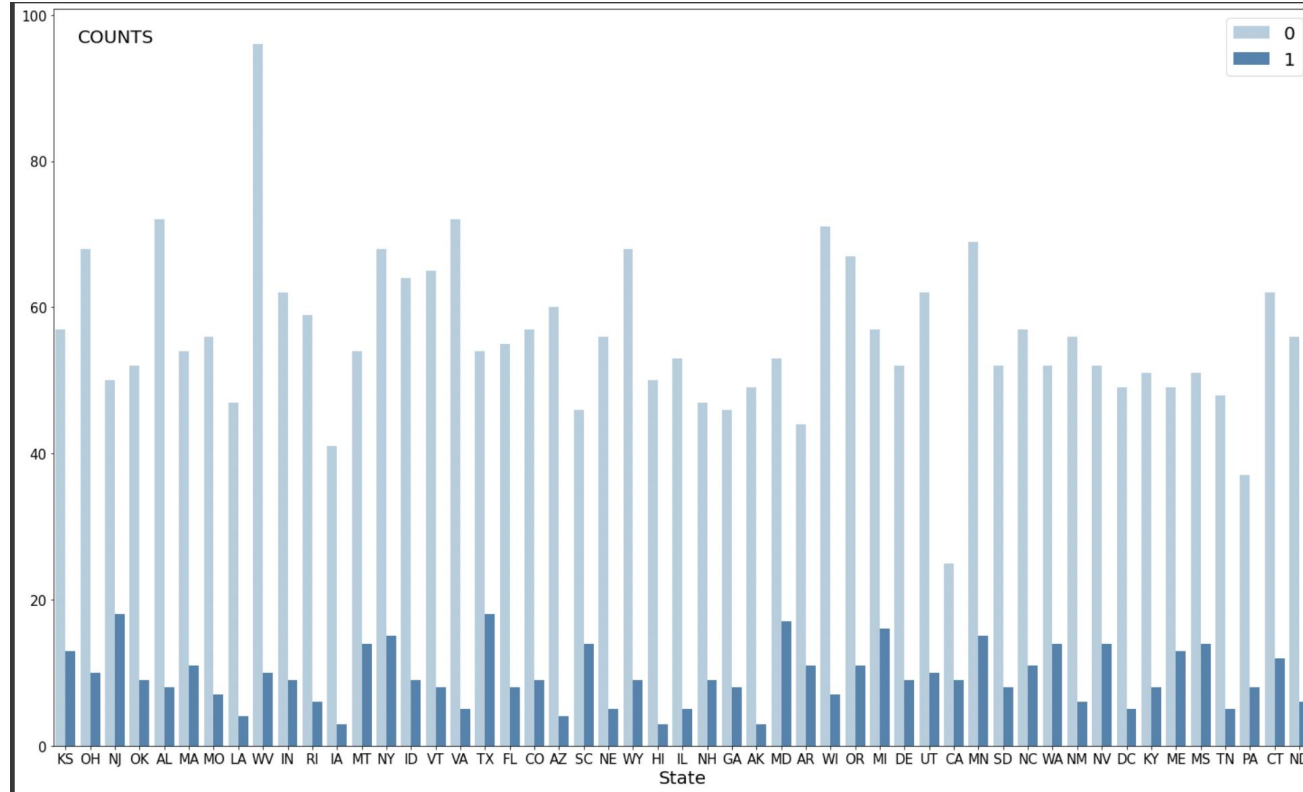
# Total Night Charges

Optimization of The Charges would lead to a loyal customer.





