Reducing Telecom Customer Churn: A Six Sigma Approach

A PROJECT REPORT SUBMITTED BY

Group 2

in partial fulfillment of the requirement for the course of

APS4742 - Statistical Applications in Industry and Project Presentation

to the

Department of Environmental and Industrial Sciences

of the

FACULTY OF SCIENCE

UNIVERSITY OF PERADENIYA

SRI LANKA

2025

Contents

Table Of Figures

Abstract

Maintaining a customer churn rate at or below industry standards is essential for the long-term profitability of telecommunication companies. According to McKinsey's Telecom Churn Reduction Strategies (2023), GSMA Intelligence benchmarks, and filings from major competitors such as Verizon and Vodafone, controlling churn is a strategic priority. In this project, we applied the DMAIC methodology of Six Sigma to identify the key drivers of churn and recommend practical improvements aligned with industry practices.

Our objective was to reduce the churn rate by 20% within 12 months. For this analysis, we used the Telco Customer Churn dataset published by IBM on Kaggle and extracted a stratified sample of 1,000 records, proportionally reflecting customer contract types. Data preprocessing involved removing invalid rows and converting necessary fields (such as TotalCharges) to numeric form. Analysis was conducted using Python with Pandas, Seaborn, Matplotlib, and SciPy.

The observed churn rate in our sample was 27.9%, exceeding industry standards and signaling the need for immediate corrective strategies. Our statistical analysis (Chi-square test) and visualizations revealed that month-to-month (short-term) customers are at the highest risk of churning. Based on these insights, we proposed actionable strategies such as encouraging long-term contracts, improving onboarding for new customers, and adjusting pricing models for high-paying clients.

While we followed the DMAIC framework throughout, we were limited in fully applying Lean methodology due to the absence of real operational process data. Nevertheless, our analysis provides a data-driven foundation for targeted churn reduction, and demonstrates how structured methods can be used to align business decisions with customer retention goals.

Introduction

Customer churn—the percentage of subscribers who discontinue their service with a telecommunication provider—is a critical indicator of business health in the telecom industry. A high churn rate not only results in direct revenue loss but also leads to increased customer acquisition costs, making customer retention an essential focus for telecom operators [1].

Several factors contribute to customer churn in the telecommunications sector. These include poor service experiences, ease of switching providers, more attractive offers from competitors, and repeated or unresolved customer support issues. Studies have highlighted the strong correlation between customer service quality and retention [2][3].

To contextualize churn performance, industry benchmarks are often used. For example, global and regional churn rate standards show that postpaid customer churn averages between 1.5% and 2.5% monthly, and 15% to 25% annually. Prepaid churn rates are typically higher, ranging from 3% to 7% [4][5][6][7][8]. These benchmarks serve as a reference point to assess whether a company's churn rate is within acceptable limits or signals the need for immediate intervention.

Given the structured nature of the problem—improving an existing service process—the Six Sigma DMAIC methodology is a suitable framework for addressing telecom churn. Six Sigma offers two main improvement strategies: DMAIC and DMADV. DMAIC, which stands for Define, Measure, Analyze, Improve, and Control, is used to enhance existing processes by identifying inefficiencies and systematically implementing solutions. In contrast, DMADV—Define, Measure, Analyze, Design, and Verify—is more appropriate for designing new products or services from scratch.

Since this analysis focuses on reducing churn within an existing telecom service framework, the DMAIC methodology offers a structured and data-driven approach for diagnosing issues, implementing changes, and sustaining improvements over time.

Data Set

The dataset used in this analysis is the **Telco Customer Churn** dataset from Kaggle [9], which focuses on understanding customer retention in the telecommunications sector. It contains **7,043 unique customer records**, each providing demographic, service usage, and billing information, along with a churn label indicating whether the customer has discontinued the service.

