Telecom Customer Churn Analysis

By

1- Eslam Mohamed Adly Hegazy

2- Samah Soliman Mahmoud Mohamed

3- Abdulrahman Mohamed Reda Amin

4- Ibrahim Reda Ibrahim Eldesouky

5- Ahmed Samir Mohamed Elrefae

Supervisor

Dr. Amal Mahmoud

2025,R3

Abstract

This project examines the Telecom Customer Churn dataset to identify the main factors that lead customers to leave telecom services.
By cleaning, exploring, and visualizing the data using Excel, MySQL, Power BI, and Tableau, the analysis uncovers key relationships between churn and factors such as contract type, payment method, tenure, and service usage.
The results include dashboards and reports that highlight overall churn rates, the most influential churn drivers, and high-risk customer segments—providing clear insights that support better retention strategies and improved customer satisfaction.

Index

Introduction

Customer churn, or the rate at which customers stop using a company's services, is a critical challenge in the telecommunications industry. With increasing competition and similar service offerings, retaining customers has become just as important—if not more—than acquiring new ones. High churn rates can significantly reduce a company's profitability and market share, making it essential to understand the factors that drive customers to leave.

This project focuses on analyzing customer data from the Telco Customer Churn dataset, which contains information about customers' demographics, subscription types, billing details, and service usage. By examining these factors, the analysis aims to uncover patterns and insights that help explain why some customers remain loyal while others decide to discontinue their service.

Through structured data analysis and visualization, this project provides a comprehensive view of churn behavior, highlighting which customer segments are most at risk and what business factors contribute to their decisions. The ultimate goal is to support data-driven decision-making that enables telecom providers to enhance customer satisfaction, reduce churn rates, and improve long-term business performance.

Objective

The objective of this project is to analyze customer data to know high risk customers to Churn using statistical analysis.This analysis helps businesses to:

- **Identify key factors influencing customer churn**, such as pricing, service quality, and subscription duration.
- **Generate data-driven insights** and churn indicators to identify high-risk customer segments based on data trends and patterns.
- **Improve customer retention strategies** by offering personalized deals and reducing churn rates.

Methodology

Dataset Description

This dataset contains information about 7043 customers of a telecom company based in California and includes 24 columns covering a wide range of customer attributes. The goal is to analyze customer retention and identify high-risk customers likely to leave (churn). The dataset provides diverse demographic, billing, and service usage features that help in understanding the factors influencing customer behavior.

- **CustomerID**: Unique identifier for each customer.
- **Gender**: Customer's gender (Male / Female).
- **Age**: Customer's age.
- **Tenure**: Number of months the customer has been subscribed.
- **Contract Type**: Type of contract (Monthly, One year, Two years).
- **MonthlyCharges**: Monthly billing amount.
- **TotalCharges**: Total amount paid by the customer.
- **InternetService**: Type of internet service (DSL, Fiber optic, None).
- **PaymentMethod**: Payment method (Credit card, Bank transfer, Check).
- **Churn**: Target variable (Yes if the customer left, No if they stayed).

Source:https://www.kaggle.com/datasets/beatafaron/telco-customer-churn-realistic-customer-feedback

The project follows these steps to analyze customer churn data:

1. **Data Understanding**
    - Review the dataset structure, including the number of rows, columns, and data types.
    - Identify key features such as demographic details, service usage, billing information, and churn status.
2. **Data Cleaning & Preprocessing**
    - Check for missing or inconsistent values and handle them appropriately.

    - Remove duplicates if any exist.
    - Ensure categorical and numerical data are correctly formatted.
3. **Exploratory Data Analysis (EDA)**
    - Analyze the distribution of churn vs. non-churn customers.
    - Examine relationships between churn and factors like contract type, tenure, monthly charges, and payment method.
    - Use visualizations (bar charts, histograms, pie charts, heatmaps) to identify patterns and trends.
4. **Insights & Interpretation**
    - Summarize key findings from the analysis.
    - Highlight the most influential factors contributing to customer churn.
    - Provide recommendations for reducing churn based on observed patterns.

Tools

This project utilizes four main tools—Excel, MySQL, Power BI, and Tableau—each serving a specific purpose in the data analysis workflow:

1. Excel

- Used for initial data cleaning and formatting.
- Handled missing values, removed duplicates, corrected data types, and prepared the dataset for deeper analysis.
- Performed basic statistical summaries and preliminary exploratory analysis.

2. MySQL

- Used to store the dataset in a structured relational format.
- Enabled efficient querying and filtering using SQL commands.
- Performed aggregations, joins, and calculations to uncover patterns such as churn rates by contract type, payment method, and tenure groups.

3. Power BI

- Used to build interactive dashboards that show key performance indicators (KPIs).
- Enabled dynamic filtering to analyze different customer segments and churn distributions.

4. Tableau

- Used for advanced visual storytelling and deep visual exploration.
- Helped uncover hidden relationships and patterns in churn behavior through interactive charts and filters.

Method

The method outlines how the tools and techniques were applied throughout the project to achieve clear and accurate churn analysis:

1. Data Integration

- Imported the cleaned dataset into Excel, MySQL, Power BI, and Tableau.