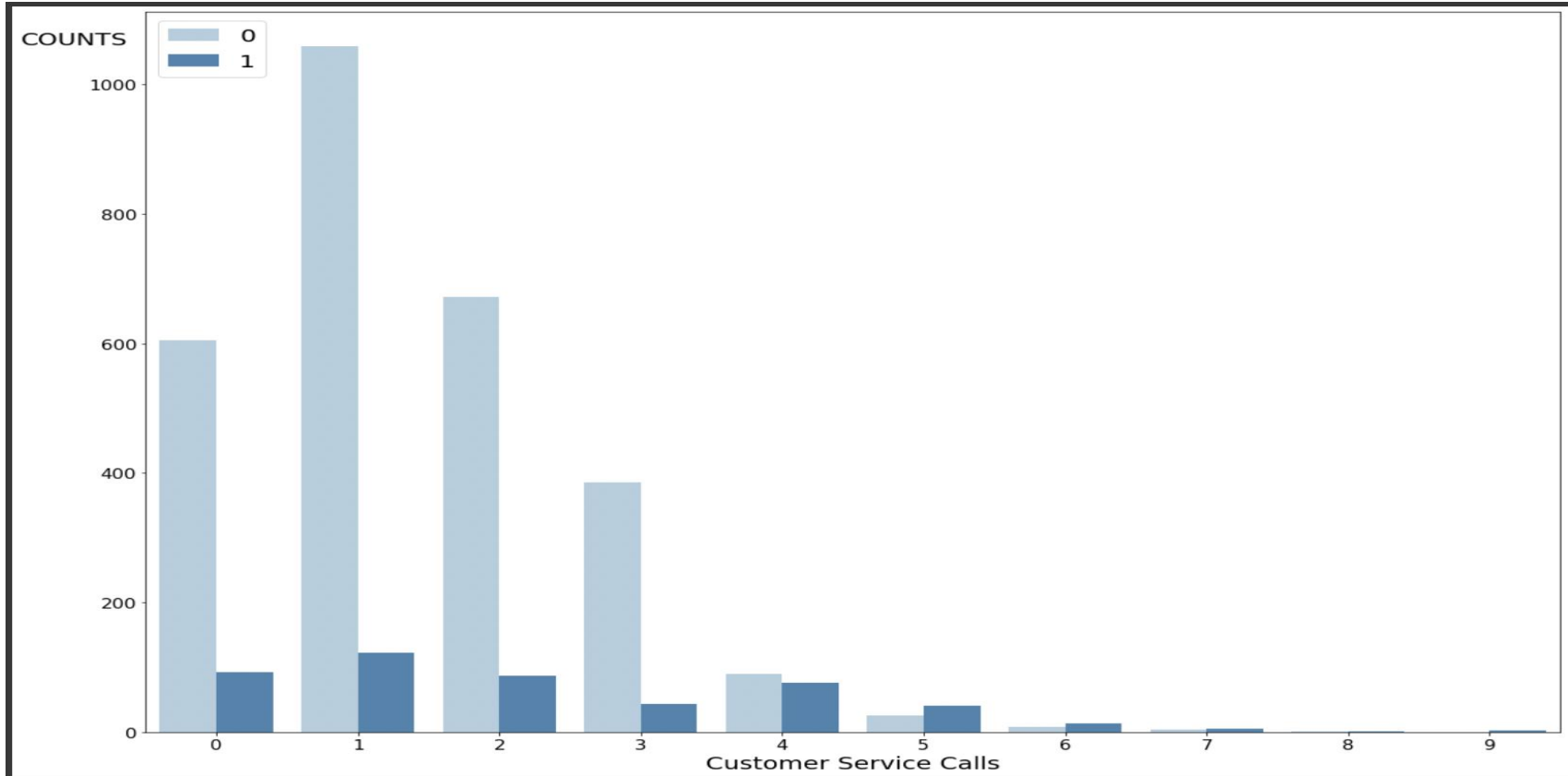
# International Calls Affecting the Churn Rate



