



CAPSTONE PROJECT-01 EDA ON

Telecom Churn Analysis



AI

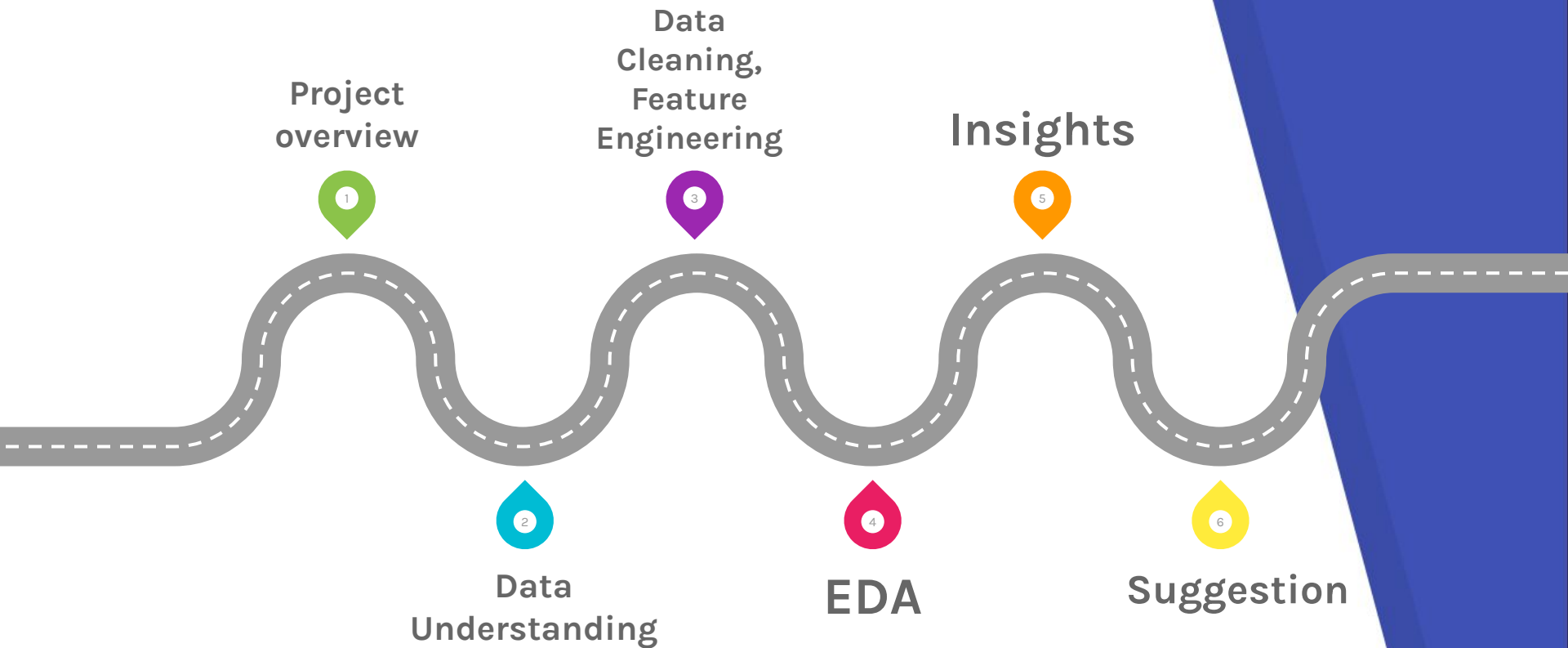


HELLO!

I am Gaurav Yogeshwar

You can find this project at -
[Github link for this project](#)

Roadmap



Project Overview



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 , along with a churn label specifying whether a customer canceled the subscription.

Goal:



Explore and analyze the data to discover key factors responsible for customer churn and come up with ways/recommendations to ensure customer retention.

Customer churn?

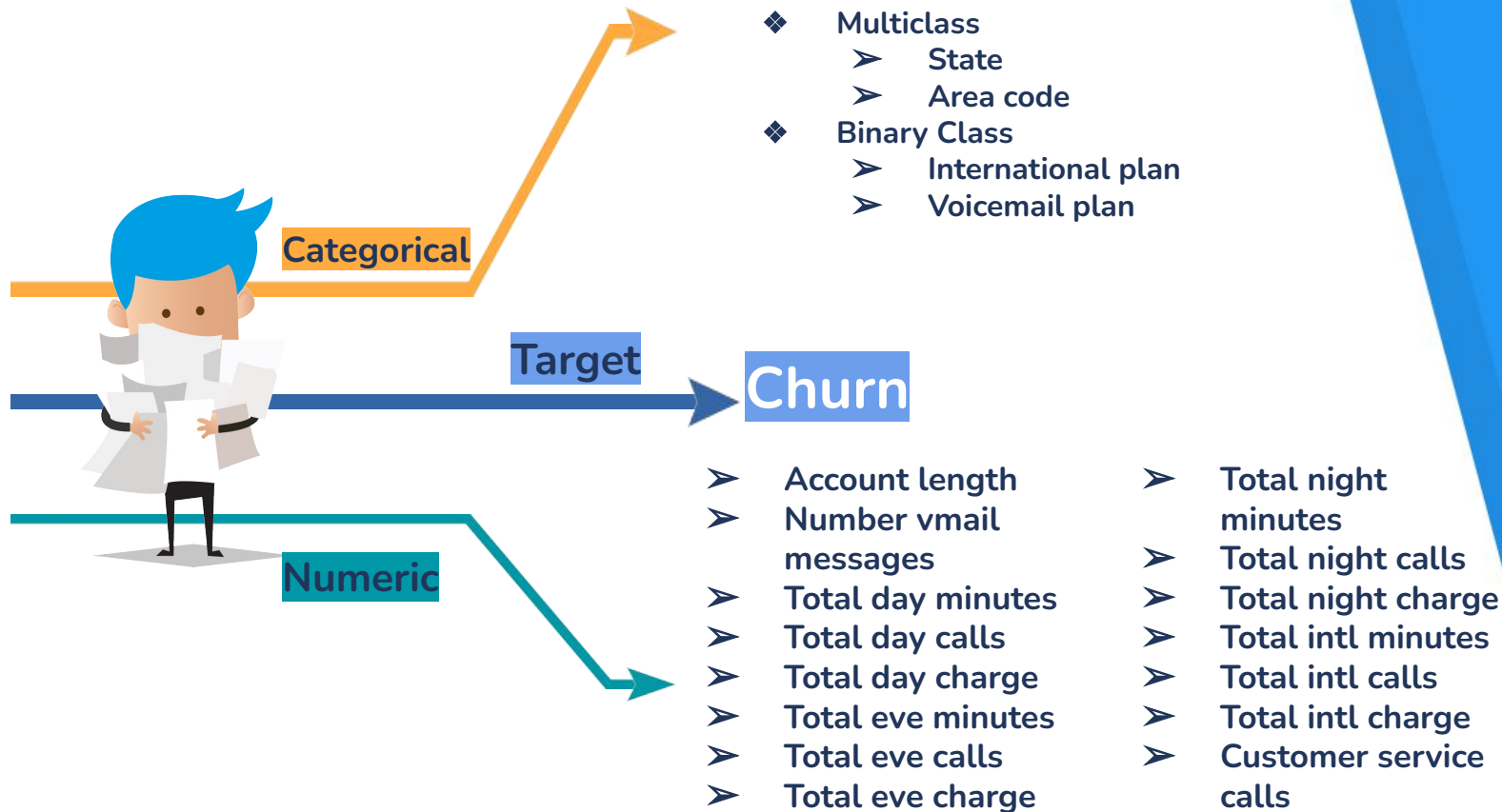
Customer churn is the term used when an existing customer stops using a company's services and/or stops buying their products. In other words, the customer chooses to cut his ties with the company.

Churn rate is defined as the proportion of customers who stopped using a particular company's products or services during a definite time frame

$$\text{Churn Rate \%} = \frac{\text{Number of Customers lost}}{\text{Total no. of Customers}} \times 100$$



Data Summary





Data Summary

Numeric features:

1. **Account Length**: Length of The Account
2. **Number vmail messages**: Number of Voicemail Messages.
3. **Total day minutes**: Total Number of Minutes Spent By Customers in Day (before evening).
4. **Total day calls**: Total Number of Calls made by Customer in Day (before evening).
5. **Total day charge**: Total Charge to the Customers in Morning.
6. **Total eve minutes**: Number of Minutes Spent By Customers in Evening.
7. **Total eve calls**: Total Number of Calls made by Customer in Evening.

Data Summary

Numeric features:

8. **Total eve charge:** Total Charge to the Customers in Evening.
9. **Total night minutes:** Total Number of Minutes Spent By Customers in the Night.
10. **Total night calls:** Total Number of Calls made by Customer in Night.
11. **Total night charge:** Total Charge to the Customers in Night.
12. **Total intl minutes:** Total Number of Minutes Spent By Customers in international calls.
13. **Total intl calls:** Total Number of International calls made by Customer.
14. **Total intl charge:** Total charge to Customers in international calls.
15. **Customer service calls:** Total number of Calls by Customer to service Center.

Data Summary

Categorical Features

Multiclass

- 16. **State**: 51 Unique States in United States of America
- 17. **Area Code** : 3 unique codes, 408 is of San Jose, 415 is of San Francisco and 510 is of City of Oakland.

