In this File, "Churn_Analysis.ipynb," provides an exploratory data analysis (EDA) of a customer churn dataset. The analysis focuses on understanding the factors influencing customer churn through data preprocessing, visualization, and the calculation of relevant percentages.

1. Data Preprocessing:

- **Loading Data:** The analysis begins by loading a CSV file named 'cust churn.csv' into a pandas DataFrame.
- **Data Cleaning:**
  - Missing values (blanks) in the "TotalCharges" column were identified and replaced with '0'.
  - The "TotalCharges" column was converted to a float data type for numerical analysis.
  - A check for null values across the entire dataset confirmed no missing values after the initial cleaning.
  - No duplicate rows were found in the dataset.
- **Feature Transformation:** The "SeniorCitizen" column, initially represented as '0's and '1's, was transformed to 'yes' and 'no' for better interpretability in visualizations.

2. Churn Overview:

- A count plot of the "Churn" column reveals the distribution of churning and non-churning customers.
- A pie chart further illustrates the churn percentage:
  - Approximately **73.46%** of customers did **not churn**.
  - Approximately **26.54%** of customers **did churn**.

3. Churn Analysis by Demographics and Services:

- **Gender:** A stacked bar chart of "Churn by Gender" indicates that churn rates are relatively similar between male and female customers, with no significant gender-based disparity observed.
- **Senior Citizen Status:** A stacked bar chart of "Churn by SeniorCitizen" shows a clear difference:
  - Among non-senior citizens, **3,638** customers did not churn, while **1,393** churned.
  - Among senior citizens, **1,009** customers did not churn, while **476** churned.
  - This suggests that **senior citizens have a higher propensity to churn** compared to non-senior citizens when considering the proportions of churn within each group.
- **Tenure:** A histogram showing "tenure" distribution by churn status highlights that:
  - Customers with **shorter tenures** exhibit a **higher churn rate**.

    - As tenure increases, the number of non-churning customers significantly outweighs churning customers.
- **Contract Type:** A stacked bar chart of "Count of Customer by Contract Type and Churn" indicates a strong correlation:
    - **Month-to-month contract** customers show the **highest churn rate** (1,655 churned vs. 2,220 not churned).
    - Customers on **one-year contracts** have a much lower churn rate (1,307 not churned vs. 166 churned).
    - **Two-year contract** customers exhibit the **lowest churn rate** (1,647 not churned vs. 48 churned), suggesting long-term contracts are effective in retaining customers.
- **Payment Method:** The "Churned Cust ny PaymentMethod" bar chart reveals:
    - **Electronic check** is associated with the **highest number of churned customers** (1,071), despite also having a large number of non-churned customers (1,294).
    - **Credit card (automatic)** and **Bank transfer (automatic)** methods show significantly lower churn counts (232 and 258 respectively) compared to Electronic check and Mailed check (308).
    - This indicates that customers using **electronic checks are more likely to churn**.
- **Other Services (PhoneService, MultipleLines, InternetService, OnlineSecurity, OnlineBackup, DeviceProtection, TechSupport, StreamingTV, StreamingMovies):** A 3x3 subplot grid visualizes churn distribution across various services. While specific numbers are not labeled, the visual trend indicates that:
    - Services like **Online Security, Online Backup, Device Protection, and Tech Support** (where customers tend to have "No" for these services) are generally associated with a **higher churn propensity**.
    - This suggests that customers without these additional security or protection features are more likely to churn.

4. Correlation Analysis (Numerical Features):

- A heatmap of the correlation matrix for numerical features (likely tenure, MonthlyCharges, and TotalCharges) is presented.
- While the exact correlation values are not explicitly stated in the summary, the heatmap visually represents the strength and direction of linear relationships between these numerical variables. Strong positive correlations are often observed between tenure and TotalCharges, and MonthlyCharges and TotalCharges.

Overall Conclusion:

The analysis highlights several key drivers of customer churn. Customers on **month-to-month contracts**, those using **electronic checks**, and non-senior

citizens (though the rate is proportionally higher for senior citizens), as well as customers lacking additional security and protection services, demonstrate a higher likelihood of churning. Conversely, longer contract terms and automatic payment methods (credit card and bank transfer) are associated with higher customer retention. These insights provide valuable information for developing targeted retention strategies.