

Customer Churn EDA Report

1. Introduction

This report presents an Exploratory Data Analysis (EDA) of a customer churn dataset from a telecom provider. The dataset includes various customer attributes such as demographic information, services subscribed, and billing information. The main objective is to understand customer behavior and identify factors contributing to churn. Churn is a key performance indicator for businesses and refers to customers who have stopped using the service. By exploring the data, we aim to identify correlations, visualize important trends, and draw insights to improve customer retention.

2. Churn Distribution

The dataset contains 7,032 entries in total, with 1,869 customers who churned and 5,163 who did not. This corresponds to an overall churn rate of 26.59%. The class imbalance is notable and should be considered in further modeling stages.

3. Numerical Feature Summary

The dataset includes numerical features such as 'tenure', 'MonthlyCharges', and 'TotalCharges'. Key statistics:

- Tenure ranges from 0 to 72 months, with an average of 32.37 months.
- Monthly charges vary from \$18.25 to \$118.75, with a mean of \$64.76.
- Total charges range widely, suggesting long-term customers with high cumulative charges.

These features show expected variability and have notable relationships with churn.

4. Correlation with Churn

The correlation analysis reveals the following:

- MonthlyCharges has a weak positive correlation (0.19) with churn.

Customer Churn EDA Report

- TotalCharges and tenure have moderate negative correlations with churn, indicating that longer and high-value customers are less likely to churn.

These insights will be crucial for developing churn prevention strategies.

5. Churn by Categorical Features

Categorical features like contract type, internet service, and payment method show strong patterns:

- Customers with month-to-month contracts have much higher churn rates than those on one- or two-year contracts.
- Customers using electronic check payment methods have a churn rate of over 45%.
- Fiber optic internet users and those without dependents are also more prone to churn.

These patterns help segment customers and prioritize retention efforts.

6. Conclusion

The EDA highlights the key drivers of churn and segments that require attention. Understanding these insights helps target interventions effectively. This report forms the basis for feature engineering and model development in subsequent phases.