**Telecom Churn Analysis**

**Md Shahrukh Khan, Shadab Ali Khan**

**Data science trainees,**

**AlmaBetter, Bangalore**

**Abstract:**  In recent years, the telecom market has been very competitive. The cost of retaining existing telecom customers is lower than attracting new customers. It is necessary for a telecom company to understand customer churn through customer relationship management (CRM).

Feature construction of the churn dataset includes equidistant grouping of customer behavior features to expand the space of features and discover latent information from the churn dataset. The original and new churn datasets are analyzed in the stacking ensemble model with four evaluation metrics. The experimental results show that the proposed customer churn predictions have accuracies of 96.12% and 98.09% for the original and new churn datasets, respectively. These results are better than state-of-the-art churn recognition systems.

**1. 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.

In this problem statement we were required to explore and analyze the data to discover key factors responsible for customer churn and come up with ways/recommendations to ensure customer retention.

The main objective is to build a predictive model, which could help them in predicting the surge pricing type proactively. This would in turn help them in matching the right cabs with the right customers quickly and efficiently.

**Variable Breakdown**

1. State - 51 Unique States in United States of America
2. Account Length - Length of The Account
3. Area Code - 415 relates to San Francisco,408 is of San Jose and 510 is of City of Okland
4. International Plan - Yes Indicate International Plan is Present and No Indicates no subscription for International Plan
5. Voice Mail Plan - Yes Indicates Voice Mail Plan is Present and No Indicates no subscription for Voice Mail Plan
6. Number vmail messages - Number of Voice Mail Messages ranging from 0 to 50
7. Total day minutes - Total Number of Minutes Spent By Customers in Morning
8. Total day calls - Total Number of Calls made by Customer in Morning.
9. Total day charge - Total Charge to the Customers in Morning.
10. Total eve minutes - Total Number of Minutes Spent By Customers in Evening
11. Total eve calls - Total Number of Calls made by Customer in Evening.
12. Total eve charge - Total Charge to the Customers in Morning.
13. Total night minutes - Total Number of Minutes Spent By Customers in the Night.
14. Total night calls - Total Number of Calls made by Customer in Night.
15. Total night charge - Total Charge to the Customers in Night.

**2. Introduction**

Owing to fierce competition among telecom companies, customer churn is inevitable. Customer churn is the act of a customer ending a subscription to a service provider and choosing the services of another company. Companies must reduce customer churn because it weakens the company. A survey showed that the annual churn rate in the telecom industry ranges from 20% to 40%, and the cost of retaining existing customers is 5–10 times lower than the cost of obtaining new customers. The cost of predicting churn customers is 16 times lower than that for obtaining new customers. Decreasing the churn rate by 5% increases the profit from 25% to 85%. This shows that customer-churn prediction is important for the telecom sector.

Telecom companies consider customer relationship management (CRM) an important factor in retaining existing customers and preventing customer churn. To retain existing customers, CRM analyzers must predict which customers will churn and analyze the reasons for customer churn. Once the at-risk customers are identified, the company must perform marketing campaigns for churn customers to maximize the churn-customer retention. Therefore, customer-churn prediction is an important part of CRM.

The accuracy of the prediction systems used by CRM analyzers is important. If analyzers are inaccurate in predicting customer churn, no campaigns can be performed. Owing to recent advancements in data science, data mining and machine learning technologies provide solutions to customer churn. However, there are several limitations in existing models. For example, logistic regression, a common churn-prediction model based on older data-mining methods, is relatively inaccurate.

In this project we are –

1. Analyzing the given data set to know the reasons of churn
2. Suggesting ways/recommendations to ensure customer retention.

## **3. Types of Day Interval**

### Day

### Evening

### Night

### International

### During all the day interval we have further classified data as -

### No of minutes

### No of call

### Charges,

## **4. Reasons for Customer Churn**

## Basic reasons for Customer Churn are –

1. Number of customer care calls
2. Voice Mail Message
3. Poor service during day time , evening time, and nigh time for churn customers
4. Poor service for customers having International Calls
5. Churn because of problem in some specific States
6. Churn because of Customer Service
7. **Customer care calls**

We have found that from analysis that no of Calls to Churn Customer is high as compared to Non Churn Customer.

Recommendation - Customers should not be called more than 7 times.

1. **Voice Mail Message Analysis**

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 or setting a limit and check whether a customer is retained.

