Capstone Project

Telecom Churn Analysis

By

Baishanvee Mahato

Data Science Trainee- Almabetter, Bengaluru

Contents

- Defining problem statement
- Project Scope
- Data Summary
- Dataset Columns
- Exploratory Data Analysis
- Conclusion
- Recommendation

Problem Statement

Orange S.A., formerly France Telecom S.A., is a French multinational telecommunication corporation. The Orange Telecom's Churn Dataset, consists of cleaned customer activity data (features), along with a churn label specifying whether a customer cancelled the subscription. The idea of the project is to identify the factors responsible for the churn .

PROJECT SCOPE

1. The project is an **Exploratory Data Analysis** of churn dataset of Orange S.A., which is a French Multinational Telecommunication Corporation.

2. Libraries such as Numpy, Pandas, Matplotlib, Seaborn are used to **Analyze** and **Visualise** the data.

Data Summary

churn_data : It contains information regarding the 3333 customer of **The Orange Telecom** in which Account length, Area code, whether they have active international plan or not , minutes spent in call and many such information are present.

churnd_cust_data : It contains information as of churn_data only of those Customer who are churned.

not_churnd_cust_data : It contains same information as of churn_data only of those customer who are not churned.

Dataset Column

- State: States name(in code).
- Account Length: period for which the Account is active.
- Area Code: Area code having States
- International Plan: Yes: Means International Plan is subscribed
  No: Means the customer has not subscribed for any international plan.
- Voicemail Plan: Yes: Voicemail Plan is subscribed,
  No: voicemail Plan is not subscribed by the Customer.
- Number vmail messages: Number of Voicemail Messages
- Total day minutes: Total Number of Call Minutes Spent by customer during Morning hours.
- Total day calls: Total Number of Calls made by customer during the Morning hours.
- Total day charge: Total Charge for all Calls made by customer during the Morning hours.
- Total eve minutes: Total Number of Call Minutes Spent by customer during Evening hours.

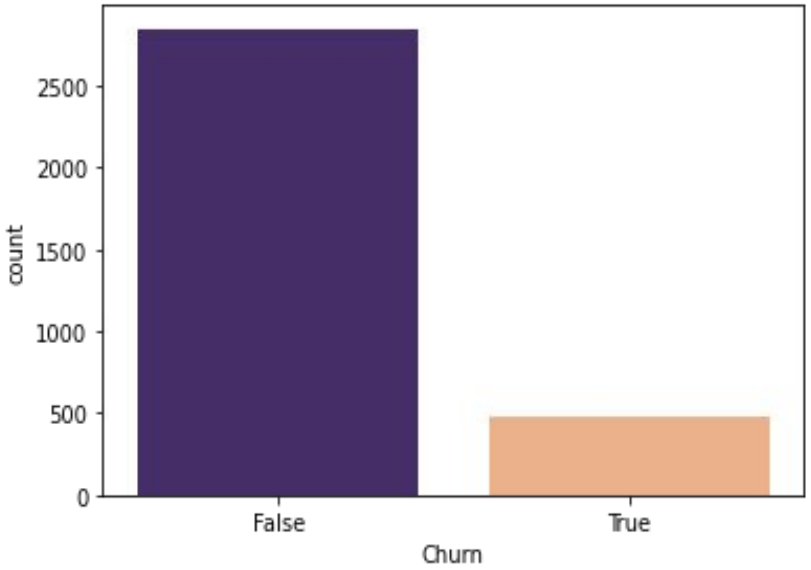
- Total eve minutes: Total Number of Call Minutes Spent by customer during the Evening hours.
- Total eve calls : Total Number of Calls made by customer during the Evening hours.
- Total eve charge: Total Charge for all the Calls made by customer during the Evening hours.
- Total night minutes: Total Number of Call Minutes Spent by customer during the Night hours
- Total night calls: Total Number of Calls made by during the night hours.
- Total night charge: Total Charge for Calls made by customer during the Night hours.
- Total intl minutes: Total Number of Call Minutes spent by customer on international calls.
- Total intl calls: Total Number of International Calls made by customer.
- Total intl charge: Total charge for all the international calls made by customer.
- Customer service calls : Total Number of customer service calls made by customer.
- Churn : True : Churned customer

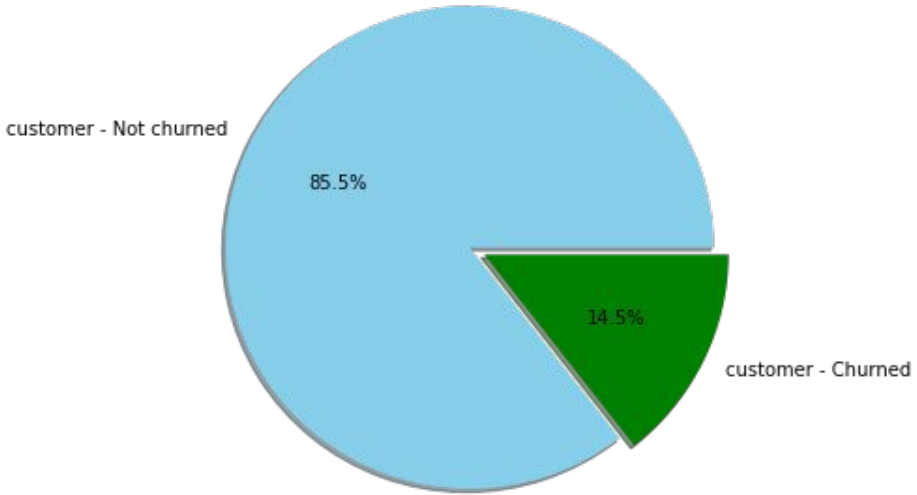
  False: means retained customer.

