

Capstone Project Submission

Instructions:

- i) Please fill in all the required information.
- ii) Avoid grammatical errors.

Team Member's Name, Email and Contribution:

1. Bhaskar Subanji

Email- bysubanji@gmail.com

- Data Understanding
- Feature Analysis
 - States
 - Area Codes
 - Call Data(night)
 - Churn
- Feature Engineering
 - Duplicate Value
- Data Visualization
 - Box plot
 - Dist plot
 - Crosstab
 - Count plot
 - Distplot
- Research Analytics
 - Technical documentation

2. Harshjyot Singh Chawla

Email- hs9158695878@gmail.com

- Data Understanding
- Feature Analysis
 - Calls Data(day, evening)
 - International Calls
 - International Charges
 - Customer Service Calls
- Feature Engineering
 - Null Value Check
- Data Visualization
 - Heat Map
 - Box plot
 - Correlation Matrix
 - Scatter plot
 - Count plot
- Research Analytics
 - Technical documentation

3. Ajinkya Dakhale

Email- ajinkya.dakhale2408@gmail.com

- Data Understanding
- Feature Analysis
 - Account Length
 - International Plan
 - Voice Mail Plan
 - Number Vmail Messages
- Feature Engineering
 - Missing Value
 - Duplicate Value
- Data Visualization
 - Pie chart
 - Box plot
 - Dist plot
 - Crosstab
 - Count plot
- Research Analytics

Please paste the GitHub Repo link.

Github Link:- <https://github.com/bysubanji/Exploratory-Data-Analysis-of-churn-dataset>

Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)

PROBLEM

Orange S.A is a French MNC telecommunications corporation with many different customer subscribed to its services and key task is to find the causes that influence customers to churn and suggesting the factors for retention of customers.

In this EDA project, we were provided with Orange's Telecom churn dataset which consists of various details about area, account, calls and minutes etc.

APPROACH

- Initially, we imported the data set and carried out the descriptive analysis over the data set to understand the information of data available.
- Checking for missing and duplication of values in the data set provided.
- Analyzing the all the variables of the data set (such as Area-code, Account, Plan, Calls, Minutes, Charges etc) with respect to churn data, to determine the possibilities and factors for Churn.
- Used data visualization with different kind plots to explore the correlation with churn and different variables.
- Predicted and explored some factors that are influencing customer to churn.

CONCLUSION

- Some states have higher churn rate compared to other, for which network issues might be the reason so it is recommended that company should increase the coverage area and solve the network problems in local as well as international level.
- Area, account length, evening & night calls has no relation with churn rate but the churn rate increases as the number of call to the customer service center increases so its evident that company must provide good customer service and resolution to customers to prevent churn by gathering feedback.
- Customers with international plan has higher churn rate compared to customers without international plan the reason could be high tariff cost or network issues so company should work on these factors and provide lower or competitive international plans and coverage.
- Customers with voicemails more than 20 (approx.) and customers with higher day call min's has higher churn rate compared to others so company should focus on providing discounts or offers, also they can offer customized service with different plans based on customer needs.