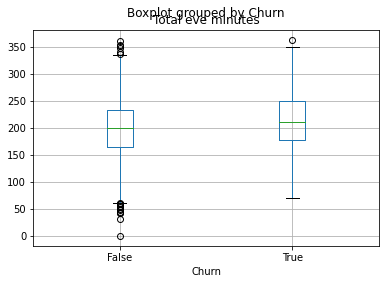
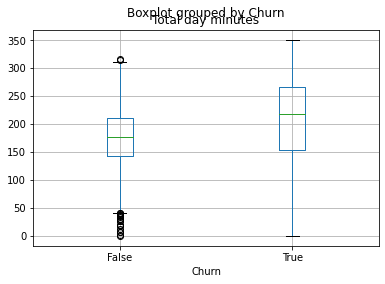
Recommendations:

1. Voice-Mail Service Up gradation
2. Setting up a limit on Voice-Mail service strictly no more than 25 voice mails.
3. Quality Drop in Voice-Mail after 25 voice mails.

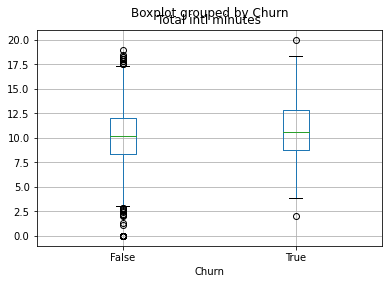
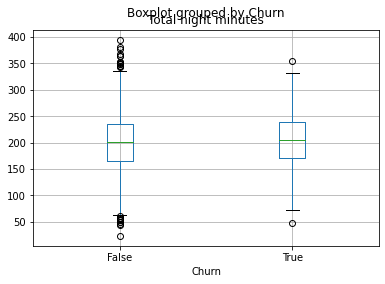
# **Complete Day Analysis**

1. **No of minutes**

We can infer from the box-plot that with users spending more 220 minutes or more i.e. approx 4hrs tend to switch to other operator.



Total Day Minutes Total Evening Minutes

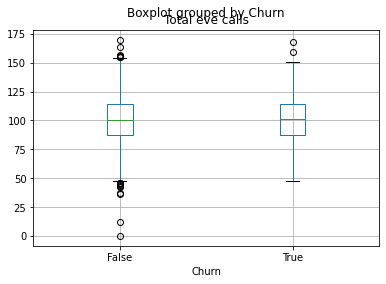
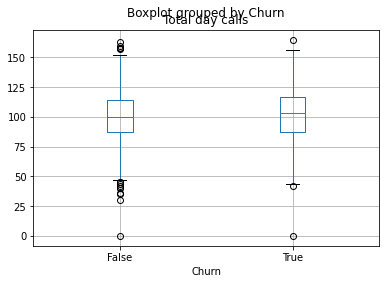


Total Night Minutes Total International Minutes

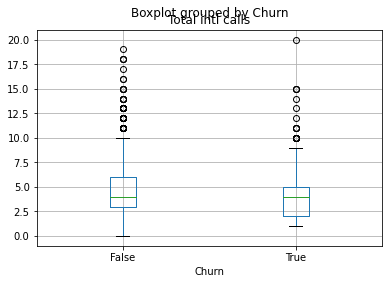
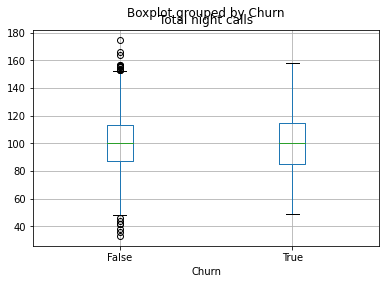
**Recommendations**

Following could be the points that should be implemented –

1. Network Disturbance during a Call should be reduced
2. Need to Upgrade or make smarter use of technologies like VoLTE to improve Voice Quality.
3. Network Up gradation
4. **No of calls**



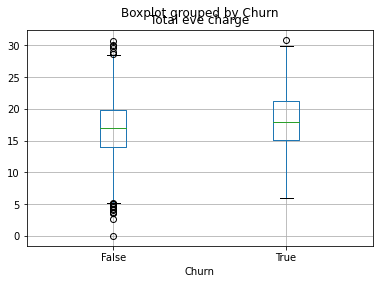
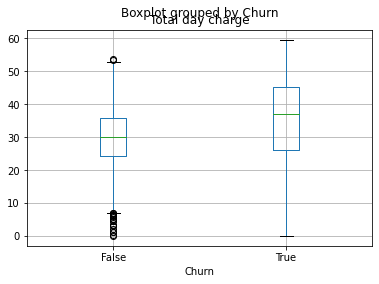
Total Day Calls Total Evening Calls