EXPLORATORY DATA ANALYSIS

Churn

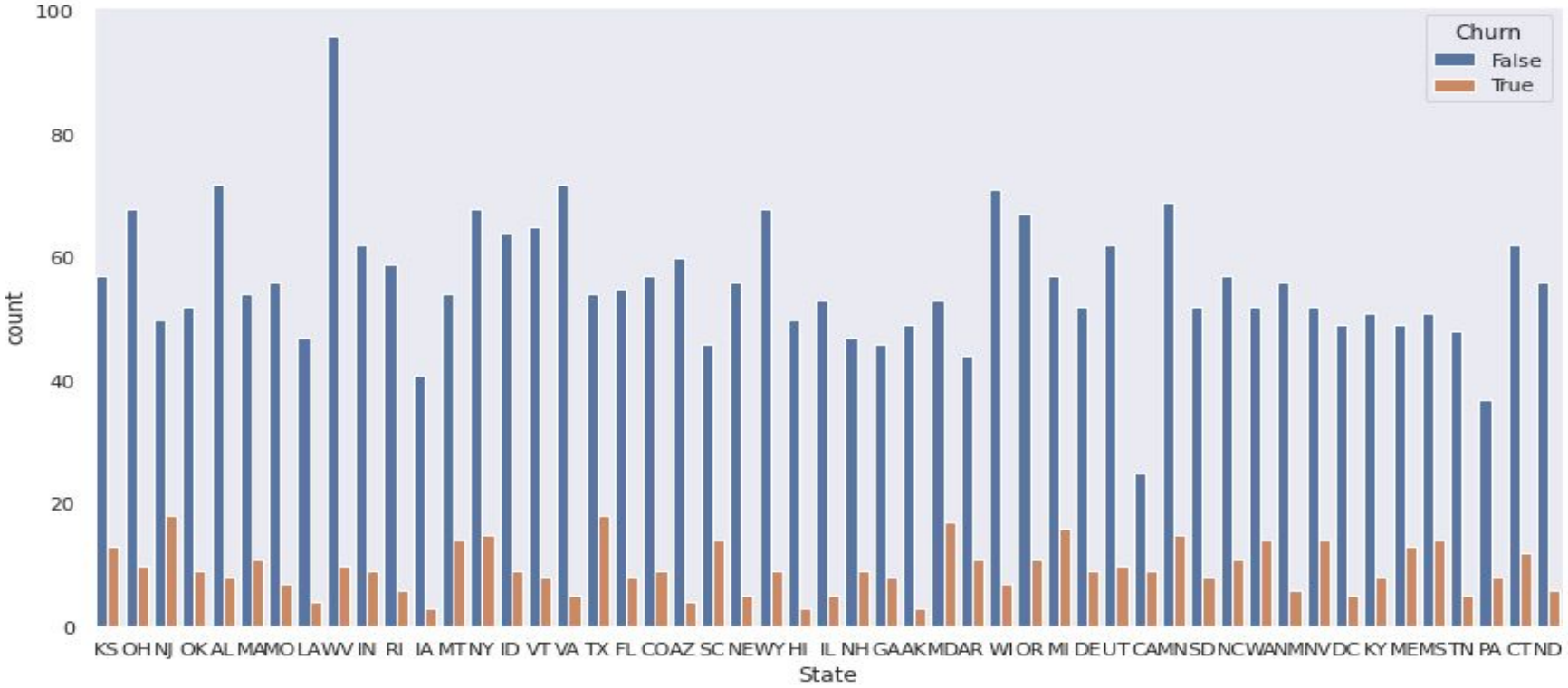
The data consists of record of 3333 customers of the orange telecom customers. Out of which 483 customer churned and the remaining 2850 did not.





Statewise analysis of churned customer data:

The data consist of customers from 51 states.