Binary class

- 18. **International Plan**: Yes Indicate International Plan is Present and No Indicates no subscription for International Plan
- 19. **Voicemail Plan**: Yes Indicates Voicemail Plan is Present and No Indicates no subscription for Voicemail Plan
- 20. **Churn**: Whether he customer churned or not(True or False)

“

Data Cleaning, Preprocessing and Feature engineering

It is process of using domain knowledge to extract features from raw data via data mining technique.

There are Three general approaches:

- ▶ Extracting Information
- ▶ Combining Information
- ▶ Transforming Information



”

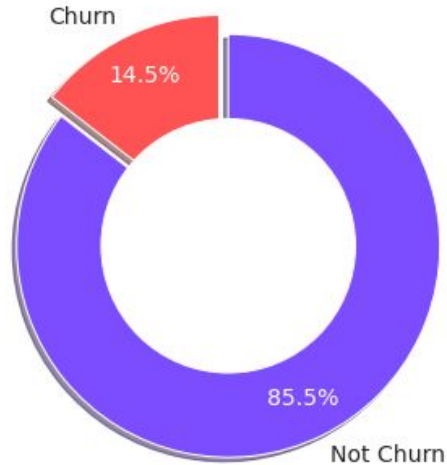


Exploratory Data Analysis

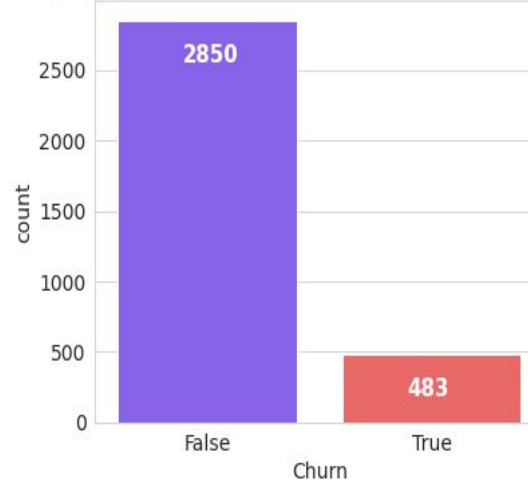


Observation 1

Display the percentage of the class labels (Churn) with a Donut Chart.



Display the balance of the class labels (Churn)



Insights:

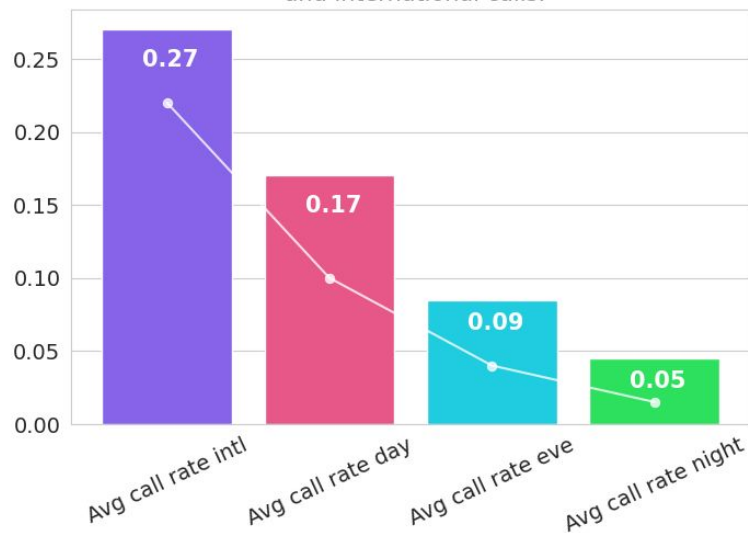
- ❖ The target variable is highly imbalanced, 14.5% of Orange Telecom's customers have churned, but from a customer churn perspective it's still high. Our job will be to reduce it as much as possible.
- ❖ For every single churn customer we have 5.9 retain customers.





Observation 2

Display the average call rate different-different time-zone and international calls.

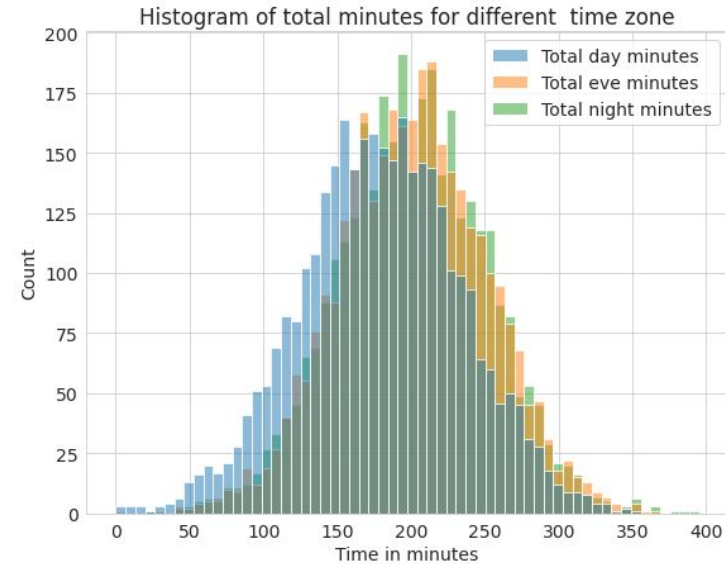
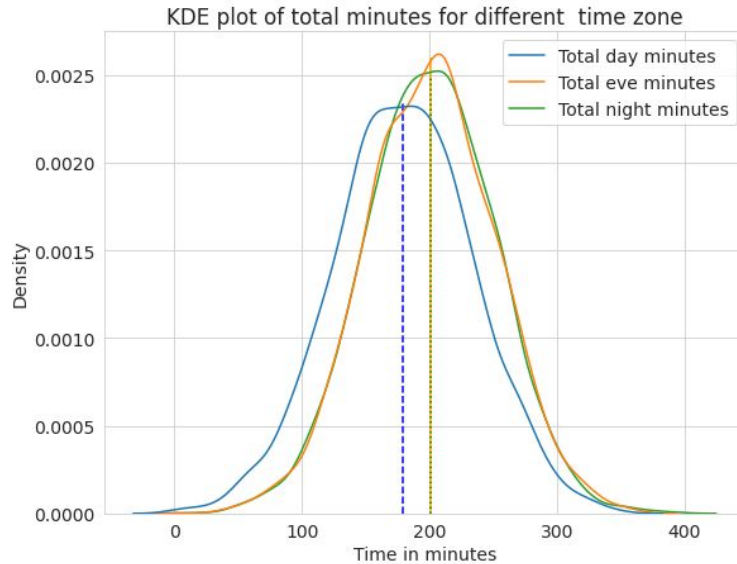


$$\text{Average call Rate} = \frac{\text{Total charge}}{\text{Total minutes}}$$

Insights:

- ▶ Charges for International calls are highest, followed by day, then evening and night calls are cheapest.
- ▶ This is obvious that international calls are expensive compared to domestic calls.





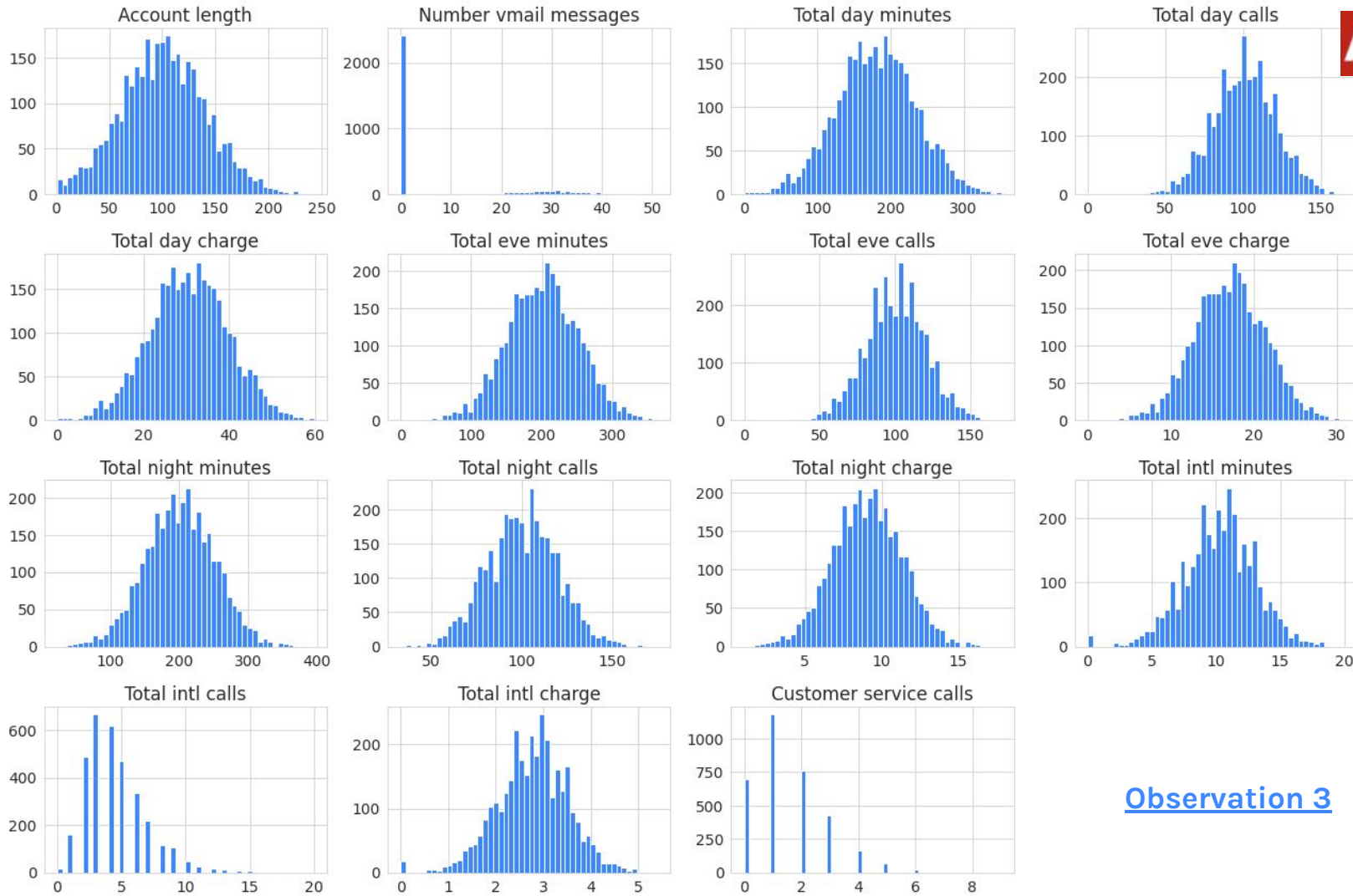
Insights:

- ▶ On average, our customers prefer to talk more at night and evening than during the day.
- ▶ Why? (Is it because day calls are expensive....)

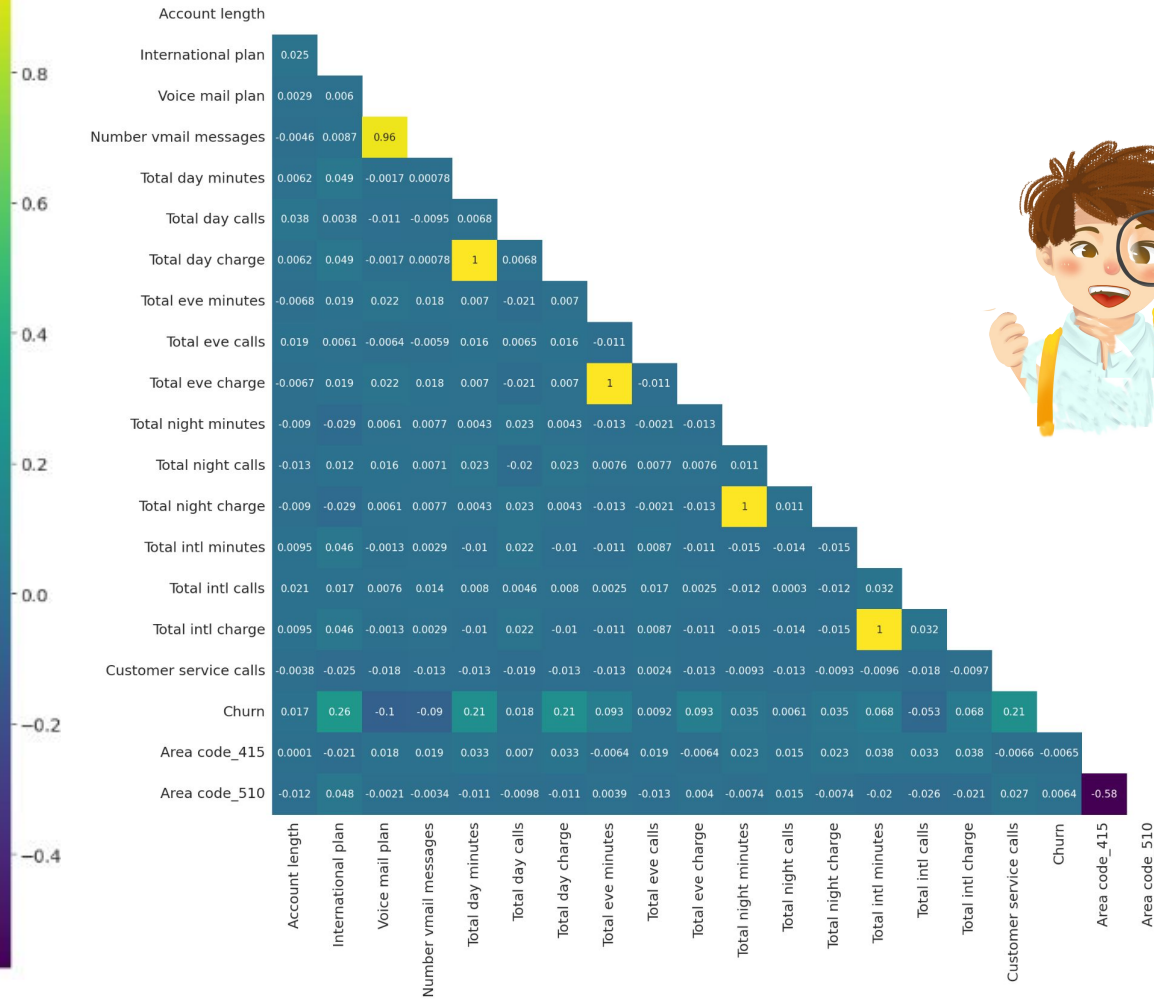


A quick look of Histogram of each numerical

attributes

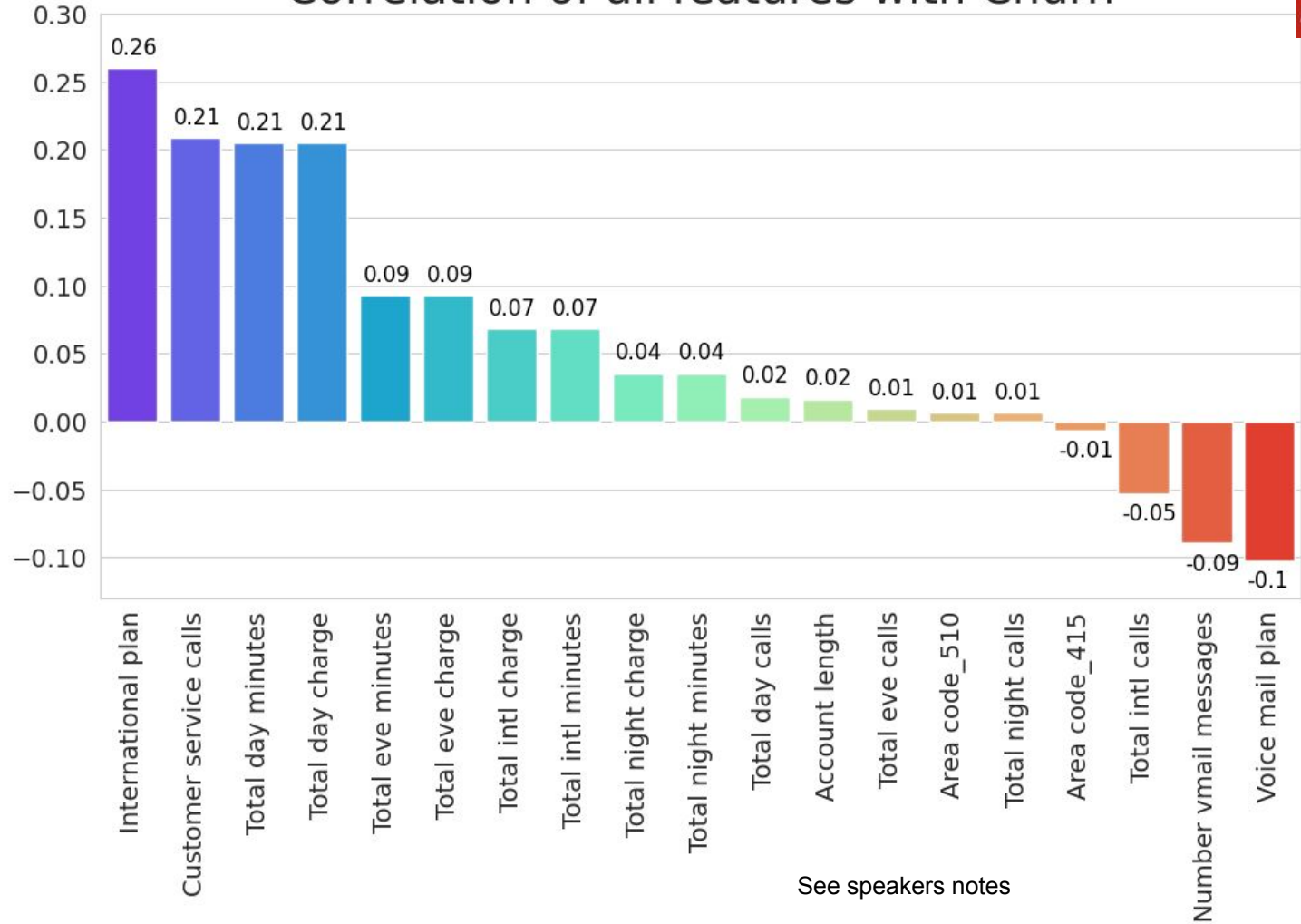


Display Heat Map of correlation matrix



Observation 4

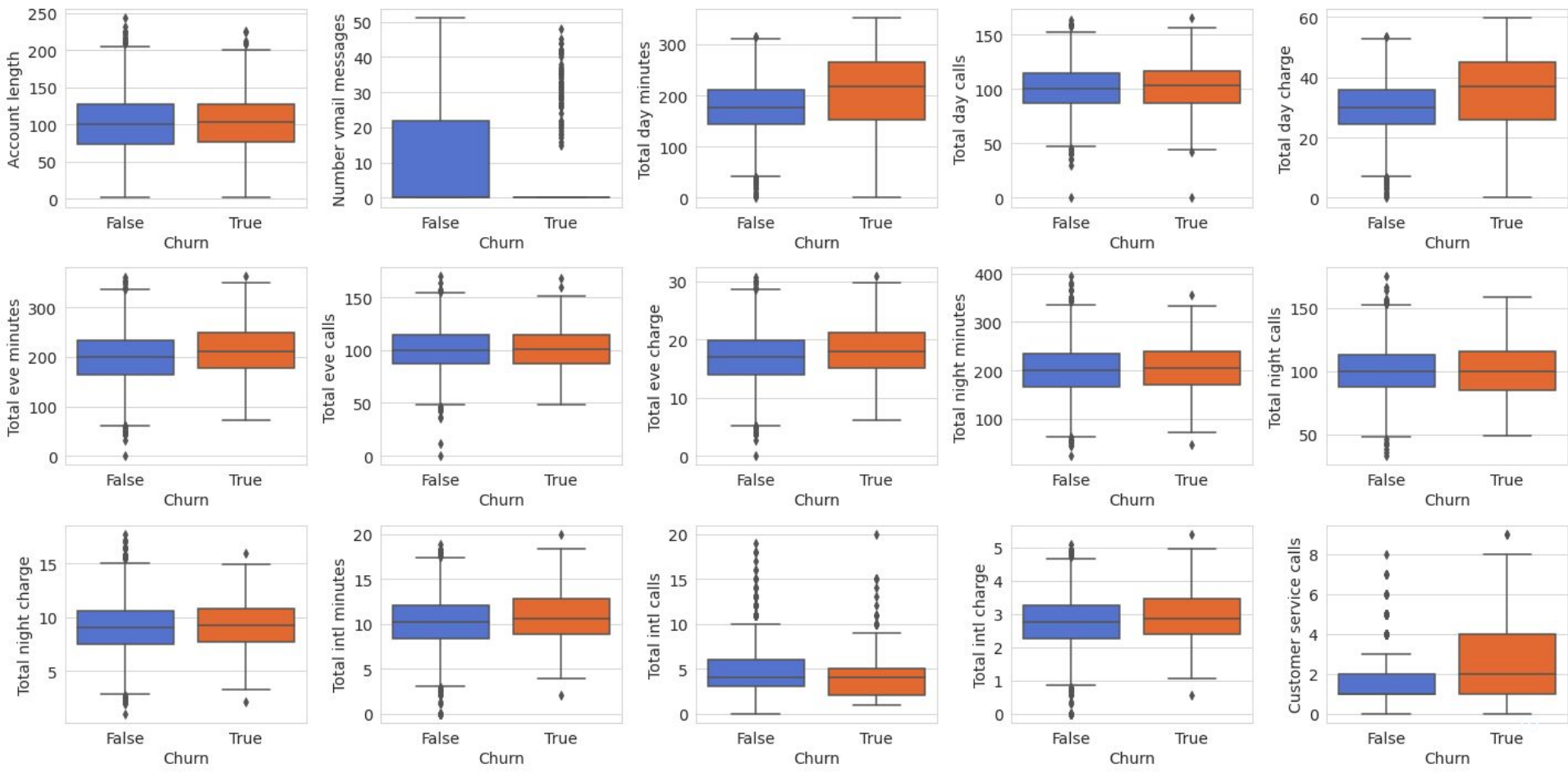
