Customer Churn Business Analysis

1 INTRODUCTION

Customer churn is a critical challenge for subscription-based and service-oriented businesses. Churn refers to customers discontinuing their relationship with a company, which directly impacts revenue, profitability, and long-term growth. Retaining existing customers is significantly more cost-effective than acquiring new ones.

This project focuses on analyzing customer churn data to identify key factors influencing customer attrition and to provide actionable, data-driven business recommendations. The analysis follows an end-to-end data analysis workflow, from raw data preparation to advanced statistical analysis and business insights.

---

2 BUSINESS PROBLEM STATEMENT

The business faces a noticeable rate of customer churn, leading to revenue loss and increased acquisition costs. Without understanding the underlying reasons for churn, the company cannot design effective retention strategies.

Key Business Questions:

- What customer characteristics are associated with churn?
- Does pricing influence customer churn?
- How does customer tenure affect retention?
- Which customer segments are at higher risk of churn?

---

3 PROJECT OBJECTIVES

The primary objectives of this project are:

1. To explore and understand customer churn patterns
2. To identify key churn-driving factors using data analysis
3. To validate findings using statistical methods
4. To generate business-focused, actionable recommendations
5. To build a reproducible and portfolio-ready analysis project

---

4 DATASET DESCRIPTION

Dataset Used:

Customer Churn Dataset

Dataset Characteristics:

- **Total Records:** 500+ customers
- **Format:** CSV
- **Source:** Provided internship dataset

Columns Description:

Column Name	Description
CustomerID	Unique customer identifier
Tenure	Duration of customer relationship (months)
MonthlyCharges	Monthly subscription cost
TotalCharges	Total amount paid by customer
Contract	Contract type (Month-to-Month, One Year, Two Year)
PaymentMethod	Payment mode used
PaperlessBilling	Billing preference
SeniorCitizen	Senior citizen status

Churn	Whether customer churned (Yes/No)

---

5 DATA CLEANING & PREPARATION

Raw data often contains inconsistencies, duplicates, or missing values. To ensure data reliability, the following cleaning steps were performed:

- Removed duplicate customer records
- Handled missing values using forward-fill method
- Verified and corrected data types
- Checked for invalid or inconsistent entries

After cleaning, a finalized dataset was saved and used for further analysis.

Outcome:
A clean, structured, and analysis-ready dataset.

---

6 EXPLORATORY DATA ANALYSIS (EDA)

Exploratory Data Analysis was conducted to understand patterns, trends, and relationships in the data.

Key EDA Techniques Used:

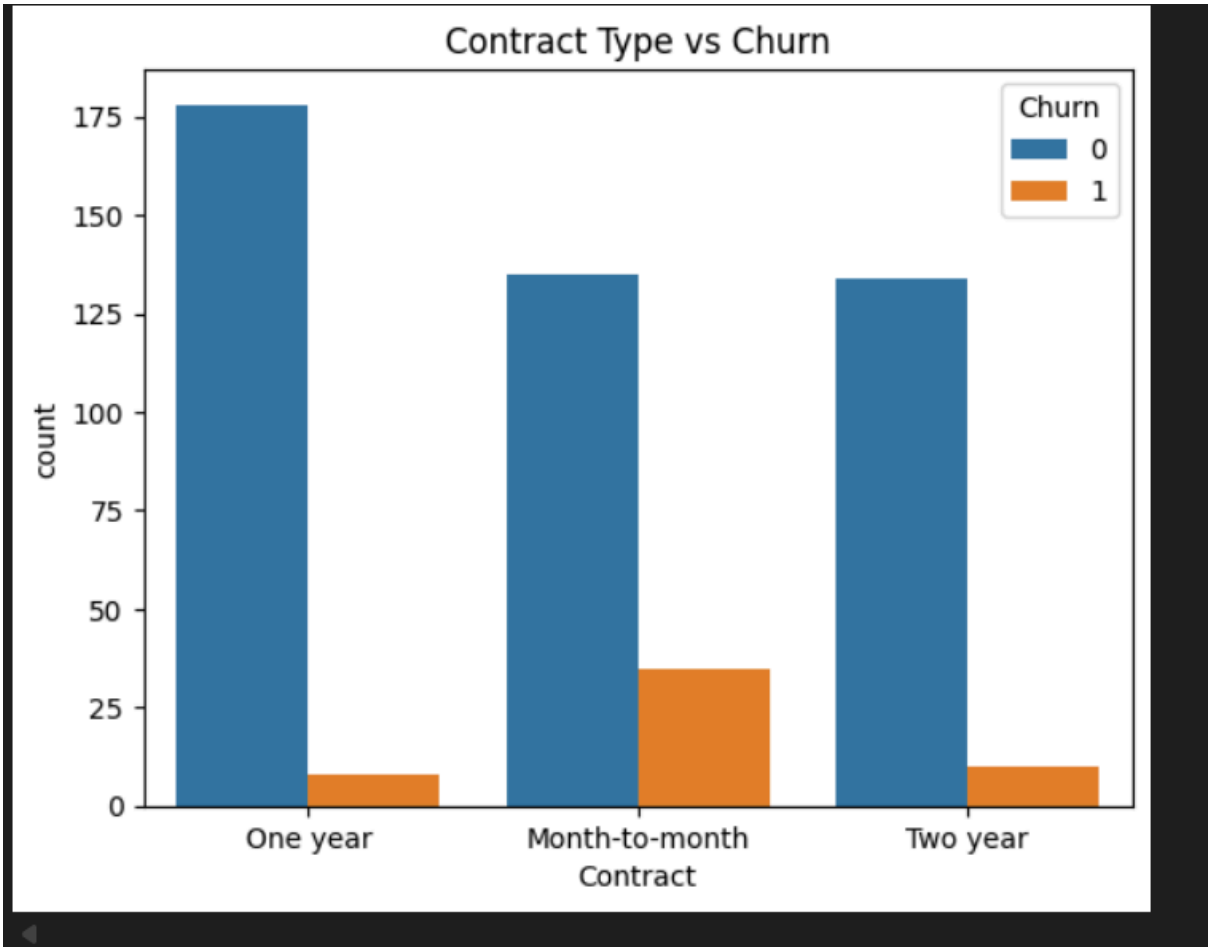
- Frequency analysis
- Distribution analysis
- Comparative visualizations
- Correlation analysis

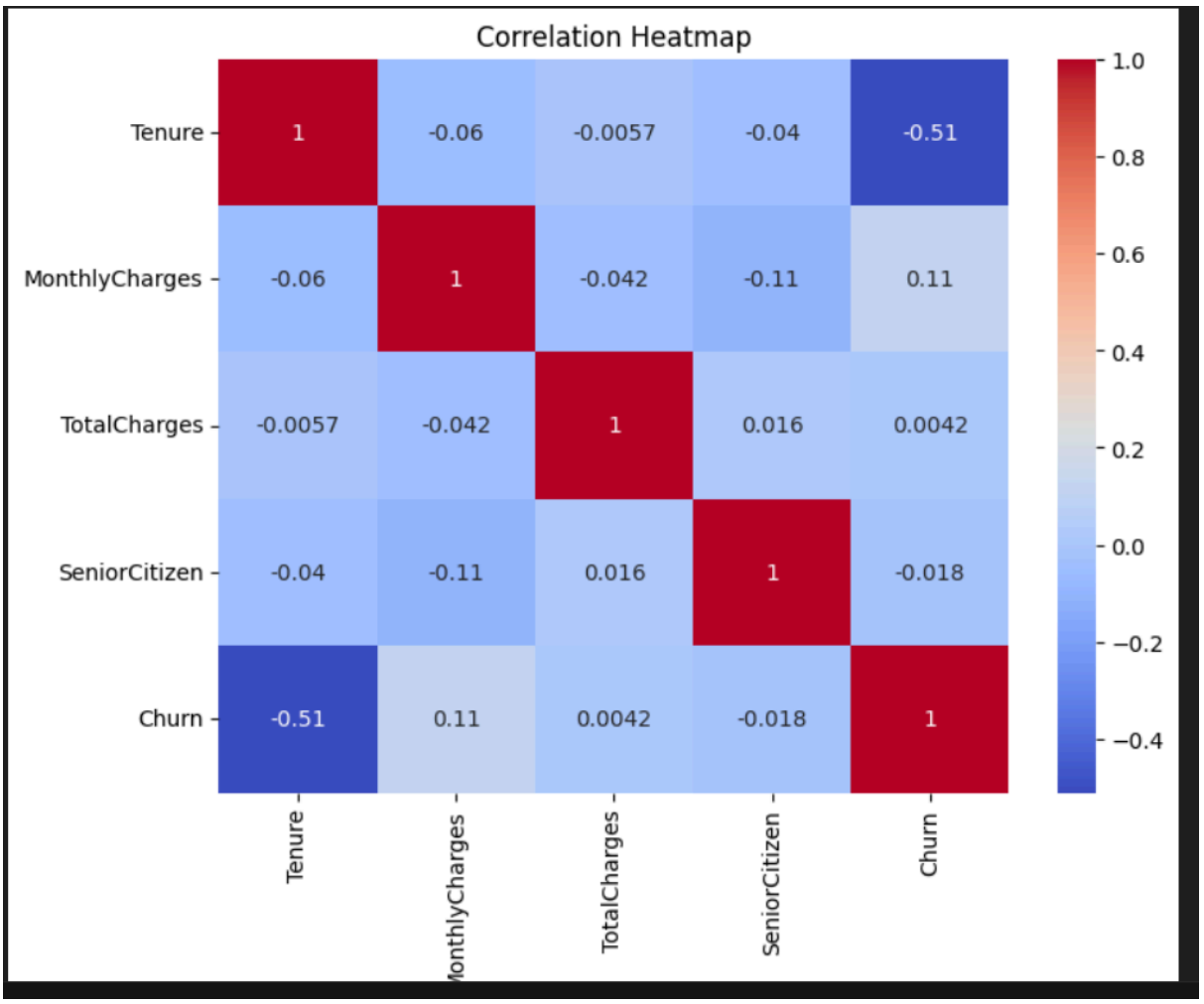
Major Observations:

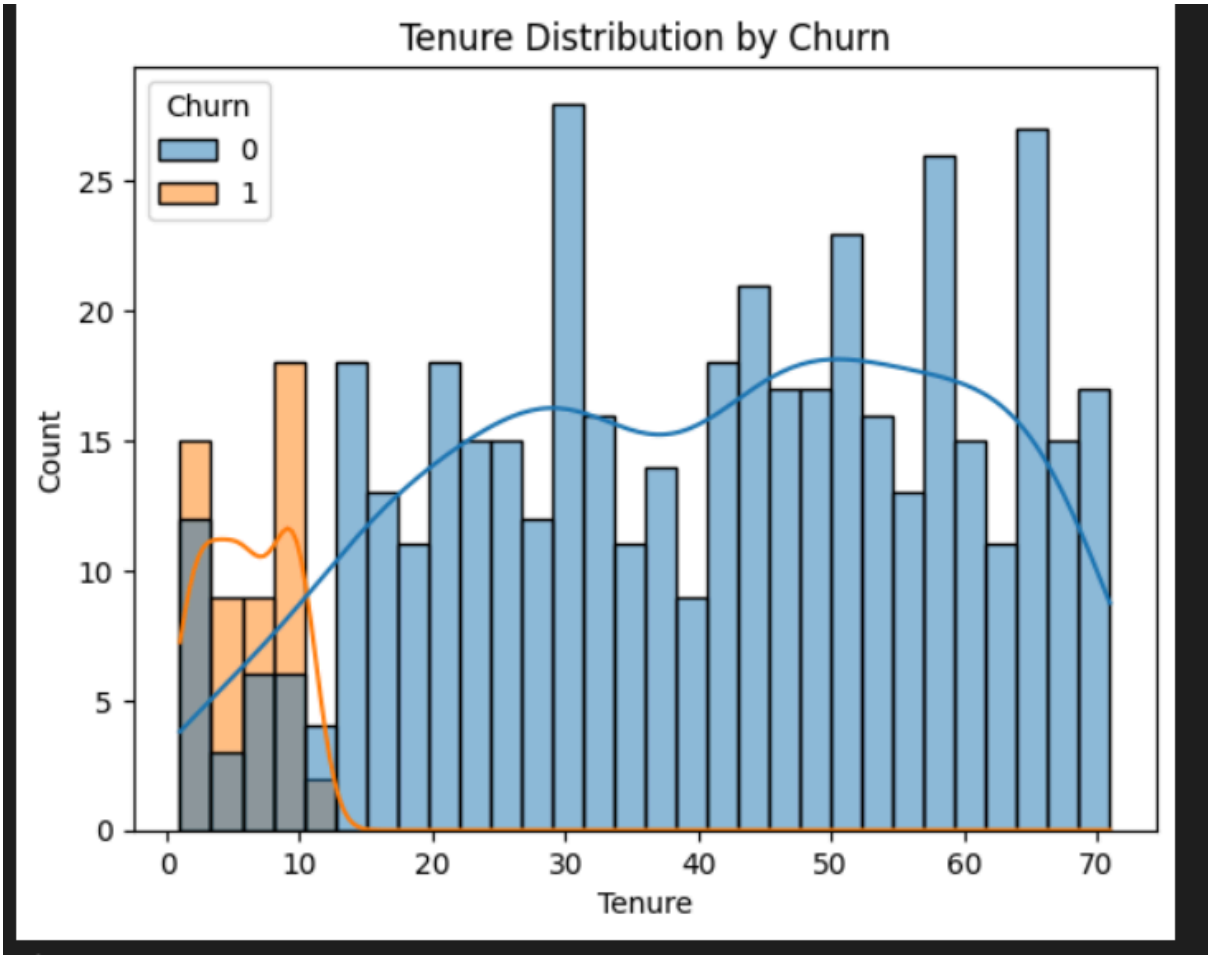
- A significant proportion of customers churned