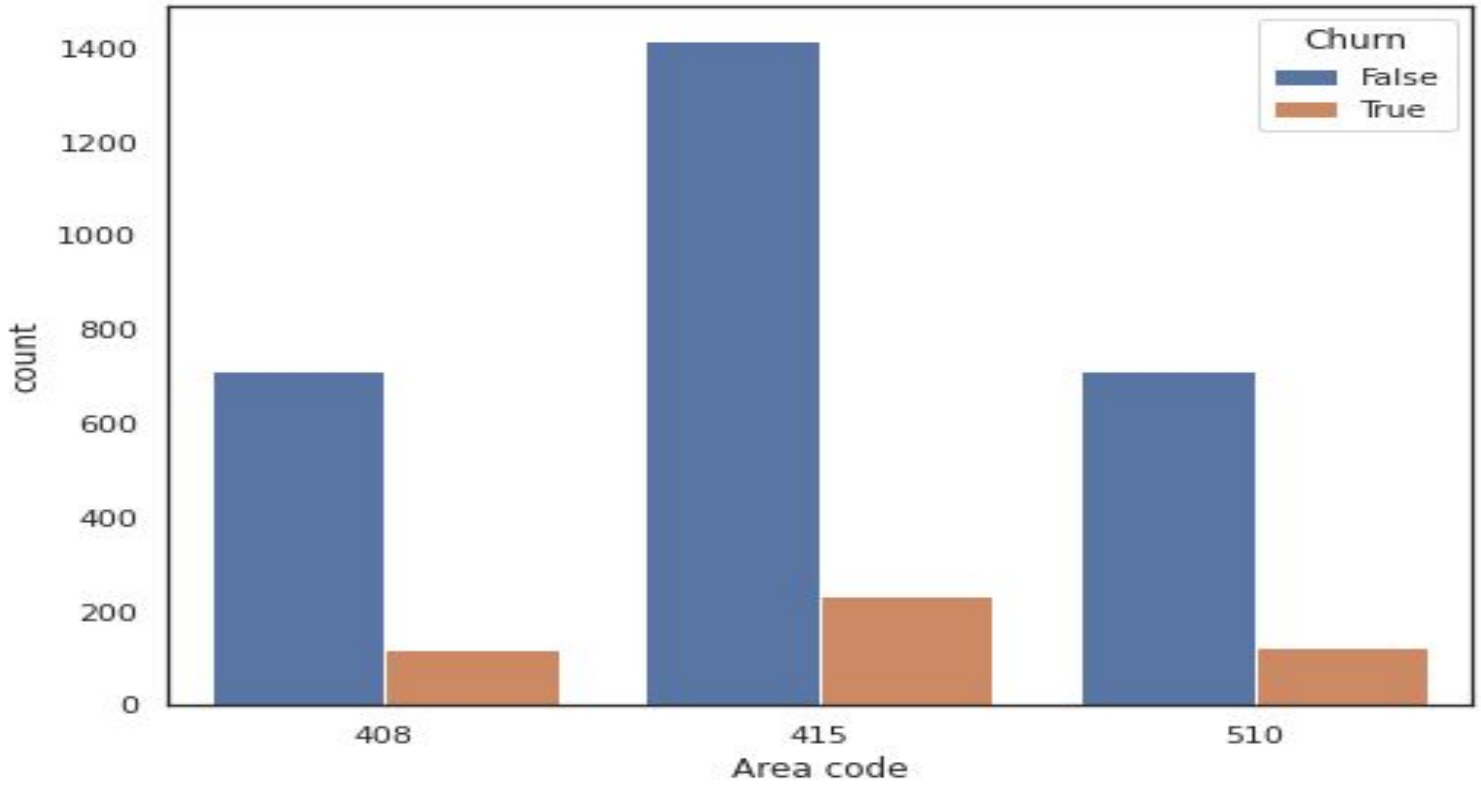
TOP 5 states with highest churn rate.

1. CA
2. NJ
3. TX
4. MD
5. SC

Churn State	False	True	%_Churn
AK	49	3	5.769231
AL	72	8	10.000000
AR	44	11	20.000000
AZ	60	4	6.250000
CA	25	9	26.470588
CO	57	9	13.636364
CT	62	12	16.216216
DC	49	5	9.259259
DE	52	9	14.754098
FL	55	8	12.698413
GA	46	8	14.814815
HI	50	3	5.660377
IA	41	3	6.818182
ID	64	9	12.328767
IL	53	5	8.620690
IN	62	9	12.676056
KS	57	13	18.571429
KY	51	8	13.559322
LA	47	4	7.843137
MA	54	11	16.923077
MD	53	17	24.285714
ME	49	13	20.967742
MI	57	16	21.917808
MN	69	15	17.857143
MO	56	7	11.111111
MS	51	14	21.538462
MT	54	14	20.588235
NC	57	11	16.176471
ND	56	6	9.677419
NE	56	5	8.196721
NH	47	9	16.071429
NJ	50	18	26.470588
NM	56	6	9.677419
NV	52	14	21.212121
NY	68	15	18.072289
OH	68	10	12.820513
OK	52	9	14.754098
OR	67	11	14.102564
PA	37	8	17.777778
RI	59	6	9.230769
SC	46	14	23.333333
SD	52	8	13.333333
TN	48	5	9.433962
TX	54	18	25.000000
UT	62	10	13.888889
VA	72	5	6.493506
VT	65	8	10.958904
WA	52	14	21.212121
WI	71	7	8.974359
WV	96	10	9.433962
WY	68	9	11.688312

