

# Exploratory Data Analysis (EDA) Using Matplotlib

## Churn Prediction Project Documentation

### 1. Objective of the Project

The primary objective of this Exploratory Data Analysis (EDA) is to understand customer churn behavior in a telecom dataset by analyzing how demographic factors, services, contracts, billing, and usage patterns influence customer churn. The analysis aims to:

- Identify key factors associated with customer churn
  - Understand service usage patterns across different demographics
  - Compare behavior across SIM operators
  - Derive business insights that can help reduce churn and improve customer retention
- 

### 2. Tools and Technologies Used

- **Programming Language:** Python
  - **Libraries:**
    - Pandas – data manipulation and analysis
    - NumPy – numerical operations
    - Matplotlib – data visualization
  - **Environment:** Google Colab / Jupyter Notebook
  - **Dataset Format:** CSV file (`Churn_Updated.csv`)
- 

### 3. Dataset Description

The dataset contains customer-level information from a telecom company. Each row represents one customer.

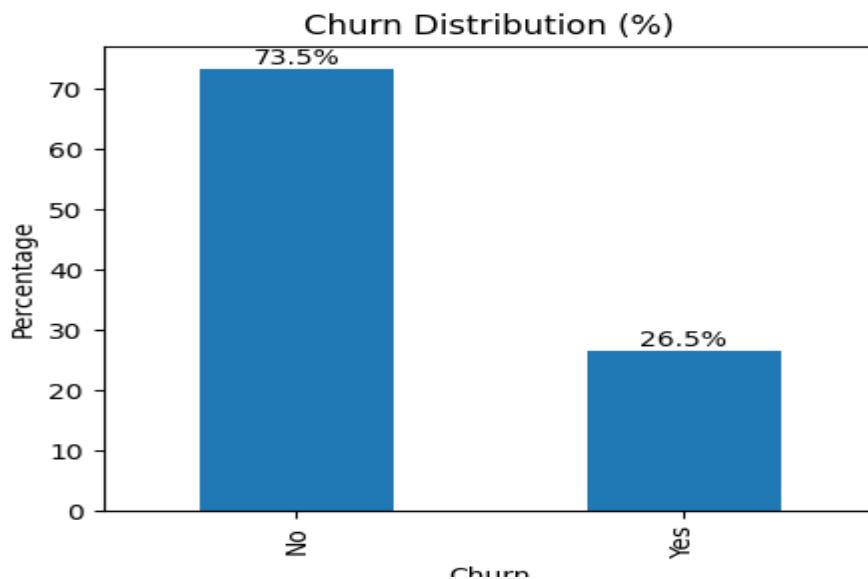
#### Key Columns:

- **Churn:** Whether the customer left the service (Yes / No)
- **gender:** Male / Female
- **SeniorCitizen:** Whether the customer is a senior citizen (0 / 1)
- **SIM:** SIM operator used by the customer
- **InternetService:** DSL / Fiber optic / No
- **PhoneService:** Yes / No
- **MultipleLines:** Yes / No
- **OnlineSecurity, OnlineBackup, DeviceProtection, TechSupport, StreamingTV, StreamingMovies:** Service subscriptions
- **Contract:** Month-to-month / One year / Two year
- **PaperlessBilling:** Yes / No
- **PaymentMethod:** Mode of payment
- **tenure:** Number of months the customer stayed
- **MonthlyCharges:** Monthly bill amount
- **Region:** Customer location

## 4. Visualizations, Graph Type & Understanding from Charts

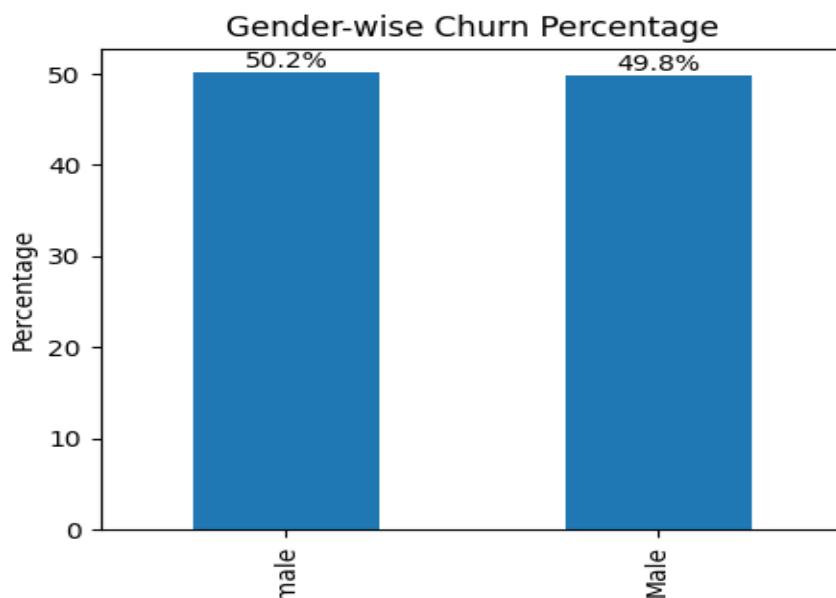
### 1. Churn Distribution

- **Graph Used:** Bar chart
- **Understanding:** The chart shows that while most customers stay with the company, a noticeable percentage leave. This indicates churn is a real business problem and cannot be ignored.



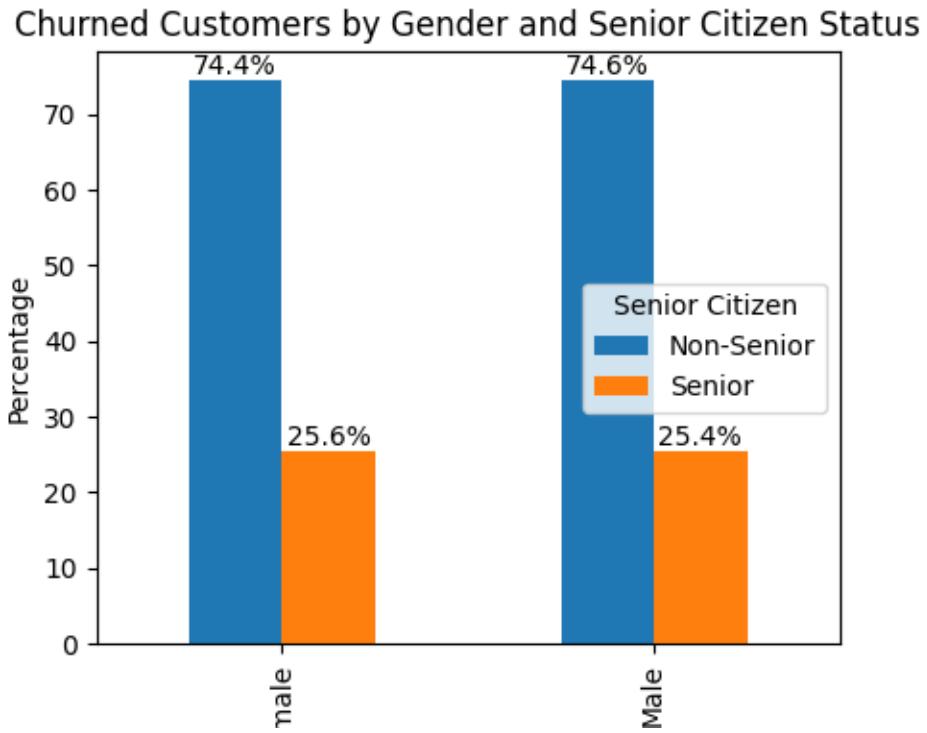
### 2. Gender-wise Churn Distribution

- **Graph Used:** Bar chart
- **Understanding:** Both male and female customers contribute to churn. The difference between genders is not extreme, indicating gender alone is not a strong churn driver but still useful when combined with other factors.



