Customer Retention Prediction System

Project Summary:

This project analyses historical telecom customer data to understand customer behaviour and predict churn. The dataset includes customer demographics, service subscriptions, billing details, tenure, and churn status. Using Exploratory Data Analysis (EDA) with Matplotlib visualizations, the study identifies patterns and trends that influence customer churn. The insights gained help the organization make data-driven decisions to improve customer retention and reduce churn.

Summary – Key Insights & Objectives:

The analysis reveals that customer churn is strongly influenced by contract type, tenure duration, payment methods, service add-ons, and billing preferences. Customers on month-to-month contracts, in early tenure stages, without additional services, and using manual payment methods show a higher likelihood of churn. In contrast, customers with long-term contracts, bundled services, dependents, and multiple service lines exhibit higher retention. The overall objective is to leverage visual insights from EDA to identify churn drivers and support effective customer retention strategies.

Dataset Overview:

The dataset used in this project contains historical telecom customer information designed to analyse customer behaviour and predict churn. Each record represents an individual customer and captures a combination of demographic details, service subscriptions, billing information, tenure, and churn status. The dataset enables a comprehensive exploratory data analysis (EDA) to identify factors influencing customer retention and churn.

Dataset Description:

The dataset includes the following major categories of features:

1. Customer Demographics:

These attributes describe the personal characteristics of customers and help analyse churn patterns across different demographic groups.

- Gender
- Senior citizen status
- Dependents

2. Service Subscription Details:

These features indicate the telecom services subscribed to by customers and help assess the impact of service usage on churn.

- Type of SIM / network provider
- Phone service and multiple lines
- Internet service type
- Add-on services such as:
    - Online security

- Online backup
- Device protection
- Technical support
- Streaming services

3. Contract & Tenure Information:

These attributes capture the duration and nature of customer relationships with the organization.

- Contract type (Month-to-month, One-year, Two-year)
- Customer tenure (number of months with the service)

4. Billing & Payment Information:

These features describe customer billing behaviour and payment preferences.

- Monthly charges
- Total charges
- Payment method (manual or automatic)
- Paperless billing status

5. Target Variable:

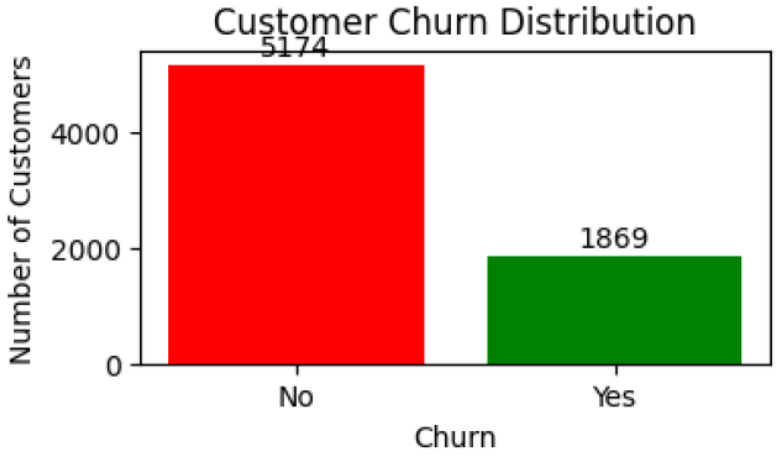
- **Churn**: Indicates whether the customer has discontinued the service (Yes/No). This is the primary variable used to analyse customer attrition behaviour.

Visualizations and Interpretation:

Exploratory Data Analysis (EDA) was performed using **Matplotlib** to visually analyse customer behaviour and identify factors influencing customer churn. The visualizations help uncover trends, patterns, and relationships within the data that are not easily observable through raw numbers.

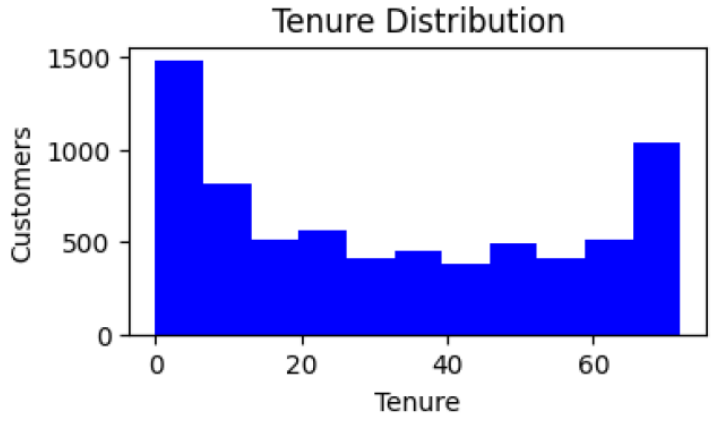
Customer Churn Distribution:

- Bar chart (Churn: Yes / No)
- Compares customers who churned vs those who stayed.
- Majority of customers are retained (No).
- A significant portion has churned (Yes), indicating a real retention problem.
- Even a smaller churn percentage can lead to major revenue loss due to the large customer base.