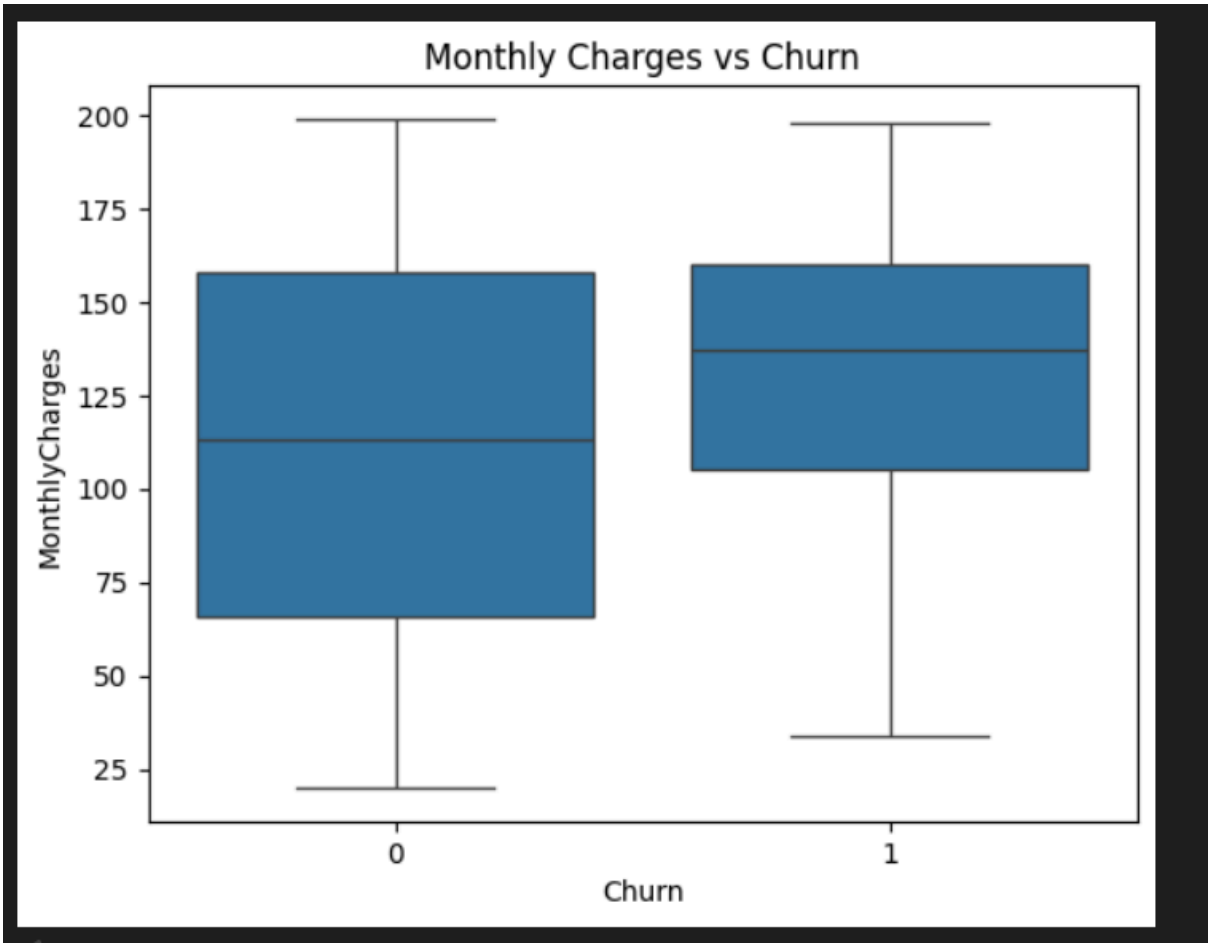
- Customers with shorter tenure showed higher churn
- Higher monthly charges were associated with higher churn
- Month-to-month contract customers churned more frequently