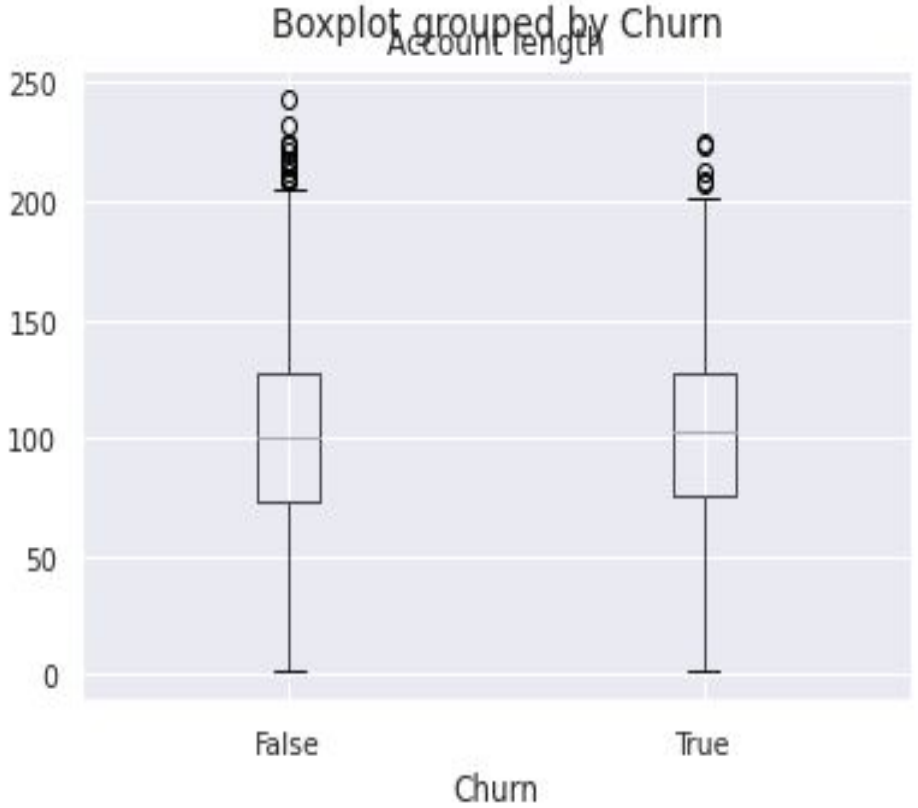
Area wise churn rate :

Given dataset consists of 3 unique Area code, under which different state fall.

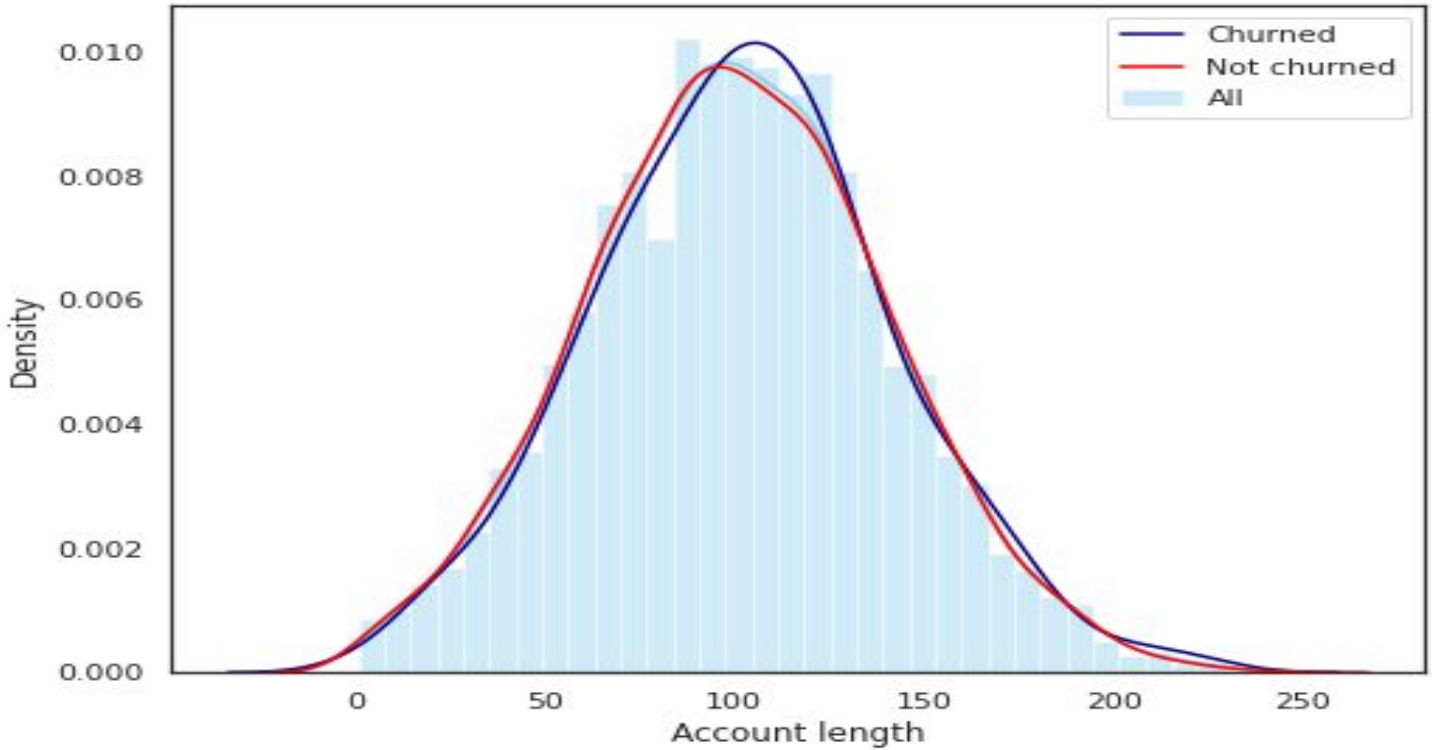


Account length:

It is the time period between the day on the sim was activated and the day on which the data was recorded. From the boxplot plotted, we can observe that the median of account length for both churned and those who did not churn is approximately same.

