EDA Report: Telco Customer Churn

1. Introduction

Customer churn refers to when customers stop doing business with a company. In the telecommunications industry, identifying factors that lead to churn is crucial for improving customer retention strategies.

This report provides an in-depth exploratory data analysis of the Telco Customer Churn dataset , aiming to uncover patterns, correlations, and insights that can be used to build a predictive model.

2. Dataset Overview

- Source : IBM Watson Analytics Lab

- Size : 7,043 rows, 21 columns

- Target Variable : **Churn** (Yes/No)

- Data Types :

    - Categorical : Gender, Partner, Dependents, InternetService, Contract, PaymentMethod, etc.

    - Numerical : MonthlyCharges, TotalCharges, Tenure

    - Binary : SeniorCitizen, PhoneService, PaperlessBilling, etc.

Key Features:

FEATURE	DESCRIPTION
CUSTOMERID	Unique identifier for each customer
GENDER,SENIORCITIZEN, PARTNER, DEPENDENTS	Demographic information
TENURE	Number of months the customer has stayed with the com
PHONESERVICE,MULTIPLELINES,INTERNETSERVICE	Service-related features
ONLINESECURITY,TECHSUPPORT, STREAMINGTV, ETC.	Additional service subscriptions

CONTRACT,PAPERLESSBILLING,PAYMENTMETHOD	Account & billing details
MONTHLYCHARGES,TOTALCHARGES	Financial metrics
CHURN	Target variable – whether the customer left (Yes/No)

3. Data Cleaning Summary

❖ **Missing Values :**
- ▪ **TotalCharges:**
  - • -Some entries were blank (). Replaced with MonthlyCharges values.
  - • -No missing values after cleaning.

❖ **Data Type Corrections :**
- ➢ TotalCharges was originally read as object → converted to numeric using pd.to_numeric().

❖ **Dropped Columns :**
- ➢ customerID: Not relevant for modeling.

4. Exploratory Data Analysis (EDA)

❖ **Numerical Features Analysis**
- ➢ **Tenure**
  - ▪ Most customers leave within the first few months.
  - ▪ Indicates early-stage churn issues.
- ➢ **Monthly Charges**
  - ▪ Customers paying more than $70/month show the highest churn rate.
  - ▪ Suggests dissatisfaction among high-value customers.
- ➢ **Total Charges**
  - ▪ Distribution skewed towards lower values.
  - ▪ Correlated with tenure and monthly charges.

❖ Categorical Features vs Churn

FEATURE	INSIGHT
CONTRACT TYPE	Month-to-month contracts have the highest churn rate. Long-term contracts (One year, Two year) significantly reduce churn.
PAYMENT METHOD	Electronic check users have the highest churn. Credit card and bank transfer users are more loyal.
PAPERLESS BILLING	Customers with paperless billing are more likely to churn.
INTERNET SERVICE	Fiber optic users have higher churn compared to DSL users. Possibly due to higher cost or poor service quality.
ADDITIONAL SERVICES	Lack of services like Online Security, Tech Support, or Device Protection increases churn likelihood. Bundling these services may improve retention.

Demographic Insights

FEATURE	INSIGHT
GENDER	No significant difference in churn between male and female customers
SENIOR CITIZEN	Senior citizens (16% of total) show slightly higher churn rates. May require special retention programs.
PARTNER / DEPENDENTS	Customers without partners or dependents are more likely to churn. Lonelier customers may be less committed to long-term plans.

5. Visualizations Highlights

- **Churn Rate by Monthly Charges Category**

MONTHLY CHARGE RANGE	CHURN RATE
LOW ($0–$30)	~15%
MID ($30–$70)	~30%
HIGH (> $70)	~55%

Insight : High-paying customers are at the highest risk of churn.

- **Contract Type vs Churn**

CONTRACT TYPE	CHURN RATE
MONTH-TO-MONTH	~50%
ONE YEAR	~10%
TWO YEAR	~5%

Insight : Encouraging month-to-month customers to upgrade contracts could reduce churn.

- **Payment Method vs Churn**

PAYMENT METHOD	CHURN RATE
ELECTRONIC CHECK	~35%
MAILED CHECK	~20%
BANK TRANSFER	~15%
CREDIT CARD	~15%

Insight : Electronic check users are at higher risk of leaving.

6. Feature Engineering

The following feature engineering steps were performed:

Binning Monthly Charges :

- Created categories: 'Low', 'Mid', 'High'.

Encoding Categorical Variables :

- Binary variables encoded as 0/1.
- Multi-class variables one-hot encoded (e.g., InternetService, Contract, PaymentMethod).

7. Key Findings

- Churn Rate : Approximately 27% of customers have churned.
- Month-to-month contracts are highly correlated with churn.
- High monthly charges increase churn risk.
- Lack of additional services (e.g., no online security) correlates with higher churn.

- Electronic check payment method is associated with higher churn.
- Demographics play a minor role compared to service and account features.

8. Recommendations

- Promote Longer Contracts : Offer incentives to switch from month-to-month to one or two-year contracts.
- Bundle Services : Encourage adoption of multiple premium services to improve retention.
- Target High-Paying Customers : Introduce loyalty programs or personalized discounts for high-value users.
- Improve Experience for Electronic Check Users : Investigate why this group churns more frequently.
- Early Retention Programs : Focus on customers in the first few months of their contract.