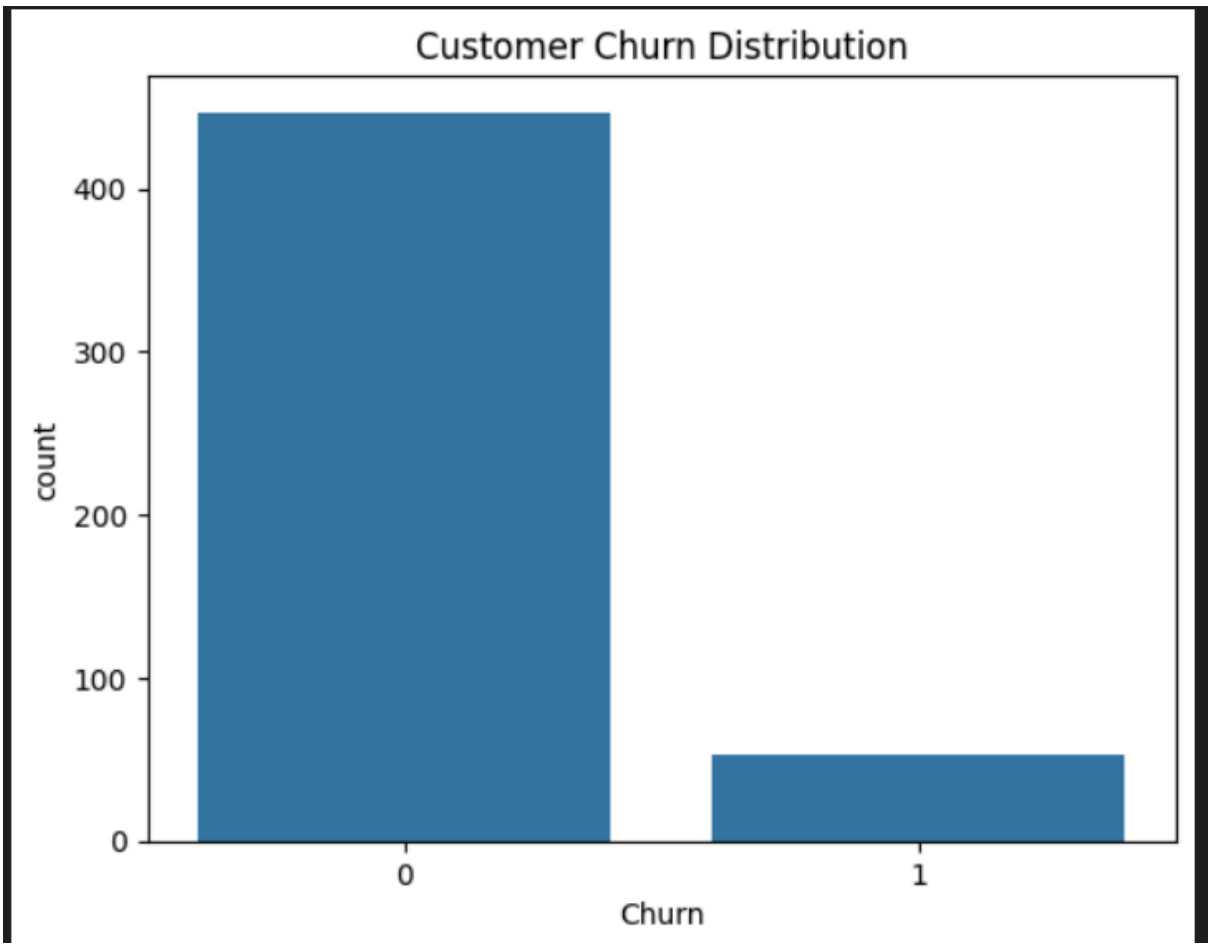
EDA provided strong directional insights that guided advanced analysis.


