Capstone Project Submission

Instructions:

i) Please fill in all the required information.

ii) Avoid grammatical errors.

Team Member's Name, Email and Contribution:

1. 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

2. 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

- Technical documentation

3. 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

Please paste the GitHub Repo link.

Github Link:- https://github.com/Harshjyot-Singh-Chawla/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 STATEMENT:

French multinational telecommunications corporation, Orange S.A. Telecom's Churn Dataset has been evaluated to find out the reasons for a customer to churn. The dataset contains customer activity data as well as churn variables stating retained customers. Factors responsible for churning, variables correlated to churning, recognizing customers that might churn will be analyzed in this EDA project.

APPROACH:

- The first step involves data wrangling to have a clean data set for effortless access and analysis.
- In the second step, we try to make sense of the dataset by dividing variables into small modules and analyzing each module with help of different plots. We performed univariate and bivariate analysis to get a better understanding of the relation shared between each variables.
- The First module involves analysis of churned customers based on their state

and area code.

- The Second module involves analysis based on plans (international and voice mail) belonging to churned customers.
- The Third module contains an evaluation based on the correlation between calls, minutes, charges, and churned customers.
- The Last module involves analysis based on customer service calls attended by churned customers.

CONCLUSION:

- Some states have a higher churn rate as compared to other states. So the company must look out for any network issues faced by the customer in those states.
- Customers having international plan churn more in comparison to customers not having an international plan. The reason could be a high tariff.
- Churn rate increases as the voice mails exchanged between customers exceed 20.
- Customers with high day call minutes churn more. Spending more minutes on call results in more charges resulting in more bill charges.
- Customer with four or more customer service calls churn more. So the company should improve on their services and work on providing accurate solutions to the problem faced by the customer.
- Variables such as area, evening calls, night calls have no relation with churn rate and can be omitted.