- Ensured consistency in column names, formats, and data types across all platforms.

2. Querying & Statistical Analysis

- Used SQL in MySQL to perform statistical calculations like churn percentages, tenure averages, and revenue metrics.
- Conducted cross-tab analysis to understand relationships between churn and variables such as contract duration and service type.

Contract Type vs. Economic Value and Churn
the strong correlation between commitment (contract) and total value/low risk.

Contract	avg_monthly_charge	avg_total_charges	avg_tenure_months	churn_rate
Two year	60.87	3728.93	57.1	2.85
One year	65.08	3034.68	42.1	11.28
Month-to-month	66.40	1369.25	18.0	42.71

Correlation : Churn Rate by Tenure Group
Measures the strength of the inverse relationship between loyalty (tenure) and churn

Tenure_Risk_Group	total_customers	avg_tenure_months	churn_rate
0-1 Year (High Risk)	2175	4.8	47.68
1-2 Years (Medium Risk)	1024	18.4	28.71
2-4 Years (Lower Risk)	1594	36.2	20.39
4+ Years (Low Risk)	2239	63.0	9.51

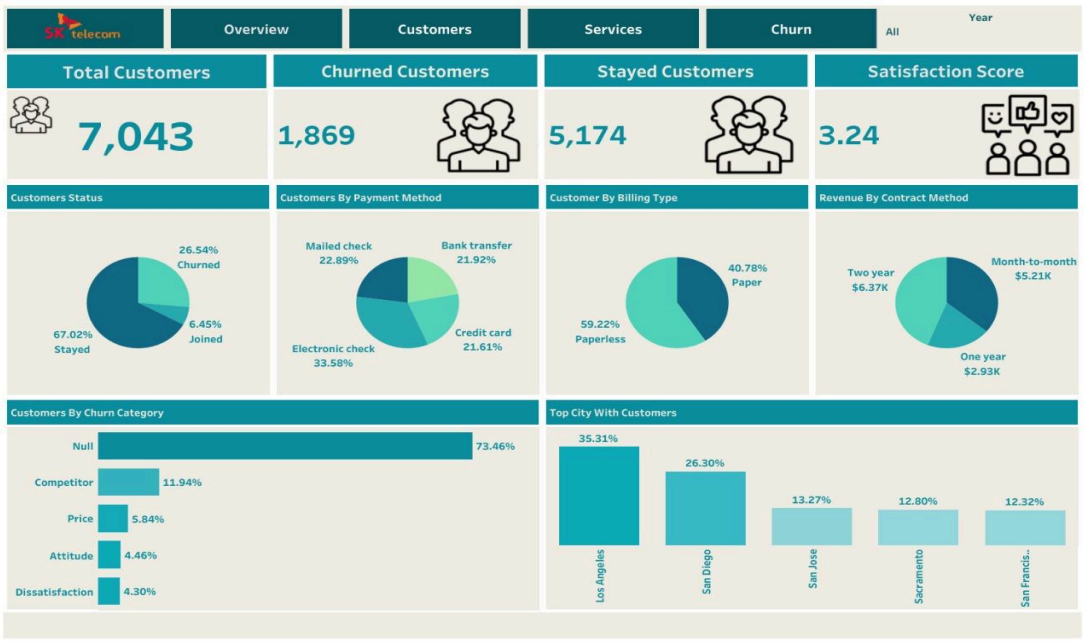
the relationship between price sensitivity and customer churn.

Charge_Bracket	total_customers	avg_charge	churn_rate
Low Charge (Less than $30)	1647	21.50	9.84
Medium Charge ($30 - $70)	1804	53.82	24.39
High Charge (More than $70)	3581	90.24	35.38

3. Visualization & Dashboard Building

- Designed dashboards in Power BI and Tableau to present insights clearly and interactively.

- Applied filtering, drill-downs, and segmentation to highlight high-risk customers and key churn drivers.



4. Insight Extraction

- Compared results across tools to ensure analytical consistency.
- Identified patterns such as higher churn among month-to-month contracts and customers with higher monthly charges.
- Formulated data-driven recommendations for improving retention and customer satisfaction.

Results and Discussion

The analysis revealed several key insights regarding customer churn:

- **Churn Rate**: Approximately **26.5%** of customers in the dataset have churned.
- **Contract Type Impact**: Customers with **month-to-month contracts** show the highest churn rate compared to those with one-year or two-year contracts.
- **Tenure Effect**: Customers with shorter tenure (less than 12 months) are more likely to churn.
- **Billing and Charges**: Higher monthly charges correlate with increased churn, especially among customers without long-term contracts.
- **Internet Service**: Fiber optic users exhibit a higher churn rate compared to DSL users.
- **Payment Method**: Electronic check users have the highest churn rate, suggesting possible dissatisfaction or payment issues.

Discussion:
These findings indicate that churn is strongly associated with contract flexibility, service type, and billing patterns. Short-term customers and those paying higher monthly fees are more likely to leave, which suggests that pricing strategies and contract incentives could significantly reduce churn.