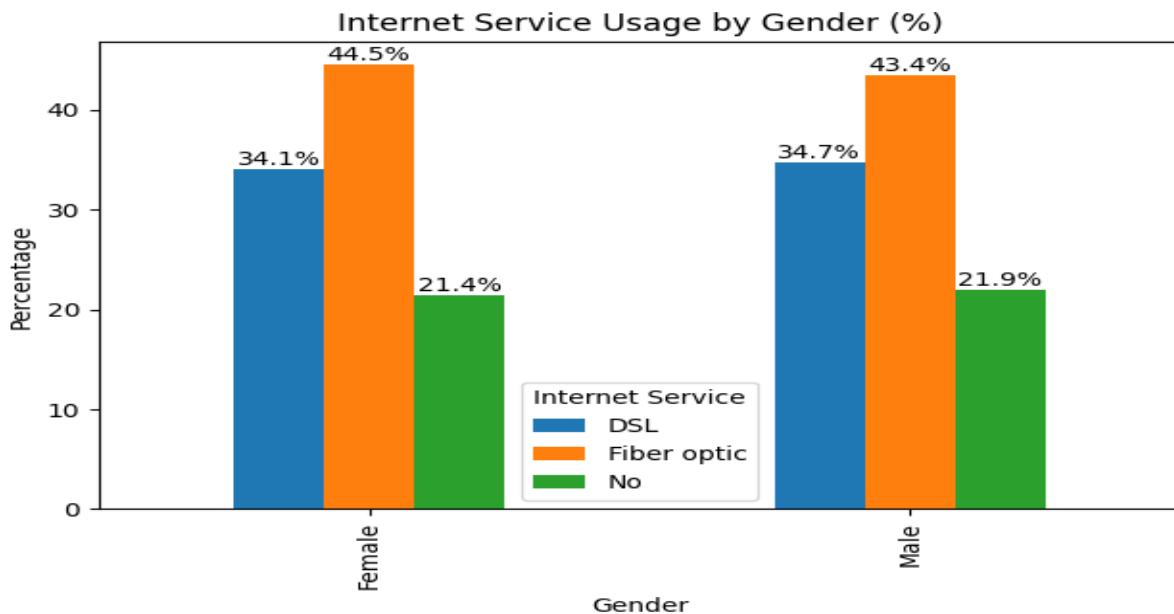
### 3. Churned Customers by Gender & Senior Citizen

- **Graph Used:** Grouped bar chart
- **Understanding:** Senior citizens within both genders churn more than non-senior customers, suggesting age-related needs and service expectations are not fully met.



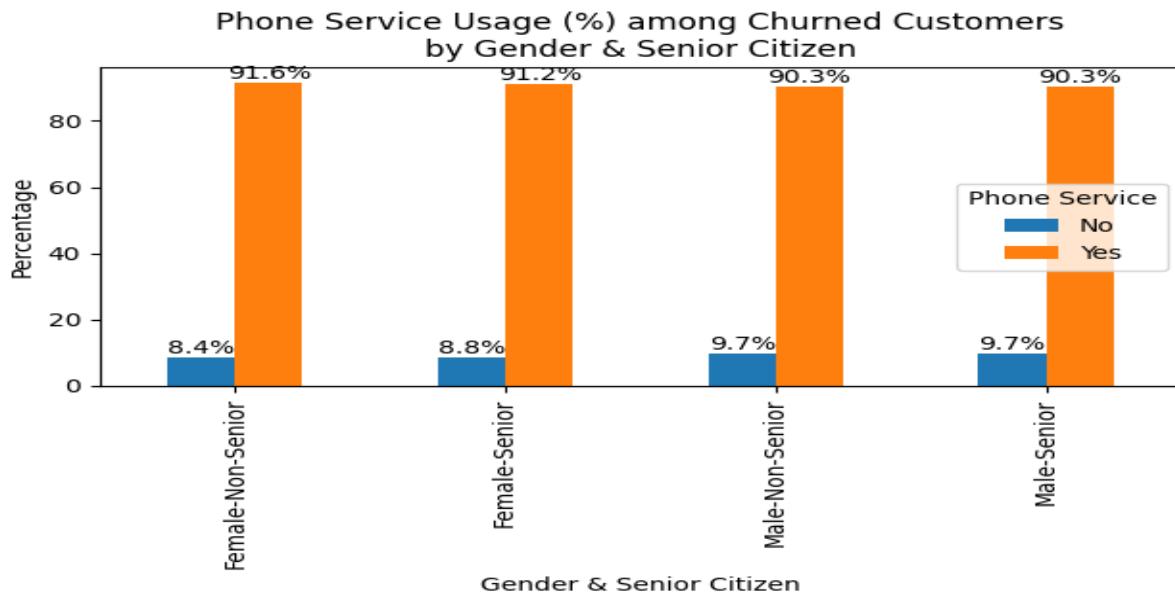
### 4. Internet Service Usage by Gender

- **Graph Used:** Bar chart (percentage)
- **Understanding:** Fiber optic service is the most commonly used internet service across genders. However, reliance on high-speed services may influence dissatisfaction and churn.



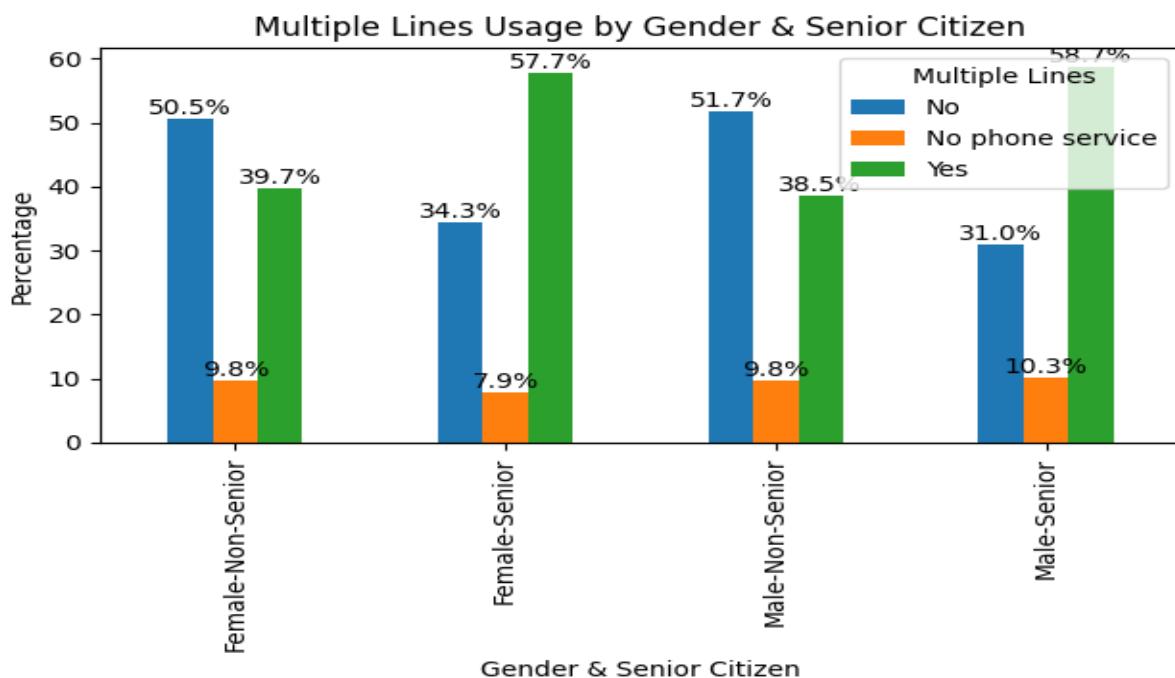
## 5. Phone Service Usage by Gender & Senior Citizen (Churned Customers)

