Capstone Project - 1
Telecom Churn Analysis

Team Members

Arunav Goswami
Ashraf Ali Korbu
Mohammed Saad Pasha
Nayanjyoti Sharma
Pratik Vishwakarma

Introduction

- **Churn :** It is a problem for Telecom companies as because it is more expensive to acquire new customers than to keep your existing ones from leaving.
- **Company :** Orange S.A. formerly France Telecom S.A. is a French Multinational telecommunications corporation.
- It has 266 million customers worldwide, and is the 11th largest mobile network operator in the world.
- **Data set :** The data set given is of the region **USA**, Consisting of all 51 states. And the data set is of **.csv** format.

Project Objectives

- To discover key factors responsible for customer churn.
- Perform EDA (Exploratory Data analysis) on different Variables present in the data set.
- To come up with ways or recommendation to ensure customer retention.

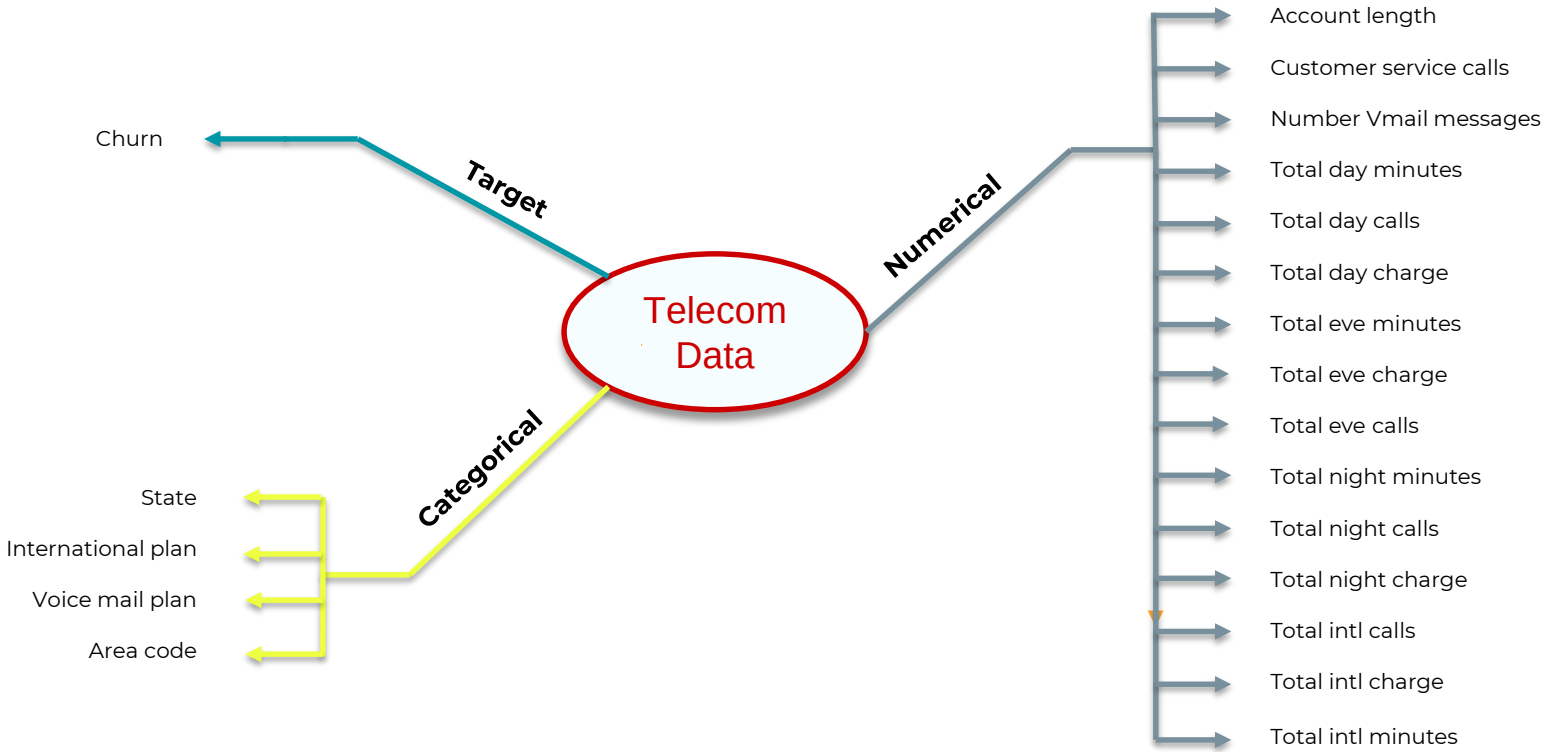
Data set description

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3333 entries, 0 to 3332
Data columns (total 20 columns):
 #   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
```

- There are a total of **3333** numbers of Rows present in the data set.
