Project 2

Project Title: Exploring Customer Churn in a Telecom Company

Project Description: A telecom company is interested in exploring customer churn, which is the rate at which customers leave their service. The company has provided a dataset with information on customer demographics, services subscribed to, and usage patterns, among other things. The goal of this project is to analyze this dataset, perform data cleaning and preprocessing, conduct feature selection and extraction, and visualize the data in order to identify factors that may be driving customer churn.

Tasks:

1. Load the dataset into a Pandas DataFrame.
2. Conduct an initial exploratory data analysis of the dataset, including calculating summary statistics, visualizing distributions of variables, and identifying missing values.
3. Preprocess the data by cleaning missing values, converting categorical variables to numerical form, and scaling variables as necessary (HINT: In this stage first Fill in missing values with column means and then #drop customer id column and then complete what is asked at part 3).
4. Conduct feature selection and extraction, including identifying correlated Variables.
5. Visualize the dataset

    i. Create a correlation matrix to explore the relationship between variables
    ii. Use a heatmap to visualize the correlation matrix
    iii. Create a bar chart to show the count of customers who churned and those who did not
    iv. Use a pie chart to show the percentage of customers who churned and those who did not