Correlation of all features with Churn



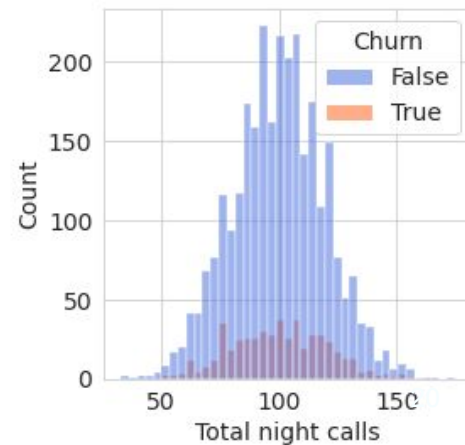
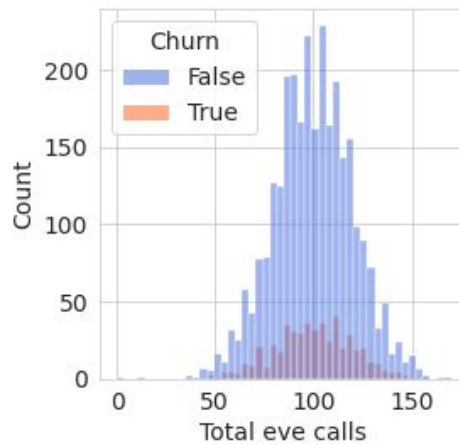
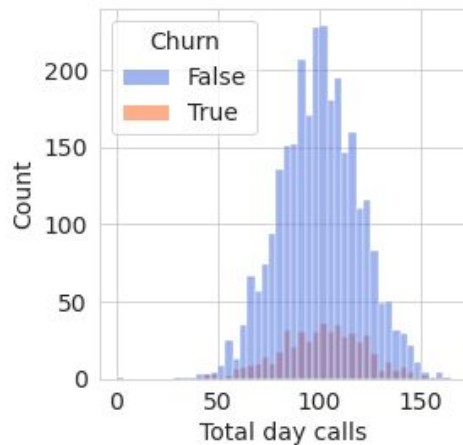
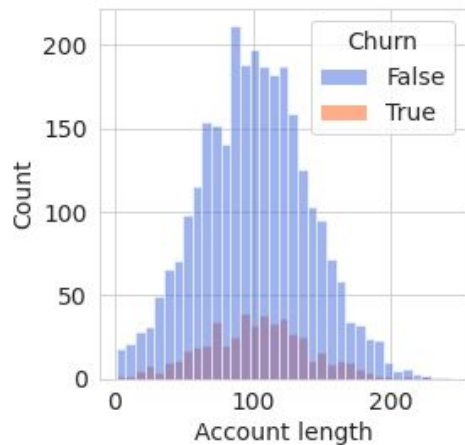
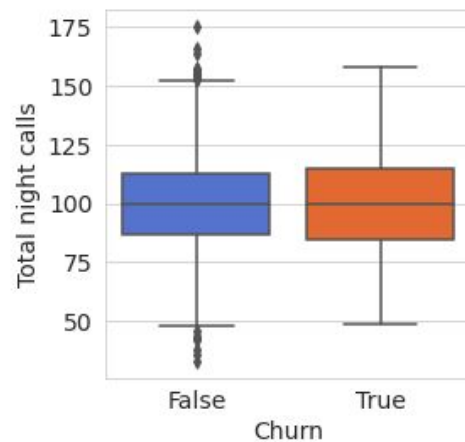
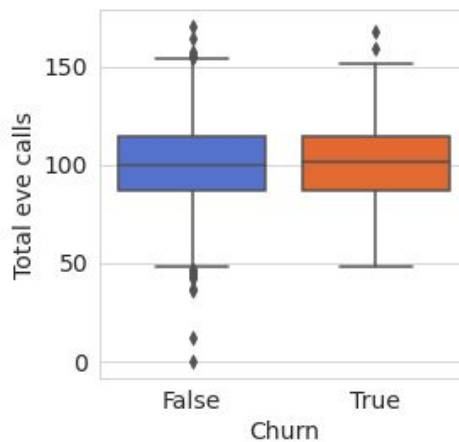
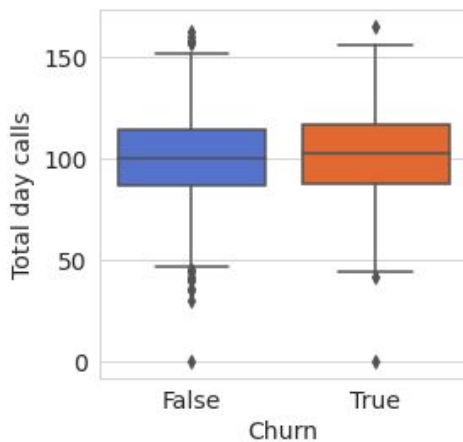
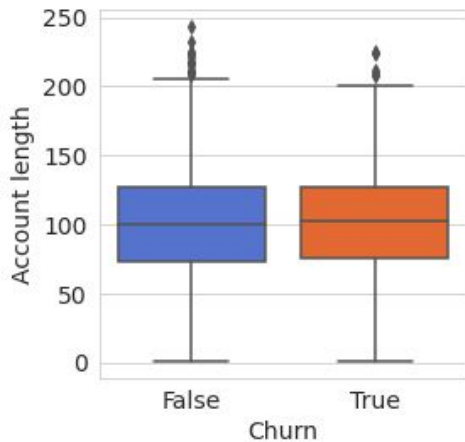
See speakers notes