- The total number of columns present are **20**.
- There are **no missing values**.

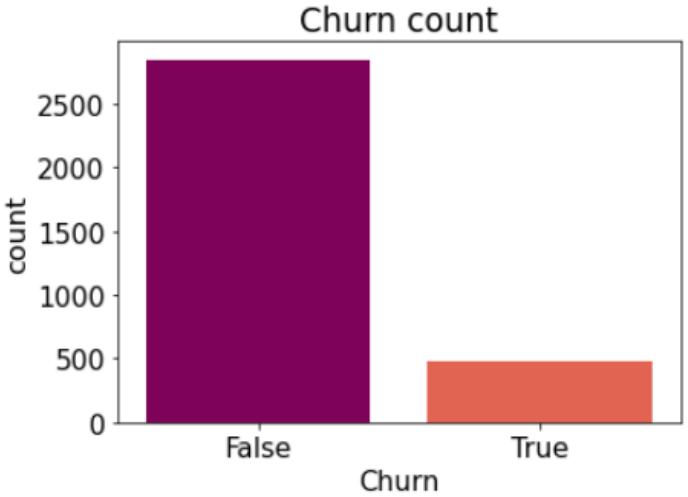
Data Summary

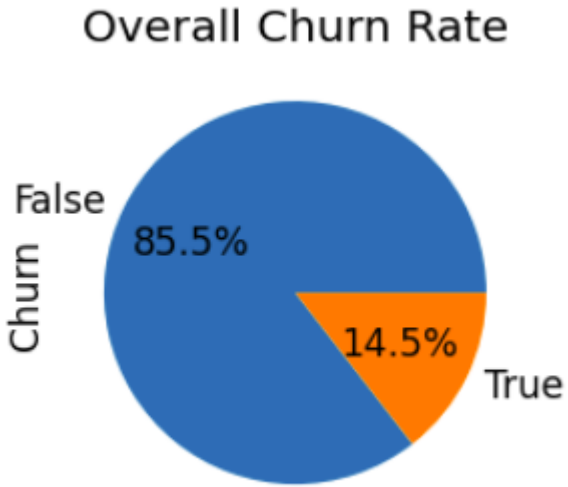


EDA (Exploratory Data Analysis)

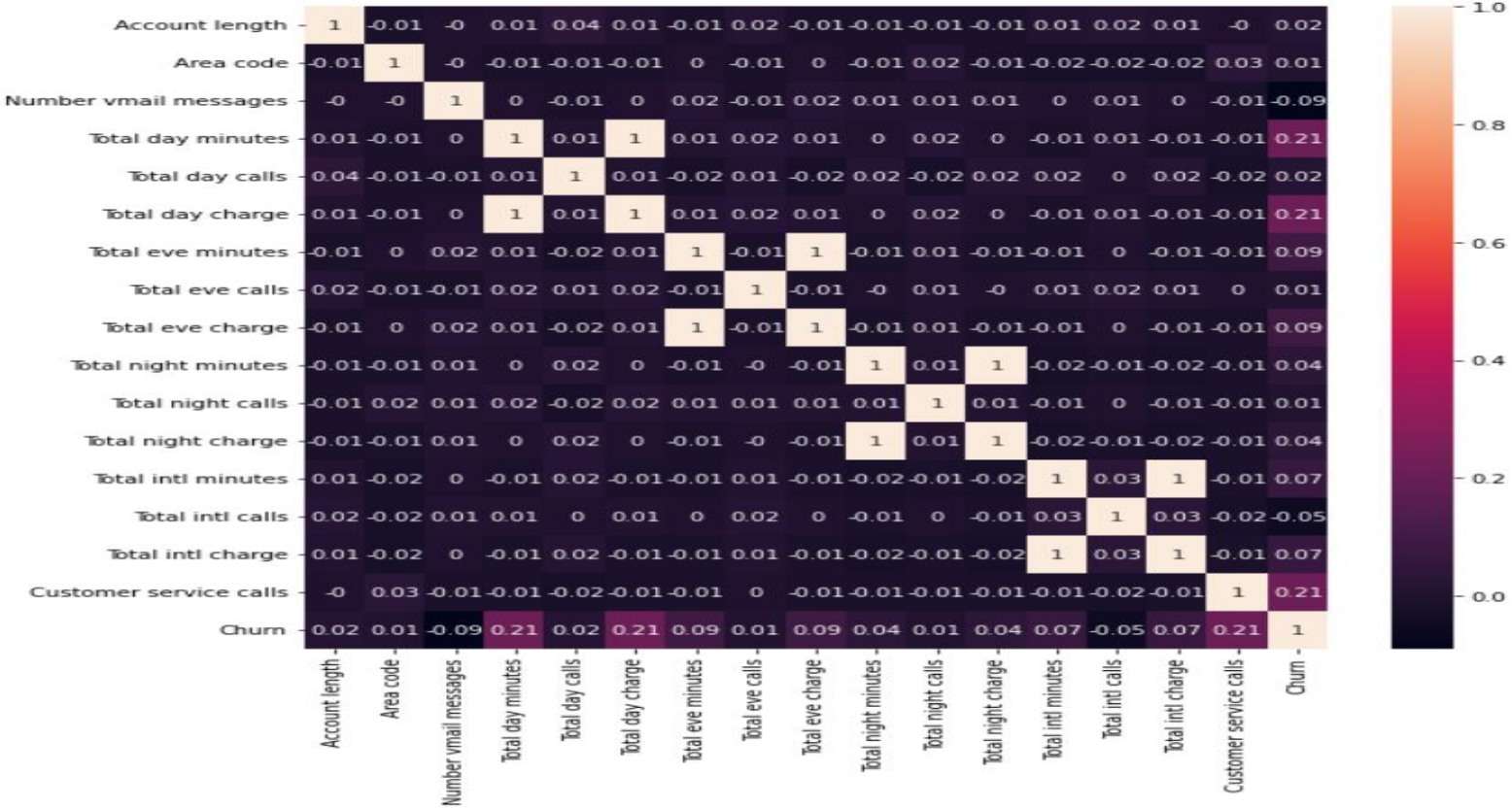
- Exploratory Data Analysis is an approach to analyze data sets and to summarize their main characteristics, often with visual methods.
- Data visualization using **Matplotlib** and **Seaborn.**
- Performing EDA helps in seeing what the data can tell us.

EDA (continued)

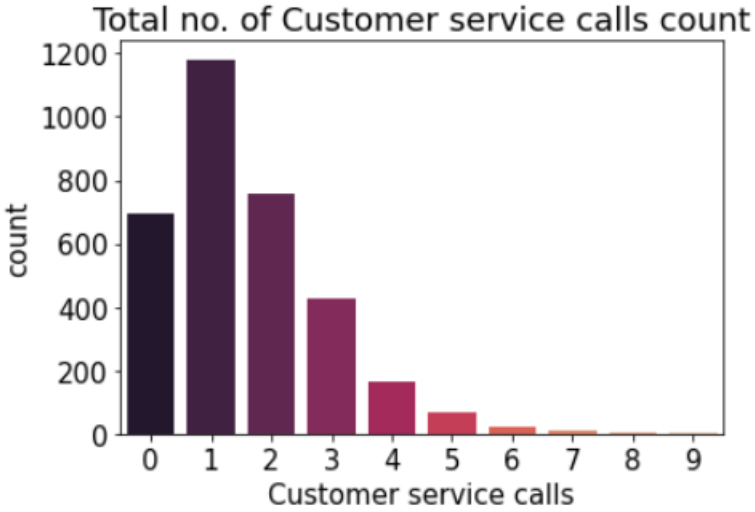


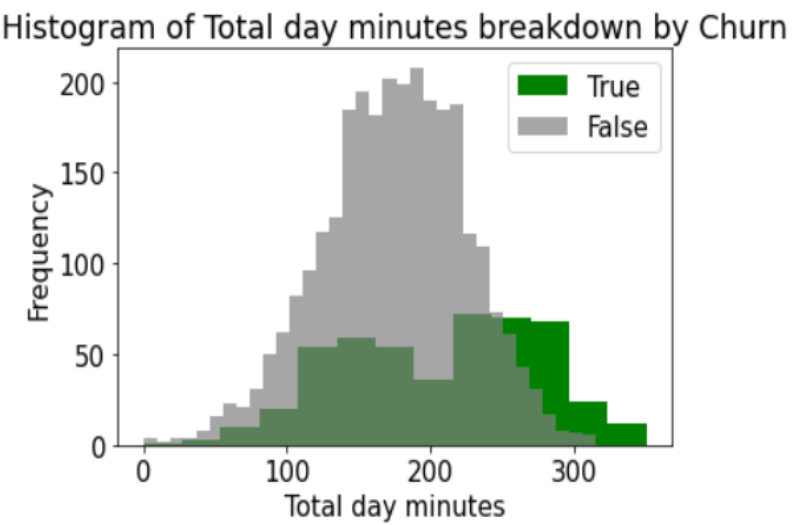


EDA (continued)