- **Graph Used:** Bar chart
- **Understanding:** Most churned customers still use phone services, indicating churn is not due to lack of basic services but possibly due to pricing or support issues.



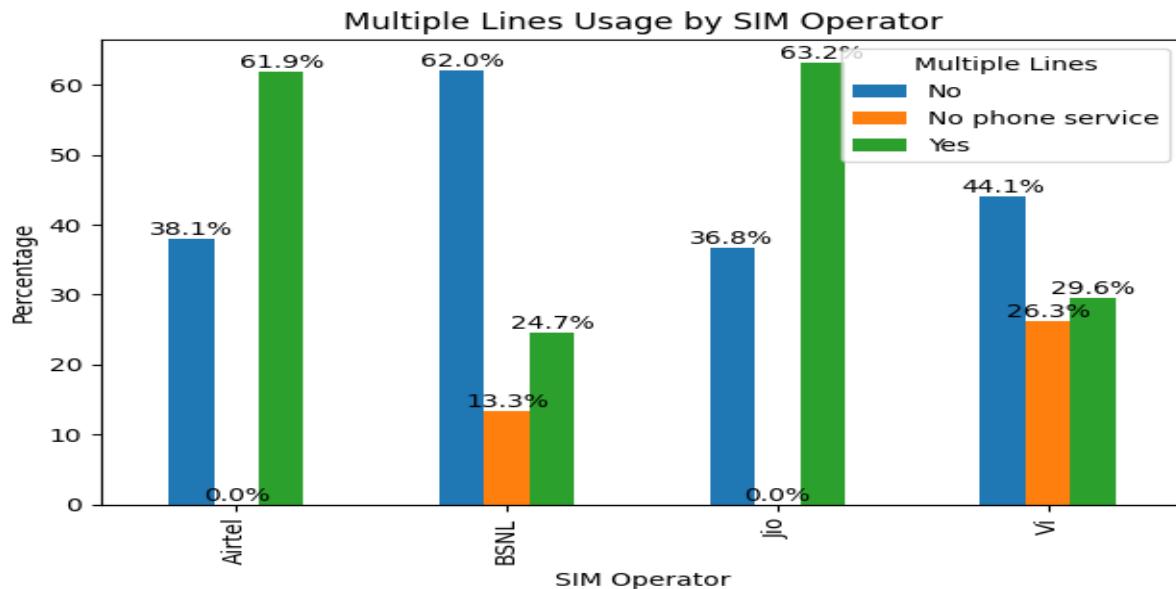
## 6. Multiple Lines Usage by Gender & Senior Citizen

- **Graph Used:** Bar chart
- **Understanding:** Customers with multiple lines are fewer among senior citizens, indicating simpler service needs for this group.



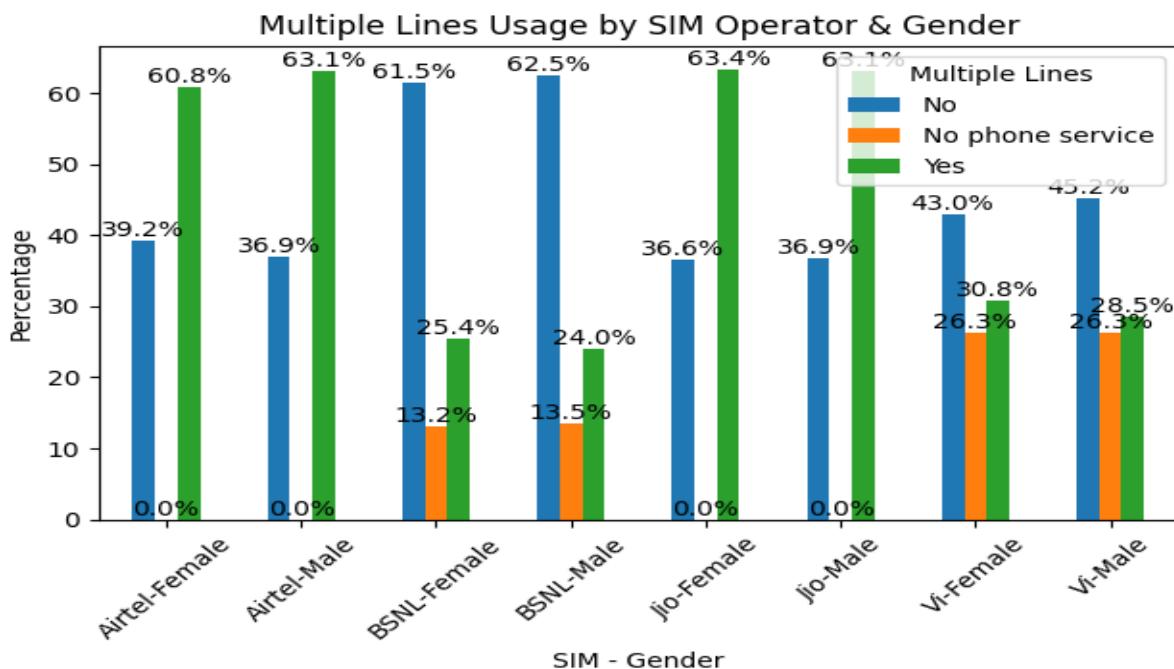
## 7. Multiple Lines Usage by SIM Operator

- **Graph Used:** Bar chart
- **Understanding:** Multiple line adoption varies by SIM operator, showing differences in plan offerings and customer preferences.



