PROJECT REPORT

Customer Churn Prediction – Exploratory Data Analysis (EDA)

1. Abstract

Customer churn refers to the loss of customers over a period of time and is a critical challenge for businesses,
especially in subscription-based industries. This project focuses on performing Exploratory Data Analysis (EDA)
on a customer dataset to identify patterns, trends, and factors influencing customer churn. The analysis helps
businesses understand customer behavior and enables data-driven decision-making to reduce churn.

2. Introduction

In today’s competitive market, retaining existing customers is more cost-effective than acquiring new ones. Customer churn directly impacts revenue and long-term business growth.
The goal of this project is to analyze customer data and identify key attributes that influence churn behavior using Exploratory Data Analysis (EDA) techniques.

EDA helps in:

Understanding data structure

Detecting missing values and outliers

Identifying relationships between features

Gaining insights before applying machine learning models

This project focuses only on EDA, not on predictive modeling.

3. Problem Statement

Businesses face difficulty in identifying why customers leave their services. Without understanding the contributing factors, retention strategies become ineffective.

This project aims to analyze customer data to answer the following questions:

- What percentage of customers are churning?
- Which customer segments are more likely to churn?
- Which services and attributes are strongly related to churn?

4. Objectives of the Study

- Analyze customer demographic data
- Study churn distribution
- Identify key features influencing churn
- Visualize churn relationships
- Provide actionable insights

5. Dataset Description

The dataset consists of customer-related information such as:

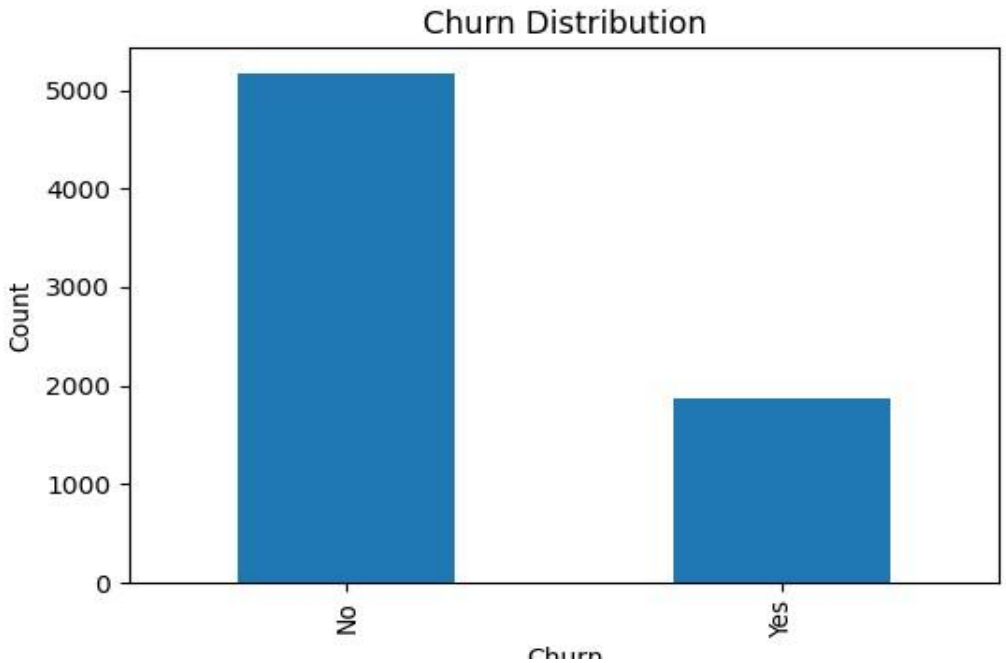
- Demographic details (gender, senior citizen status)
- Account information (tenure, contract type, payment method)
- Service usage (internet service, phone service, add-on services)
- Target variable: **Churn** (Yes / No)