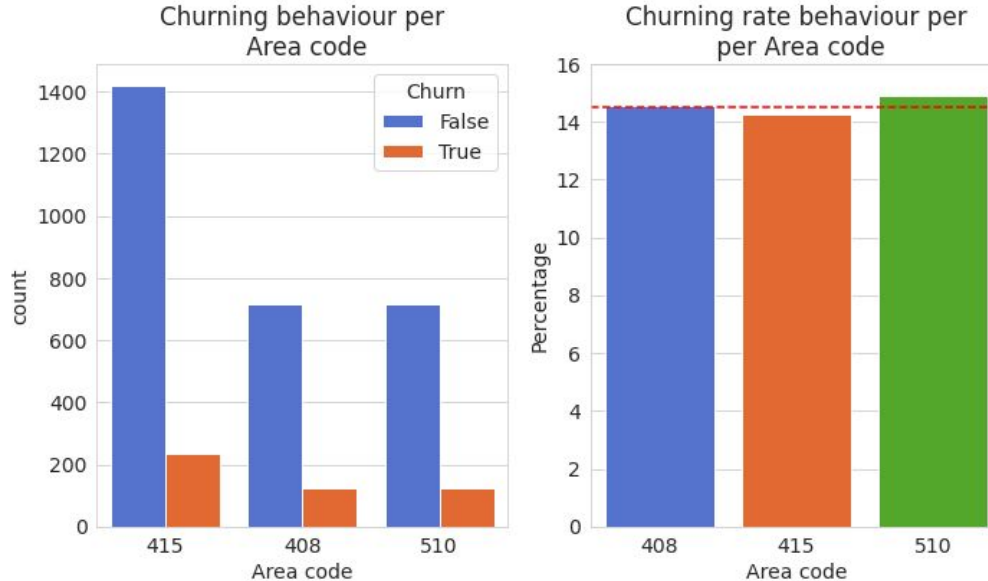
Quick view of Churn behavior with all numerical features



Observation 5: Behavior of features that have no correlation with the target variable.



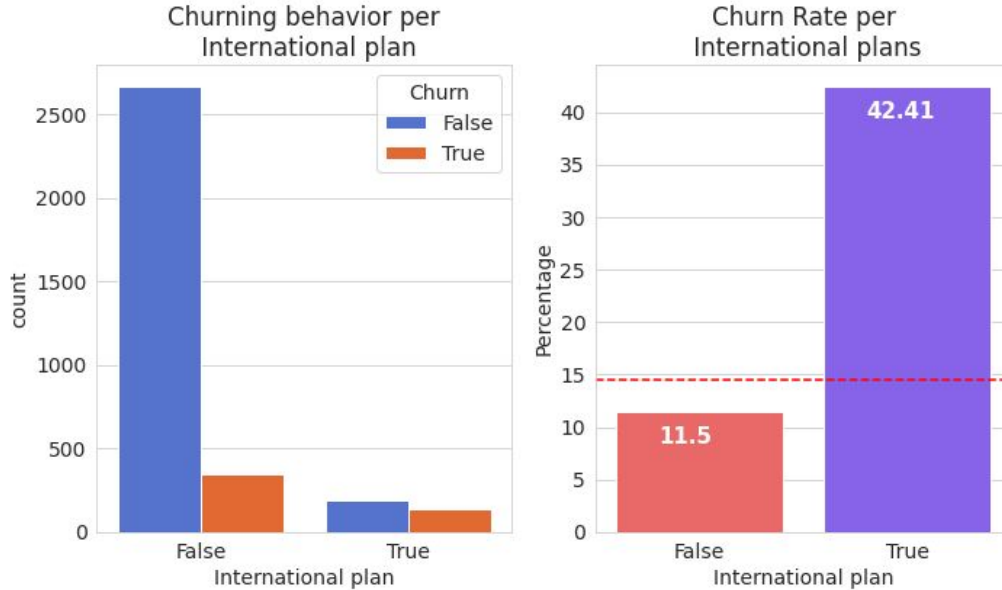
Area Code



Insights:

- ▶ As we can see that Account length, Area code, total calls at different time does not affect the churn behavior of the customers.
- ▶ Area code 415 has twice as many customers as others.

Churn analysis with International plans and International minutes

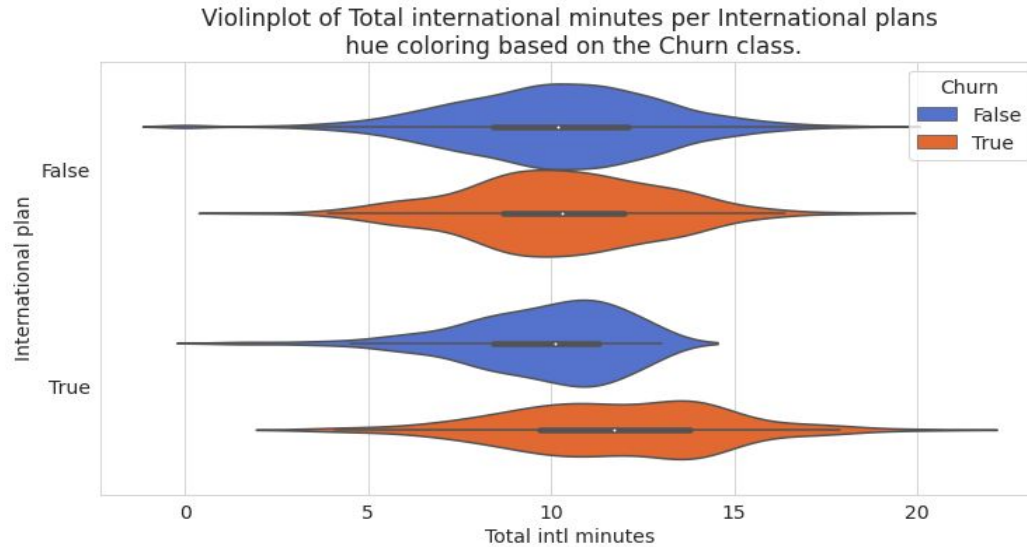


Observation 6



- 9.69% of users have international plans.
- Customers having International plans have overall high churn rate.

With International plan and Total intl minutes



Customer with international plan

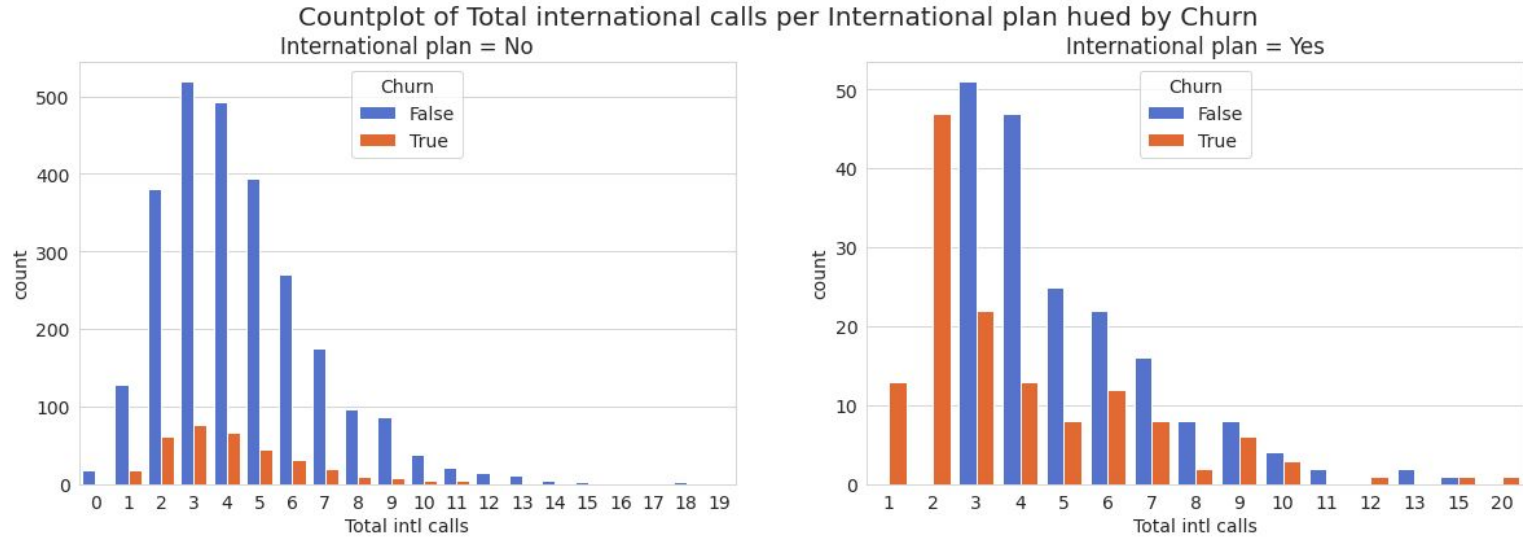
- ▶ High churn rate for high value of total international minutes.

Customers with no International plan

- ▶ Churn rate is independent of the total international minutes.



With With International plan and Total intl calls



Users without international

- ▶ Churn and retain behavior plans follow same pattern

Users with International Plans

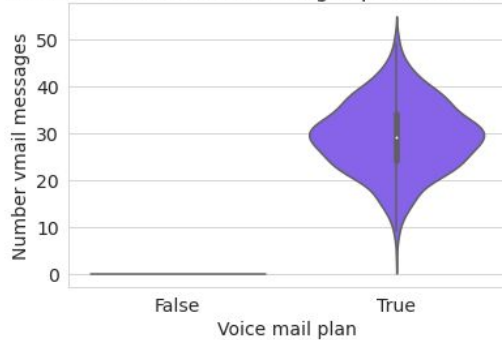
- ▶ Regardless of any number total international calls Churn rate is very high, not a single **user retained till 2 calls**.