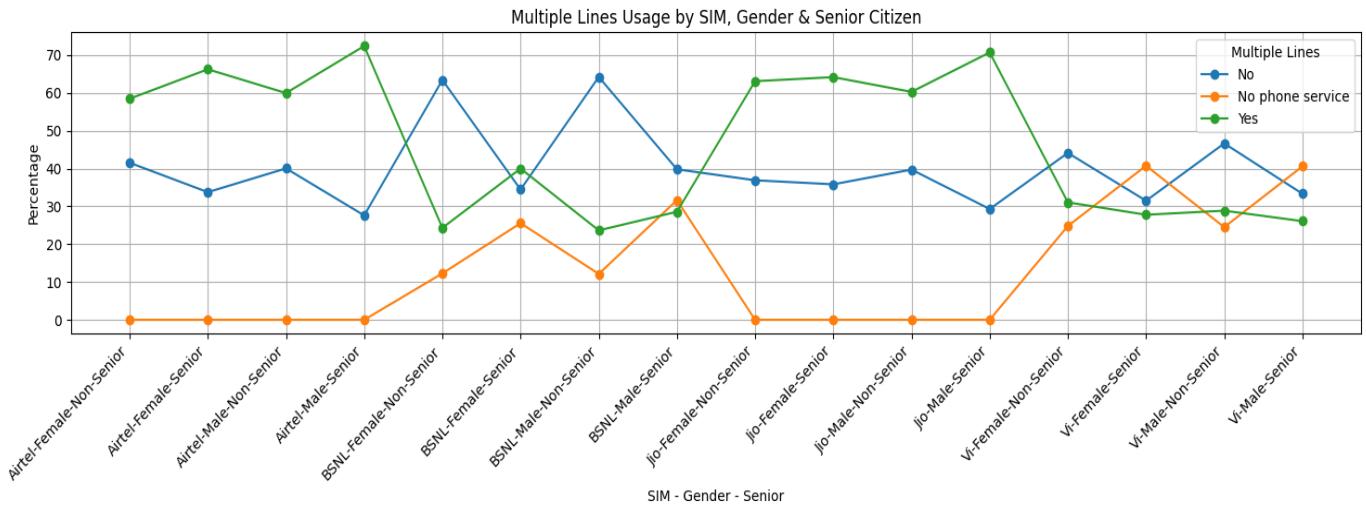
## 8. Multiple Lines Usage by SIM Operator & Gender

- **Graph Used:** Bar chart
- **Understanding:** Gender-based differences in multiple line usage are minor, but SIM operator policies strongly influence adoption.



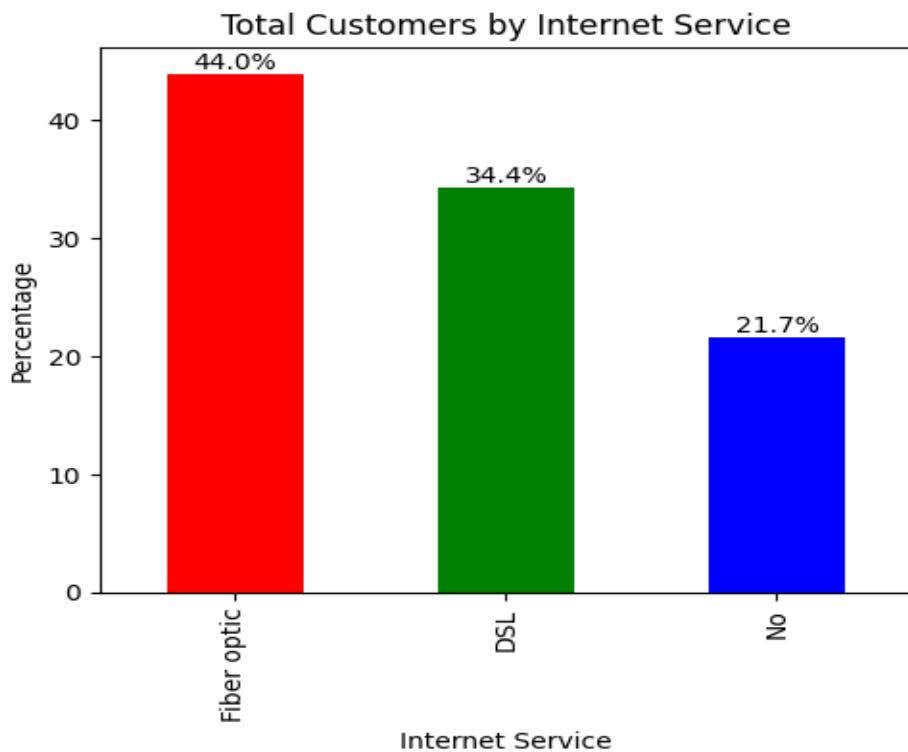
## 9. Multiple Lines Usage by SIM, Gender & Senior Citizen

- **Graph Used:** Line chart
- **Understanding:** This combined view helps identify specific demographic segments where multiple line usage is high or low, supporting targeted marketing strategies.



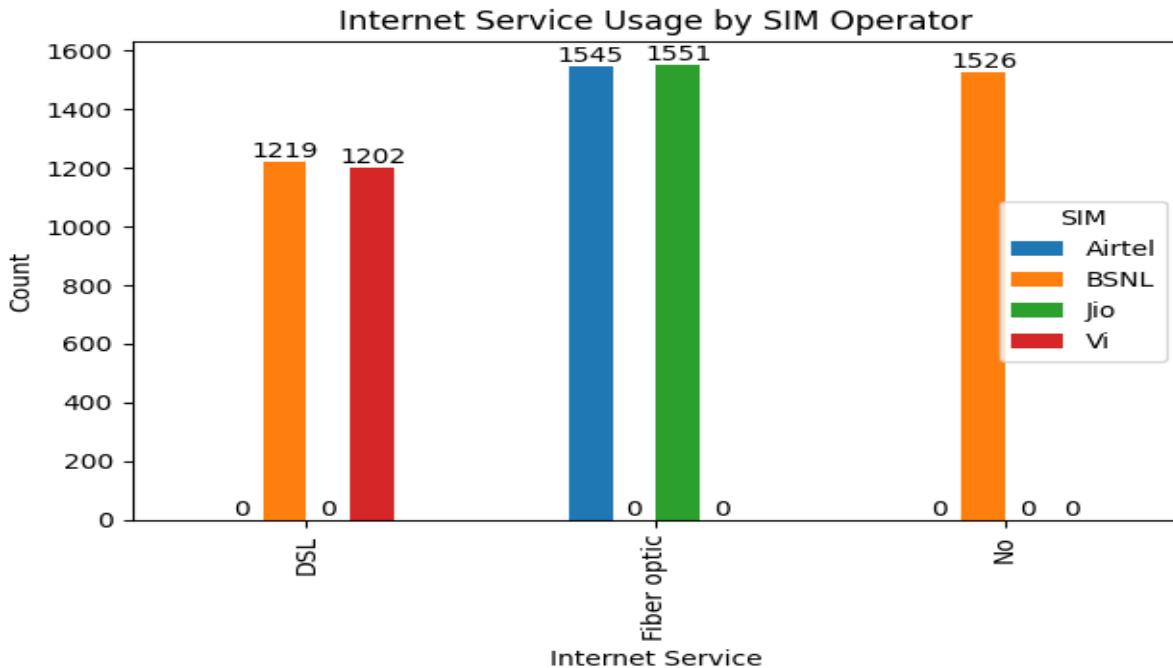
## 10. Total Customers by Internet Service

- **Graph Used:** Bar chart
- **Understanding:** Fiber optic dominates customer subscriptions, indicating a strong demand for high-speed internet services.



