Capstone Project Submission

Instructions:

i) Please fill in all the required information.

ii) Avoid grammatical errors.

Team Member's Name, Email and Contribution:
Pooja Rathore (pooja.rathore090297@gmail.com) • Data Wrangling • Frequency distribution of Total day charge • Frequency distribution of Total evening charge • Frequency distribution of Total night charge • Churn distribution across the states • States having lowest churn rate • Area code wise churn distribution • International plan is positively or negatively affect the churning rate • Voice mail plan across the churn • Total number of customer service call is affect the churn rate • Overall minutes is how much affect the overall charges • Find out the correlation between all the features • Determine the highly correlated variables
Please paste the GitHub Repo link.
Github Link:- https://github.com/Pooja-2709/Capstone-Project1---Telecom-Churn-Analysis Drive Link:- https://colab.research.google.com/drive/1D6wQzJlE7QujGk_Iy2OIH1HpCP7zm5Tj?usp=sharing
Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)

France Telecom S.A. is a telecommunication company which is trying to identify the customer churn. The goal is to identify customers churn, that is, customers most likely to cancel subscription to a fictitious Telecom company. That is a really interesting problem, because if you could predict in advance which customers are at risk of leaving, you could reduce customer retention efforts by directing them toward such customers, providing adequate intervention to encourage them to stay and minimize customer exit. In this EDA project I was provided the csv file of Telecom which consists 3333 rows and 20 columns.

The first step we took was to import necessary libraries of Exploratory data analysis like numpy, pandas, malplotlib, seaborn. After that we look forward to Data Wrangling for data cleaning, identify and removing null and duplicated values in the give data set, changing data types, handling outliers.

Now the data is ready to discover the key understanding like Univariant analysis of total day charge, total evening charge, total night charge, total international charges, after that we move on to the Bivariant analysis means looking all the features to the churn column, identify the overall time spending by the customer and accordingly overall charges and how much both are associated with each other. Finally we saw the correlation between all the features and identify the particular features which are highly correlated.

Through EDA of this dataset we brought many conclusions and key understanding of consumer perception towards the services and many more which will help the company while taking futuristic decision.