Capstone Project

Telecom Churn Analysis

By,
Kumar Anand
Data Science Trainee at AlmaBetter

Agenda

- Introduction
- Problem Statement
- Attributes
- Steps Involved in this EDA
- Visualization
- Conclusion

Introduction

- Churn:- Churn is the attrition of customers from their telecom providers. Churn is a problem for telecom companies because its expensive to acquire new customers rather than keeping an existing one.
- Analyze telecommunication company customer data to predict whether or not a customer is likely to leave the platform (churn)
- Highlighting main variables/factor influencing customer churn
- Exploratory Data Analysis (EDA) is an approach of analyzing data sets to summarize their main characteristics, often using statistical graphics and other data visualization methods.
- The main goal or objective of this project is to perform EDA on telecom churn data to discover the customers who are likely to churn and factors affecting it.

Problem Statement

Orange S.A., formerly France Télécom S.A., is a French multinational telecommunications corporation. The Orange Telecom's Churn Dataset, consists of cleaned customer activity data (features), along with a churn label specifying whether a customer canceled the subscription.
Explore and analyze the data to discover key factors responsible for customer churn and come up with ways/recommendations to ensure customer retention.

Attributes

STATE: States in United States of America
AREA CODE: Code of area
INTERNATIONAL PLAN: Yes Indicate International Plan is Present and No Indicates no subscription for International Plan
VOICE MAIL PLAN: Yes Indicates Voice Mail Plan is Present and No Indicates no subscription for Voice Mail Plan
TOTAL DAY MINUTES: Total Number of Minutes Spent By Customers in Morning
TOTAL DAY CALLS: Total Number of Calls made by Customer in Morning.
TOTAL DAY CHARGE: Total Charge to the Customers in Morning.
TOTAL EVE MINUTES: Total Number of Minutes Spent By Customers in Evening
TOTAL EVE CALLS: Total Number of Calls made by Customer in Evening.
TOTAL EVE CHARGE: Total Charge to the Customers in Morning.
TOTAL NIGHT MINUTES: Total Number of Minutes Spent By Customers in the Night.
TOTAL NIGHT CALLS: Total Number of Calls made by Customer in Night.
TOTAL NIGHT CHARGE: Total Charge to the Customers in Night.
TOTAL INTL MINUTES: Total Number of Minutes Spent By Customers for international calls
TOTAL INTL CALLS: Total Number of International Calls made by customer
TOTAL INTL CHARGE: Total charge to customers for international calls

Steps involved in this EDA

Step 1

Exploring Dataset

Checking summary, checking columns and rows and shape of data-frame

Step 2

Data Cleaning

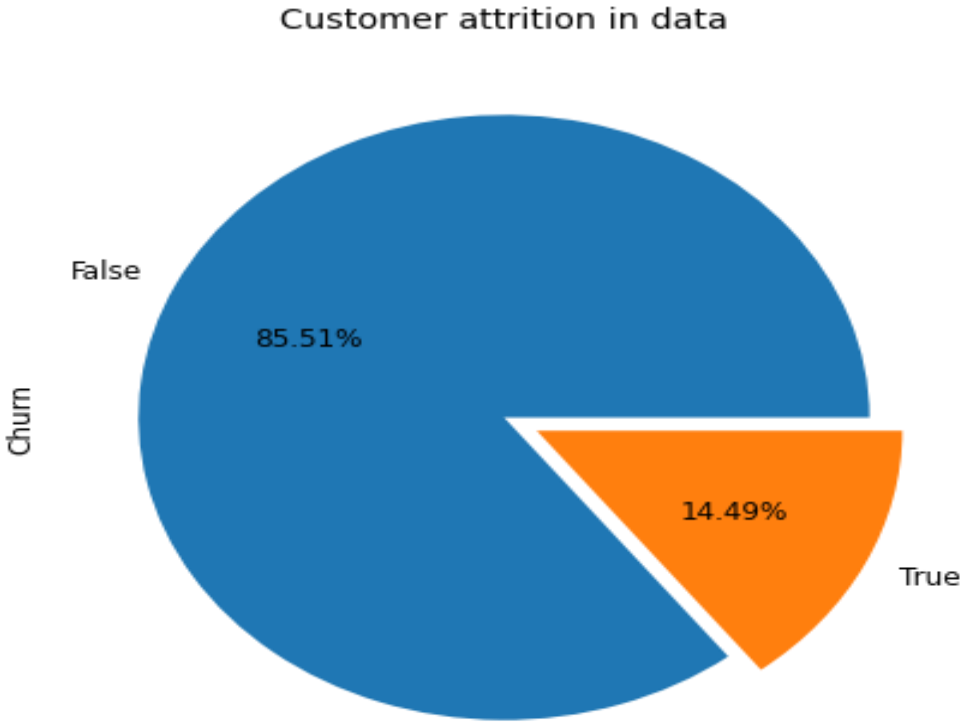
Removing null values if any, converting ‘Yes’ and ‘No’ values to ‘1’ and ‘0’ respectively