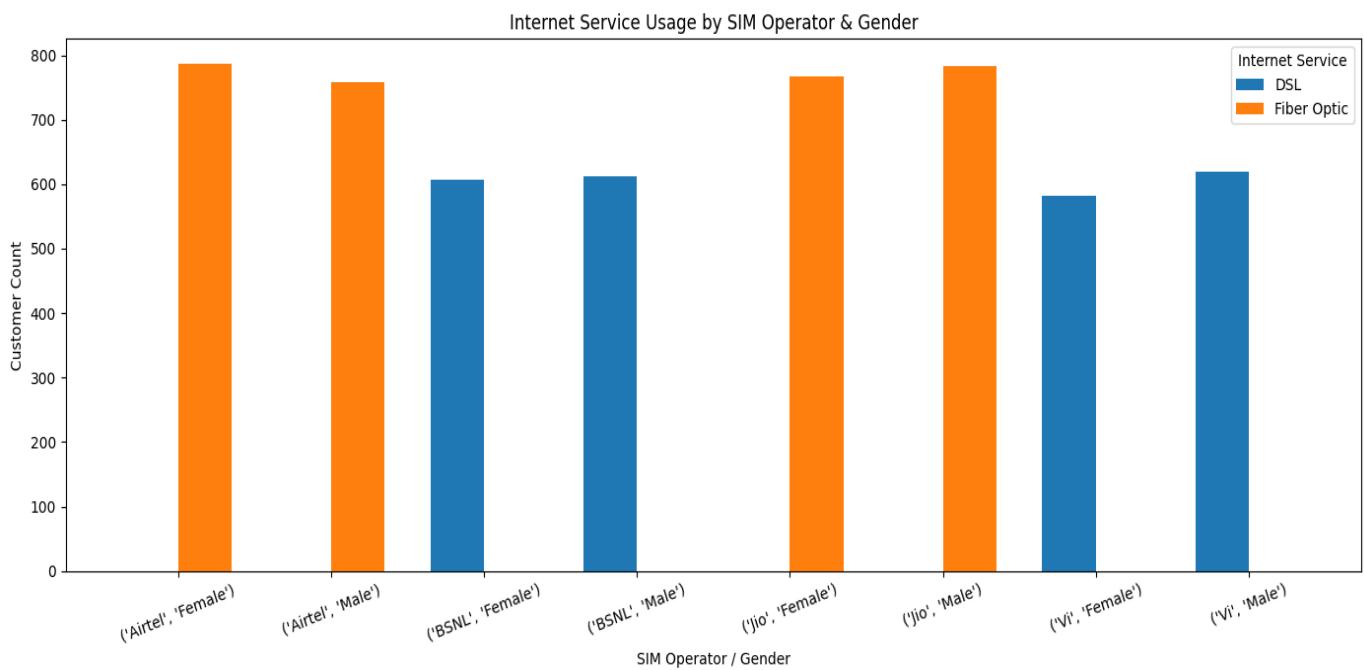
## 11. Internet Service Usage by SIM Operator

- **Graph Used:** Bar chart (count)
- **Understanding:** Different SIM operators specialize in different internet services, affecting customer experience and churn.



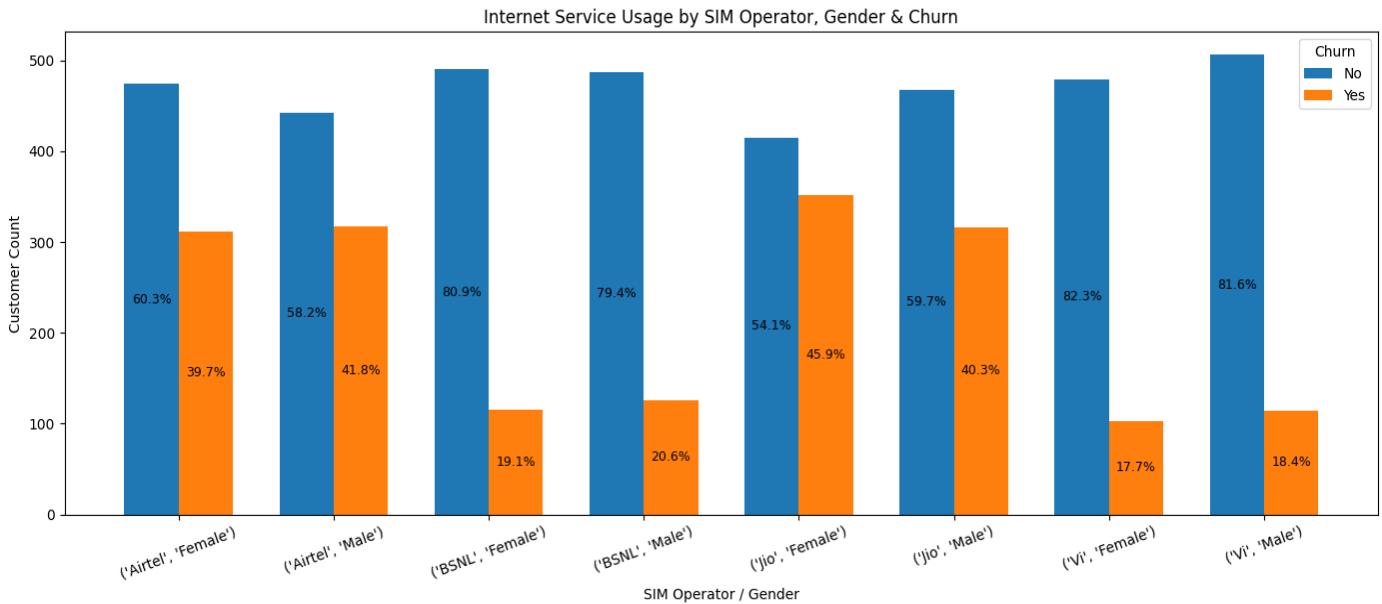
## 12. Internet Service Usage by SIM Operator & Gender

- **Graph Used:** Grouped bar chart
- **Understanding:** Service preference varies slightly across genders, but SIM operator influence remains stronger than gender.



