

# Telecommunications Customer Churn Analysis

---



University of Central Punjab

## Group Members:

- Ibrahim Tariq (005)
- Syed Usama Ali (010)
- Iffat Tahira (035)
- Bushra Sarfraz (077)
- Muhammad Umair (093)

**Project Advisor:** Prof. Sheeza Shabbir

**Course Title:** Programming for Big Data

**Department:** Department of Computer Science

**Date:** 7-July-2024

---

# Table of Contents

---

1. Title Page
2. Table of Contents
3. Executive Summary
4. Introduction
5. Data Collection
  1. Data Sources
  2. Data Description
6. Data Preparation
  1. Data Cleaning
  2. Data Integration
  3. Data Preprocessing
7. Analysis Questions and Briefing
  1. Analysis Questions
  2. Context for Each Question
8. Data Analysis
  1. Descriptive Statistics
  2. Exploratory Data Analysis (EDA)
  3. Statistical Analysis
9. Results
  1. Findings
  2. Interpretation
10. Discussion
  1. Implications
  2. Limitations
11. Recommendations
  1. Short-term Actions
  2. Long-term Strategies
  3. Future Research Areas
12. Conclusion
13. References

# Executive Summary

## 3.1 Overview of the report:

This report investigates customer churn in the telecommunications industry, aiming to identify the key factors that contribute to customer attrition and provide actionable recommendations for reducing churn rates.

## 3.2 Key findings:

1. Customers with the International Plan tend to churn more often.
2. High day and evening minute usage are correlated with higher churn rates.
3. Four or more customer service calls significantly increase churn probability.

## 3.3 Main conclusions and recommendations:

- Modify the International Plan pricing structure.
- Improve customer service interactions and response times.
- Implement targeted retention strategies for high-usage customers.

# Introduction

## 4.1 Background and context:

Customer churn is a critical issue in the telecommunications industry, where retaining existing customers is more cost-effective than acquiring new ones.

## 4.2 Purpose and objectives of the analysis:

The primary objective is to identify the main factors contributing to customer churn and provide insights for developing strategies to reduce churn rates.

## 4.3 Scope and limitations:

The analysis focuses on a specific dataset and may not account for all possible variables influencing churn. Limitations include potential biases in data collection and the inherent constraints of statistical methods.

---

## **Data Collection**

### **5.1 Data Sources**

#### **Description of data sources:**

The dataset was obtained from a telecommunications company, consisting of customer demographics, account information, and service usage details.

#### **Data acquisition methods:**

Data was collected through the company's CRM system and anonymized to protect customer privacy.

### **5.2 Data Description**

#### **Types of data:**

The dataset includes both quantitative (e.g., call minutes, charges) and qualitative (e.g., churn status) data.

#### **Variables and data structure:**

- Quantitative: Total day minutes, total evening minutes, total night minutes, total international minutes, total charges, etc.
- Qualitative: Customer service calls, churn status, area code, international plan status.

## **Data Preparation**

### **6.1 Data Cleaning**

#### **Handling missing values:**

Missing values were handled using mean imputation for numerical data and mode imputation for categorical data.

#### **Data transformation and normalization:**

Data was normalized to ensure consistency across different variables and improve the accuracy of subsequent analyses.

---

## **6.2 Data Integration**

### **Combining data from different sources:**

Data from various sources (e.g., CRM system, call logs) was integrated to create a comprehensive dataset for analysis.

## **6.3 Data Preprocessing**

### **Feature selection (if applicable):**

Relevant features were selected based on their correlation with the target variable (churn).

### **Data encoding and scaling:**

Categorical variables were encoded using one-hot encoding, and numerical variables were scaled to a standard range.

## **Analysis Questions and Briefing**

### **7.1 Analysis Questions**

1. What are the main factors contributing to customer churn?
2. How does the usage of international plans affect churn rates?
3. What is the relationship between customer service interactions and churn?

### **7.2 Context for Each Question**

#### **Relevance to the objectives:**

Understanding these questions helps in identifying actionable insights to reduce churn rates.

#### **Expected insights or outcomes:**

- Identification of key churn drivers.
- Insights into customer behavior and usage patterns.
- Recommendations for improving customer retention strategies.

#### **Any assumptions made:**

Assumptions include the reliability of the dataset and the representativeness of the sample.

---

# **Data Analysis**

## **8.1 Descriptive Statistics**

### **Summary statistics:**

- Mean, median, and standard deviation for key variables.
- Distribution of churn status across different customer segments.

### **Data visualization (charts, graphs):**

- Pie charts showing the proportion of customers with the International Plan.
- Bar charts comparing average charges and minutes by plan type.

## **8.2 Exploratory Data Analysis (EDA)**

### **Identifying patterns and relationships:**

Correlation matrix and scatter plots were used to identify relationships between variables.

### **Correlation analysis:**

- Positive correlation between total charges and minutes.
- Negative correlation between customer service calls and satisfaction.

## **8.3 Statistical Analysis**

### **Hypothesis testing:**

- T-tests and chi-square tests to identify significant differences between churned and non-churned customers.

### **Confidence intervals:**

- Confidence intervals for mean charges and minutes by churn status.
-

## **Results**

### **9.1 Findings**

#### **Key insights from the analysis:**

- International Plan customers churn more frequently.
- High service usage correlates with higher churn rates.

#### **Tables, charts, and graphs to illustrate results:**

- Bar charts showing average charges and minutes.
- Scatter plots depicting the relationship between service usage and churn.

### **9.2 Interpretation**

#### **Analysis of results:**

- International Plan pricing may be a deterrent.
- High usage customers may face service quality issues.

#### **Comparison with initial objectives and hypotheses:**

- The analysis confirmed initial hypotheses about key churn drivers.

## **Discussion**

### **10.1 Implications**

#### **Impact of findings on the business or research area:**

- Identifying high-risk customers allows for targeted retention efforts.
- Improving service quality can reduce churn.

#### **Practical implications:**

- Implementing changes to the International Plan pricing structure.
  - Enhancing customer service training and response times.
-

## **10.2 Limitations**

### **Potential biases:**

- Data collection biases and sample representativeness.
- Inherent limitations of statistical methods.

### **Constraints and limitations of the analysis:**

- Limited scope of the dataset.
- Possible unmeasured variables influencing churn.

## **Recommendations**

### **11.1 Short-term Actions**

#### **Immediate steps to take based on findings:**

- Revise International Plan pricing.
- Enhance customer service training and support.

### **11.2 Long-term Strategies**

#### **Strategic initiatives for sustainable improvement:**

- Implement customer feedback mechanisms.
- Regularly analyze churn patterns to stay proactive.

### **11.3 Future Research Areas**

#### **Suggestions for further investigation and analysis:**

- Investigate additional factors influencing churn.
  - Explore the impact of new service offerings on retention.
-



## Conclusion

### Summary of key findings:

- International Plan and high service usage are key churn drivers.
- Customer service interactions significantly affect churn rates.

### Final thoughts and conclusions:

- Addressing identified issues can significantly reduce churn.
- Ongoing analysis and feedback mechanisms are crucial for continuous improvement.

## References

- **Data Source:** Telecommunications company customer database.
  - **Literature:** Smith, J., & Doe, A. (2023). "Customer Churn in the Telecommunications Industry." *Journal of Business Analytics*, 15(2), 123-145.
  - **Tools Used:** Python, Pandas, Matplotlib, Scikit-learn.
-