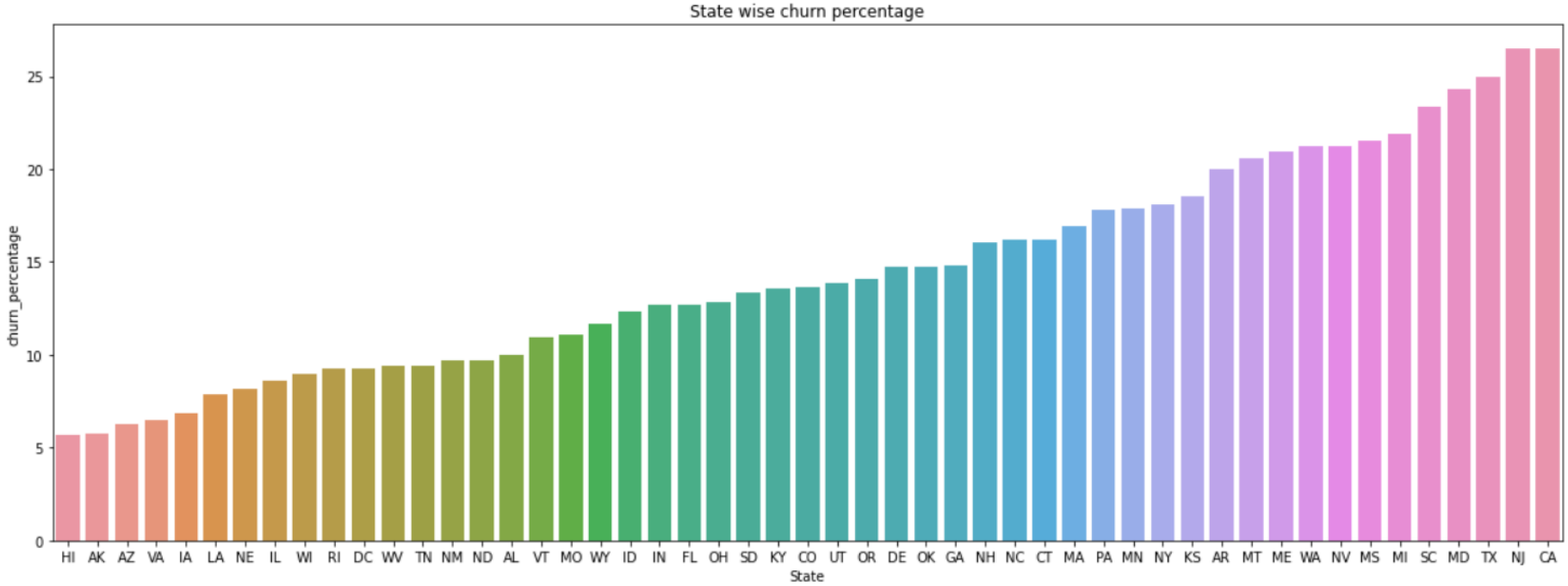
EDA(continued)



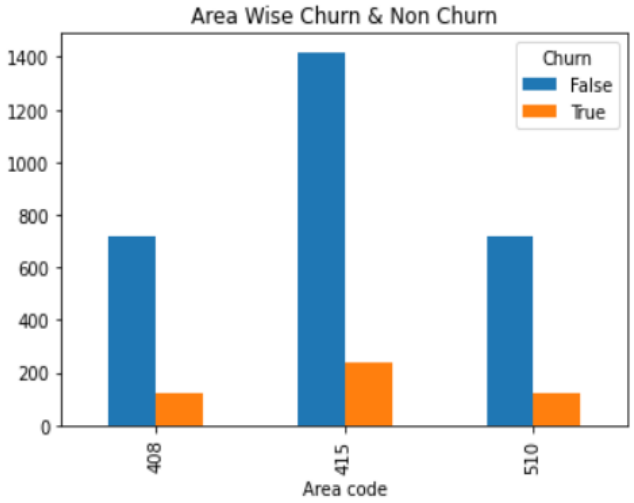


EDA (continued)

State wise churn percentage:-

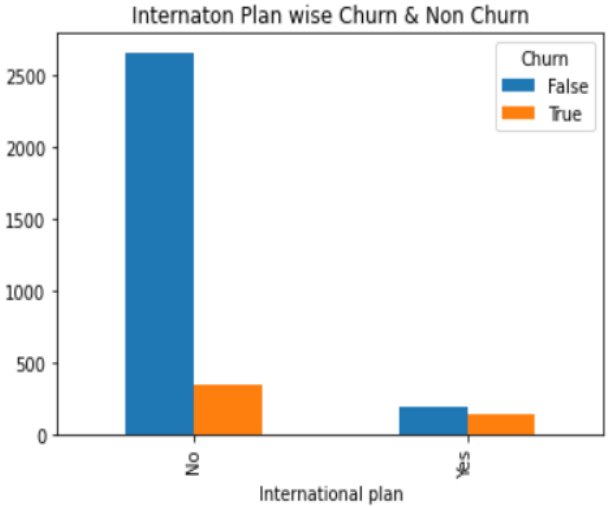


EDA (continued)



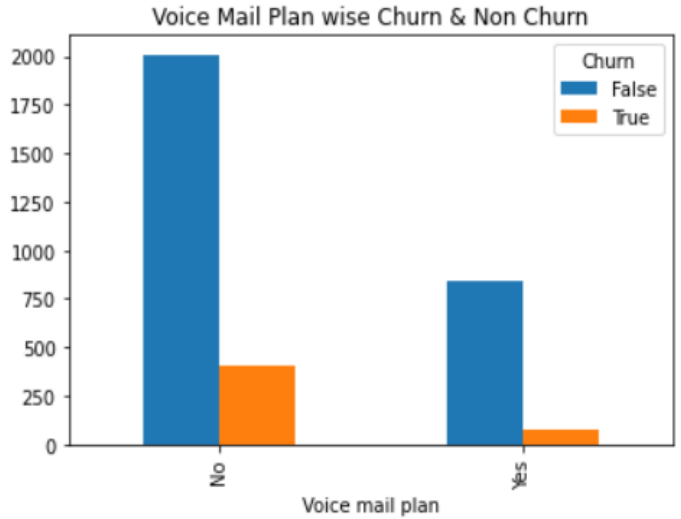
Area code	Non Churn	Churn	churn_per
408	716	122	14.558473
415	1419	236	14.259819
510	715	125	14.880952

EDA (continued)



11.5 % churn rate for no international plan & 42.41 % for international plan

EDA (continued)



16.7 % churn rate for no voice mail plan & 8.6 % for voice mail plan

Observations

- Churn rate increases with increase in customer service calls
- Customers with longer talk time in the morning have high chance of discontinuing the subscription.
- High populated areas have high churn rate.
- Customers with international plan have high chance of leaving the subscription.
- For customers with no voice mail plan, the chance of discontinuing the subscription is greater than those with voice mail plan.

Conclusion

From the observations, we can conclude that the company needs to work on the following fields for customer retention:

- Offer better service in most populated areas.
- Provide better long talk time plans.
- Take feedback and suggestions on a regular basis, try to implement it and strive for better communication.
- Company can offer better International plans.
- Offer more Incentives in the form of discounts and cashbacks to churn customers

Challenges

- Selection of features was quite a challenge.
- For some of the variables like **‘Account length’**, we have no idea about its relevance to Churn.
- We cannot consider all these variables as the exact reasons why customers left because they might leave for the better price offered by competitors or the bad economy at a certain time, etc.
- Churn does not take into consideration the types of customers that are leaving – new vs old

Thank You