**Capstone Project Submission**

**Instructions:**

i) Please fill in all the required information.

ii) Avoid grammatical errors.

|  |
| --- |
| **Team Member’s Name, Email and Contribution:** |
| Mohd Muttalib  [muttalib1326@gmail.com](mailto:muttalib1326@gmail.com) |
| **Please paste the GitHub Repo link.** |
| Github Link:- <https://github.com/MMuttalib1326/Telecom-Churn-Analysis---Capstone-Project./tree/main> |
| **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****In this project ,we are working on the customer churn dataset of Orange S.A., formerly known as France Télécom S.A. here we are analyzing customer churn and we will be doing a complete EDA process to determine if the customer from Orange S.A. company will cancelled the subscription of the telecom service or not. we will also draw some insights from data visualization and analysis so that we could get the factors which are affect the output(churn of the customer)** **Tabel of content**  * **Step 1: Importing the Relevant Libraries** * **Step 2: Data Inspection** * **Step 3: Data Cleaning** * **Step 4: Exploratory Data Analysis** * **Observing and Exploring Dataset**   After loading the dataset we performed some basic functions and methods for knowing the data type. This process helped us figuring out various aspects and relationships among the users and the data factors. It gave us a better idea of which factor behaves in which manner compared to the churner.  After observing the data we would say that there are 20 columns and 3333 rows. According to churn column there are 86% non churned and 14% churned users. We can definitely make some suggestions for churned users for their retention.   * **Null, Missing and Duplicate Values Treatment**   It is an important aspect of Data Cleaning because there can be some null, missing and duplicate values in our dataset. But our dataset doesn't contains a null or missing values which might tend to disturb our accuracy, if it has null or missing values then we have to drop them at the beginning of our project in order to get a better results.   * **Dropping unneccessary Column**   Our dataset does contains some unnecessary column like Account length etc. we didn't need them for analysis so we can drop them out from dataset.   * **Encoding of categorical columns**   We changed no and yes to 0 and 1 to encode our categorical column because categorical column that are in string format cannot be understood by the machine and needs to be converted to numerical format for visualization and good analysis.   * **Analyzing and Visualization of data**   In these steps we used plots like count plot to check the results of each churned customer and we used box plot for observing which column is more important compared to churned column and which is of less importance.  Next we used bar plot for finding top states in which customer have churned a lot.We are using pie plot and cat plot also as well as heat map for correlation.   * **Using different visualization methods**   For visualization we tried various plots like:   * **Count Plot** * **Bar Plot** * **Pie Plot** * **Box Plot** * **Cat Plot** * **Heat map**  **Conclusion**In this project ,we tried to analyze customer churn. First we did inspection of dataset on a basic level.We looked for missing values and check the outlier.Then we used the matplotlib and seaborn to do Exploratory Data Analysis on sample data by plotting different graphs like count plot, pie chart,lmplot,barplot,boxplot, subplot and heat map from this we got useful insights like: customer having more daily charge will be more chances of churn, states like New Jersey,Texas and Maryland have higher churn rate, customer having international plan have more churn rate, customer having less customer service call have more churn rate and 14% percentage of customer who has been churnedHere is our suggestion to prevent churn :1. Upgrading network to improve services for long duration users.2. Improving Pricing Strategies.3. Optimizing and Updating International Call Rates.4. Implmenting a better network infrastructure in New Jersey,Texas and Maryland Areas where there is more Churn Rate.6. Improvement in the customer service can be done to reduce the number of calls which cause the churn7. Decreasing the prices as the talk-time increases can be an effective way to reduce the churn. |