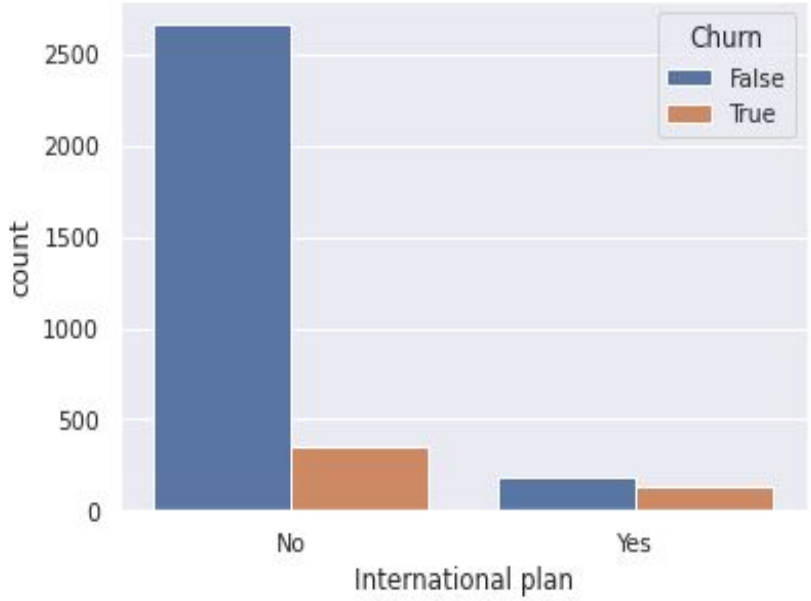


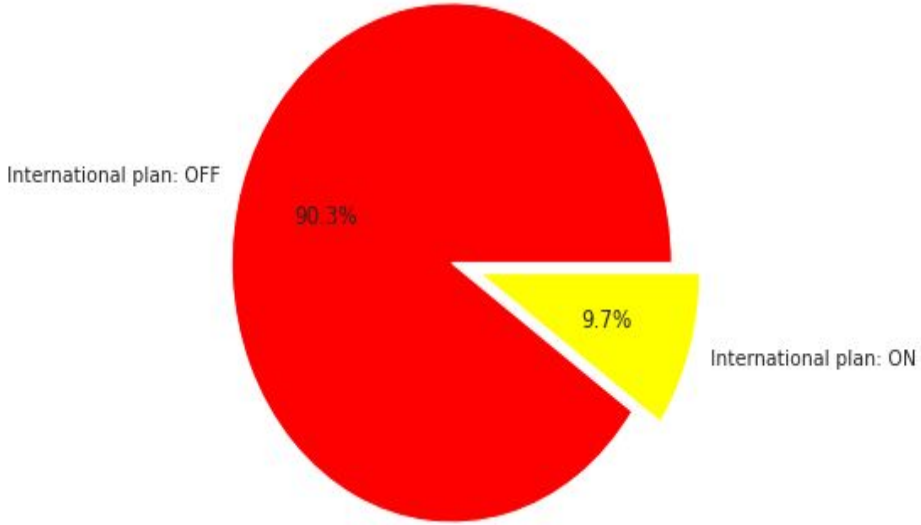
Comparing of not churned account length and churned account length



International plan Analysis:

International plan allows customers to make call to the person from other country. It should be noted that the tariff of international are always high as compared to domestic, not many people opt for plan . Out of 3333 customers only 323 had active international plan.

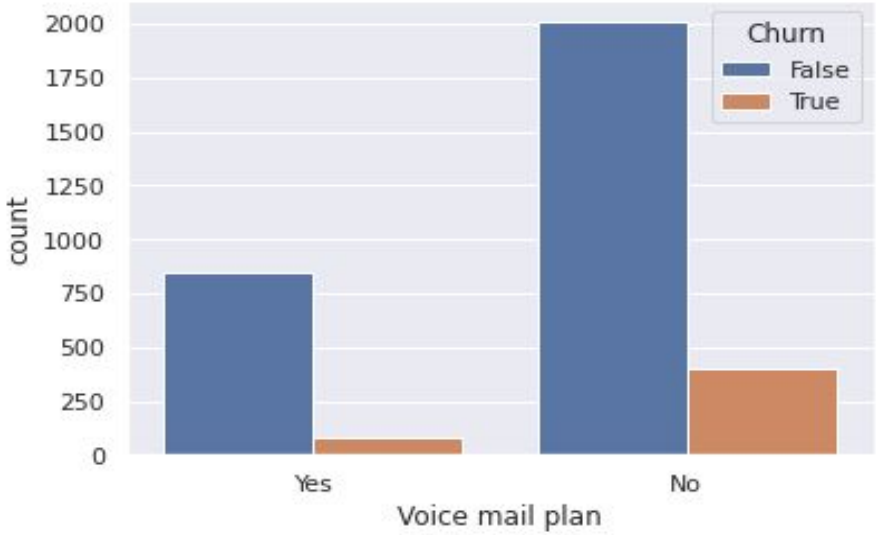




Voicemail plan

A **voicemail** is a computer-based system that allows users and subscribers to Exchange personal voicemail messages.

Churn	False	True	Churn_%
Voice mail plan			
No	2008	403	16.715056
Yes	842	80	8.676790



Analysis - Number of voicemail messages:

The number of voicemail exchanged by the customer.