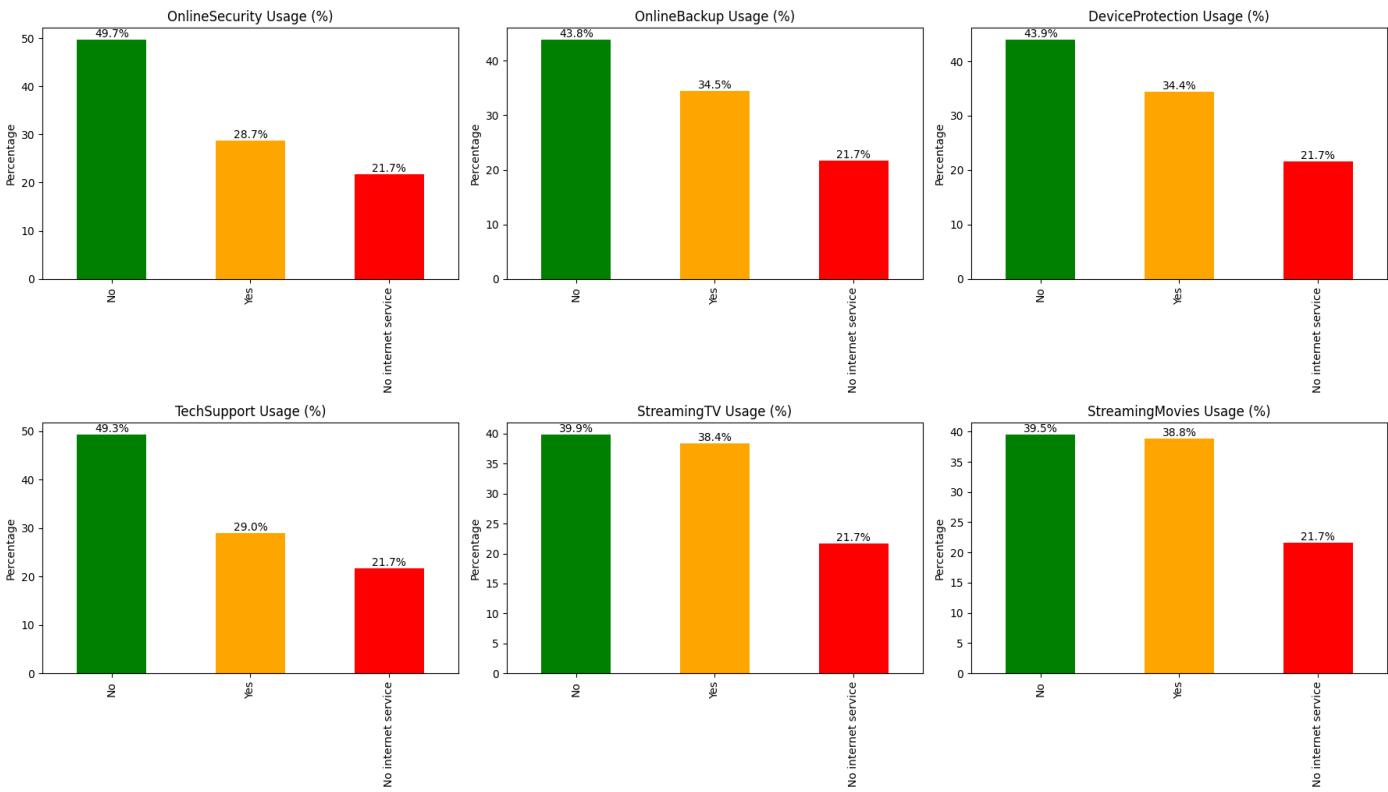
## 13. Internet Service Usage by SIM Operator, Gender & Churn

- Graph Used:** Grouped bar chart
- Understanding:** Certain SIM and gender combinations show higher churn, indicating service quality or pricing differences.



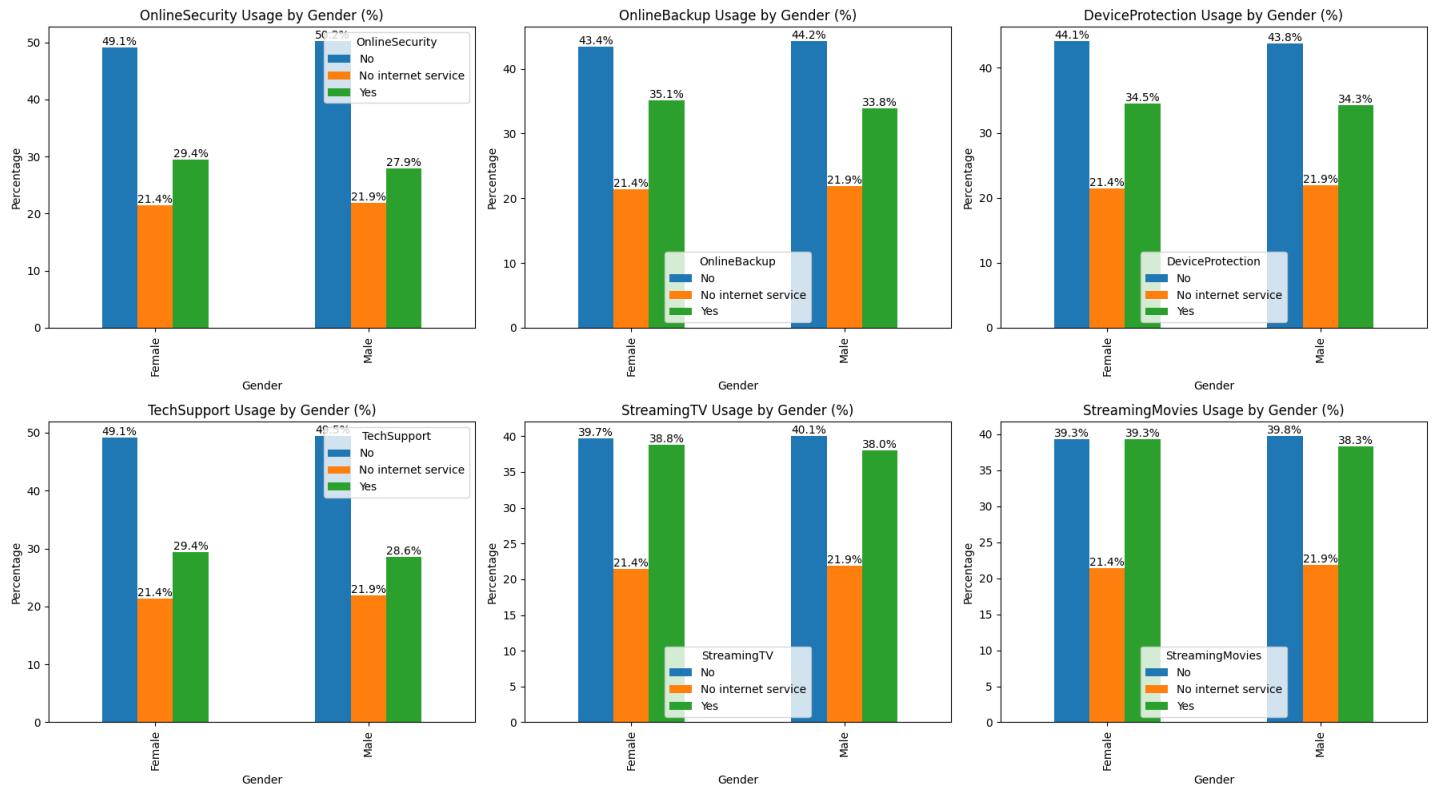
## 14. Service Columns Distribution

- Graph Used:** Subplot bar charts
- Understanding:** Value-added services such as OnlineSecurity and TechSupport have lower adoption, which may increase churn risk.



