Customer Churn Analysis - End-to-End Data Analytics Case Study

Overview & Business Objective

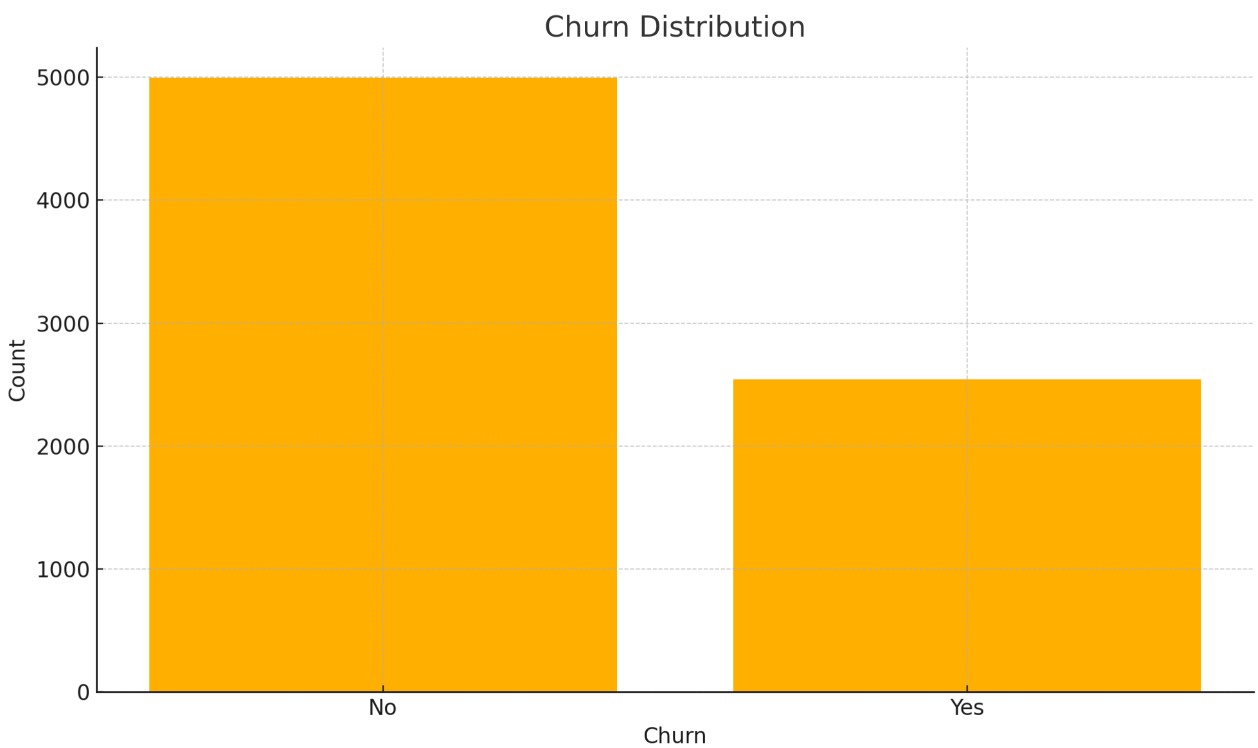
This case study investigates telecom customer churn to diagnose why customers leave and how to reduce attrition. We transform raw customer, service, and billing records into insights that guide retention strategy-identifying high-risk cohorts, quantifying price sensitivity, and assessing the effect of contractual commitment.

Dataset Description

Raw dataset: 7537 rows x 23 columns (data/raw/telecom_customer_data.csv)
Processed dataset: 7500 rows x 33 columns (data/processed/final_data.csv)
Key fields include: tenure (customer longevity), MonthlyCharges (recurring bill), TotalCharges, Contract type, InternetService and add-ons (e.g., security, backup), and the Churn label (Yes/No). Processed data standardizes types and encodes categories.