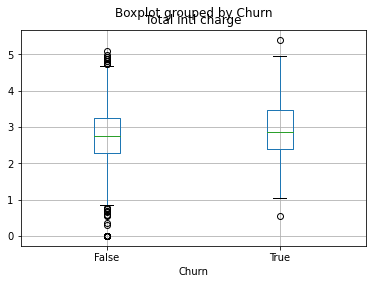
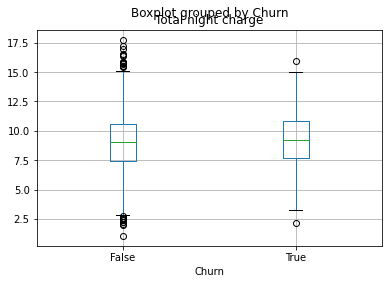
Total Night Calls Total International Calls

We can infer here that on an average 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 indicate there are call drops which may lead to more calls in a morning.

1. **Service Charge**



Total Day Charges Total Evening Charges



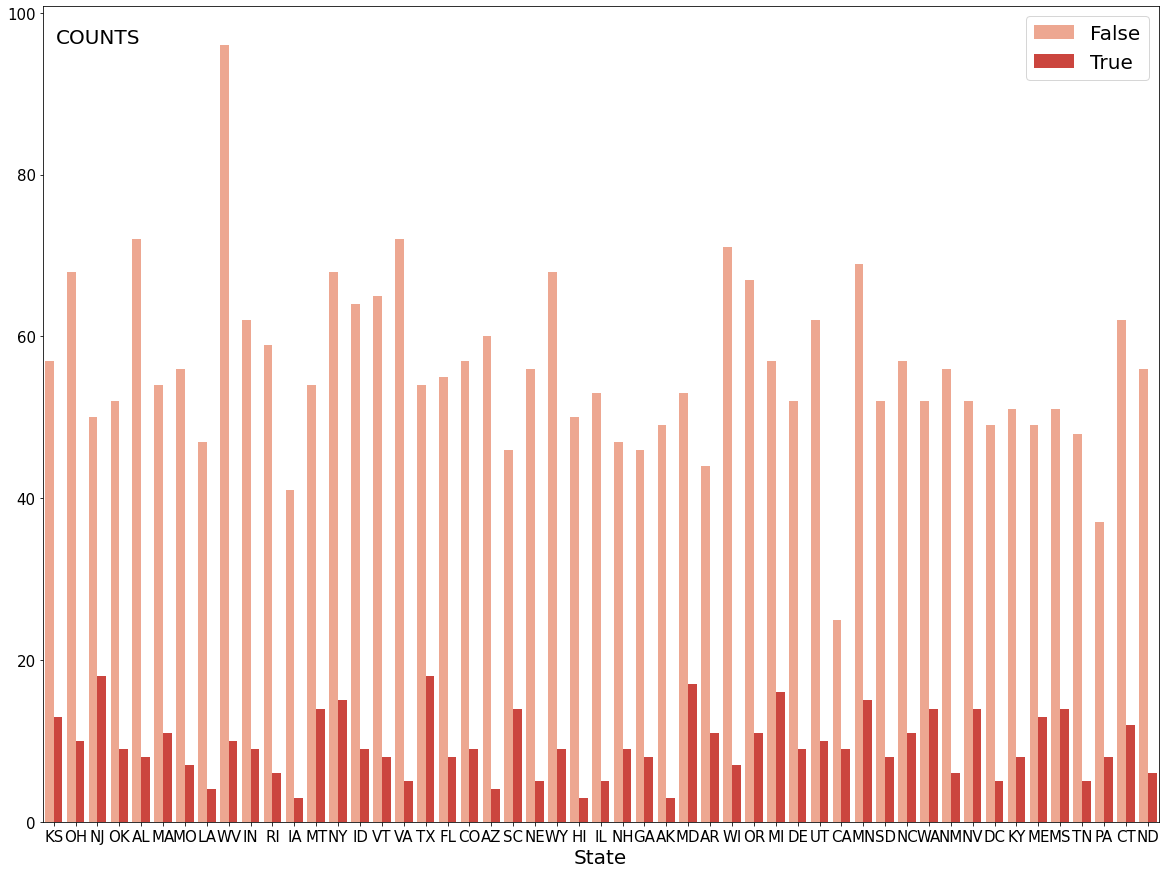
Total Night Charges Total Intern. Charges

Well, here we can clearly indicate a strategy 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 strategy of the company.