The dataset includes the following variables:

- customerID: Unique customer identifier
- gender: Gender of the customer (Male, Female)
- SeniorCitizen: Indicates if the customer is a senior citizen (0 = No, 1 = Yes)
- Partner: Whether the customer has a partner (Yes/No)
- Dependents: Whether the customer has dependents (Yes/No)
- tenure: Number of months the customer has stayed with the company
- PhoneService: Whether the customer has a phone service (Yes/No)
- MultipleLines: Whether the customer has multiple phone lines (Yes, No, No phone service)
- InternetService: Type of internet service (DSL, Fiber optic, No)
- OnlineSecurity: Whether the customer has online security (Yes, No, No internet service)
- OnlineBackup: Whether the customer has online backup (Yes, No, No internet service)
- DeviceProtection: Whether the customer has device protection (Yes, No, No internet service)
- TechSupport: Whether the customer has technical support (Yes, No, No internet service)
- Contract: Type of contract (Month-to-month, One year, Two year)
- PaperlessBilling: Whether the customer uses paperless billing (Yes/No)
- PaymentMethod: Payment method (Electronic check, Mailed check, Bank transfer (automatic), Credit card (automatic))
- MonthlyCharges: Monthly billing amount
- TotalCharges: Total amount charged to the customer
- Churn: Target variable indicating whether the customer churned (Yes/No)

METHODOLOGY

Define phase

In the Define phase, the primary focus is to clearly understand and articulate the problem. Analysis of the dataset revealed that the customer churn rate is significantly high, which negatively impacts the company's profitability. In the context of the telecom business, this has emphasized the need for focused customer retention programs.

Accordingly, the problem was defined as:

"The customer churn rate is high, adversely affecting profitability."

To ensure the project goals are clear and actionable, we applied the SMART goal principles [10]. According to these principles:

The goal should be measurable and achievable; thus, we target a 20% reduction in churn rate.

The goal must be time-bound; therefore, we set a timeline of 12 months, aligning with typical customer contract and billing cycles. This timeframe will also facilitate monitoring and adjustments during the Control phase.

The scope includes all customers covered by the dataset.

Our specific project objective is defined as:

"Reduce the customer churn rate by 20% within 12 months across the entire customer base."

Measure Phase

In the Measure phase, we aimed to understand the current performance of the customer retention process through data analysis and baseline metric calculations.

The dataset publisher verified that the data was already clean and free from inconsistencies, allowing us to proceed directly with exploratory and statistical analysis.

In this analysis, each customer ID represents a single connection, which is treated as one opportunity. Therefore, each customer corresponds to one unit, and each churned customer is considered a defect. As such, the metric Defects Per Opportunity (DPO) is equivalent to Defects Per Unit (DPU), which in this context is equal to the churn rate.

$$\text{Churn Rate (DPU)} = \frac{Number\ of\ Churned\ Customers}{Total\ Number\ of\ Customers}$$

From our stratified sample of 1,000 customers, the churn rate is:

$$\text{DPU} = \frac{279}{1000} = 0.279$$

To express this in Six Sigma terms, we calculated the Defects Per Million Opportunities (DPMO) as:

$$\text{DPMO} = \text{DPU} \times 1\ 000\ 000 = 0.279 \times 1\ 000\ 000 = 279\ 000$$

This high DPMO value reflects poor process capability and corresponds approximately to 2σ level, which is below industry expectations.

Our project goal is to reduce churn by 20%, which would bring the churn rate down to approximately 22.3%

$$0.279 \times (1 - 0.20) = 0.223$$

This improvement would not only reduce customer attrition but also elevate the Sigma level, indicating better process performance and alignment with industry benchmarks.

The churn rate is calculated as:To understand the structure of the population, we analyzed the distribution of customers by contract type and the corresponding churn rates per contract category. These patterns were visualized in Figure 1, revealing notable differences in churn behavior based on contract type.