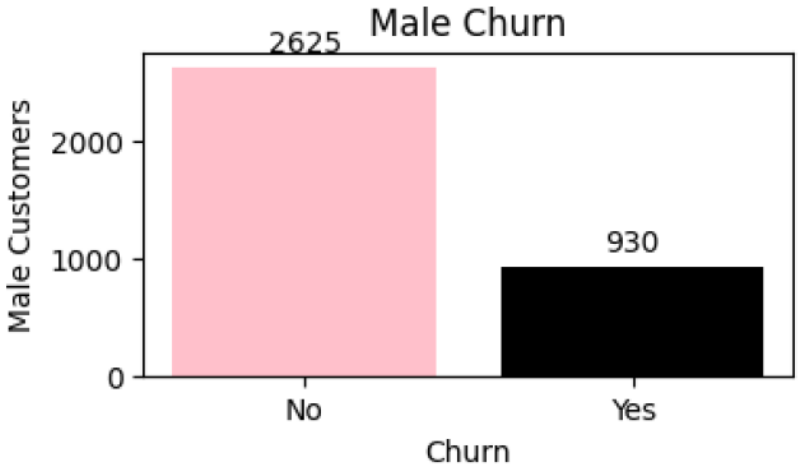
Tenure Distribution:

- Histogram
- Distribution of customers based on how long they stayed.
- Many customers churn within early tenure, while long-term customers are more stable.
- Retention strategies should focus heavily on the first few months.



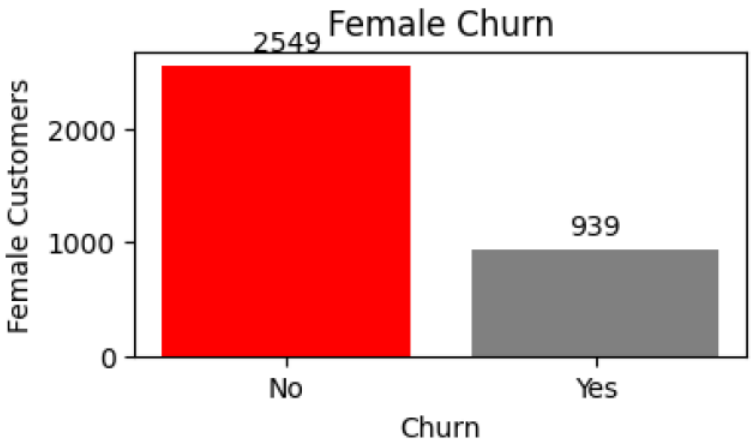
Male Churn Analysis:

- Bar chart (Male – Churn Yes / No).
- Churn behaviour specifically for male customers.
- More males are retained than churned.
- Male customers may respond differently to offers or pricing.
- Gender alone is not a strong churn predictor.



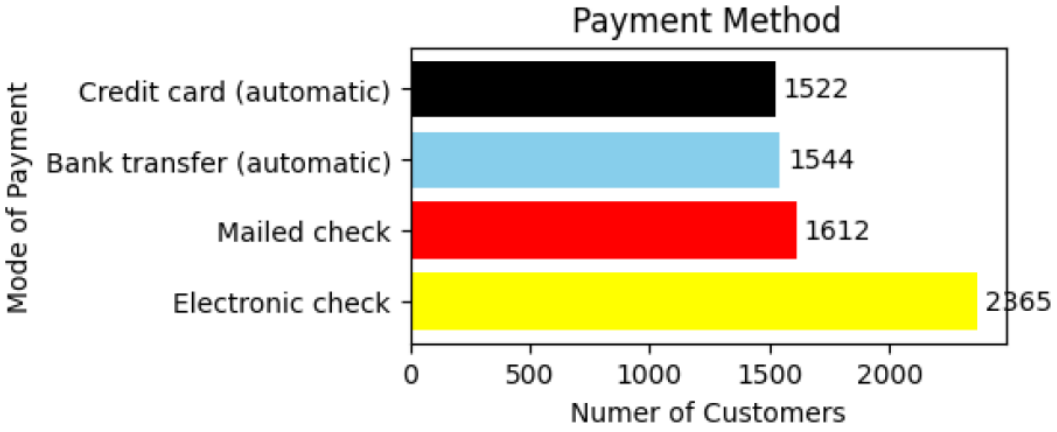
Female Churn Analysis:

- Bar chart (Female – Churn Yes / No).
- Churn behaviour among female customers.
- Similar to males, retention is higher than churn.
- Female churn count is close to male churn count.
- Churn is not heavily gender-biased**.**
- Gender alone may not be a strong predictor but still relevant in combination.



