Exploratory Data Analysis (EDA) Summary Report – Telco Customer Churn

# 1. Introduction

This report provides an exploratory data analysis (EDA) of the uploaded Telco Customer Churn dataset to assess data quality, uncover patterns, and highlight factors associated with churn risk. The goal is to ensure data readiness for modeling and decision-making.

# 2. Dataset Overview

**Records:** 7043  
**Columns:** 22  
**Numeric columns:** 5; **Categorical columns:** 17  
**Overall churn rate:** 26.54%

# 3. Missing Data Analysis

We examined missing values across all columns. The attached table ('Missing Values by Column') details counts and percentages. Columns with substantial missingness should be imputed thoughtfully (e.g., median for numeric; most frequent or separate level for categorical).

# 4. Key Findings and Risk Indicators

Top numeric features correlated with churn (positive implies higher churn with higher values):

• MonthlyCharges: correlation 0.193

• SeniorCitizen: correlation 0.151

• TotalCharges: correlation -0.199

• tenure: correlation -0.352

# 5. AI & GenAI Usage

GenAI tools were used to summarize patterns, detect missing values, and highlight features most associated with churn. Findings were cross-checked with standard telco churn heuristics (e.g., higher churn among month-to-month contracts, higher charges with lower perceived value).

# 6. Conclusion & Next Steps

• Overall churn rate observed: 26.54%.

• Top numeric correlation with churn: MonthlyCharges (0.193).

• Recommended next steps: feature engineering (e.g., tenure buckets, service bundles), robust imputation for missing values, and stratified modeling with explainability (e.g., SHAP) to validate drivers.