📊 Customer Segmentation Visualization & Advanced Analysis Project

---

Project Overview

The project aims to analyze customer churn in a telecommunications company and develop predictive insights to identify at-risk customers. The ultimate goal is to provide actionable insights and data-driven recommendations that help reduce customer churn and improve long-term customer retention.

Customer churn is a major challenge in the telecom industry due to intense competition and low switching costs. By analyzing customer demographics, service usage, billing behavior, and contract details, this project helps uncover the key factors influencing churn and supports strategic decision-making.

Objectives

The primary objectives of this project are:

- To understand the structure and quality of the customer churn dataset
- To clean and prepare data for accurate analysis
- To perform exploratory data analysis (EDA) to identify trends and distributions
- To segment customers based on tenure and lifecycle stages
- To analyze churn patterns across demographics, payment methods, and contract types
- To visualize key insights for clear interpretation and business communication

---

Tools & Technologies Used

- **Programming Language:** Python
- **Libraries:**
    - Pandas – Data manipulation and analysis
    - Matplotlib – Static data visualization
    - Seaborn – Statistical data visualization
    - Plotly – Interactive visualizations

Dataset Description

The dataset contains customer-level data from a telecommunications company. Each record represents an individual customer along with their service usage, billing information, and churn status.

Key Features

- **Demographic Variables:** Gender, Senior Citizen status
- **Account Information:** Tenure, Contract type, Payment method
- **Billing Information:** Monthly charges
- **Target Variable:** Churn (Yes / No)

Task 1: Understanding the Dataset

The dataset was loaded using Pandas and examined to understand its structure. The first few rows were inspected to identify the nature of the variables. Data types of each column were reviewed to differentiate between numerical and categorical features. Missing values were checked to identify potential data quality issues.

This step ensured a strong foundation for further analysis by providing clarity on data composition and potential preprocessing needs.

Task 2: Data Cleaning

Data cleaning was performed to ensure the dataset was suitable for analysis. Missing values were handled by removing incomplete records to maintain consistency. Duplicate records were removed to avoid biased analysis.

Column names were standardized by converting them to lowercase and replacing spaces with underscores. This improved readability and made the dataset easier to work with during analysis and visualization.

Task 3: Exploratory Data Analysis (EDA)

Exploratory Data Analysis was conducted to understand distributions, detect anomalies, and uncover initial patterns.

Statistical Analysis

- Mean, median, and mode were calculated for numerical variables
- Summary statistics provided insights into data spread and central tendency

Visual Analysis

- **Histograms** were used to analyze distributions of tenure and monthly charges
- **Box plots** helped identify spread and potential outliers
- **Churn distribution plots** showed the proportion of churned vs retained customers

Key Observations

- Most customers are non-senior citizens
- A large portion of customers have low tenure
- Monthly charges vary significantly across customers
- Approximately one-fourth of customers have churned

Task 4: Customer Segmentation Visualization

Customers were segmented based on tenure to analyze behavior across lifecycle stages.

Tenure Segments

- 0–12 months (New customers)
- 13–36 months (Mid-term customers)
- 37+ months (Long-term customers)

Visualizations

- **Donut chart:** Showed customer distribution across tenure groups
- **Clustered bar chart:** Compared average monthly charges across tenure segments

Insights

- Most customers fall into the early tenure group
- Long-term customers tend to have higher average monthly charges
- Retention increases as tenure increases

Task 5: Advanced Analysis

Advanced analysis focused on understanding churn behavior across different customer dimensions.

Tenure-Based Analysis

- New customers (0–12 months) have the highest churn rates
- Long-tenure customers show strong retention

Demographic Analysis

- Senior citizens exhibit slightly higher churn rates
- Gender shows minimal impact on churn

Payment Method Analysis

- Customers using electronic checks show higher churn
- Automatic payment methods are associated with better retention

Contract Type Analysis

- Month-to-month contracts have the highest churn
- Long-term contracts significantly reduce churn

Lifecycle Visualization

Bar charts highlighted churn trends across customer lifecycle stages, clearly showing decreasing churn with increasing tenure.

Key Business Insights

- Customer churn is highest during the first year
- Contract type is a strong indicator of churn
- Payment methods influence customer retention
- Long-term customers generate higher revenue and show lower churn

---

Recommendations

- Focus retention efforts on customers within the first 12 months
- Encourage long-term contracts through incentives
- Promote automatic payment methods
- Design targeted engagement programs for senior citizens
- Improve onboarding experience for new customers

Conclusion

This project successfully analyzed customer churn using statistical analysis, segmentation, and visualization techniques. By identifying high-risk customer segments and key churn drivers, the analysis provides actionable insights to improve customer retention.

The results demonstrate how data-driven decision-making can help telecommunications companies reduce churn, improve customer lifetime value, and strengthen long-term business performance.

Skills Demonstrated

- Data cleaning and preprocessing
- Exploratory data analysis
- Customer segmentation
- Statistical and comparative analysis
- Data visualization and storytelling