Capstone Project Submission

Instructions:

i) Please fill in all the required information.

ii) Avoid grammatical errors.

Team Member's Name, Email and Contribution:

1. 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

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. 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/Ajinkya-dak/Exploratory-Data-Analysis-of-telecom-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:

Orange S.A., formerly France Telecom 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 cancelled the subscription.

Explore and analyze the data (consisting of 3333 customers) to discover key factors responsible for customer churn and come up with ways/recommendations to ensure customer retention.

APPROACH:

- We begin with by performing data wrangling on given dataset to have a clean data set for proper access and analysis.
- we analyzed the churn data(i.e. churn column) consisting of customer those who churned and those who didn't.
- next step was to perform univariate as well as bivariate analysis on each column label(variable), in biavariate analysis one variable was common which is churn column, we also performed some pandas function to get more insights for each column of the data set .
- further, to have more clarity we performed multvariate consisting of variable using heat map.
- we also calculated the charges for each kind of call (i.e. Day, night ,evening, international) to view if the pricing of call led to churn.

CONCLUSION:

- Some states have higher churn rate than other, for which network issues could the reason because if the competitor company had low tariff for calls then most of the states would have shown the appprox. same churn rate.

- Area and Account length has no relation with churn rate, hence this columns can be omitted or it can be said that the data is redundant.
- Customers with international plan ON has higher churn rate compared to customers with international plan OFF , this could be because the customer could be unhappy with the high tariff cost or network issues.
- It could been seen that customers with vmails more than 20 (approx.) has higher churn rate.
- Customers with higher day call mins has higher churn rate compared to other , could be because of the higher charges which is quite obvious, frequent caller might have found some other company offering low tariff .
- With other variables such as evening ,night calls no relation could be found.
- The churn rate increases as the call to the service center increases. Customers who have called customer service three or fewer times have a markedly lower churn rate than that of customers who have called customer service four or more times.