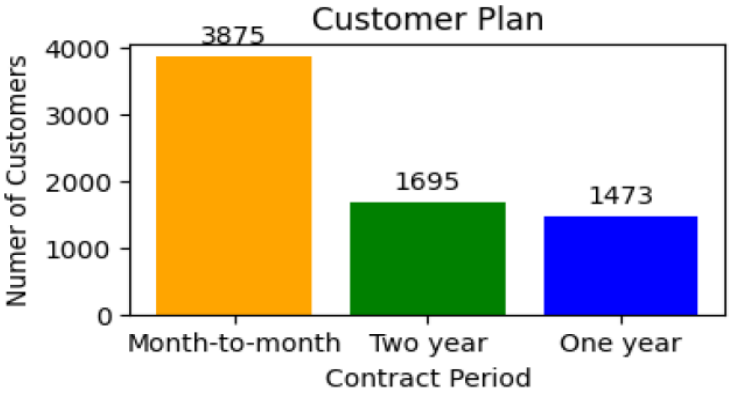
Payment Method Distribution:

- Horizontal bar chart.
- Number of customers using different payment methods.
- Electronic check is the most used payment method.
- Automatic payments (bank transfer, credit card) are lower.
- Customers using manual payment methods are often linked to higher churn.



Contract Type (Customer Plan):

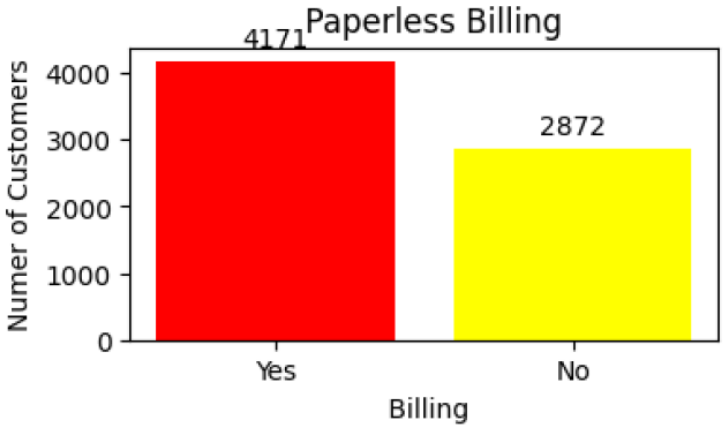
- Bar Chart
- Customers by contract duration.
- Month-to-month contracts dominate.



- Long-term contracts are fewer.
- Long-term contracts strongly improve retention.

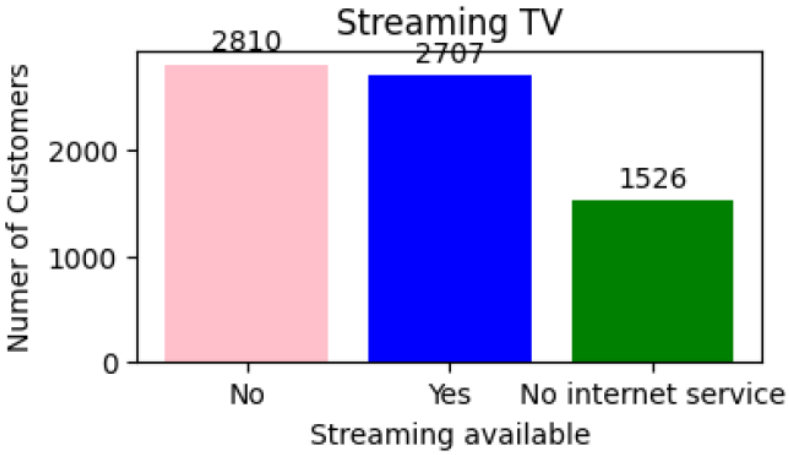
Paperless Billing:

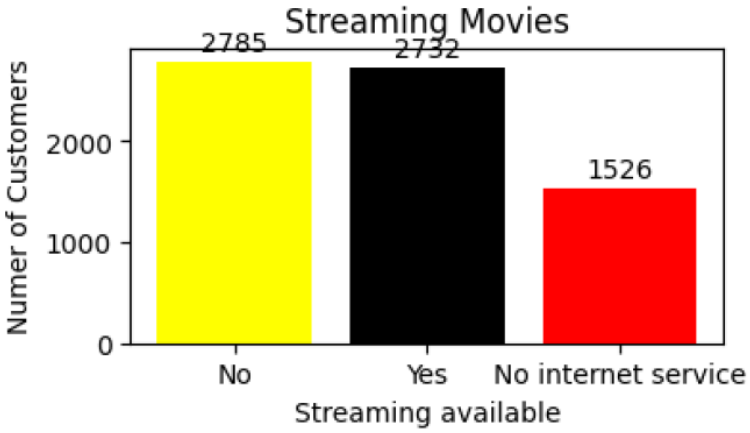
- Bar chart
- Customers using paperless vs non-paperless billing.
- Most customers prefer paperless billing.
- Digitally engaged customers are higher.
- Paperless billing customers are often more tech-savvy and potentially more loyal.



Streaming Services (TV & Movies):

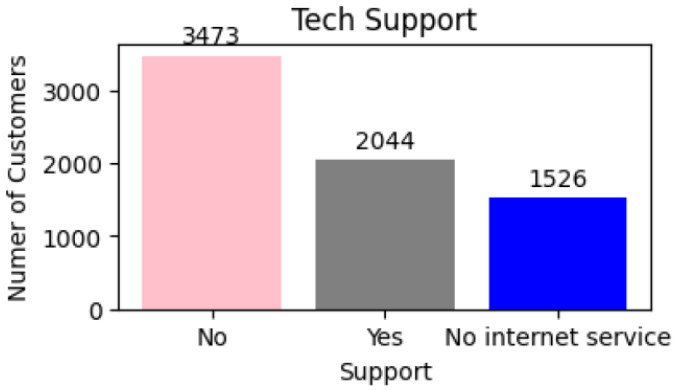
- Bar Charts
- Customers with Streaming TV, without it, or without internet service.
- Almost equal number of users have and don't have streaming TV.
- Bundling streaming may reduce churn.
- Customers without streaming services are slightly higher.
- Cross-selling entertainment services can increase engagement and retention.

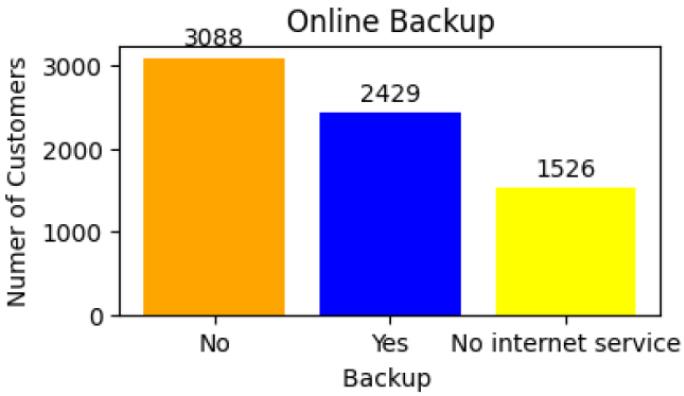


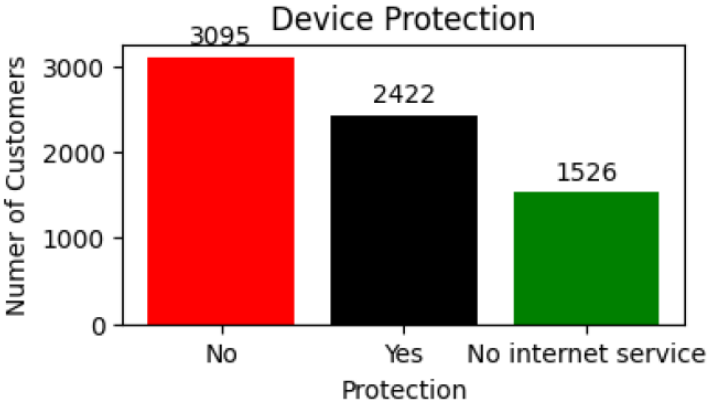


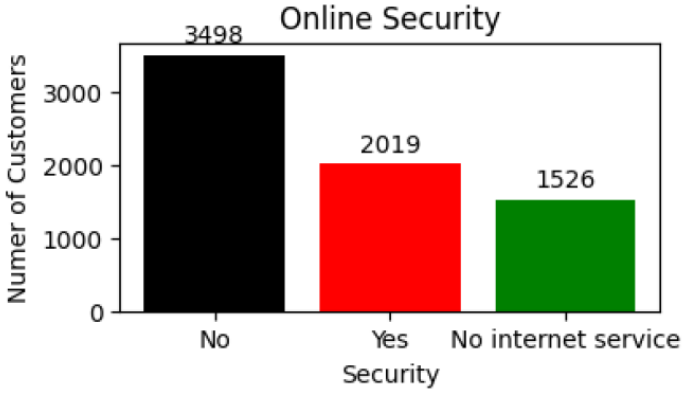
Tech Support, Online Security, Backup & Device Protection:

- Bar Charts
- Usage of value-added services.
- A large number of customers do not opt for add-on services.
- Customers with more services tend to stay longer.
- Many customers do not have tech support.
- Offering affordable tech support plans may improve retention.
- Customers without protection may churn after device issues.
- Security-conscious customers are usually long-term users.