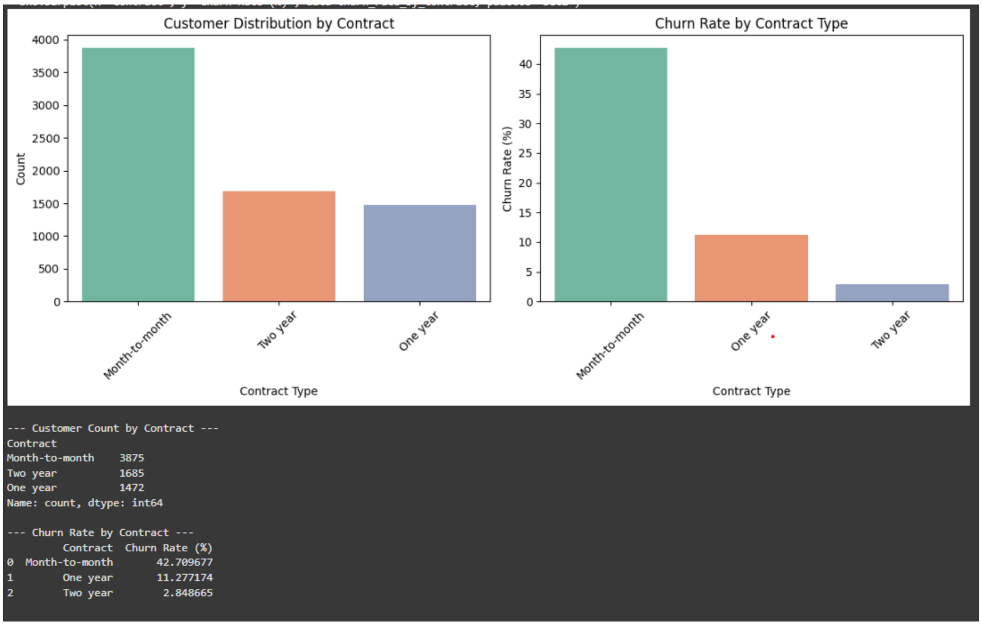

Figure 1: Distribution of Customers and Churn rate by Contact type

To ensure the representativeness of the data while maintaining manageability for the analysis, we employed a stratified sampling technique. This technique preserved the proportions of contract types and churn outcomes while reducing the dataset to approximately 1,000 customers. The sampling process was performed using the train_test_split function from sklearn.model_selection, as shown in Figure 2.

```
# Take stratified sample of about 1000 rows based on 'Contract'
df_sample, _ = train_test_split(
    df,
    train_size=1000,          # Request approximately 1000 samples
    stratify=df['Contract'],
    random_state=42
)
```

