Customer Churn Analysis Using Telco Customer Churn Dataset

Name: Kartik S. Rathod

Tools Used: Excel, SQL(postgresql), Python(Pandas,Seaborn,Matplotlib), Power BI

Date: October 2025

Project Summary:

The 'Customer Churn Analysis' project explores customer behavior within a telecom company to understand the factors influencing service discontinuation. Using the Telco Customer Churn dataset, the project applies an end-to-end analytics approach—beginning with data cleaning in Excel, followed by data validation and aggregation in SQL, EDA in Python, and dashboard visualization in Power BI. The insights reveal that customers with month-to-month contracts, electronic check payments, and shorter tenures are more likely to churn. These findings can help the company implement targeted retention strategies and enhance customer satisfaction.

Problem Statement & Objective:

Customer churn is one of the most critical challenges faced by telecommunication companies. Losing existing customers not only impacts revenue but also increases marketing costs to acquire new ones. The company wants to identify the key factors that influence customer churn and build a data-driven strategy to predict which customers are most likely to leave.

This project aims to analyze customer demographics, account information, and service usage patterns to uncover insights into churn behavior and help the company take proactive retention actions.

Objectives

- Clean and prepare the dataset using Excel for accurate analysis.
- Perform data validation and aggregation using SQL to ensure consistency.
- Conduct Exploratory Data Analysis (EDA) in Python to identify churn patterns and key influencing factors.
- Create a Power BI dashboard to visualize insights effectively.
- Provide data-driven recommendations to reduce churn and improve customer satisfaction.

Dataset Overview:

The Telco Customer Churn Dataset contains information about customers of a telecom company and their subscription details. The dataset helps analyze the factors that influence whether a customer continues or discontinues their service.

Total Records: 7,043

Total Columns: 22

Missing Values: None

Data Structure:

Numerical Features: tenure, MonthlyCharges, TotalCharges, AvgChargesPerMonth.

Categorical Features: customerID, gender, SeniorCitizen, Partner, Dependents, PhoneService, MultipleLines, InternetService, OnlineSecurity, OnlineBackup, DeviceProtection, TechSupport, StreamingTV, StreamingMovies, Contract, PaperlessBilling, PaymentMethod, Churn.

Key Statistical Insights:

- Average tenure: 32 months
- Average Monthly Charges: ₹64.76
- Average Total Charges: ₹2279.73
- Average AvgChargesPerMonth: ₹64.70
- Churned customers: 26.5%
- Non-churned customers: 73.5%
- Churn Rate : 26.54%

Data Processing and Visualization Workflow:

Data Cleaning and Preprocessing With Excel:
The dataset was first imported and reviewed to ensure its quality and readiness for analysis.

Steps performed:

- Converted SeniorCitizen column values from 0/1 to No/Yes.
- Standardized all categorical fields to ensure consistent labels.
- Created a new column AvgChargesPerMonth to calculate average monthly billing.
- Checked and verified the dataset for null or blank values.
- Validated that all numerical columns contain only numeric values and are in the correct format.

- Cleaned formatting and made column names consistent and readable

⊞ Data Validation and Aggregation With SQL:

- Created a structured SQL table schema for the CustomerChurn_Dataset with appropriate data types and primary key.
- Validated total record count using COUNT(customerID) to ensure complete data load.
- Performed category-wise churn checks using SUM(CASE…) logic to validate churn distribution.
- Computed overall churn count and churn rate from base table to validate against EDA numbers.
- Aggregated churn statistics by gender, partner status, dependents, internet service, contract type, and payment method.
- Created tenure-based buckets (0–1 year, 1–2 years, 2–4 years, 4+ years) and calculated churn rate per tenure group.
- Generated grouped aggregation outputs to be used directly in Power BI dashboard visuals.
- Ensured SQL logic correctness with CASE, GROUP BY, and rounded percentage measures for consistency with final reporting