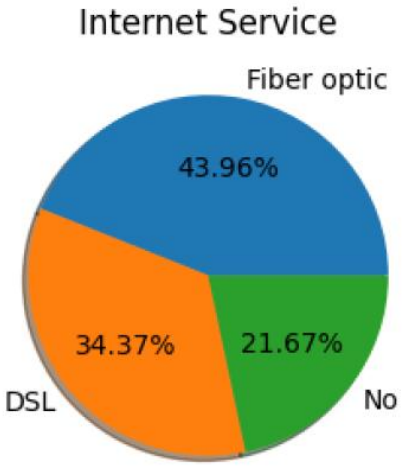


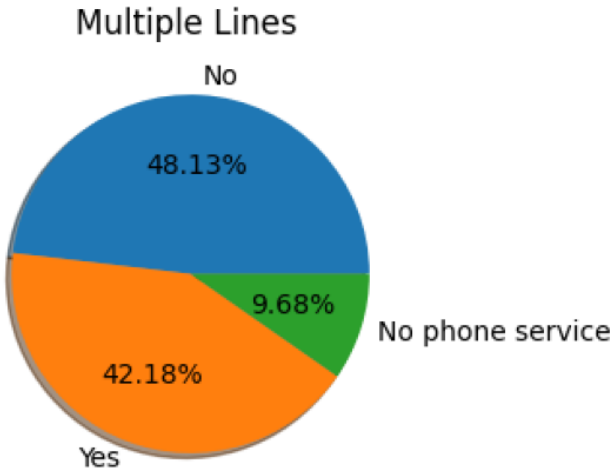




Internet Service & Multiple Lines:

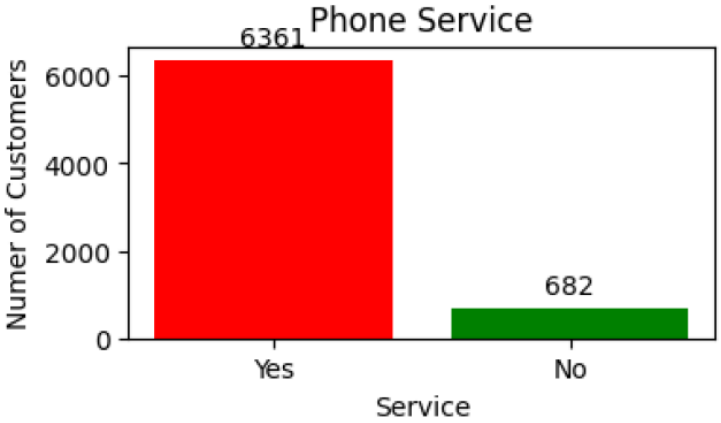
- Pie Charts
- Proportion of customers by internet service and phone line usage.
- Distribution across internet service types.
- Internet type directly impacts service usage and churn.
- Multiple-line customers are more invested.
- Fiber users usually show lower churn.





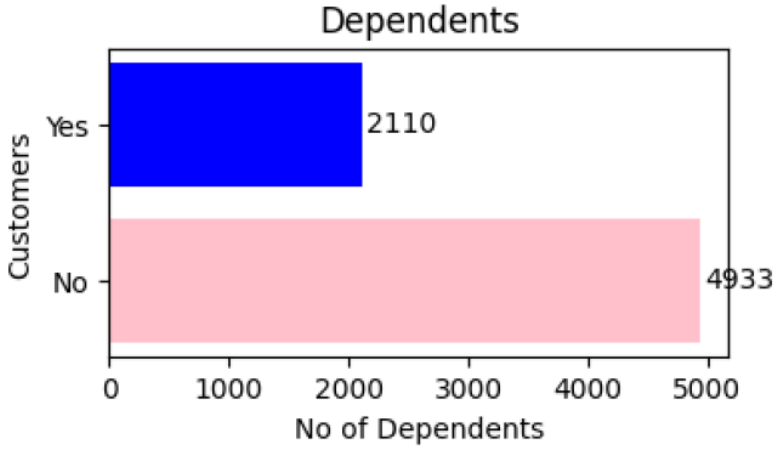
Phone Service:

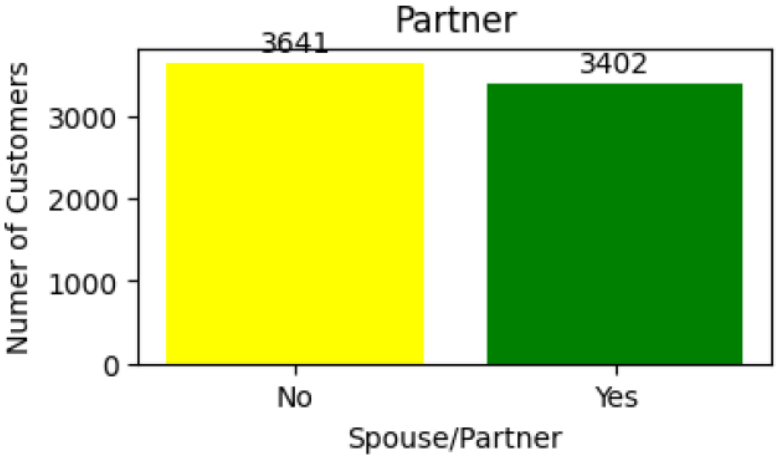
- Bar chart
- Customers with and without phone service.
- Vast majority have phone service.
- Phone service is core; churn depends more on add-ons.



Partner & Dependents:

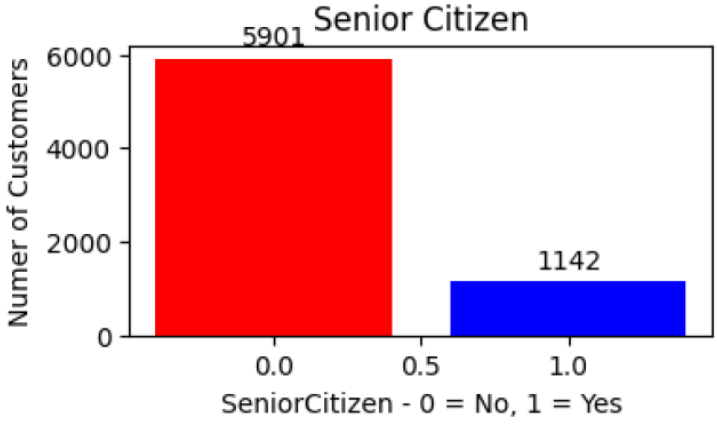
- Bar Charts
- Family status of customers.
- Customers with and without dependents.
- Customers with partners.
- Partnered customers are generally more stable.
- Customers with dependents tend to churn less.
- Family-based plans can be designed to improve retention.



- 

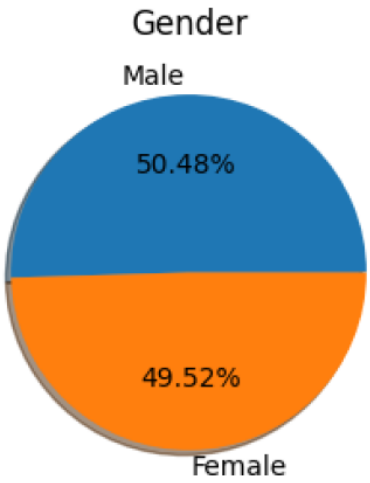
Senior Citizen Analysis:

- Bar Chart
- Senior vs non-senior customers.
- Senior citizens are fewer.
- Age-specific plans may help retention.
- Targeted plans for senior citizens may reduce churn.
- Seniors often have different churn drivers.



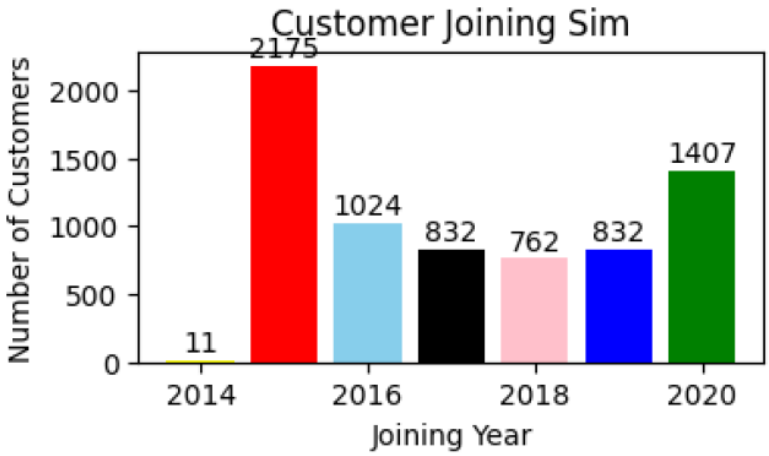
Gender Distribution:

- Pie Chart
- Overall gender proportion.
- Almost balanced gender representation.
- Gender should be combined with other features for prediction.