**Recommendations** –

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

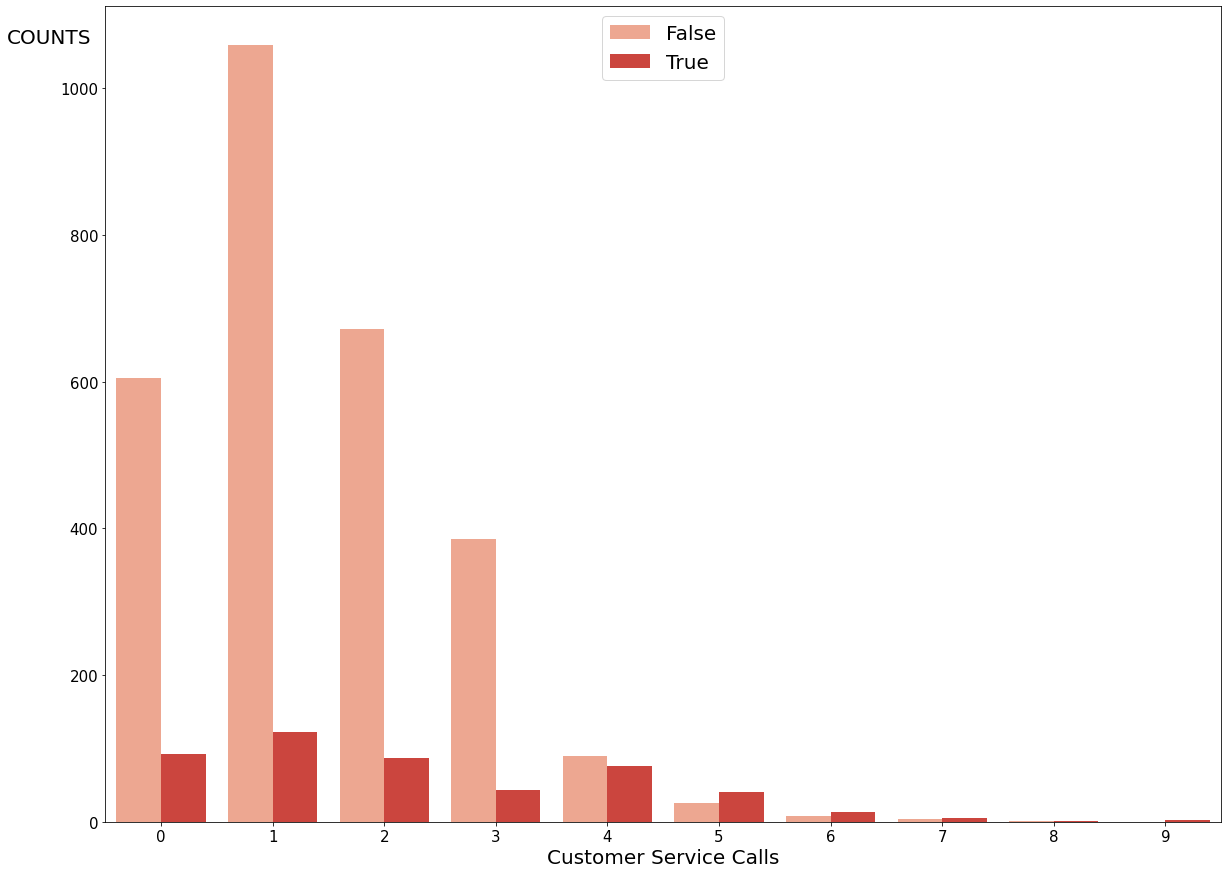
# **States having more churn than other**



We can clearly see that NJ, TX, MD have the bit more churn rate than usual, A Network Up gradation would be strongly suggested in these areas.

I believe by upgrading connectivity and overall services in these areas churn rate can be reduced significantly.

## **Churn because of Customer Service**



While some consumers are lazy and have switched to another network operator without resolving their issue, customers who have just contacted once have a high churn rate, indicating that their issue was not resolved on the first attempt. Some clients' problems remained unsolved even after multiple calls.

1. In such circumstances, feedback is required.
2. It should provide the customer assurance that their problem will be resolved in the first attempt.

**Conclusion**

The analysis of CSV file helped us determine important factors affecting customer purchasing behavior in the telecommunications industry. We have analysed all the factors carefully and came to conclusion that there were several shortcomings in the services of Orange Telecom Inc.

On the basis of our analysis we gave various recommendations. I believe that on implementing these points Churn rate can be reduced drastically.

**Reference**

[1] Applied Science Article MDPI

[2] GeeksforGeeks

[3] Wikipedia

[4] DataCamp