- **Exploratory Data Analysis With Python:**
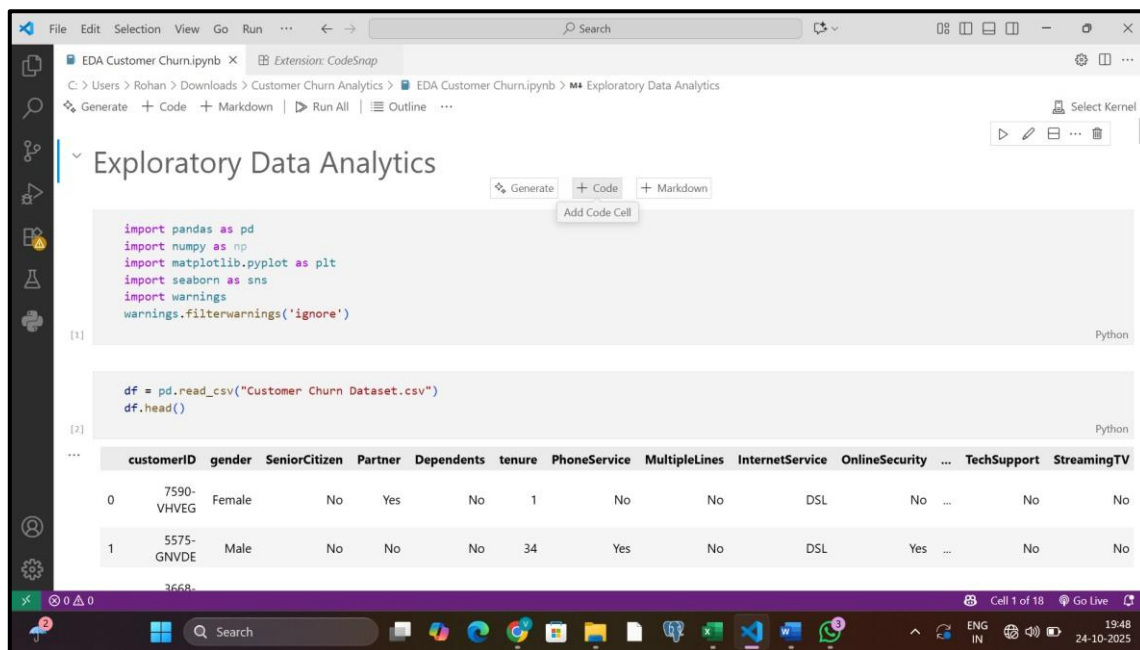
Installation of the essential liberaries and reading csv.

- Imported required Python libraries (Pandas, NumPy, Matplotlib, Seaborn) and loaded the cleaned churn dataset.
- Inspected dataset shape, structure, and column data types to understand the data schema.
- Examined basic statistics (distribution of tenure, charges, etc.) using .describe() and visual summaries.
- Analyzed overall churn rate to understand the proportion of customers who churned vs retained.
- Performed univariate EDA on key categorical variables (Gender, SeniorCitizen, Partner, Contract, PaymentMethod, InternetService) using countplots/pie charts.
- Performed churn comparison using grouped bar charts to identify categories with higher churn tendency.
- Analyzed relationship between churn and key numeric variables (MonthlyCharges, TotalCharges, Tenure) using histograms and boxplots.

- Created tenure buckets to examine churn trend across customer lifecycle segments.
- Derived insights from correlation and visual patterns to support further dashboard and recommendation steps