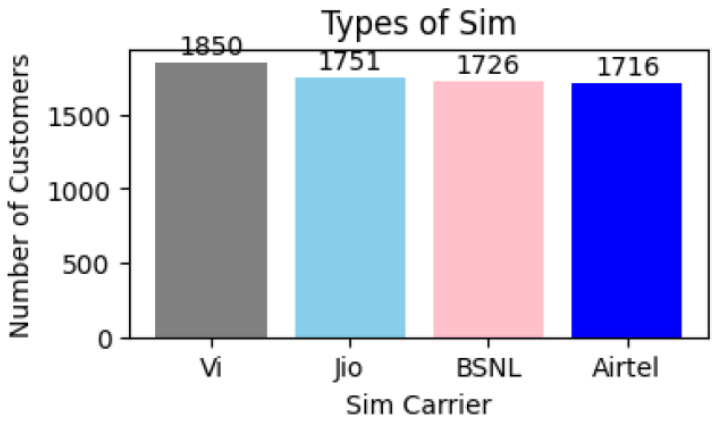
Customer Joining Year:

- Bar chart
- Certain years saw higher customer acquisition.
- Customer acquisition varies by year.
- Helps evaluate marketing effectiveness over time.
- Helps track churn trends by joining period.
- Useful for cohort analysis.



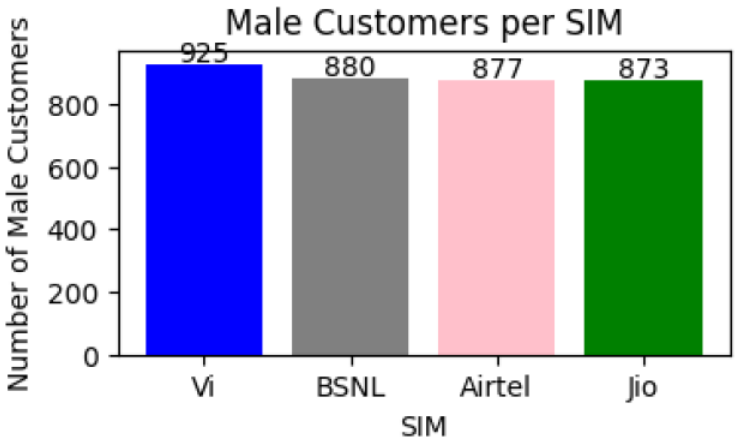
SIM-wise Customer Distribution:

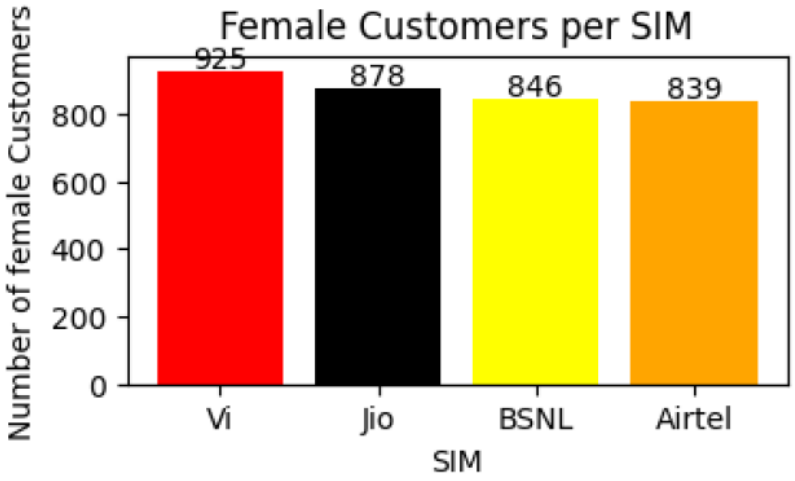
- Bar Chart
- Customer count per SIM provider.
- Some SIM providers have higher customer bases.
- Provider-wise strategy optimization is possible.
- Useful for targeted retention strategies.