Data Cleaning & Preparation

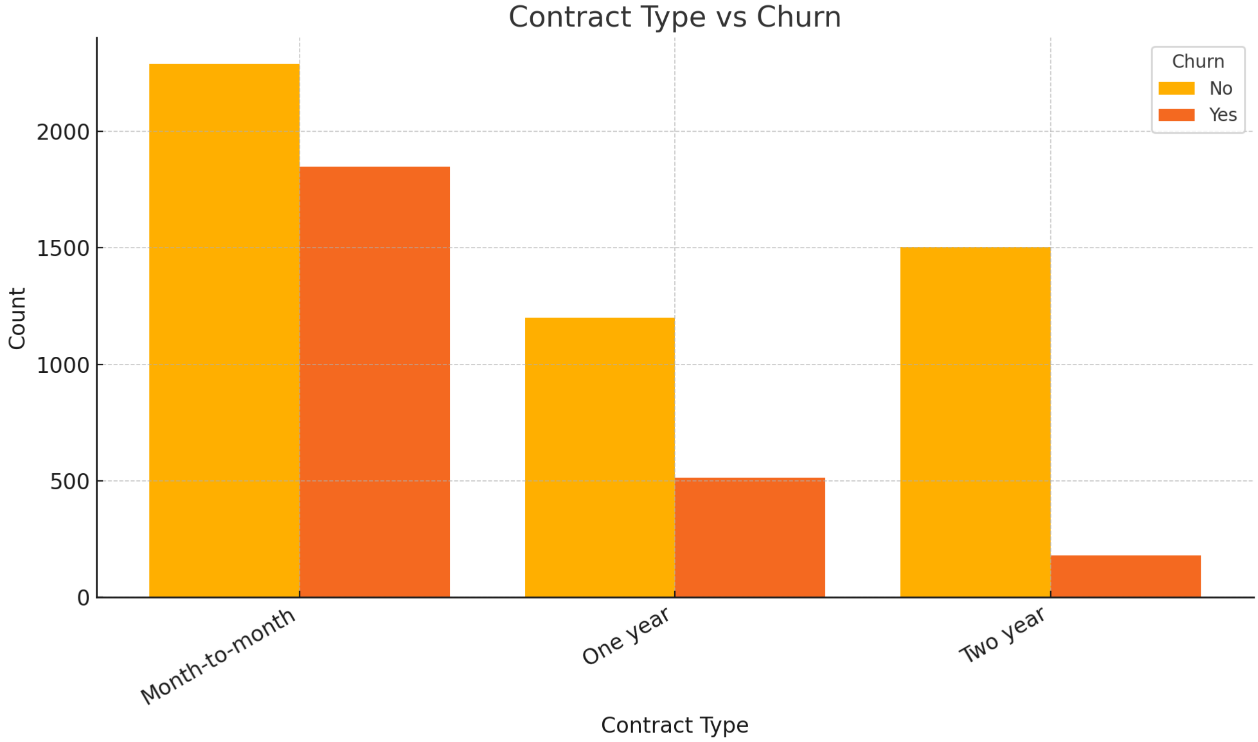
We addressed missing values (e.g., coercing numeric fields and imputing when appropriate), corrected data types, normalized category labels, and removed clear inconsistencies. These steps stabilize downstream analysis and produce reliable metrics that leadership can trust.