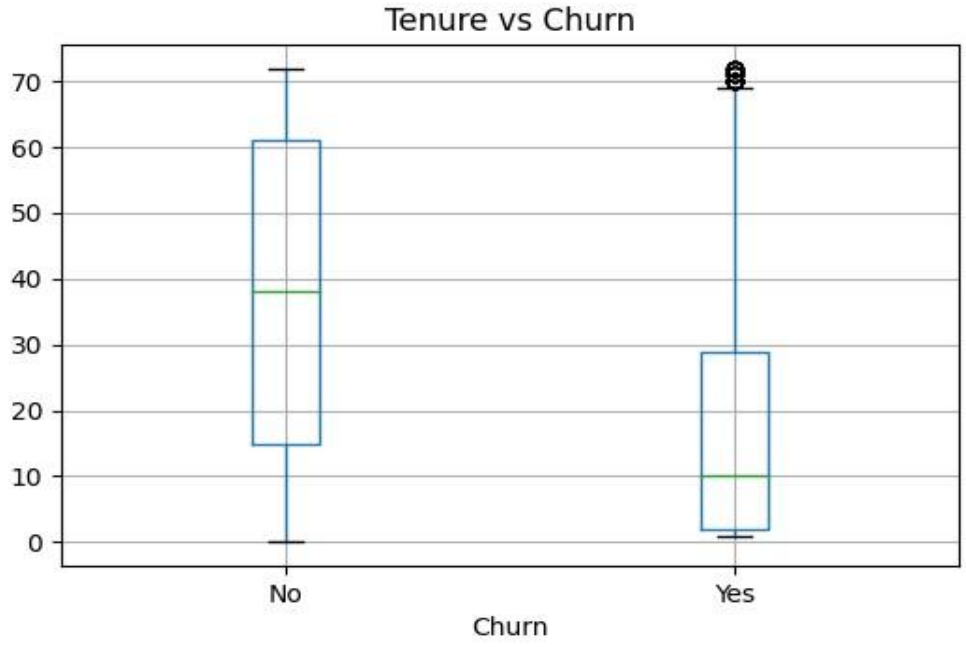
6. Methodology

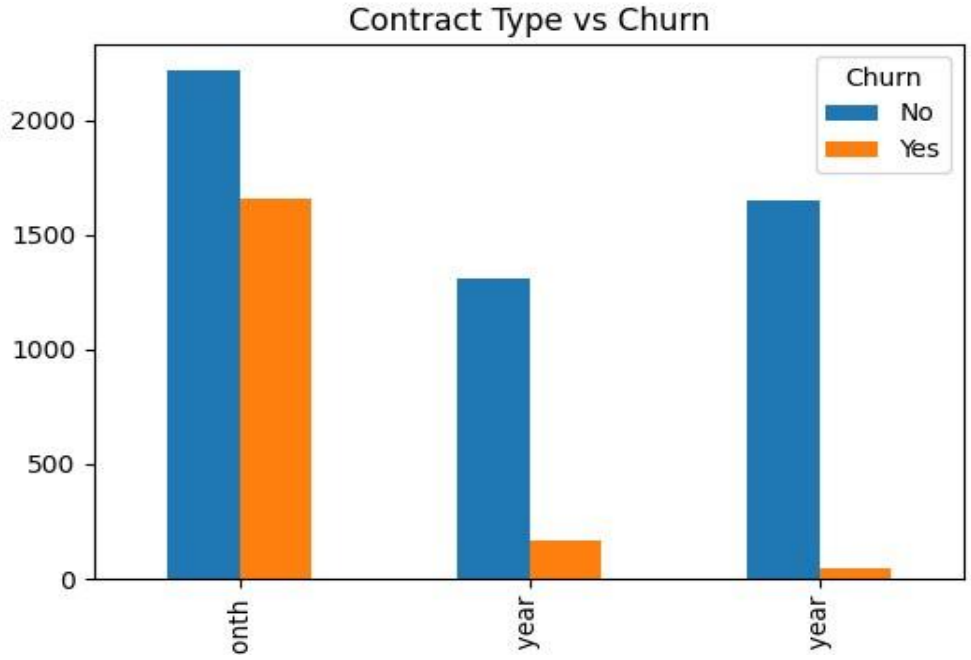
The following steps were followed in the project:

1. **Data Loading**
    - Dataset imported using Python libraries
2. **Data Cleaning**
    - Checked for missing values
    - Converted categorical variables
3. **Data Understanding**
    - Shape, data types, and summary statistics
4. **Univariate Analysis**
    - Distribution of individual features
5. **Bivariate Analysis**
    - Relationship between churn and other variables
6. **Visualization**
    - Bar charts, histograms, box plots, and count plots
7. **Insight Generation**
    - Key observations from analysis

7. Exploratory Data Analysis (EDA)






Contract Type vs Churn
Churn
No
Yes
2000
1500
1000
500
0
onth
year
year

8. Key Findings

- Churn strongly influenced by tenure and contract type
- Short-term customers are at higher risk
- Automated payments reduce churn
- Add-on services improve retention

9. Business Insights

- Encourage long-term contracts
- Offer early-stage incentives
- Promote add-on services
- Improve electronic payment experience

10. Tools & Technologies Used

Programming Language: Python
Libraries: Pandas, NumPy, Matplotlib,
Seaborn Environment: Jupyter Notebook

11. Limitations of the Study

- Limited dataset
- No predictive modeling
- External factors not included

12. Future Scope

- Implement ML models
- Feature engineering
- Real-time churn dashboards
- CRM integration

13. Conclusion

This project successfully analyzed customer churn data using Exploratory Data Analysis techniques. The insights obtained highlight key factors affecting customer retention. Businesses can leverage these findings to design effective churn reduction strategies and improve customer satisfaction.

14. References

- Customer Churn Analysis – Industry Reports
- Pandas & Seaborn Documentation
- Data Analytics Case Studies