Step 3

Data visualization

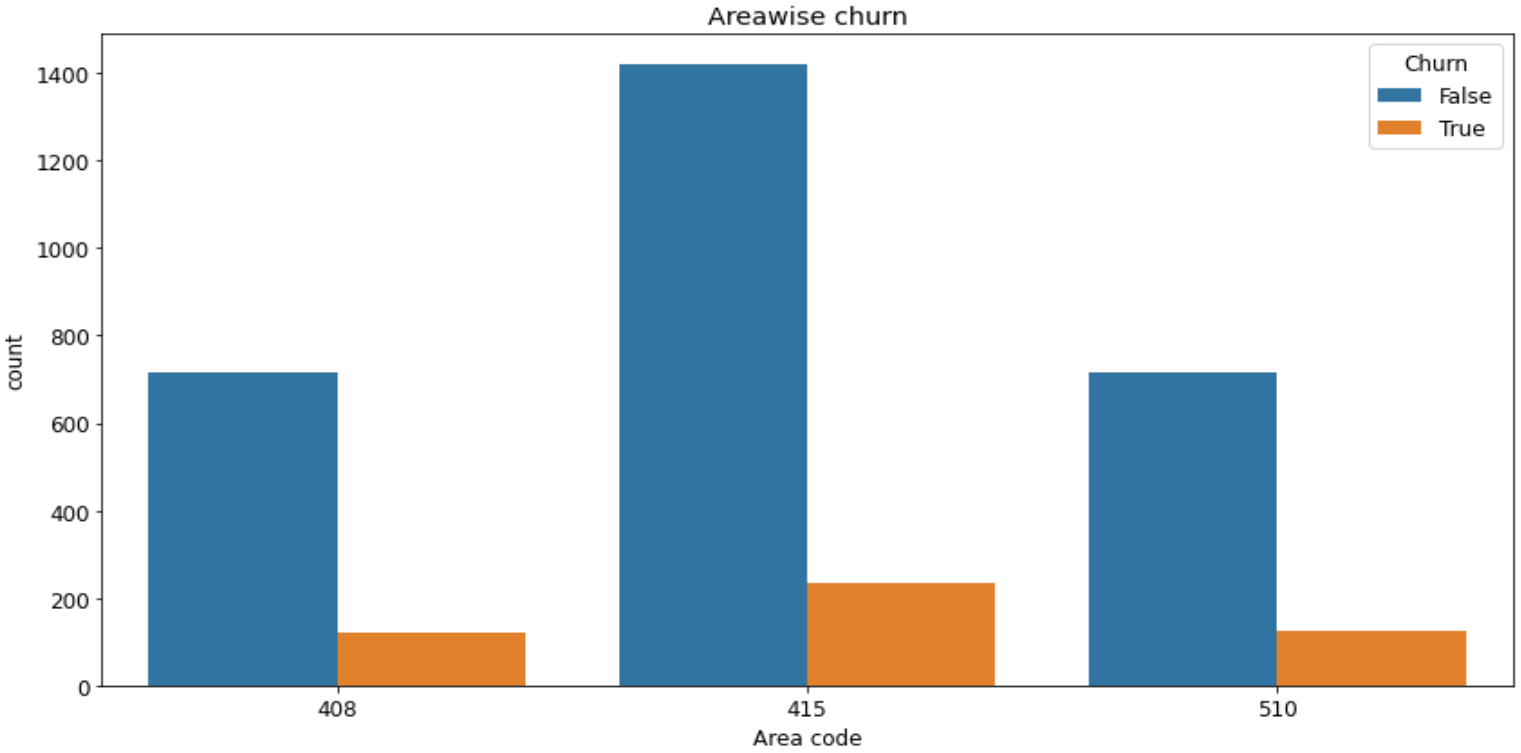
Plotting various graphs and extracting insights from dataset

Visualization



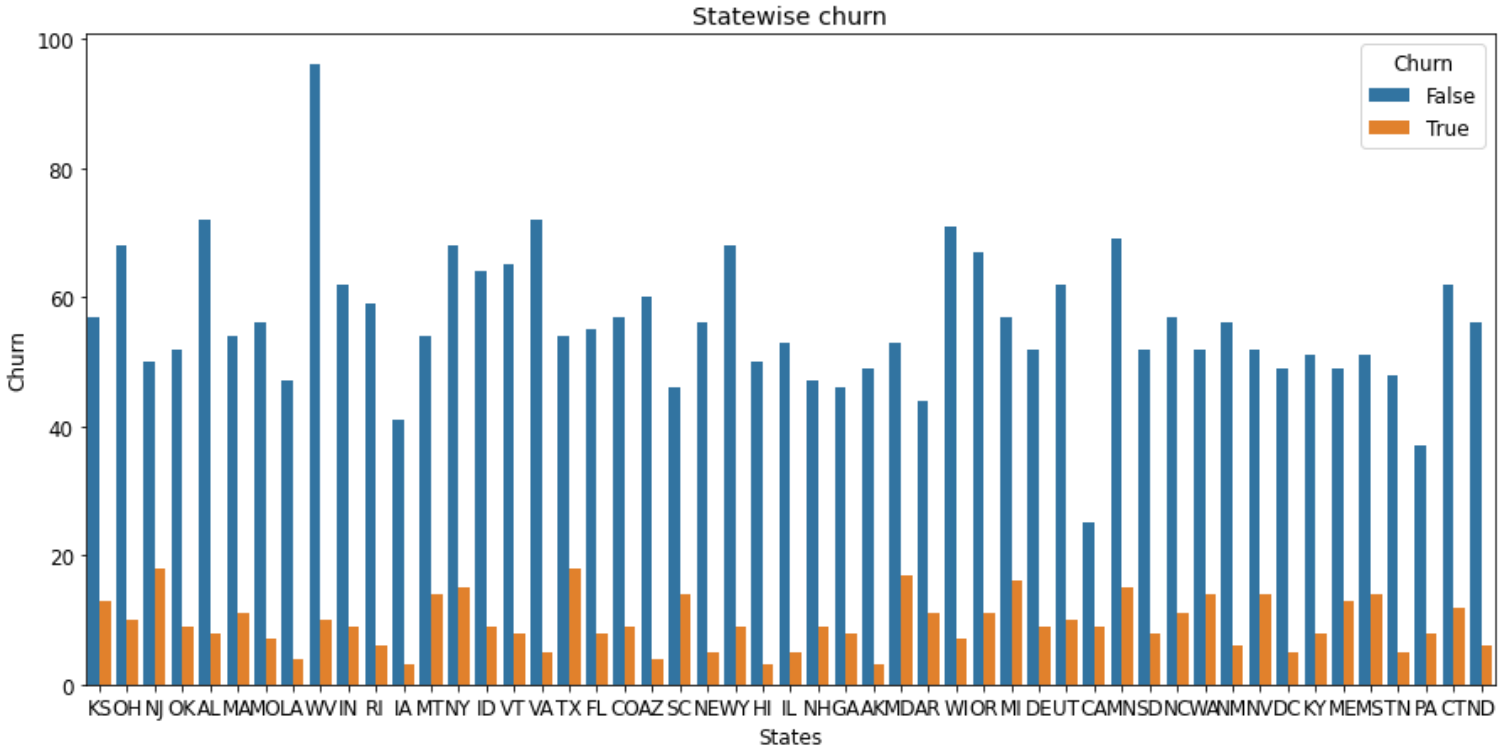
As we can see that Company’s churn rate 14.49%