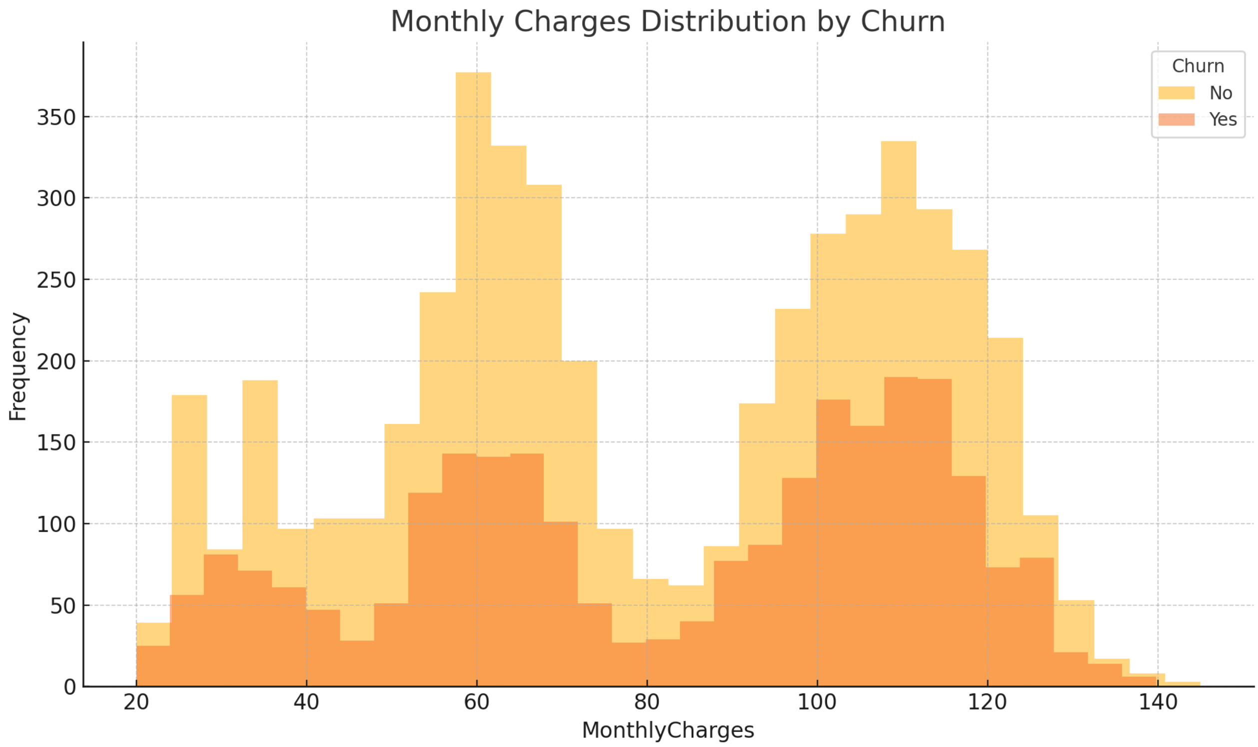
Churn distribution: establishes baseline class balance and informs evaluation strategy.

Exploratory Data Analysis (EDA)

EDA focuses on how contractual commitment, pricing, and tenure relate to churn, surfacing patterns that translate into action.