Recommendations

Based on the analysis, the following actions are recommended:

1. **Promote Long-Term Contracts**
   Offer discounts or benefits for customers who switch from month-to-month to annual contracts.
2. **Improve Customer Experience for Fiber Users**
   Investigate service quality issues for fiber optic customers and address complaints promptly.
3. **Review Pricing Strategy**
   Consider introducing tiered pricing or loyalty discounts for customers with high monthly charges.
4. **Enhance Payment Options**
   Provide incentives for customers to switch from electronic checks to more reliable payment methods (credit card, auto-pay).
5. **Target At-Risk Customers**
   Use churn indicators (short tenure, high charges, month-to-month contracts) to proactively engage with customers through personalized offers.

References

- Kaggle: Telco Customer Churn Dataset
- IBM Sample Data: Telco Customer Churn
- Research articles on churn analysis and retention strategies.

Appendix

Some SQL Code to show

Total Churn rate

```
SELECT
    COUNT(*) AS total_customers,
    SUM(CASE WHEN Churn = 'Yes' THEN 1 ELSE 0 END) AS churned_customers,
    ROUND(SUM(CASE WHEN Churn = 'Yes' THEN 1 ELSE 0 END) * 100.0 / COUNT(*), 2) AS
churn_rate_percentage
FROM telecom_churn;
```

Churn by(Contract)

```
SELECT
    Contract,
    COUNT(*) AS total,
    SUM(CASE WHEN Churn = 'Yes' THEN 1 ELSE 0 END) AS churned,
    ROUND(SUM(CASE WHEN Churn = 'Yes' THEN 1 ELSE 0 END) * 100.0 / COUNT(*), 2)
AS churn_rate
FROM telecom_churn
GROUP BY Contract
ORDER BY churn_rate DESC;
```

Churn by Tenure group

```
SELECT
    CASE
        WHEN Tenure <= 12 THEN '0-1 year'
        WHEN Tenure <= 24 THEN '1-2 years'
        WHEN Tenure <= 48 THEN '2-4 years'
        WHEN Tenure <= 72 THEN '4-6 years'
        ELSE '+6years'
    END AS tenure_group,
    COUNT(*) AS total,
    SUM(CASE WHEN Churn = 'Yes' THEN 1 ELSE 0 END) AS churned,
    ROUND(SUM(CASE WHEN Churn = 'Yes' THEN 1 ELSE 0 END) * 100.0 / COUNT(*), 2) AS
churn_rate
FROM telecom_churn
GROUP BY tenure_group
ORDER BY MIN(Tenure);
```

Churn Rate by Internet Service and Monthly Charge Bracket

This helps identify which internet services are high-risk at high charge levels.

```
SELECT
    InternetService,
    CASE
        WHEN MonthlyCharges < 50 THEN 'Low Charge'
        WHEN MonthlyCharges BETWEEN 50 AND 100 THEN 'Medium Charge'
        ELSE 'High Charge'
    END AS charge_bracket,
    COUNT(customerID) AS total_customers,
    ROUND(AVG(CASE WHEN Churn = 'Yes' THEN 1 ELSE 0 END) * 100, 2) AS churn_rate
FROM
    telecom_churn
GROUP BY
    InternetService, charge_bracket
ORDER BY
    InternetService, churn_rate DESC;
```

Revenue Churn Rate (Percentage of Total Revenue Lost)

Measures the proportion of total revenue (TotalCharges) that comes from churned customers.

```
SELECT
    ROUND(
        SUM(CASE WHEN Churn = 'Yes' THEN TotalCharges ELSE 0 END) / SUM(TotalCharges) *
100,
        2
    ) AS revenue_churn_percentage
FROM
    telecom_churn
WHERE TotalCharges > 0;
```

Customers with High Risk Profile (High Charge, Short Tenure, No Protection)
Identifies specific customers who fit the high-risk criteria for targeted retention efforts.

```
SELECT
    customerID,
    gender,
```

```
    tenure,
    MonthlyCharges,
    InternetService,
    TechSupport,
    TotalCharges
FROM
    telecom_churn
WHERE
    Churn = 'No' -- We look for customers who HAVEN'T churned yet, but are high risk
    AND tenure <= 6 -- Short tenure
    AND MonthlyCharges > 90 -- High monthly charge
    AND Contract = 'Month-to-month' -- Risky contract
    AND TechSupport = 'No'
LIMIT 20;
```

Churn Rate by High-Risk Segment (Month-to-month AND No Tech Support)
This highlights the churn rate for the customer segment with the lowest commitment and lowest protection.

```
SELECT
    Contract,
    TechSupport,
    COUNT(customerID) AS total_customers,
    SUM(CASE WHEN Churn = 'Yes' THEN 1 ELSE 0 END) AS churned_customers,
    ROUND(AVG(CASE WHEN Churn = 'Yes' THEN 1 ELSE 0 END) * 100, 2) AS churn_rate
FROM
    telecom_churn
WHERE
    Contract = 'Month-to-month'
GROUP BY
    TechSupport
ORDER BY
    churn_rate DESC;
```