Churned Customer

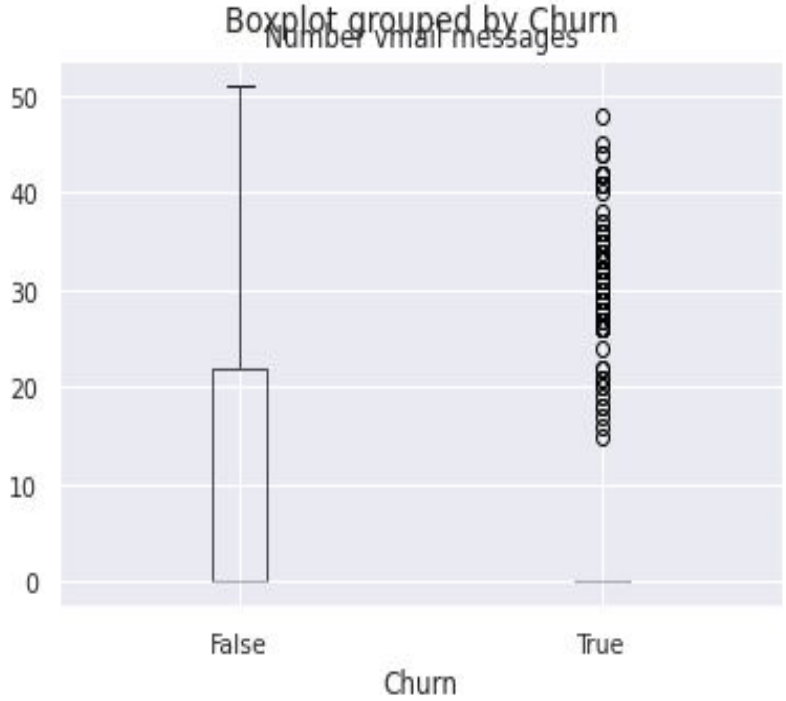
Count : 483.000000

mean :5.115942
std :11.860138
min :0.000000
25% :0.000000
50% :0.000000
75% :0.000000
max :48.000000

Customers who did not churn

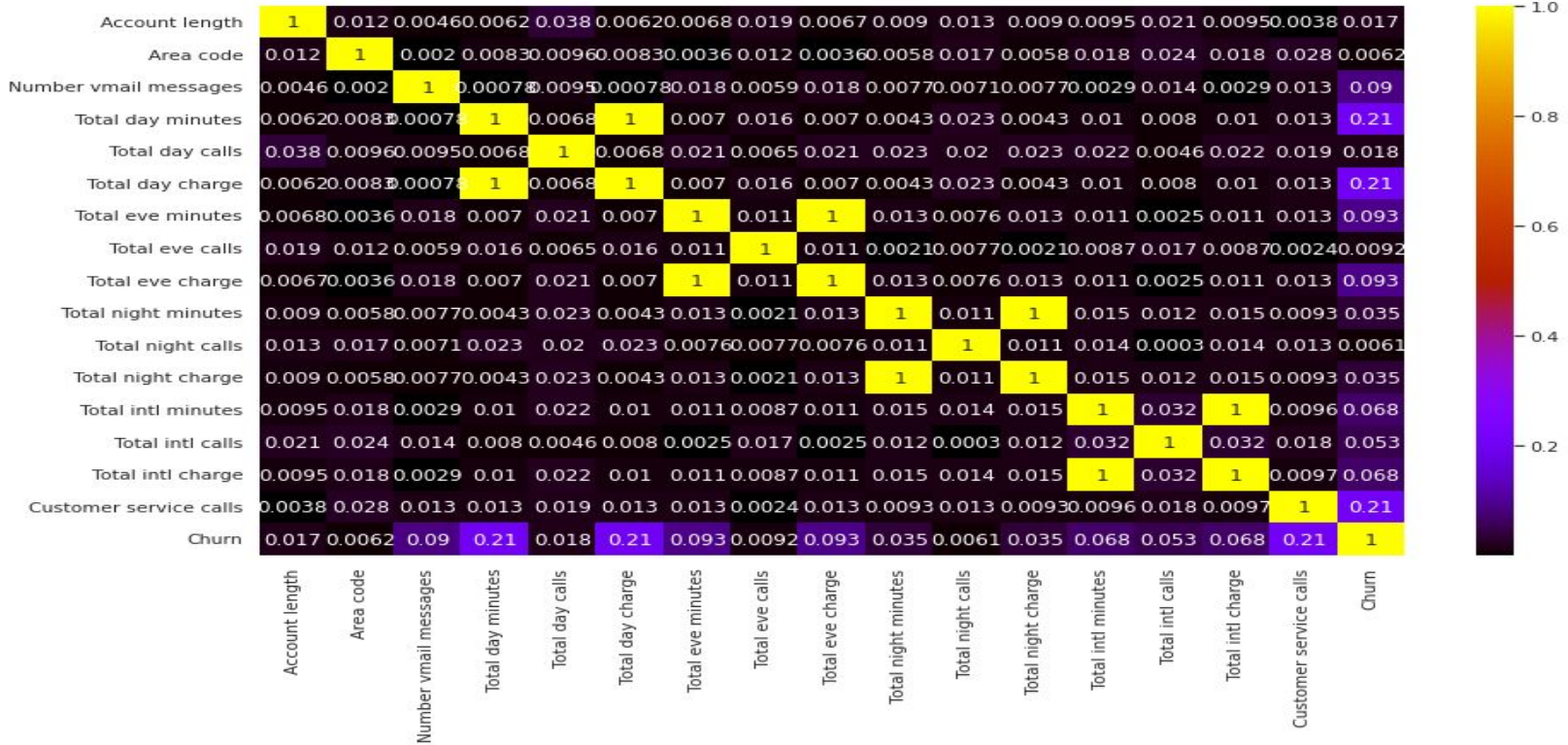
Count : 2850.000000

mean :8.604561
std :13.913125
min :0.000000
25% :0.000000
50% :0.000000
75% :22.000000
max :51.000000



