

Exploratory Data Analysis (EDA) Report on Customer Churn

Objective:

to understand the factors contributing to customer churn at a telecommunication company. By analyzing the churn data, I aim to provide insights that will help the business improve customer retention strategies.

Dataset Description:

The dataset contains information on 10,000 customers of a commercial bank. Key variables in the dataset include customer ID, Age, Tenure, gender, subscription type, number of support calls, usage frequency, contract length, and churn status.

Data Cleaning and Preprocessing:

Ensured all variables are of the correct data type. For instance, 'CustomerID' is a categorical variable, 'age' and 'total spend' are numerical variables, and 'churn' is a numerical value for integrated purposes but dealt with and interpreted as a binary categorical variable.

Churn Analysis

- **Churn Rate:**

The overall churn rate in the dataset is **56.7%**, meaning **5,670 out of 10,000** customers have churned.

- **Churn Rate by Segments:**

Gender: The churn rate for **male** customers is **49.13%**, whereas for **female** customers it is **66.67%**.

Age Groups: Customers aged 50-65 years have the highest churn rate at 100%.

SubscriptionTypes: Customers with **premium** plans have a churn rate of **48.20%**. For **basic** plans, the churn rate is **50.73%** for customers with account balances between 1k-10k.

Key Factors Contributing to Churn:

Tenure: Customers with shorter tenures (less than 12 months) tend to churn more than those with longer tenure in both genders.

Support calls: Customers with more than 4 support calls tend to churn more than those who have 0-4 calls only.

Visualization

Line Charts:

1. Finance related: Line chart showing the relation between the total amount of money spent by the customer and the churn rate. Another one showing the trend between the delay of payment by the customer and the churn rate.
2. Support and interaction related: Line chart showing the number of support calls that the customer engaged in and the churn rate. In addition, there is a line chart showing the last interaction time against the churn rate.

Bar Charts and Pie Charts:

1. Bar charts were used to display the distribution of churn across different segments such as tenure, gender and age group.
2. Pie charts provided a clear visualization of the proportion of premium vs. basic vs. standard customers' churn rate.

Insights and Recommendations

Enhanced Support: Increase support and engagement efforts for inactive customers to reactivate them and reduce churn. Additional analysis into the reasons behind the support calls leading to churn. Potential reasons

include inaccessibility, unprofessional behavior by representatives, recurring problems.

Product Improvements: Investigate reasons for high churn rates in basic plans and make necessary improvements or offer an upgrade to the current basic plan customers.

Loyalty programs: Offering the customers with longer tenures incentives using a reward system. This would also motivate customers with shorter tenures to stay to receive those incentives later on.