Figure 2: Code snippet of sratified sampling using sklearn model

Analyze Phase

Box-and-whisker plot

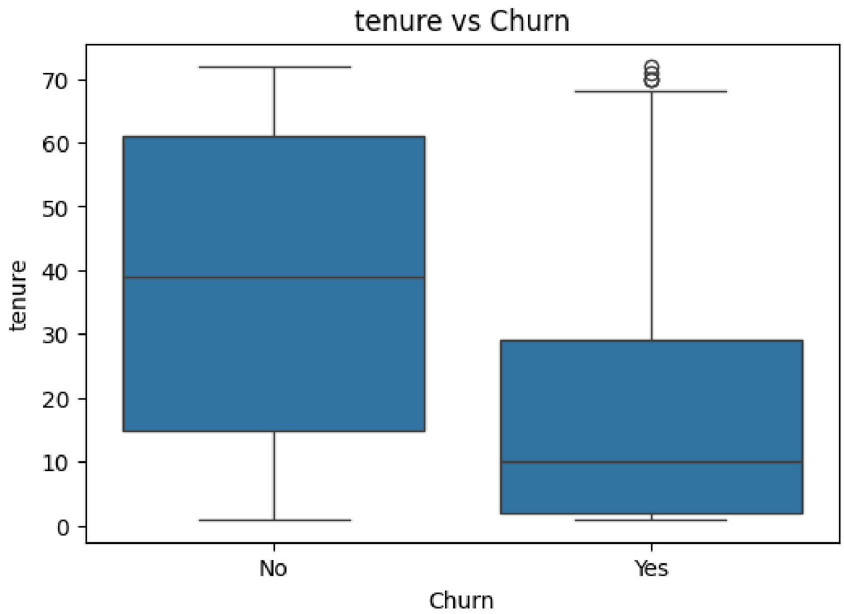

Figure 3: Boxplot of tenure vs churn

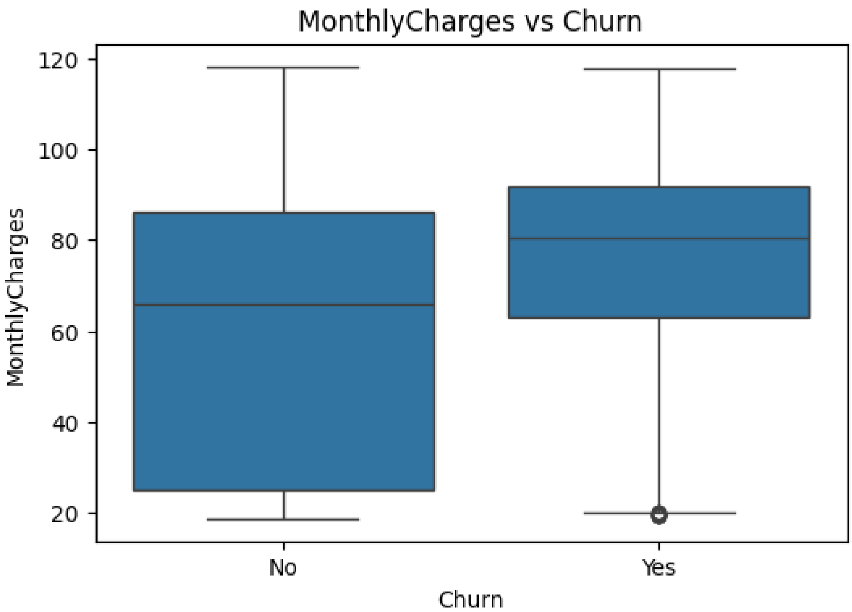

Figure 4: Boxplot of Monthly charges vs churn

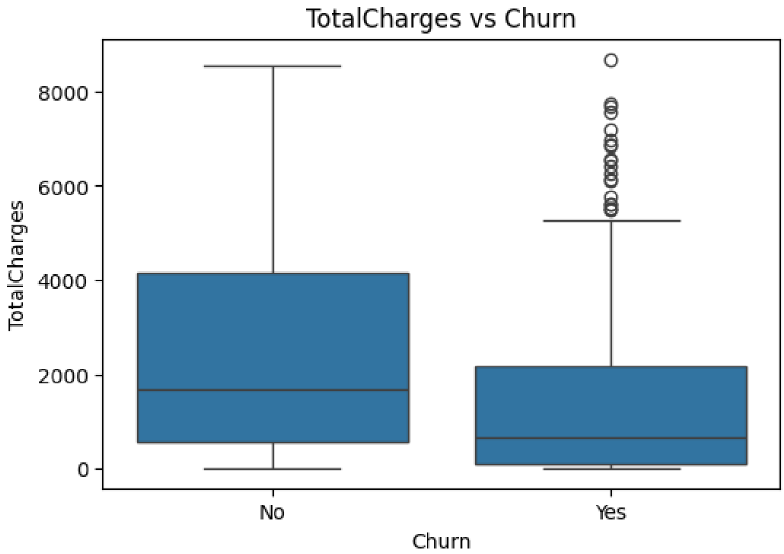

Figure 5: Boxplot of total charges vs churn

To identify key factors contributing to customer churn, several visual analyses were conducted using box-and-whisker plots.

- ✓ Figure 3: The plot of tenure vs. churn reveals that customers with shorter tenure are significantly more likely to churn. This trend is especially noticeable among month-to-month contract customers, indicating that the absence of long-term commitment increases the likelihood of churn.
- ✓ Figures 4 and 5: The plots of Monthly Charges and Total Charges against churn status show that customers with higher charges are more prone to leaving the service. This suggests that pricing or perceived value may be key drivers of dissatisfaction.