Overall it has overall negative correlation as we can see in few slide back

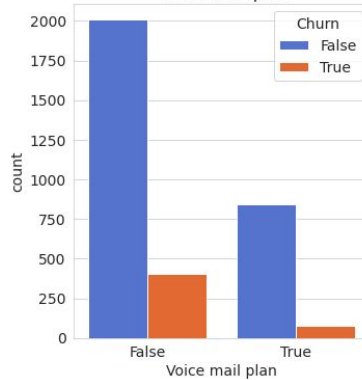


Observation 7

Number of voice mail messages per Voice mail plan



Churning behavior per Voice mail plan



Churn Rate per Voice mail plan



★ 27.66% of users have Voicemail plans.

User without voicemail plan

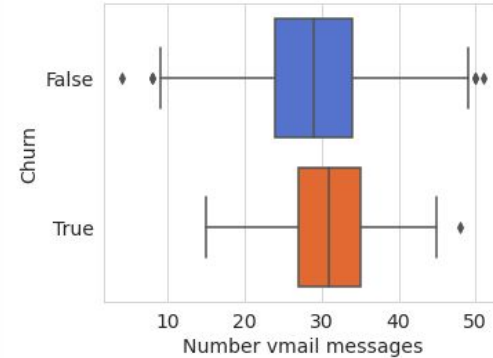
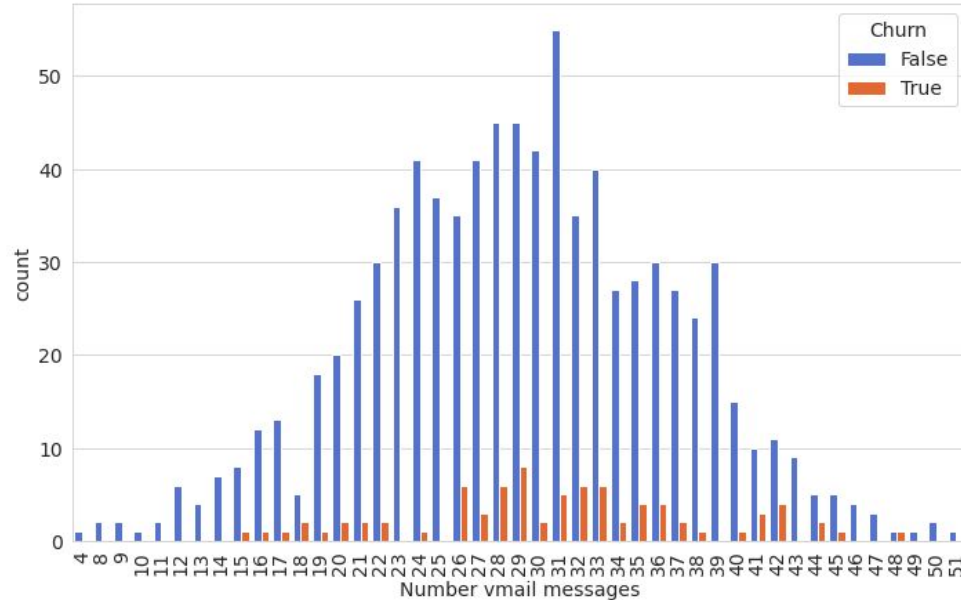
- Those who don't have a voicemail plan don't send voicemails.

User with voicemail plan

- Churn rate of users is low if they have vmail plans.

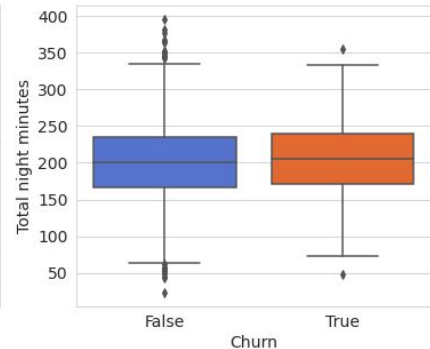
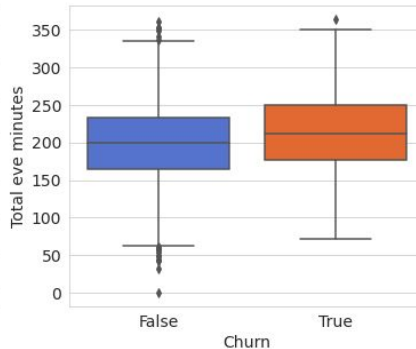
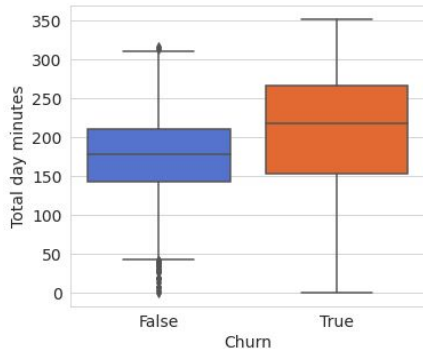
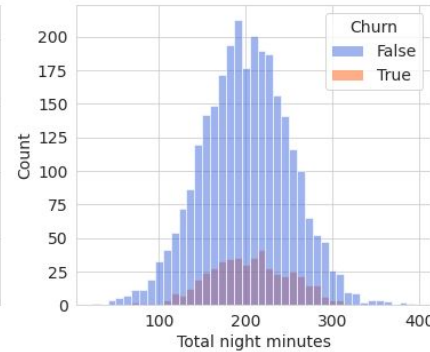
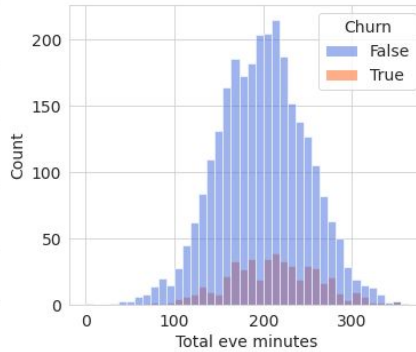
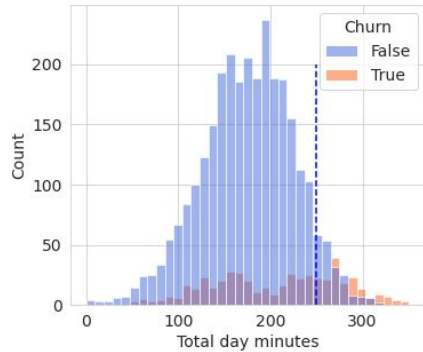
With Number of voice messages

Churn behavior per number of voice mail plan for customers who have taken vmail plans



As number of voicemail messages increases then chances of getting churned is also increasing.

Total day, evening and night minutes



For Day Users

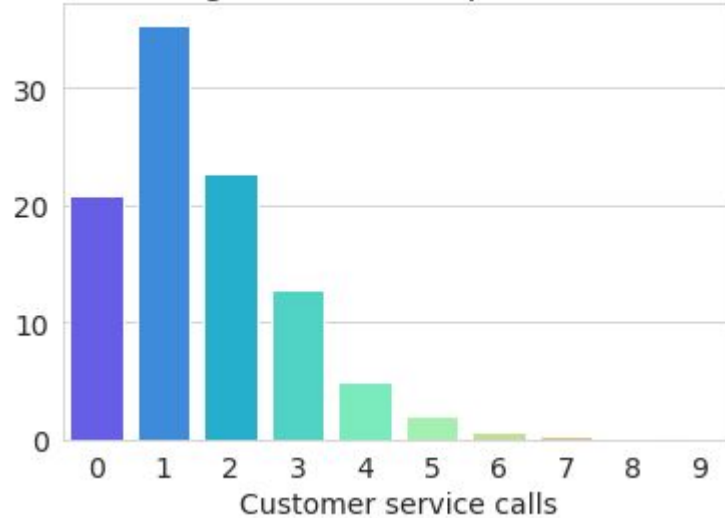
- ▶ The churn rate of the user who talks more on the phone has been seen higher.

For Evening and Night Users

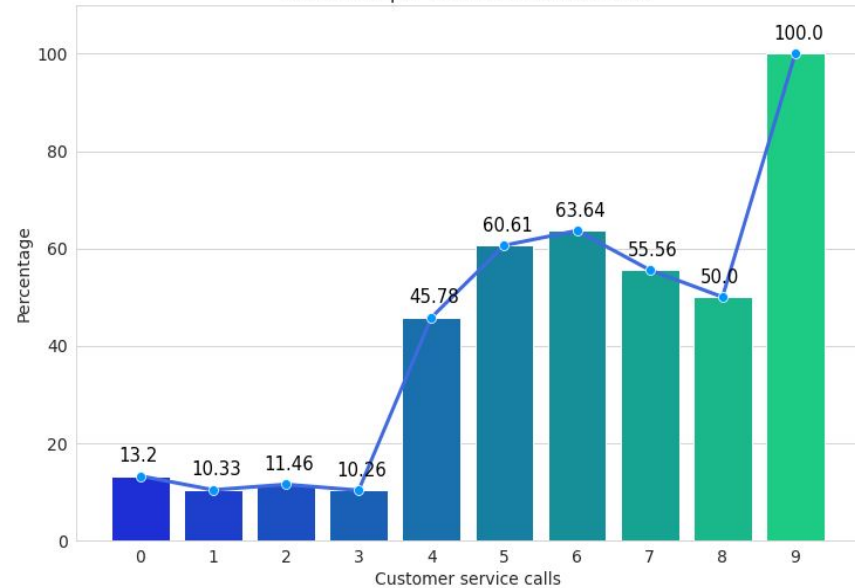
- ▶ The churn behavior is the same for all users, although the in box plot median is slightly higher for churned user in evening times, but the difference is that high compare to day users.