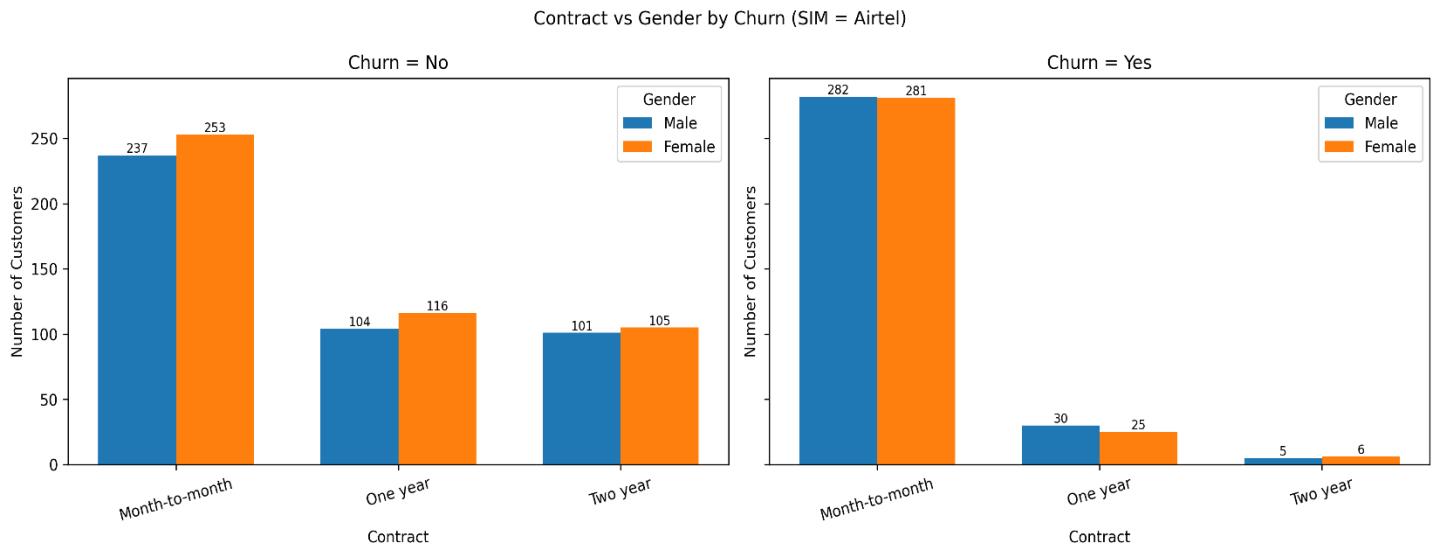
## 15. Service Usage by Gender

- Graph Used:** Subplot grouped bar charts
- Understanding:** Service adoption is similar across genders, showing that awareness and pricing matter more than gender.

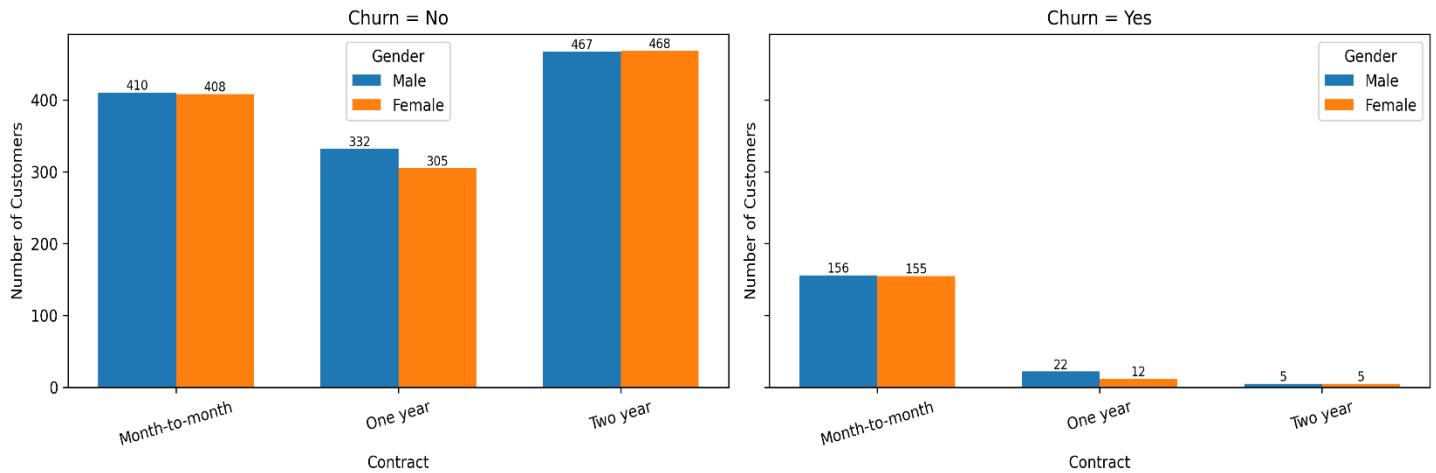


## 16. Contract Distribution by SIM, Gender & Churn

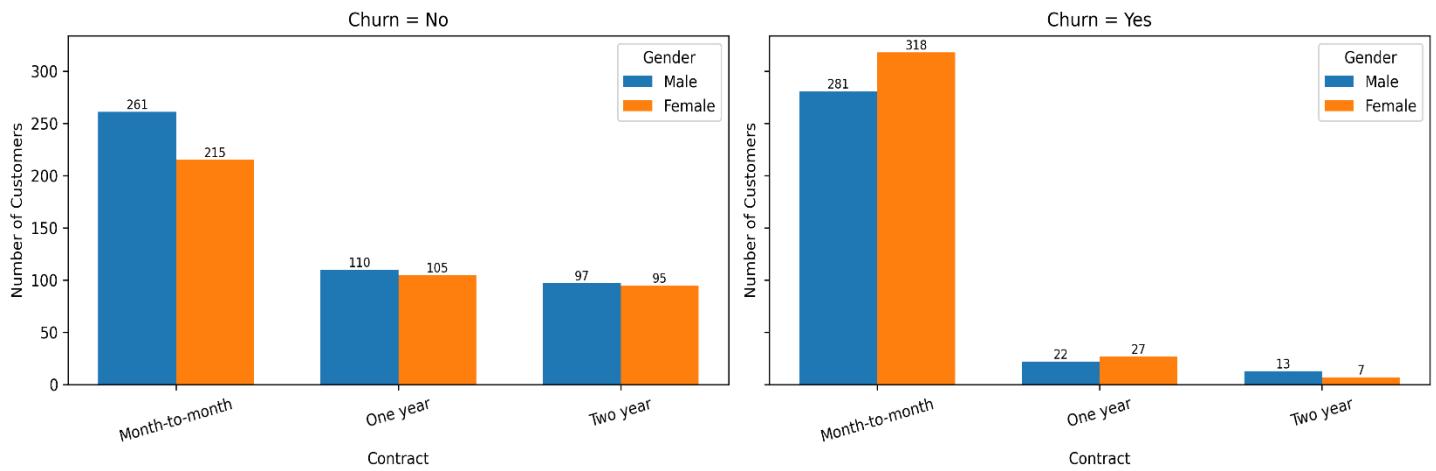
- Graph Used:** Grouped bar charts
- Understanding:** Month-to-month contracts are strongly associated with churn, while long-term contracts improve retention.



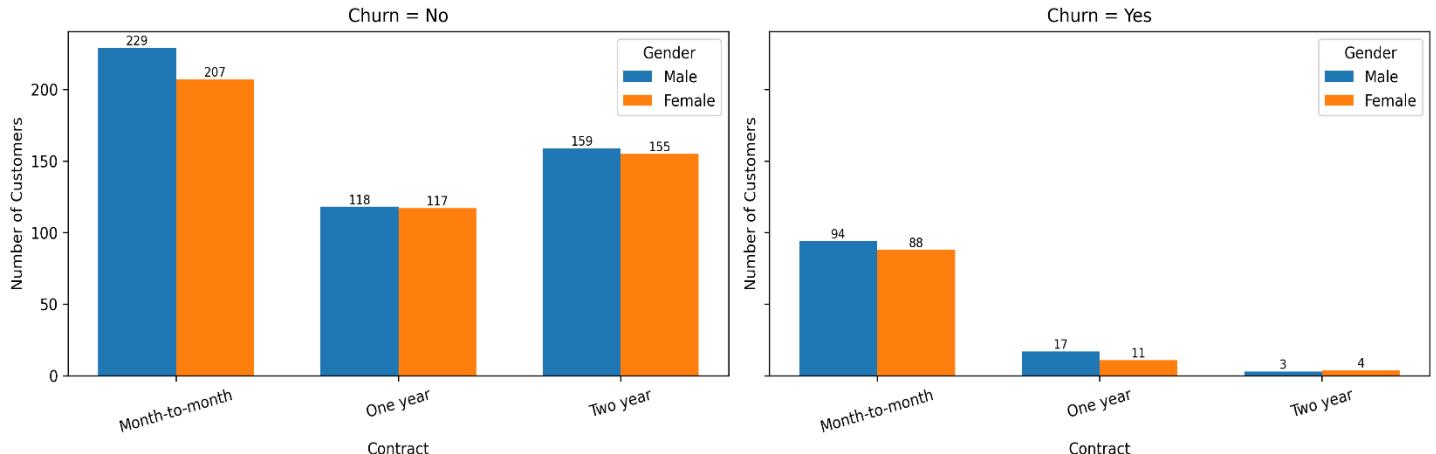
Contract vs Gender by Churn (SIM = BSNL)