Visualization



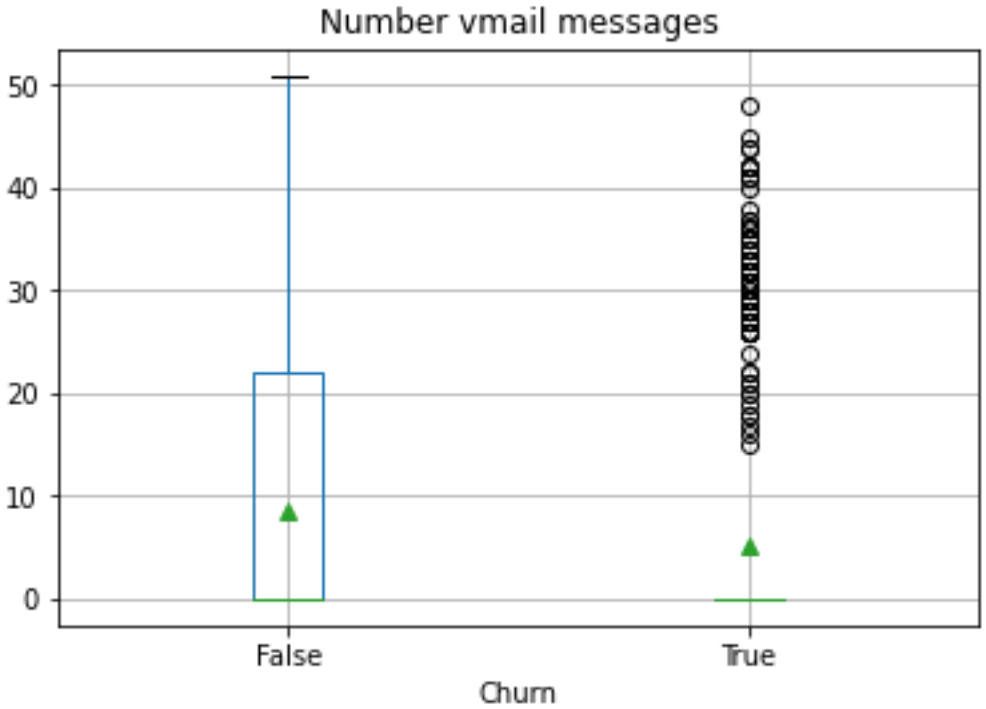
We can see area 415 has more churn in relation to other two areas

Visualization



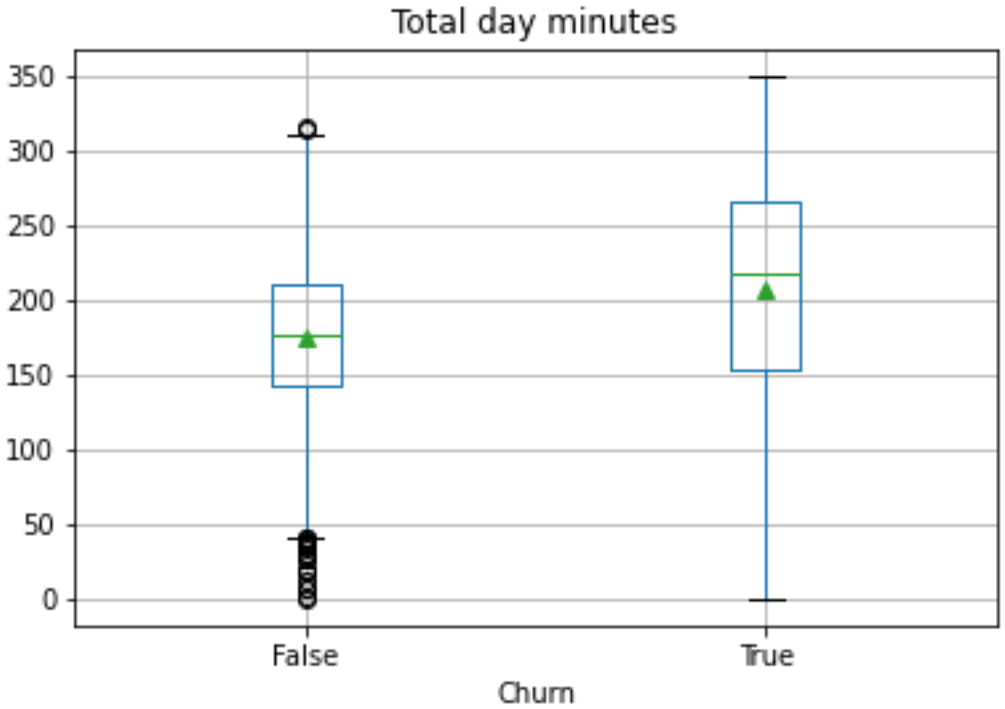
State wise churn:- As we see above States NJ, TX and MD has more churn rates and company should look into these states to decrease the churn rate