With Customer service calls

Percentage of customer per service calls

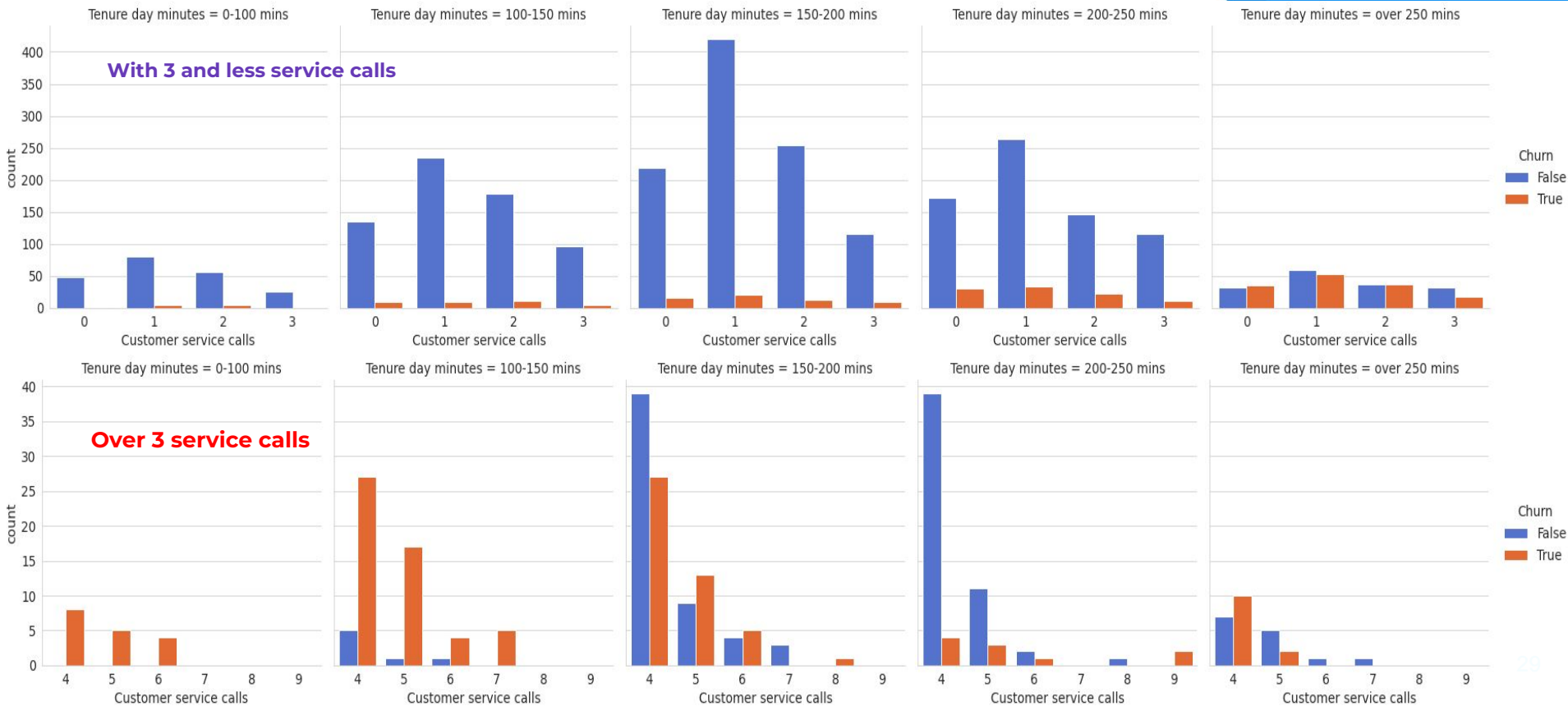


Churn rate per Customer service calls



- ▶ Churn rate increases significantly for 4 or more calls to the customer service.
- ▶ The number of customer service calls greater than 3 is significantly low.

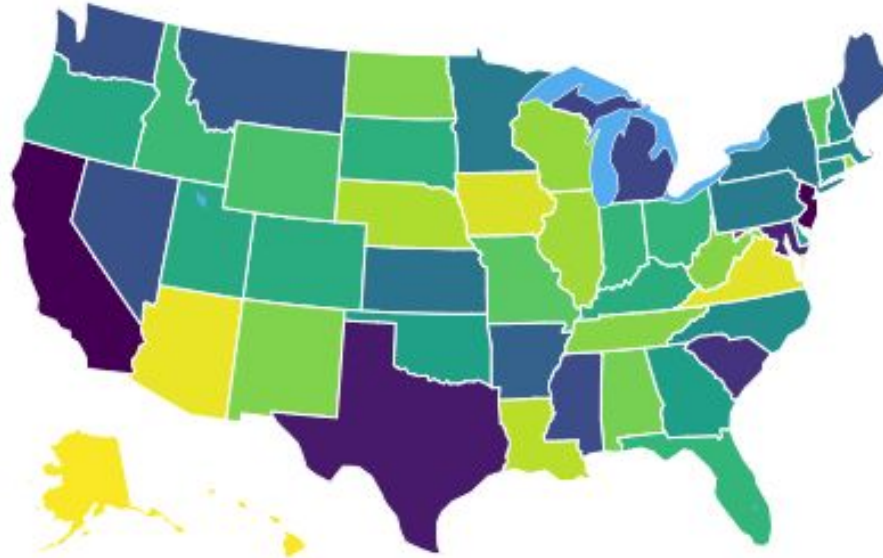
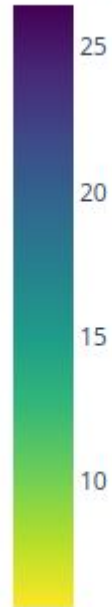
Customer Service calls, per Tenure day minutes





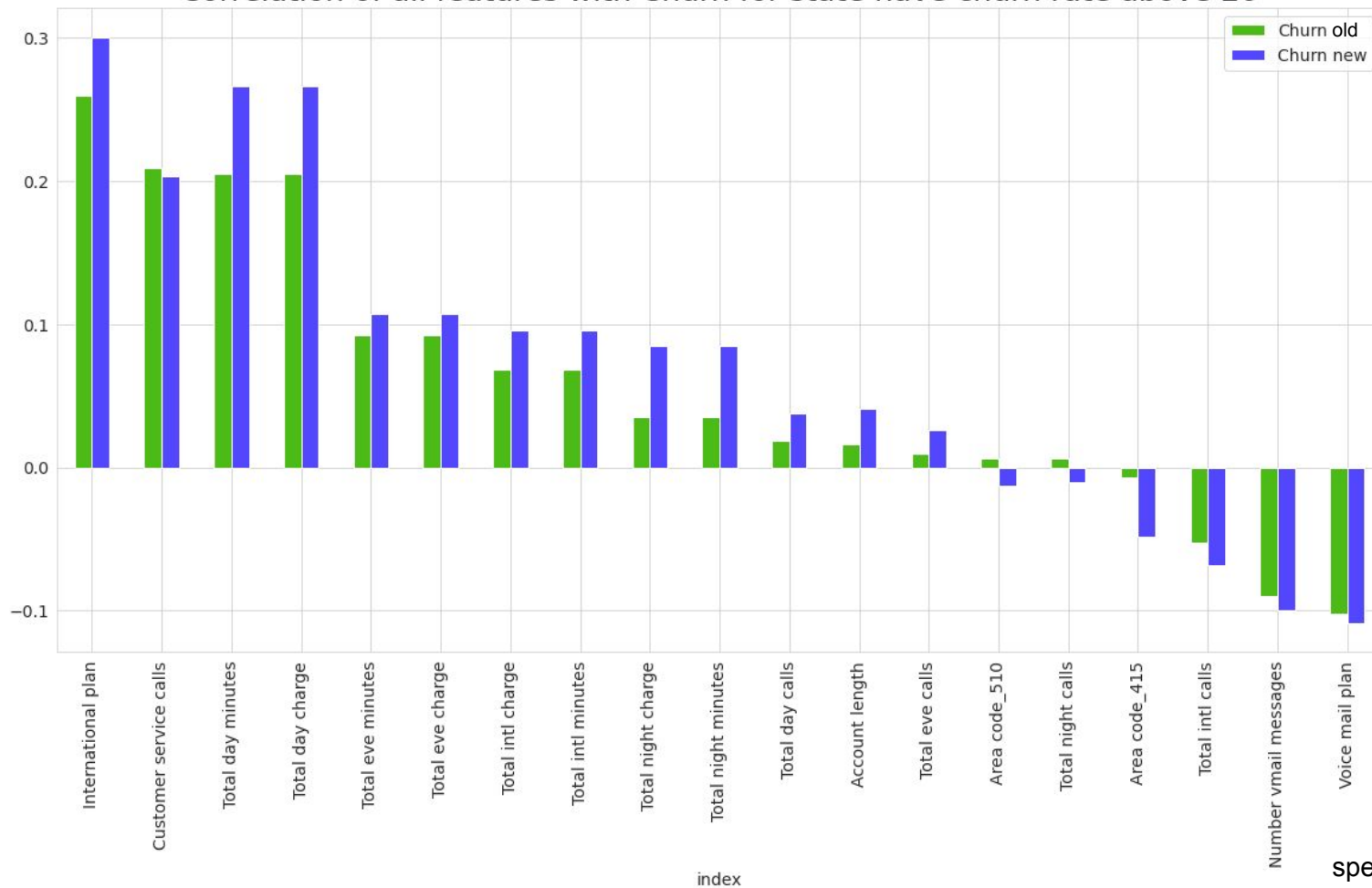
Churn rate for different states by choromap