# International Calls Affecting the Churn Rate Continue ..

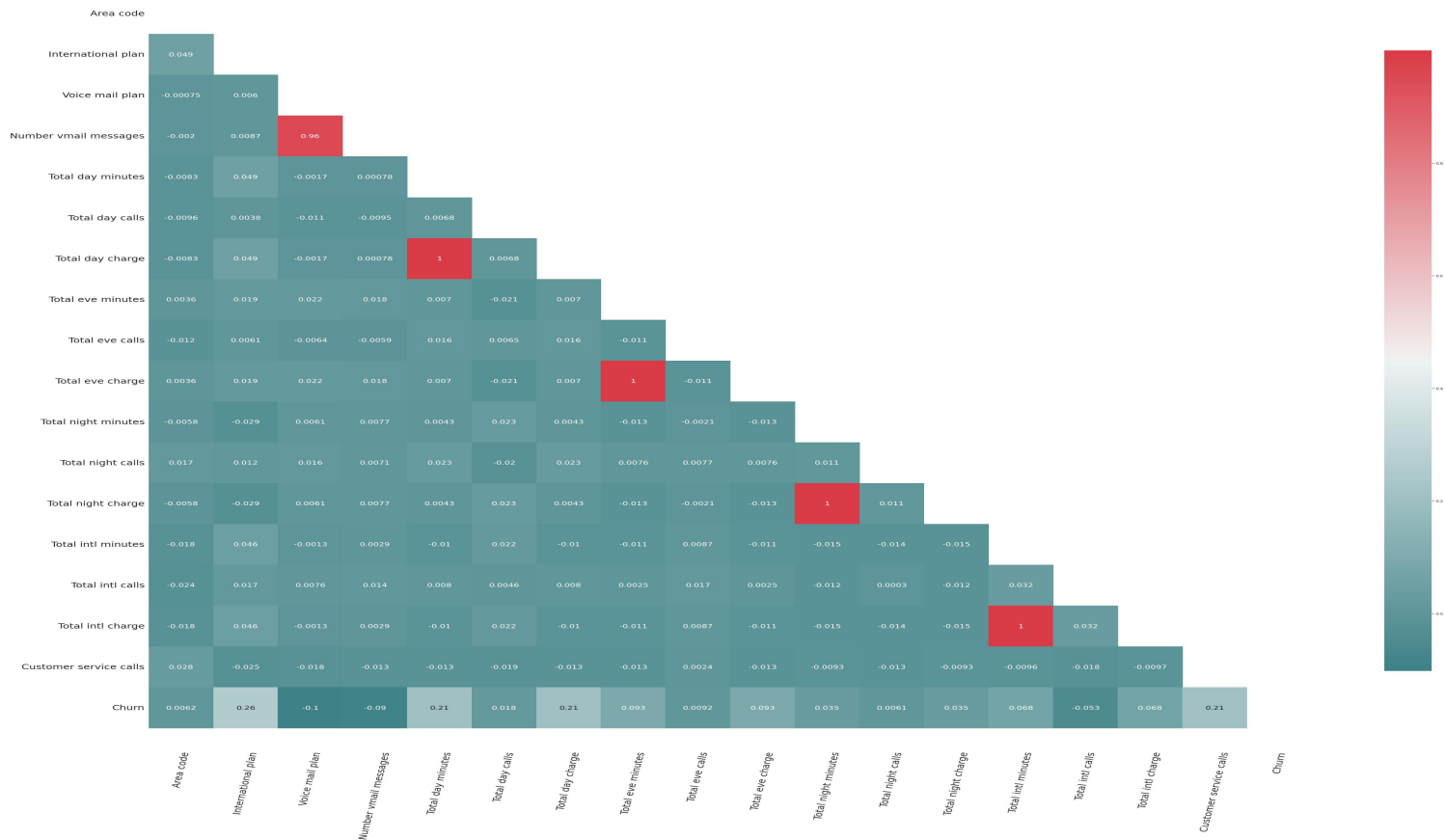
- We can see a bar graph of different states and the total customers churned or loyal and is observable that states like Texas and Maryland have the bit more churn rate than usual, A Network Upgradation would be strongly suggested in these areas.