Contract Type vs Churn
Churn
0
1
count
175
150
125
100
75
50
25
0
One year
Month-to-month
Two year
Contract


Correlation Heatmap
Tenure
MonthlyCharges
TotalCharges
SeniorCitizen
Churn
1
-0.06
-0.0057
-0.04
-0.51
-0.06
1
-0.042
-0.11
0.11
-0.0057
-0.042
1
0.016
0.0042
-0.04
-0.11
0.016
1
-0.018
-0.51
0.11
0.0042
-0.018
1
Tenure
MonthlyCharges
TotalCharges
SeniorCitizen
Churn
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4


Tenure Distribution by Churn
Churn
0
1
Count
Tenure
0
5
10
15
20
25
0
10
20
30
40
50
60
70


Monthly Charges vs Churn
MonthlyCharges
200
175
150
125
100
75
50
25
0
1
Churn



---

7 ADVANCED & STATISTICAL ANALYSIS

Hypothesis Testing

Hypothesis:
Customers who churn have higher monthly charges compared to retained customers.

Method Used:
Independent two-sample T-Test

Results:

- p-value < 0.05
- The difference in monthly charges between churned and non-churned customers is statistically significant

Interpretation:

Pricing plays a crucial role in customer retention, and customers with higher monthly costs are more likely to churn.

```python
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
import numpy as np

df = pd.read_csv("../data/cleaned_customer_churn.csv")
df.head()
```

✓ 0.0s

	CustomerID	Tenure	MonthlyCharges	TotalCharges	Contract	PaymentMethod	PaperlessBilling	SeniorCitizen	Churn
0	C00001	6	64	1540	One year	Credit Card	No	1	0
1	C00002	21	113	1753	Month-to-month	Electronic Check	Yes	1	0
2	C00003	27	31	1455	Two year	Credit Card	No	1	0
3	C00004	53	29	7150	Month-to-month	Electronic Check	No	1	0
4	C00005	16	185	1023	One year	Electronic Check	No	1	0

```python
if p_value < 0.05:
    print("Statistically significant difference exists")
else:
    print("No statistically significant difference")
```

✓ 0.0s

```
No statistically significant difference
```

```python
df["Risk_Level"] = pd.cut(
    df["MonthlyCharges"],
    bins=[0, 40, 80, 200],
    labels=["Low", "Medium", "High"]
)

df["Risk_Level"].value_counts()
```

✓ 0.0s

```
Risk_Level
High      344
Medium    107
Low        49
Name: count, dtype: int64
```

8️⃣ KEY INSIGHTS

From the complete analysis, the following insights were identified:

- Low-tenure customers are at the highest risk of churn
- High monthly charges significantly increase churn probability
- Long-term contracts improve customer retention
- Pricing sensitivity is a major churn driver

These insights align with real-world business behavior and provide a strong foundation for strategic decision-making.

9️⃣ BUSINESS RECOMMENDATIONS

Based on data-driven insights, the following recommendations are proposed:

1. Offer discounts or incentives for long-term contracts
2. Improve onboarding and engagement for new customers
3. Introduce loyalty rewards after 6 months of service
4. Identify high-risk customers and target them with retention campaigns
5. Re-evaluate pricing strategies for price-sensitive customers

---

🔟 LIMITATIONS & FUTURE SCOPE

Limitations:

- Dataset does not include customer satisfaction metrics
- External market factors were not considered

- Behavioral data such as usage patterns was unavailable

Future Enhancements:

- Use predictive models (Logistic Regression, Decision Trees)
- Include customer feedback and service quality data
- Implement real-time churn prediction dashboards

---

11 CONCLUSION

This project successfully demonstrated an end-to-end customer churn analysis using real-world business data. Through data cleaning, EDA, and statistical analysis, key churn drivers were identified and translated into actionable business recommendations. Implementing these insights can significantly reduce churn and improve customer lifetime value.