Visualization



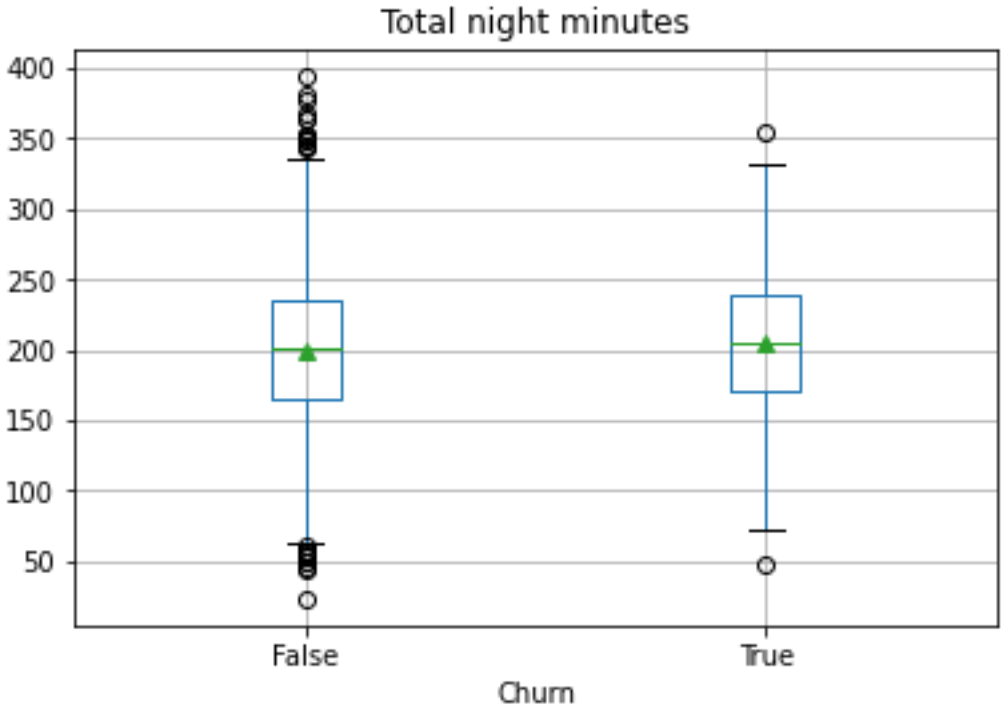
We can Notice for Voice-Mail Feature when there are more than 20 voice-mail messages then certainly there is a churn indicating improving the voice-mail feature

Visualization



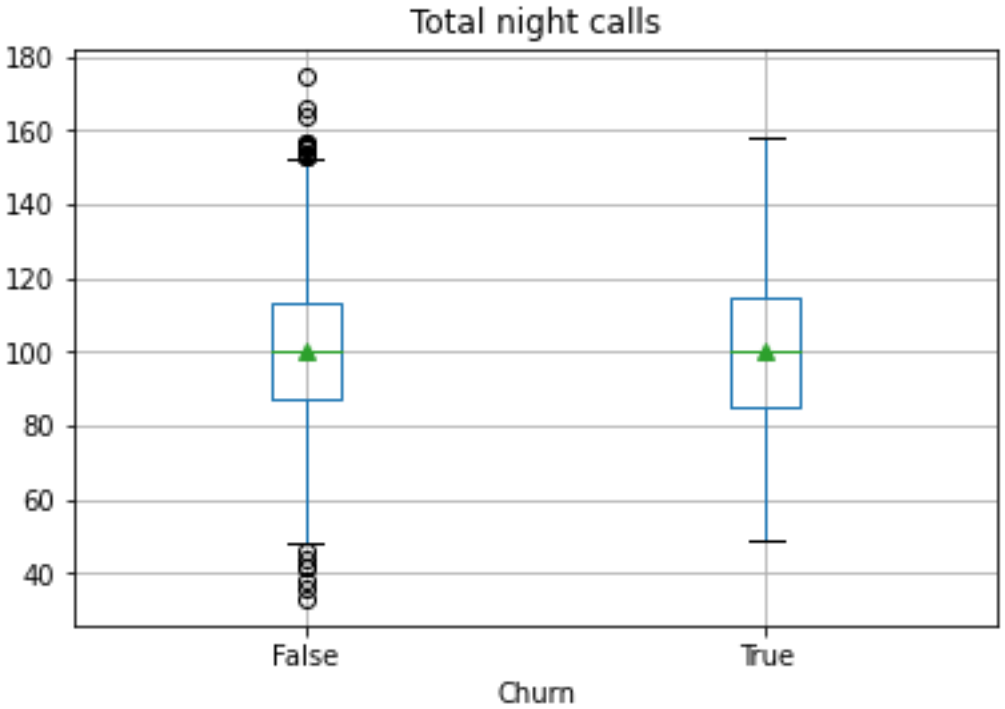
We can infer from above box-plot that with users spending more 225 minutes or more i.e. approx 4hrs tend to switch to other operator. So for long call time company should offer some discounts and upgrade the network

