Customer Churn Analysis – Technical Report

1. Introduction

Customer churn is a critical challenge for subscription-based and service-oriented businesses. This technical report documents an end-to-end data analysis project conducted to understand customer churn behavior, identify its key drivers, and support data-driven business decisions.

2. Business Problem Statement

The business is experiencing customer attrition, which negatively impacts revenue and growth. The objective is to analyze historical customer data to determine factors influencing churn and to recommend actionable retention strategies based on analytical findings.

3. Project Objectives

The primary objectives of this project are: • Measure overall customer churn rate • Identify key factors contributing to churn • Segment customers based on churn risk • Provide actionable, data-driven business recommendations

4. Dataset Description

The dataset used in this project contains approximately 500 customer records. Each record represents a unique customer and includes both numerical and categorical attributes, along with a churn indicator representing whether the customer exited the service.

5. Data Cleaning and Preparation

The raw dataset underwent multiple data preparation steps to ensure quality and consistency. Column names were standardized, duplicate records were removed, and missing values were handled using median imputation for numerical variables and mode imputation for categorical variables. The churn variable, originally encoded numerically, was converted into categorical labels to improve interpretability.

6. Exploratory Data Analysis (EDA)

Exploratory Data Analysis was performed to uncover patterns, trends, and relationships within the data. Visualizations such as bar charts, box plots, count plots, and correlation heatmaps were used. EDA revealed noticeable differences in feature distributions between churned and retained customers, highlighting potential churn drivers.

7. Advanced Analysis and Statistical Validation

Advanced analysis techniques were applied to validate churn drivers. Segment-wise churn rates were calculated to identify high-risk customer groups. Correlation analysis was conducted to quantify relationships between numerical features and churn. These techniques provided statistical support for observed EDA patterns.

8. High-Risk Customer Profiling

Customers identified as churned were analyzed separately to construct a high-risk customer profile. Descriptive statistics of this group revealed common characteristics that differentiate them from retained customers. This profiling enables targeted and proactive retention strategies.

9. Key Insights

• Customer churn is significantly influenced by specific customer attributes • Certain customer segments exhibit disproportionately higher churn rates • Numerical variables show consistent differences between churned and retained users • Analytical findings align with business intuition and can guide strategic decisions

10. Business Recommendations

Based on analytical findings, the following recommendations are proposed: • Deploy targeted retention campaigns for high-risk segments • Implement early warning systems to identify churn-prone customers • Enhance customer engagement and loyalty initiatives • Monitor churn-related KPIs regularly for continuous improvement

11. Limitations

This analysis is based on historical data and does not account for future behavioral changes. The dataset size, while sufficient for analysis, limits the complexity of predictive modeling. External factors influencing churn were not captured in the dataset.

12. Conclusion

This project demonstrates a complete data analysis workflow from raw data to actionable insights. The findings provide valuable guidance for reducing customer churn and improving retention. Future work may include predictive modeling and real-time churn monitoring systems.