# Customer Service Calls affecting Churn rate



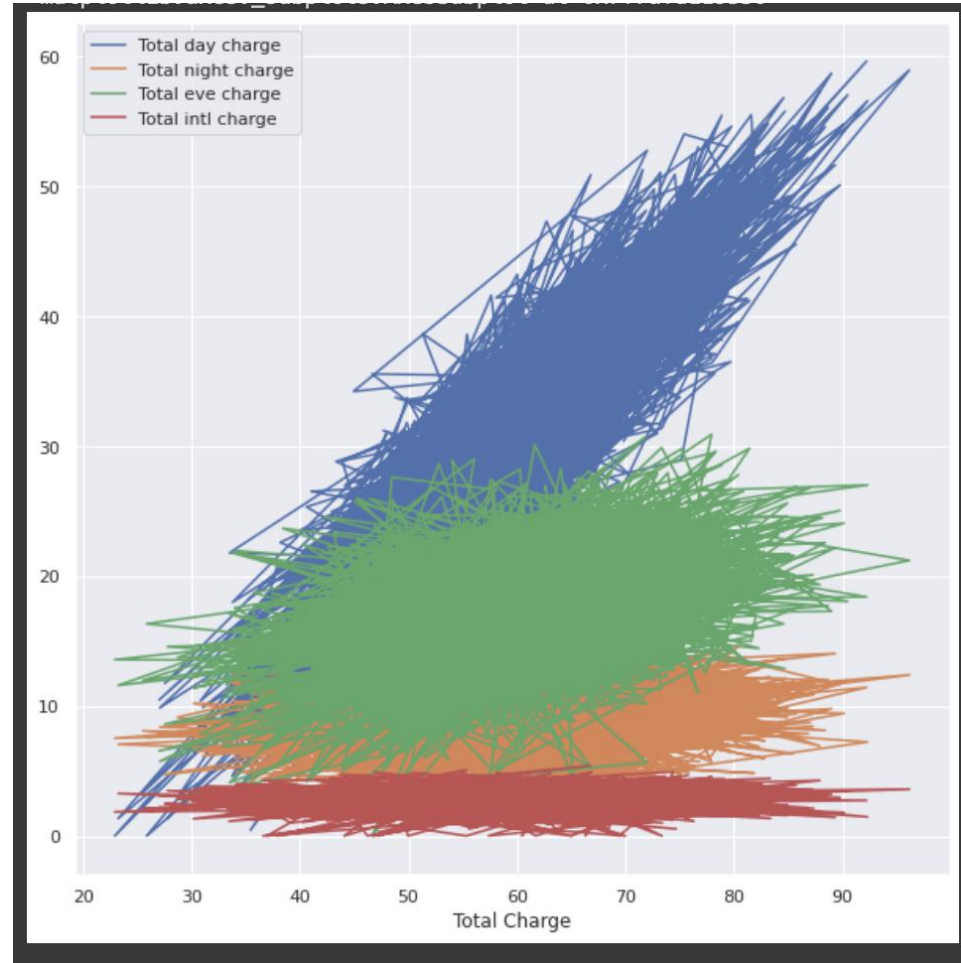
# Customer Service Calls affecting Churn rate Continue..

- While some customers are lazy and hence without resolving the issue they have jumped to other network operator, while the customers who have called once also have high churn rate indicating their issue was not solved in first attempt.
- A Feedback is necessary in such situations.
- A Confirmation to the Customer is important, that their issue would be solved in first attempt or not.



## EDA continued..

This line graph depicts the total charge against Total day, night, eve and international charges with total charge being at the x axis and other variables plotted on the y axis.



# Conlusions

We can definitely suggest to prevent churn :

1. Upgrading network to improve services for long duration users.
2. Updating Pricing Strategies.
3. Updating and Optimizing Internationall Call Rates.
4. Implementing a better network infrastructure in Maryland and Texas Areas where there is more Churn Rate.
5. Upgrading their services when in emegency only in evening period as low network traffic.
6. Consider an overall discount since the price is always one of the major factors for customers to choose among existing incumbents.

## Conclusions continue..

7. If the customer perceives the services received as positive and satisfactory, it leads to higher service quality in terms of reliability, responsiveness and other tangible factors. Higher service quality built confidence and trust with the services received.



# Challenges

1. To evaluate the techniques used in the churn prediction.
2. Churn is a complicated phenomenon, which involves deciphering each and every nerve of customer behaviour, analysing technological advancements.
3. Customer retention is very challenging in mobile telecom industry.
4. Churn Prediction is essentially predicting which clients are most likely to cancel a subscription i.e 'leave a company' based on their usage of the service. From a company point of view, it is necessary to gain this information because acquiring new customers is often arduous and costlier than retaining old ones.

# Q & A