Visualization



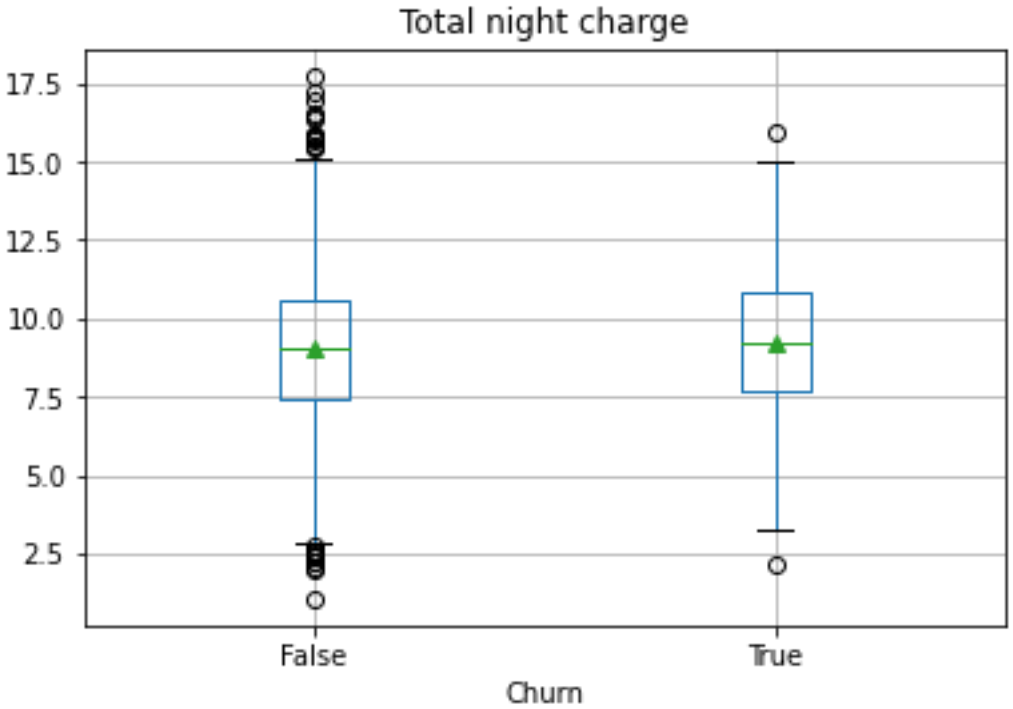
In night loyal customers spend more time calling

Visualization



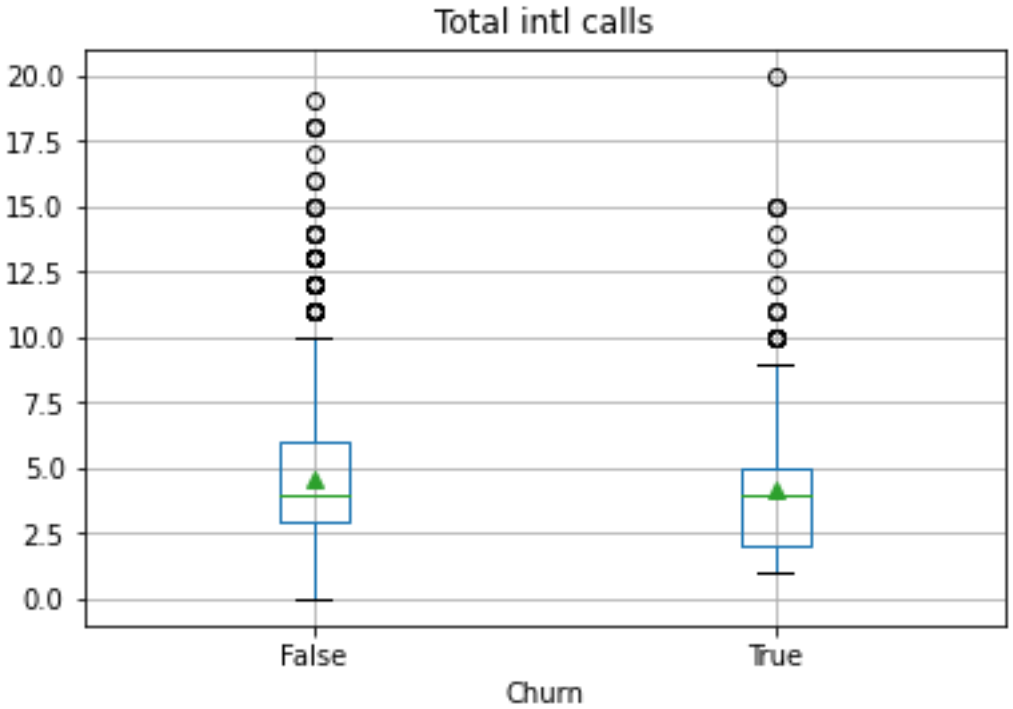
The number of calls at night made by churned customer is more

Visualization



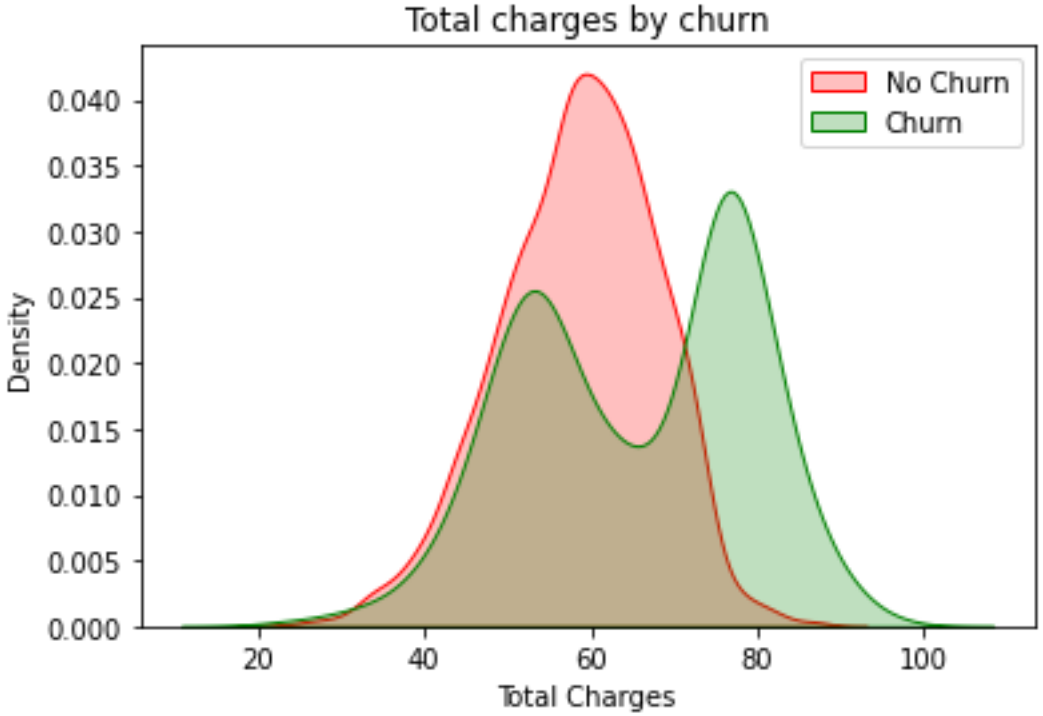
Call charge reduction required during night time