Churn Rate

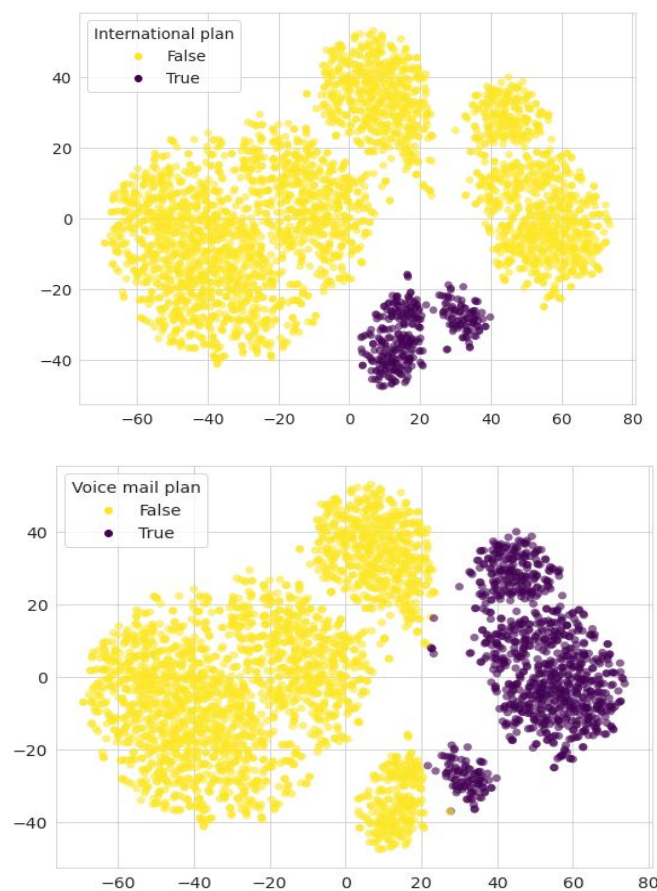
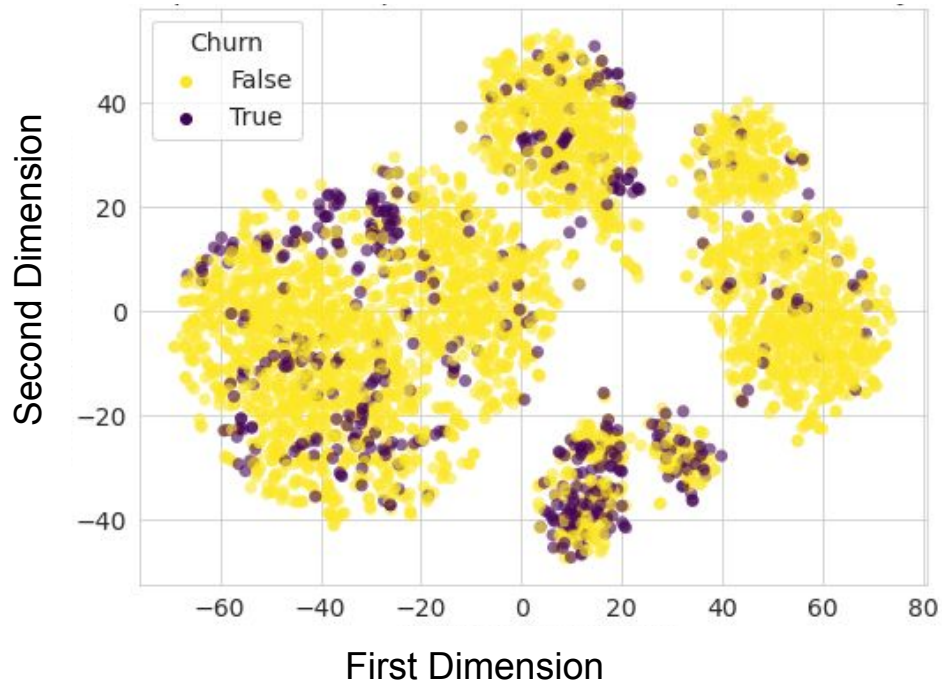


Observation 10

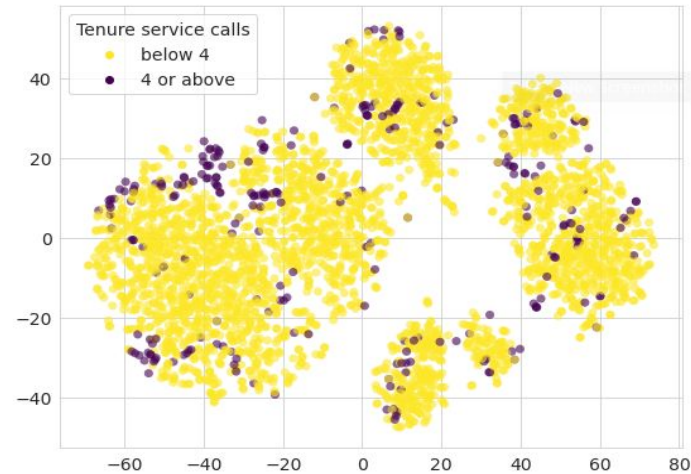
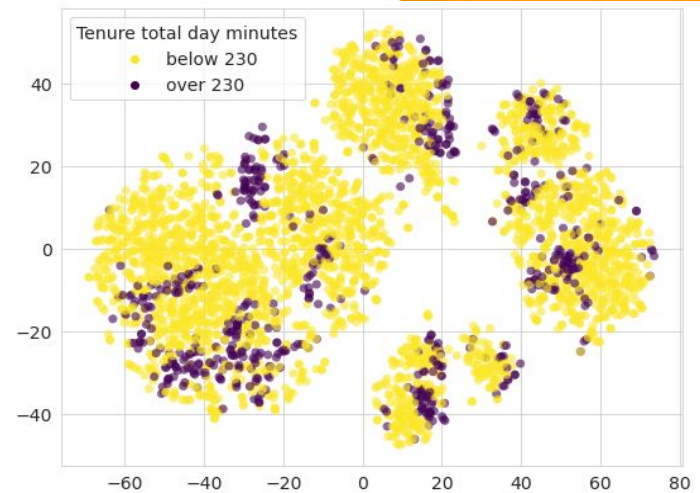
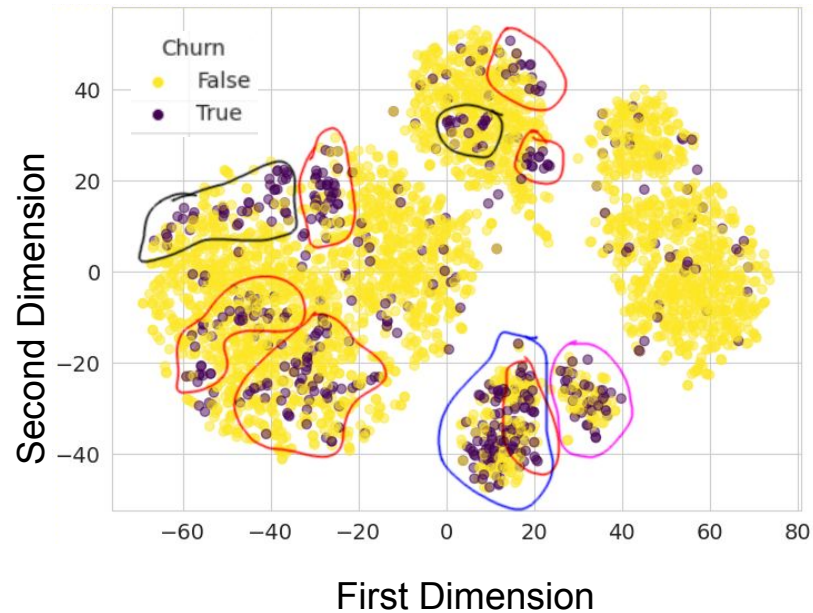
Correlation of all features with Churn for state have churn rate above 20



Scatter plot for 2 components in Reduced dimension (using t- SNE)



Churn cause in the reduced dimension by grouping on features



Recommendations

International plan

We need a very attractive **international plan**, which can provide satisfaction to the customers making international calls.

I

Tariff plan

Need to introduce better tariff plan for **day** as well as evening calling which is specially designed for users talk too much day. Day tariff will be the first priority

T

Need a new **voicemail** plan along with the old one which is specially designed keeping in mind the more voicemail senders.

Voicemail Plan

V

Need to improve feedback system that doesn't ignore customer problems

Feedback System

F



Thank You!!!

Any questions?



*Thank you
For your
Attention*