Bar Chart

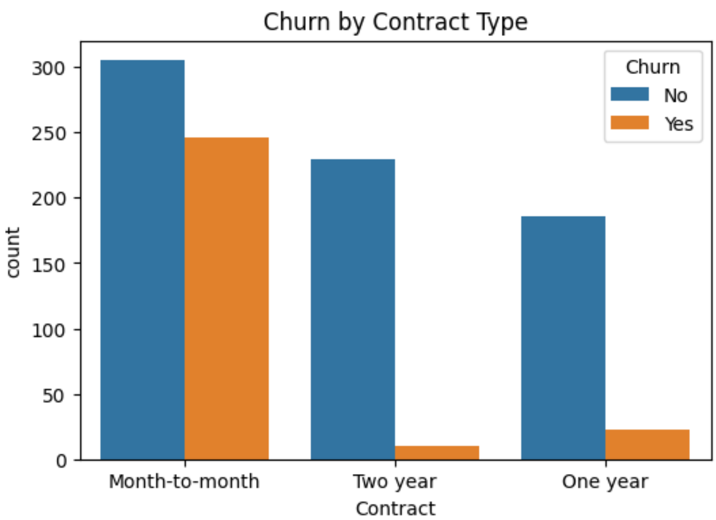

Figure 6: Churn by Contract Type

A bar chart of churn rate by contract type (Figure 6) further confirms that month-to-month contract customers exhibit the highest churn rate. In contrast, customers with one-year and two-year contracts show significantly lower churn rates, indicating a higher level of loyalty and long-term commitment.

This insight aligns with the earlier box plot findings (Figures 3–5), reinforcing the conclusion that contract type is a strong predictor of customer churn. The consistency between the bar chart and box plot results strengthens the evidence that short-term contracts contribute heavily to customer attrition.

Chi-square test

To statistically validate the relationship between contract type and customer churn, a Chi-square test of independence was conducted. The test evaluated whether churn is dependent on the type of contract a customer holds.

```
contingency = pd.crosstab(df_sample['Contract'], df_sample['Churn'])
chi2, p, dof, expected = chi2_contingency(contingency)
print(f"Chi-Square Test p-value (Contract vs Churn): {p:.4f}")
```

```
Chi-Square Test p-value (Contract vs Churn): 0.0000
```

Figure 7: Chi-square test code and result output for contract type vs. churn

- ✓ Figure 7 displays the Python code snippet used to perform the test, along with the resulting statistical output.
- ✓ The test result showed a p-value < 0.05, indicating that the association between contract type and churn is statistically significant.

This confirms that the observed differences in churn rates across contract types are unlikely to be due to random chance, further strengthening the evidence that contract type is a key factor influencing customer retention.

Improve Phase

Based on the key factors identified during the Analyze phase—particularly the strong correlation between short-term contracts, high monthly charges, and churn—we propose a set of targeted strategic actions aimed at reducing customer attrition. These interventions are designed to directly address the causes of churn and align with the goal of reducing the churn rate by 20% within 12 months.

To suggest these improvements, we studied both local and international telecommunication companies that have faced similar problem and found solutions. The following suggestions are based on the strategies they implemented to overcome those challenges

1. Encourage Long-Term Contracts Through Incentives

To reduce churn among month-to-month customers, promotional strategies can be used to incentivize longer commitments:

- ✓ 10% discount for customers switching to a 1-year contract
- ✓ 15% discount for 2-year contracts
- ✓ Promotional offer: “Sign a 2-year contract and receive 1 month of premium service free”

These incentives aim to improve retention by offering customers added value in exchange for commitment.

2. Enhance Onboarding and Early Engagement

Churn is especially high among new customers with low tenure, so improving the first-month experience is critical. A structured onboarding program can significantly increase engagement:

“First 30 Days” Welcome Kit – a personalized email series:

- ✓ Day 1: “Get Started” guide
- ✓ Day 7: “Tip of the Week” (e.g., how to optimize data usage)
- ✓ Day 30: Feedback survey with a $5 account credit for completion

This approach builds trust early in the customer journey and reduces early churn risk.

3. Offer Flexible Pricing for High-Paying Customers

Analysis showed that customers with higher monthly charges tend to churn more frequently. To retain this valuable customer segment:

- ✓ Tiered Data Plans: Allow mid-cycle adjustments without penalties

  E.g., customers can downgrade from 50GB to 30GB if usage drops

- ✓ Loyalty Rewards Program:

  Customers paying over $80/month receive an annual 5% rebate after 12 months of continued service

This flexibility improves perceived fairness and reduces churn due to cost concerns.

Control Phase

The Control phase focuses on sustaining improvements made during the Improve phase and ensuring that the churn reduction strategies continue to be effective over time. A set of monitoring tools and review processes are proposed to maintain control over the churn rate and enable timely corrective actions.