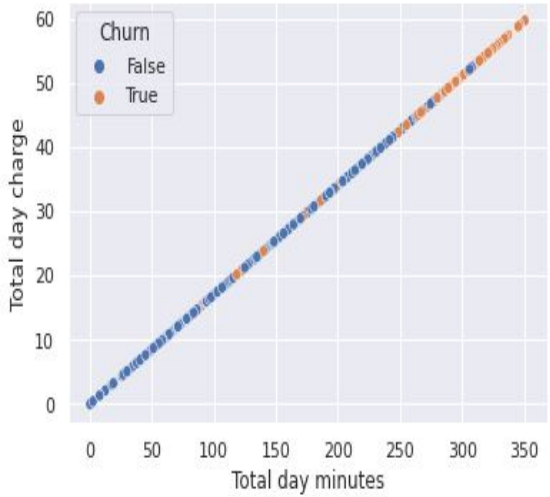
Analysis on Call Minutes and Call Charges:

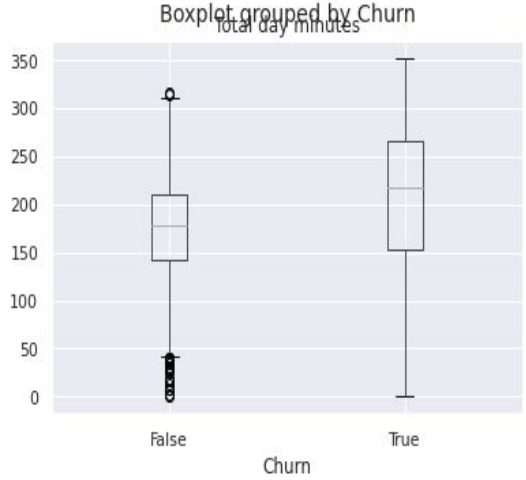
Below table shows the correlation between the columns of the dataframe.

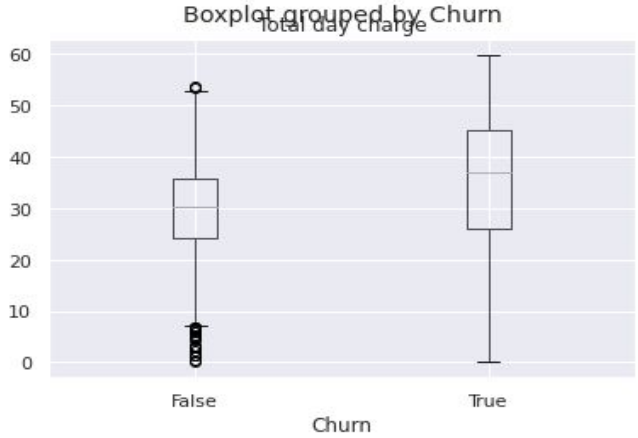


The given dataset consists of (number of calls made, duration of those calls In minutes, total charge charged for those calls) divided in different phase such as Daytime, Evening and night time.

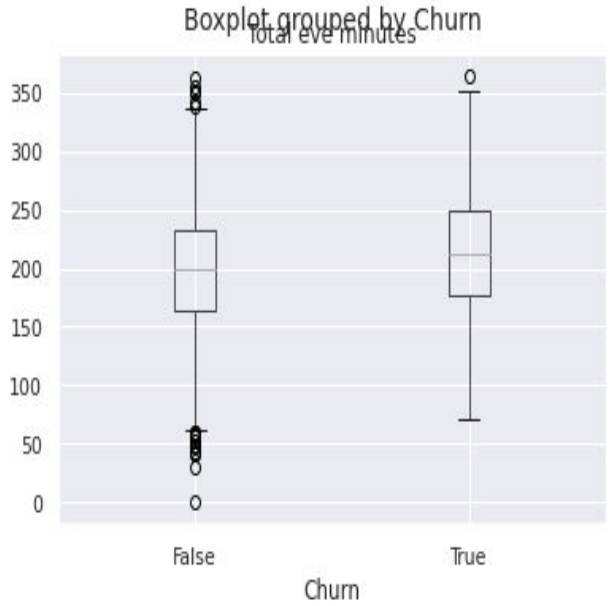
For call made during day time:

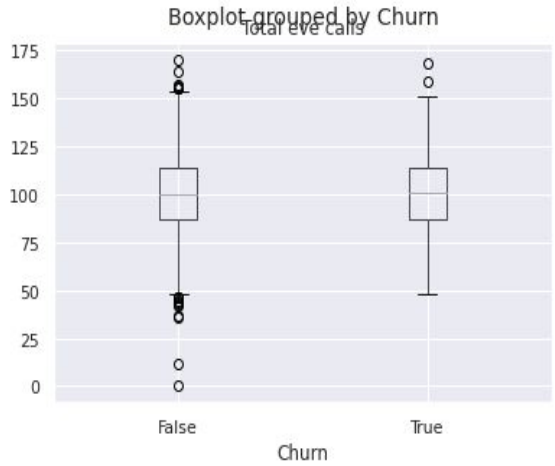


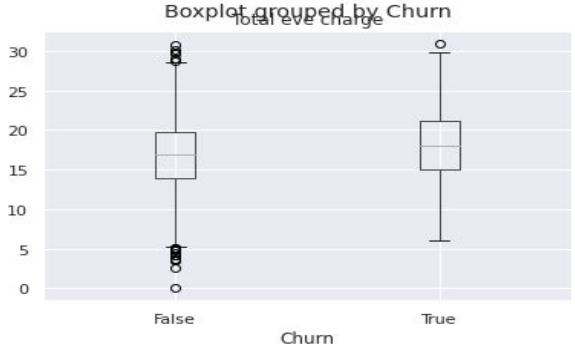




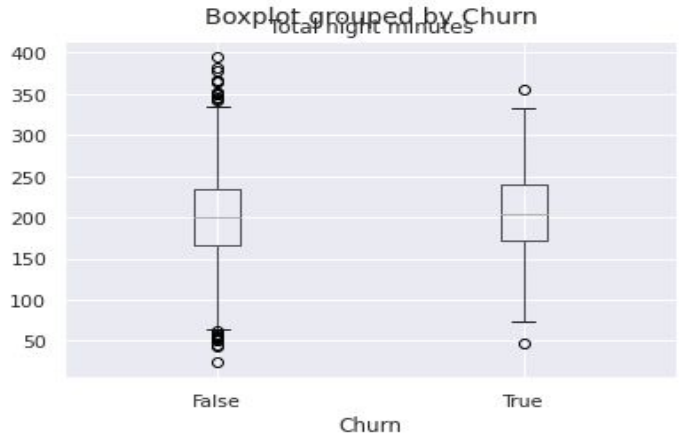
For evening calls:

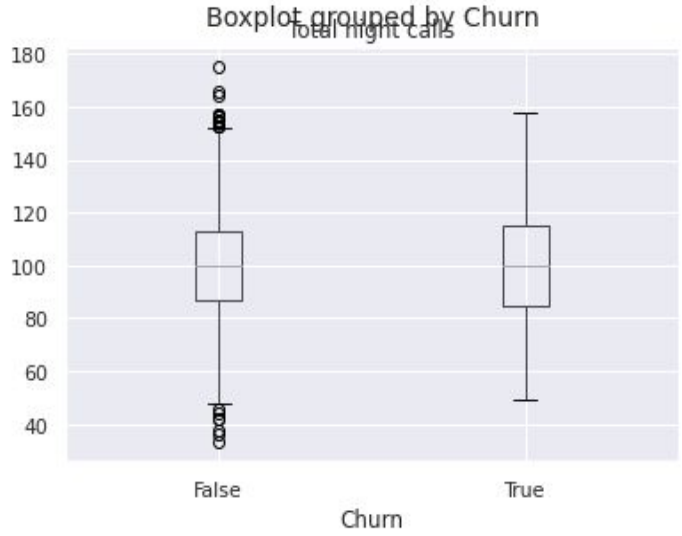


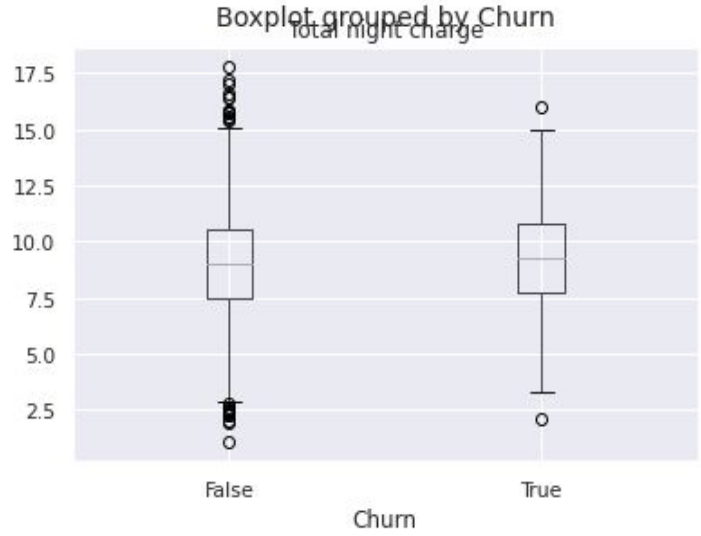




For Night calls:







Customer service calls:

This column present us with the data about the number of times a customer called customer service, mostly calls made to customer service are primarily for the reasons such as network issue, info regarding charged deducted etc.

CONCLUSION

- Customers with international plan active has higher churn rate compared to customers without international plan , reasons may be because the customers are unhappy with the higher tariff charges or network problems.
- Area and Account length have no relation with churn, hence these columns can be deleted or it can be said that the data is redundant.
- Few states have higher churn compared to others, for which network problems may be reason or competitors offering low tariff.
- Call to the service centre increases proportionally churn rate increases. Customers who called customer service less than 3 times have a lower churn rate than that of customers who have called customer service 4 or more times.
- Customers with higher day call mins has higher churn rate compared to other , could be because of the higher charges which is quite obvious.

Recommendation

- Improvement in coverage area and solving network issues.
- Give discount or create a plan or allow customised plans.
- Lower the international plan tariff or some discounts.
- Provide better customer service and problem solution, also take their feedback and work on the feedback suggested by the customers.