Visualization



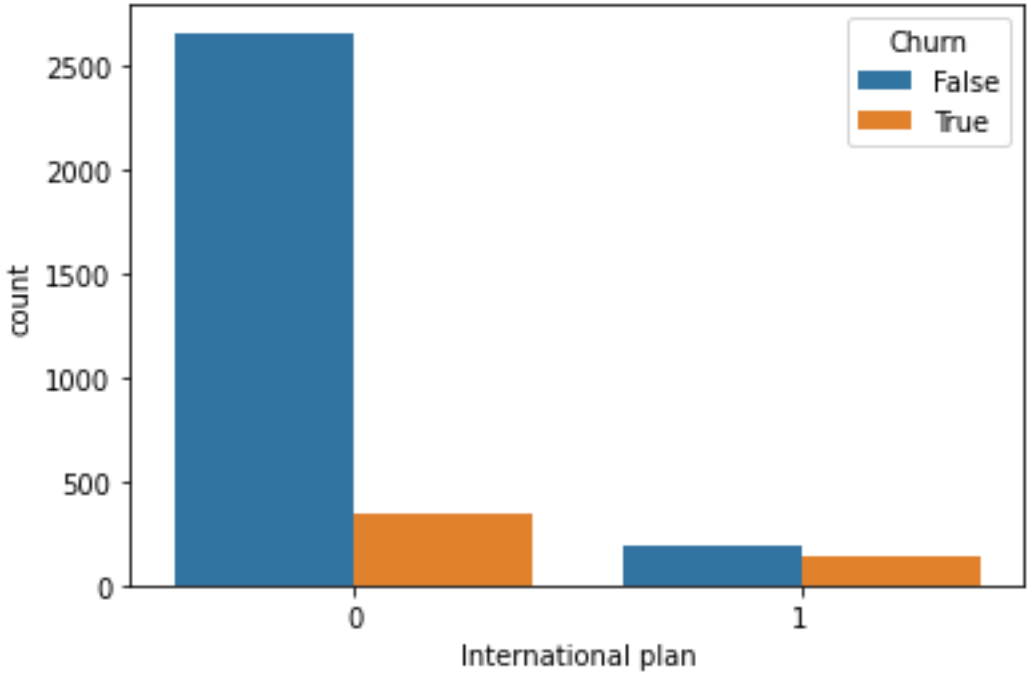
We can see higher attrition of people making more number international calls without international packs.

Visualization



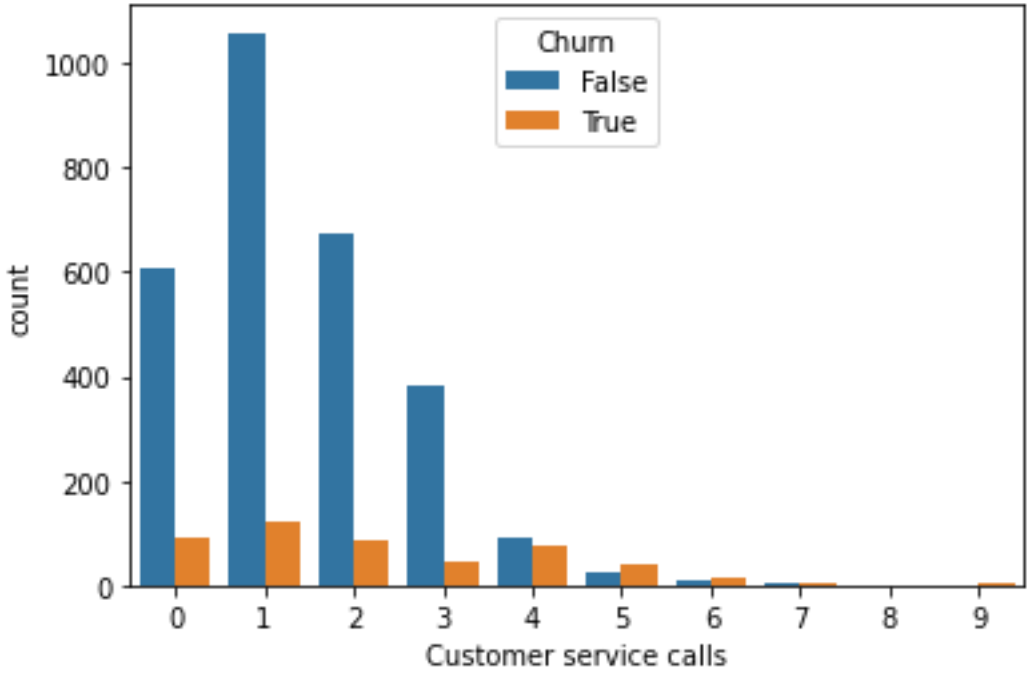
Here we see that churn is more when total call cost is more.