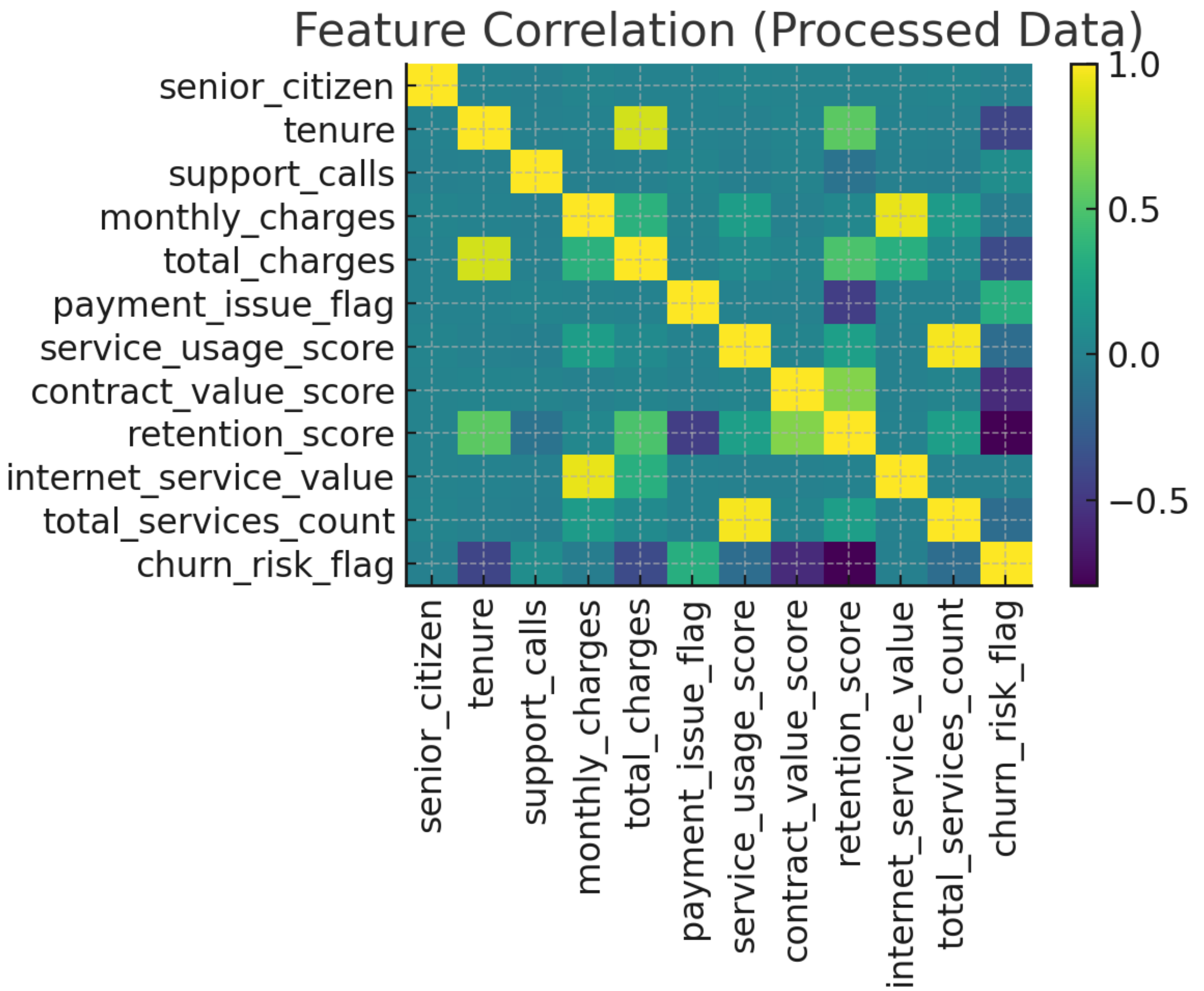
Contract type vs churn: month-to-month contracts show higher churn than longer-term agreements.



Monthly charges by churn: churned customers concentrate at higher price points, suggesting price sensitivity.

Feature Engineering & Correlation Insights

We prepared model-ready features through encoding, potential tenure banding, and interactions (e.g., pricing by contract type). Correlation analysis among numeric features highlights relationships to manage during modeling (e.g., multicollinearity).



Correlation across processed numerical features: informs variable selection and regularization choices.

Key Findings & Business Recommendations

Findings:

- Month-to-month customers are the most at-risk group.
- Higher monthly charges align with increased churn likelihood.
- Longer tenure is associated with lower churn.

Recommendations:

1) Create retention offers for month-to-month customers (loyalty credits, targeted discounts, nudges to longer-term plans).
2) Review price tiers and value messaging for high-charge cohorts to reduce bill shock and increase perceived value.
3) Implement tenure-based loyalty benefits to reinforce stickiness (anniversary rewards, bundled add-ons).
4) Establish a simple KPI dashboard to monitor churn by segment weekly and trigger early interventions.

Conclusion

The end-to-end workflow converts raw data into actionable business insight. The pipeline (ingestion, cleaning, feature engineering, EDA) is reproducible and extensible to modeling and A/B-tested retention strategies. The findings point to practical levers-contract structure, pricing, and loyalty programs-to reduce churn.

About the Author

Lakshay Sharma
Email: lakshaysharma406@gmail.com
LinkedIn: linkedin.com/in/lakshay-sharma-6a5365299
Title: Data Analyst | Power BI | Python | SQL