🔡 Data Visualization With Power BI:
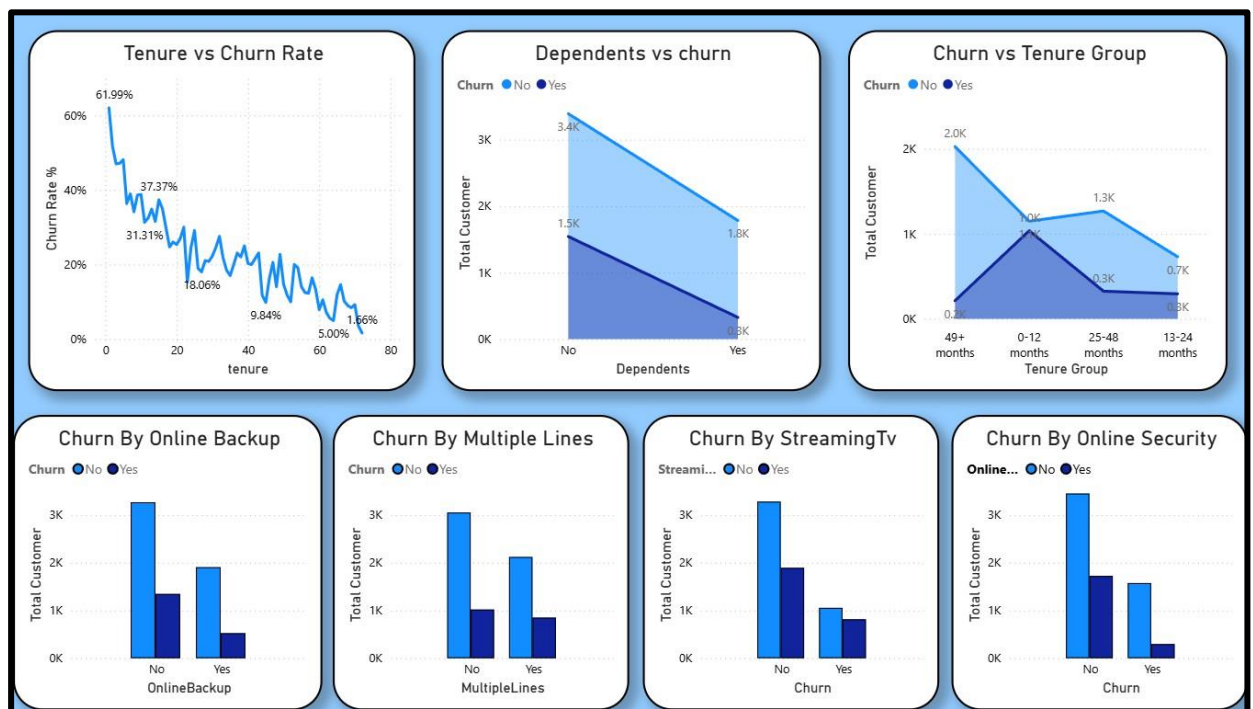
- **Customer Churn Overview :**
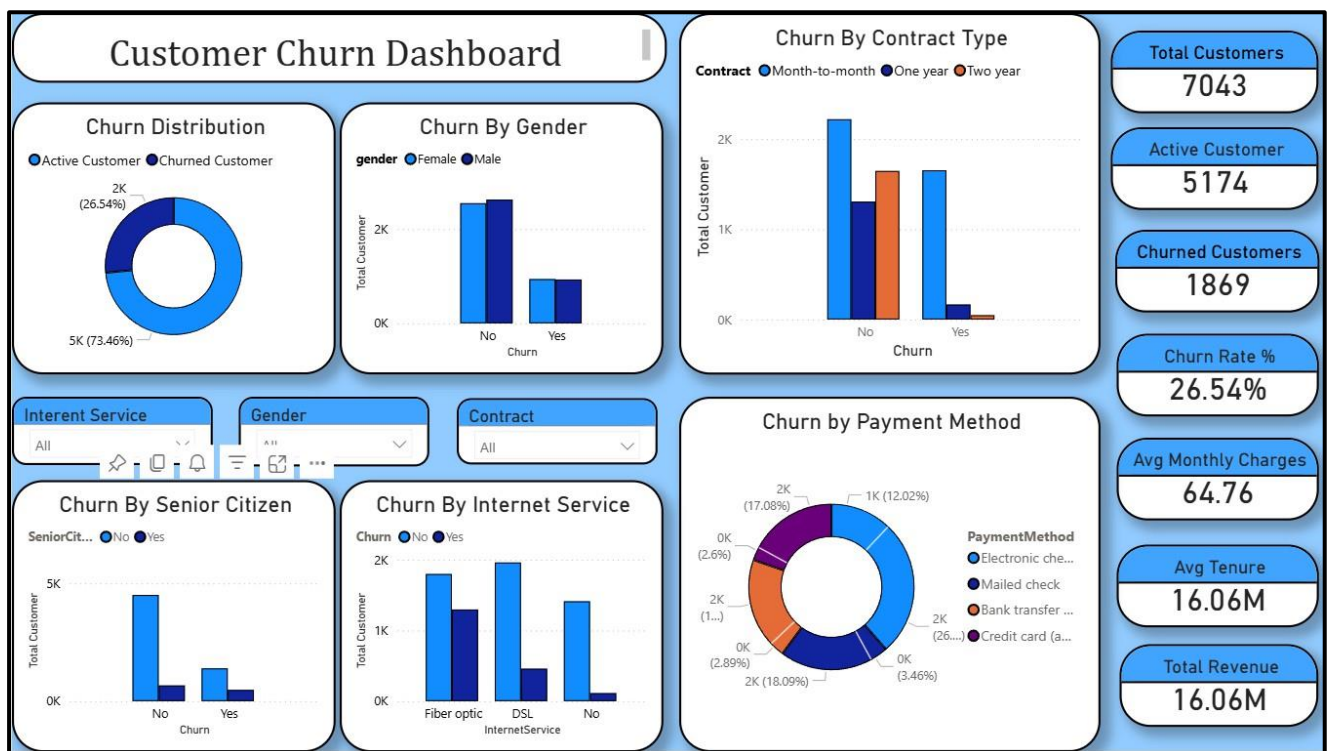
Insights:

- Out of 7,043 total customers, 26.54% (1,869) have churned, showing a moderate churn rate.
- Male and female customers show a similar churn pattern, indicating gender has minimal impact on churn.
- Senior citizens have a higher churn rate compared to non-senior citizens, suggesting they are less satisfied or face more service issues.

- Customers using fiber optic internet have the highest churn rate, followed by DSL, while those without internet service churn the least.
- Month-to-month contract customers churn the most, whereas one-year and two-year contract holders show better retention.
- Customers paying via electronic check churn the most, while those using credit card or bank transfer remain more loyal.
- The dashboard metrics show 7,043 total customers, 5,174 active, 1,869 churned, a churn rate of 26.54%, an average tenure of 16.06 months, average monthly charges of $64.17, and total revenue of 16.06M, reflecting moderate retention and good revenue stability.

- **Detailed Insights:**

Insights:

- Customers without online backup services have a noticeably higher churn rate compared to those who use it, suggesting that online backup helps improve retention.
- Customers without multiple lines churn more frequently, while those having multiple lines are more loyal and likely to stay longer.
- Streaming TV users show a higher churn rate compared to non-users, indicating that customers using entertainment services are more likely to leave.
- Customers without online security services churn significantly more than those who have it, implying that added security features improve customer satisfaction and retention.
- The tenure vs churn rate visual shows churn is highest among new customers (low tenure) and decreases as tenure increases, meaning loyal, long-term customers are less likely to leave.
- The dependents vs churn chart indicates customers without dependents churn more than those with dependents, showing that family customers tend to stay longer.
- The churn vs tenure group visual reveals that churn is highest in the 0–12 month group and drops sharply for customers with longer tenure (49+ months), emphasizing the importance of early engagement and retention strategies.

Recommendation:

- Encourage long-term contracts by offering discounts or loyalty rewards, as customers on month-to-month plans have the highest churn rate.
- Improve service quality and customer experience for Fiber Optic Internet users, since they represent the largest group of churned customers.
- Target new customers (0–12 months tenure) with engagement and onboarding programs, as early churn is significantly higher.
- Offer bundled plans (internet + security + backup) to increase stickiness; customers with online security and backup services tend to stay longer.
- Provide special assistance or senior-friendly plans for senior citizens, who show higher churn than younger users.
- Review and simplify payment methods, particularly for electronic check users, who show higher churn rates compared to other payment options.
- Develop personalized retention campaigns for high-churn segments, such as single-line users or those without dependents.
- Continuously monitor churn metrics through dashboards to identify early warning signs and take proactive actions.

Conclusion:

The Telco Customer Churn analysis reveals that around 26–27% of customers have churned, indicating moderate retention performance. The key churn drivers include short-term contracts, fiber optic service dissatisfaction, lack of security or backup services, and low customer tenure.

Long-tenure customers, those using annual contracts, and those with value-added services show stronger loyalty. Focusing on improving the early customer experience, offering better long-term deals, and maintaining service reliability can significantly reduce churn rates.

Overall, strategic retention initiatives and data-driven decision-making can help the company enhance customer satisfaction, improve loyalty, and increase long-term revenue.

Project Repository:

You can explore the complete project files, including datasets, SQL queries, EDA notebooks, and Power BI dashboard, at the following link:

https://github.com/KartikRathod2003/Customer-Churn-Analysis