Visualization



From this plot we can see that customers having no international plan will like to churn more

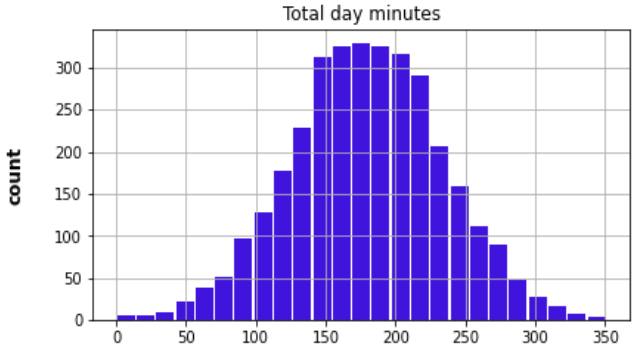
Visualization

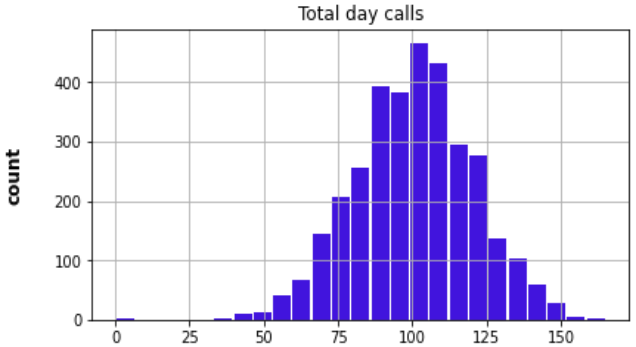


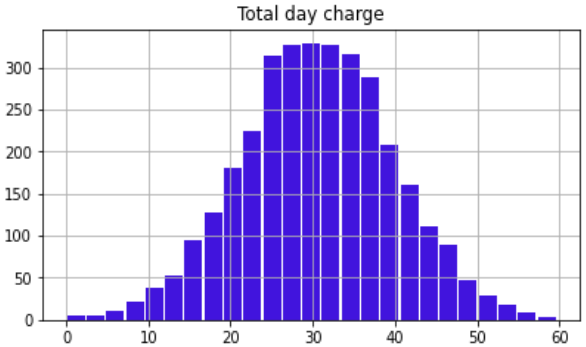
We can see from the plot that the churn rate increases sharply up to 4 customer service calls. So company should resolve their issues within that number of customer care calls

Visualization

Day call visualization



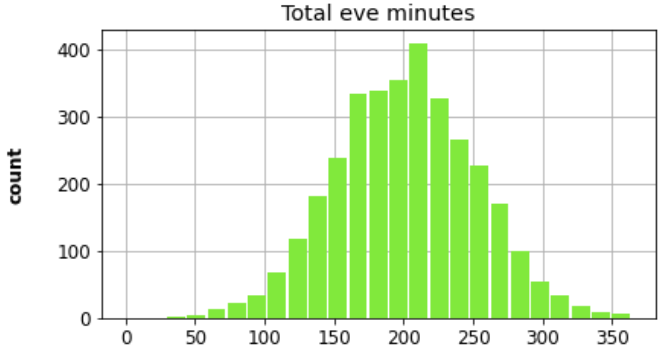


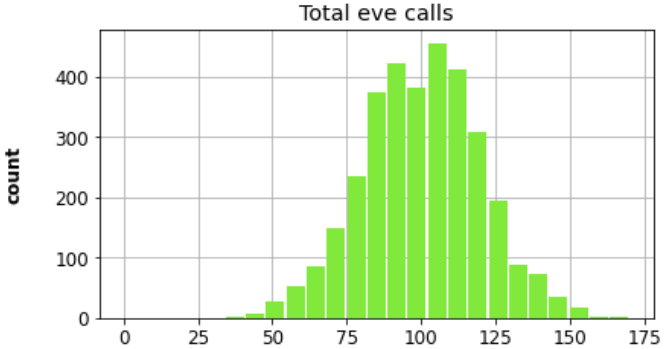


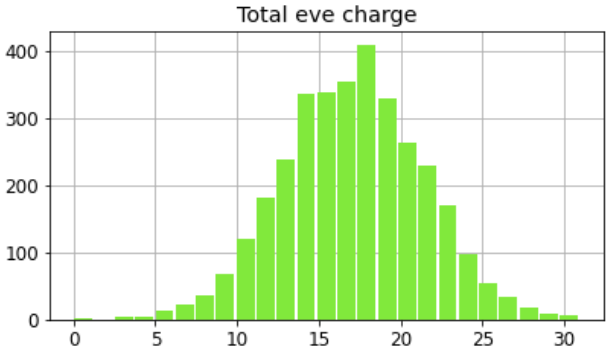
Day call distribution with respect to total day minutes, total number of day calls and its charge.