1. Real-Time Churn Tracking Dashboards

To maintain visibility into customer behavior, interactive dashboards will be developed to monitor churn metrics in real time. These dashboards will provide:

- ✓ Live churn rate trends segmented by contract type, tenure, and charges
- ✓ Visual indicators of churn spikes by demographic or service group
- ✓ Filters for custom time ranges and service attributes

This will allow stakeholders to quickly detect changes and investigate anomalies.

2. Threshold-Based Alerts

To ensure proactive response, automated threshold alerts will be set up:

- ✓ Example: If the churn rate exceeds 30% in any given customer segment or region, the system will trigger an alert for immediate review.
- ✓ Alerts will be integrated into the dashboard system and optionally linked to email notifications for management teams.

This approach prevents delay in recognizing churn escalations and enables early intervention.

3. Quarterly Churn Analysis and Strategy Review

A structured quarterly review cycle will be established to assess the effectiveness of retention strategies and adjust as needed:

- ✓ Review updated churn data and compare it against SMART goal targets
- ✓ Evaluate performance of specific strategies (e.g., discount offers, loyalty programs)
- ✓ Identify new trends or customer segments at risk
- ✓ Adjust and iterate on retention campaigns accordingly

These reviews ensure continuous improvement and long-term sustainability of the churn management process.

Conclusion

This study successfully applied the Six Sigma DMAIC methodology to investigate and propose strategies for reducing customer churn in the telecommunications sector. By analyzing a stratified sample of customer data, we identified contract type, tenure, and monthly charges as significant predictors of churn. In particular, customers with month-to-month contracts and higher billing amounts showed a substantially higher risk of attrition.

A key strength of this project lies in its structured approach: the use of data visualization, statistical validation through the Chi-square test, and practical, actionable recommendations. The integration of real-world business practices, such as loyalty programs and onboarding campaigns, also adds relevance and potential applicability in an industry context.

However, some limitations should be noted. The analysis relied on secondary data without access to operational process metrics, which restricted the use of Lean methodology and limited deeper process-level diagnostics. Furthermore, customer sentiment data (e.g., satisfaction scores or support call logs) were not available, which could have enriched the understanding of churn drivers.

Despite these limitations, the results offer a compelling foundation for informed decision-making. The proposed interventions—encouraging long-term contracts, enhancing early engagement, and introducing pricing flexibility—directly address the primary churn risks uncovered in the analysis. If implemented and monitored effectively, these strategies have strong potential to achieve the targeted 20% reduction in churn within a 12-month period.

Reference

[1] S. John, *Churn: What It Is? What Can Brands Do About It?*. Available: *https://www.verint.com/blog/churn-what-it-is-what-can-brands-do-about-it* accessed on 25th June 2025

[2] *Crash and Churn: Results of 2019 Consumer Survey into the Relationship Between Quality of Customer Service and Retention in the Telecom Industry*, TechSee, 2019.

[3] S. Crossett, *A Complete Guide to Reducing Churn in the Telecom Industry* Available: https://www.userpilot.com/blog/reducing-churn-telecom accessed on 25th June 2025

[4] *Telecom Churn Reduction Strategies*, McKinsey & Company, 2023.

[5] *Global Operator Benchmarks*, GSMA Intelligence.

[6] *Annual Report*, Deloitte. *https://www.deloitte.com/uk/en/about/governance/annual-reports.html* accessed on 25th June 2025

[7] Federal Communications Commission, *Monthly Churn Rate for Mobile Services*, Report 05-173.

[8] *Telecommunications Industry Churn Benchmarks*, FCC Reports and Ericsson.

[9] Kaggle, Telco Customer Churn Dataset. Available: https://www.kaggle.com/datasets/blastchar/telco-customer-churn, Accessed on: 23rd June 2025.

[10] P. Doran, “There’s a S.M.A.R.T. way to write management’s goals and objectives,” Management Review, vol. 70, no. 11, pp. 35–36, 1981.

Appendices

1.Python Source Code for Analysis is available :
https://colab.research.google.com/drive/1D4yKDiHPPOyPD3Sx6l2rE0MEzrxchVZR?usp=sharing

2. The used data set available: https://www.kaggle.com/datasets/blastchar/telco-customer-churn

3. used Libraries

```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from scipy.stats import chi2_contingency
from sklearn.model_selection import train_test_split
```