Male & Female Customers per SIM:

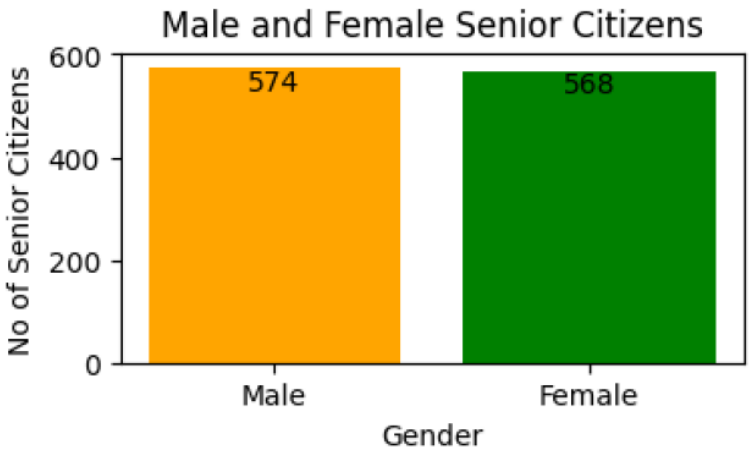
- Bar Charts
- Gender distribution across SIMs.
- Gender distribution is consistent across providers.
- No gender bias in SIM preference.





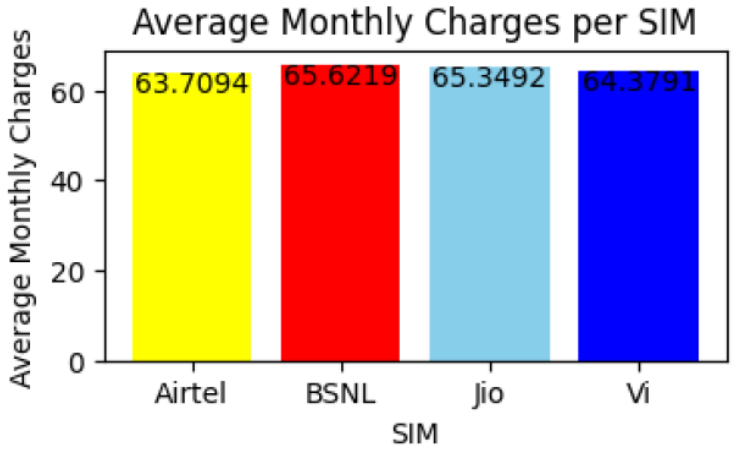
Senior Citizens by Gender:

- Male vs female senior citizens.
- Slight gender variation among seniors.



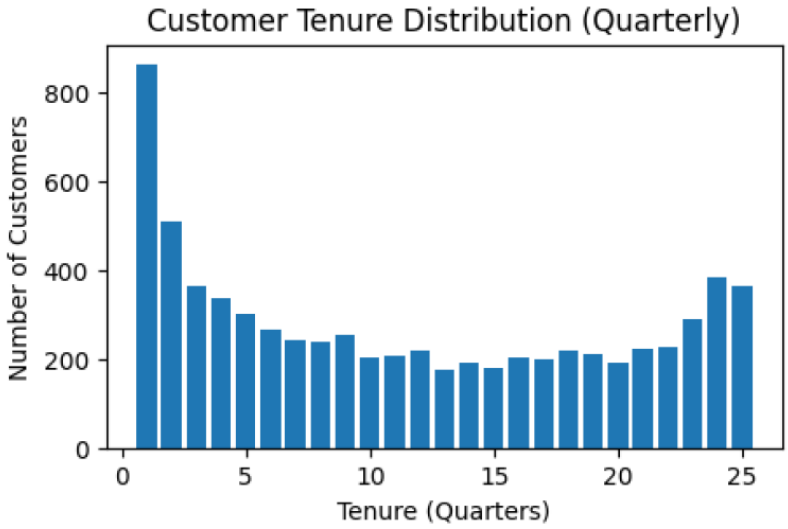
Average Monthly Charges per SIM:

- Bar Chart
- Average billing per SIM provider.
- Certain SIMs generate higher revenue.
- Pricing differs across providers.
- High-charge SIMs may have higher churn risk.



Quarterly Tenure Distribution:

- Bar chart
- Customer tenure grouped by quarters.
- Most churn occurs in early quarters.
- First year is critical for churn prevention.
- Early-stage retention strategies are critical.
- Many customers are in early quarters.



Conclusion:

The visual analysis clearly demonstrates that customer churn is strongly influenced by factors such as contract duration, tenure length, payment method, monthly charges, and value-added services. Customers with short tenure, month-to-month contracts, manual payment methods, and fewer subscribed services exhibit a higher tendency to churn, whereas customers with long-term contracts, bundled services, multiple lines, and family responsibilities show higher retention levels.

These insights not only validate the importance of feature selection for churn prediction but also provide actionable business intelligence. The findings can help organizations design **targeted retention strategies**, optimize pricing models, improve service bundling, and enhance customer engagement during critical early stages of the customer lifecycle.