Visualization

Evening call visualization



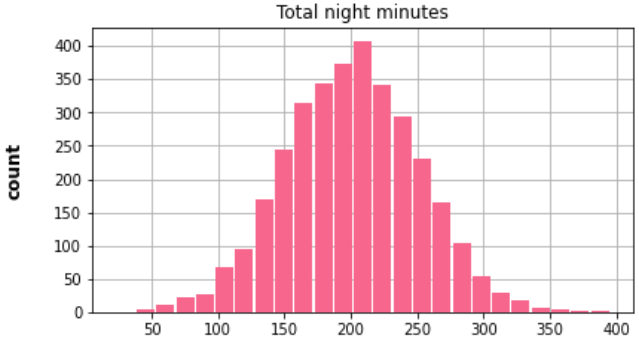


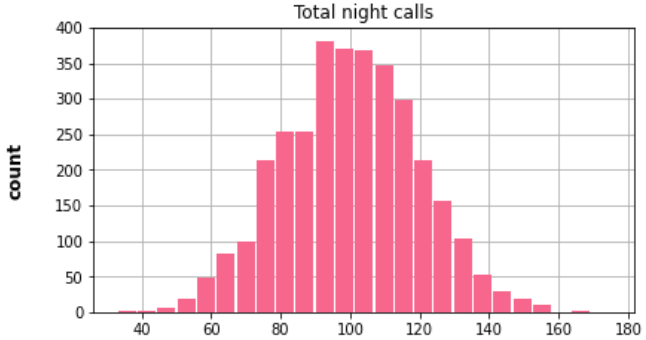


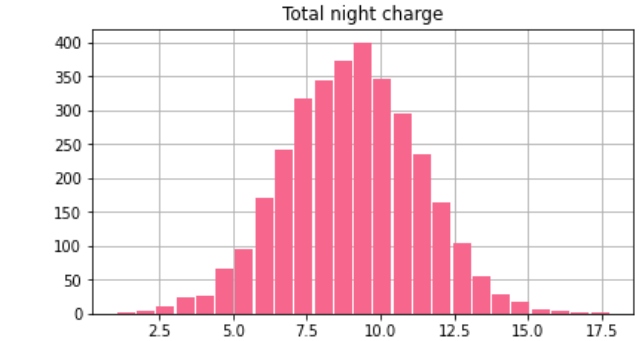
Evening call distribution with respect to total eve minutes, total number of eve calls and its charge.

Visualization

Night call visualization

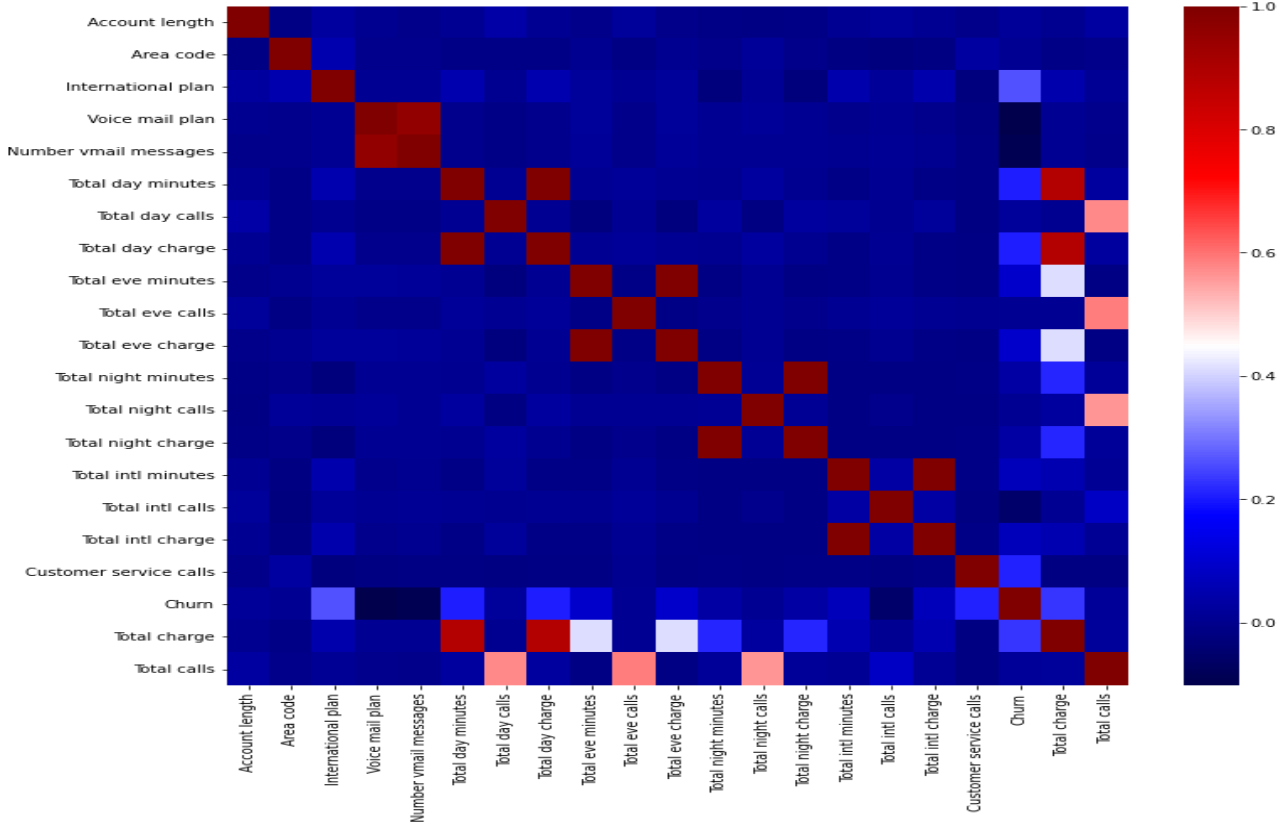






Night call distribution with respect to total night minutes, total number of night calls and its charge.

Visualization



Heat map of complete dataset

Conclusion

- Company’s churn rate is 14.49%.
- More churn is in area 415, so company needs to focus on that area.
- States NJ(New Jersey), TX(Texas) and MD(Maryland) has more churn rates and company should look into these states to decrease the churn rate.
- We can Notice 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.
- For long call time company should offer some discounts and upgrade the network.

Conclusion Cont’d

- Customers having more minutes spent on the network tend to leave the it's subscription and it clearly indicates that there is defect in the pricing strategy of the company.
- We can see customers without international plan are likely to churn more.
- We can see from that the churn rate is more up to 4 customer service calls. So company should resolve their issues within that number of customer care calls.