Contract vs Gender by Churn (SIM = Jio)

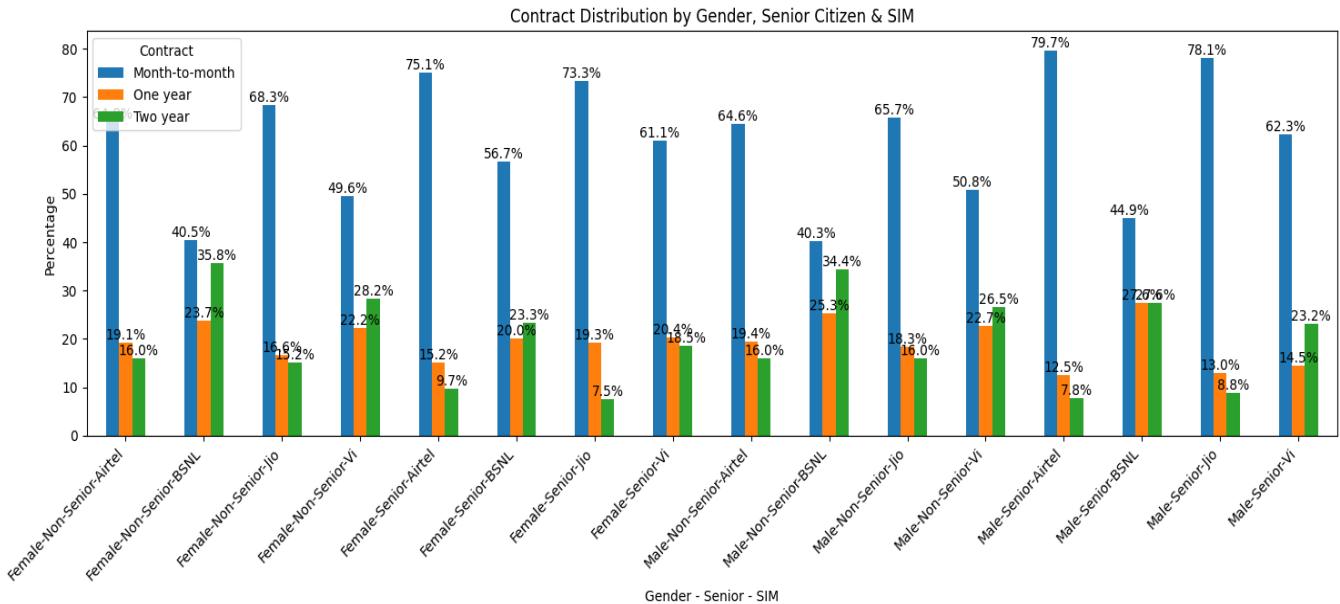


Contract vs Gender by Churn (SIM = Vi)



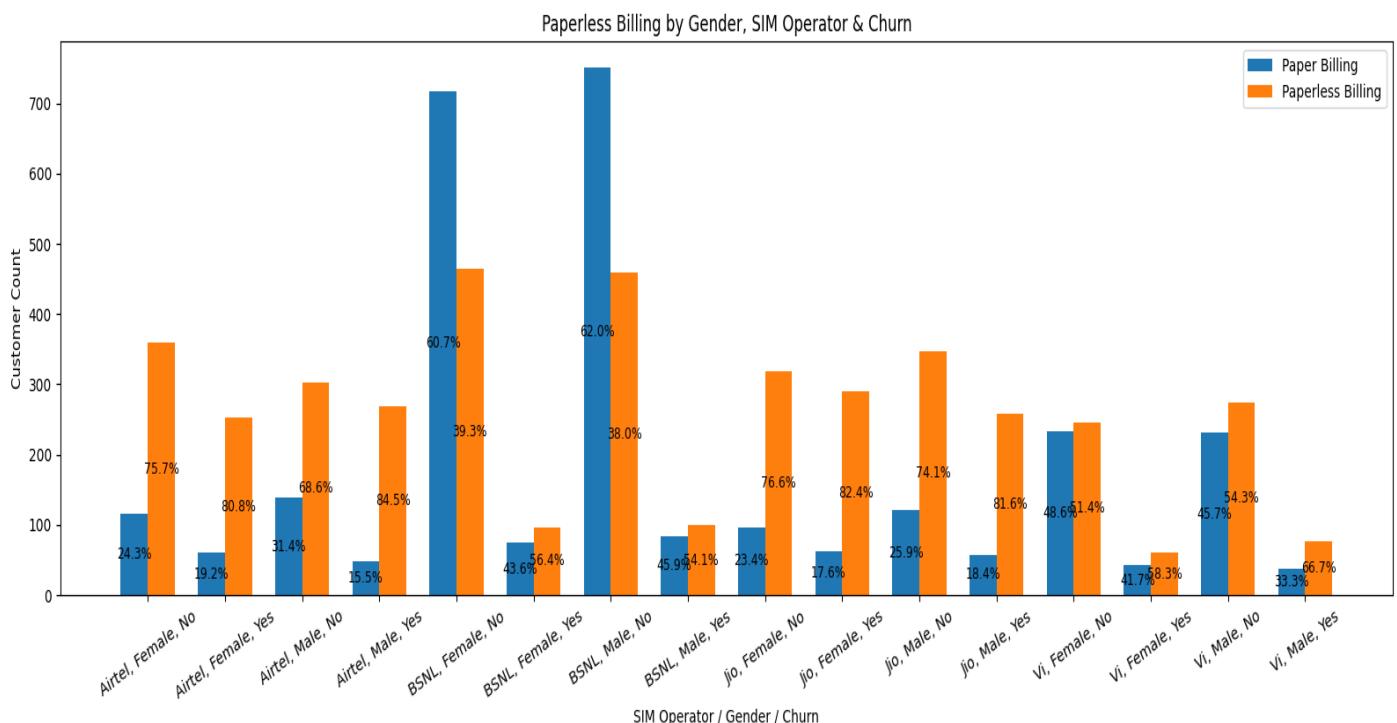
## 17. Contract Distribution by Gender, Senior Citizen & SIM

- Graph Used:** Bar chart (percentage)
- Understanding:** Senior citizens prefer shorter contracts, which aligns with their higher churn levels.



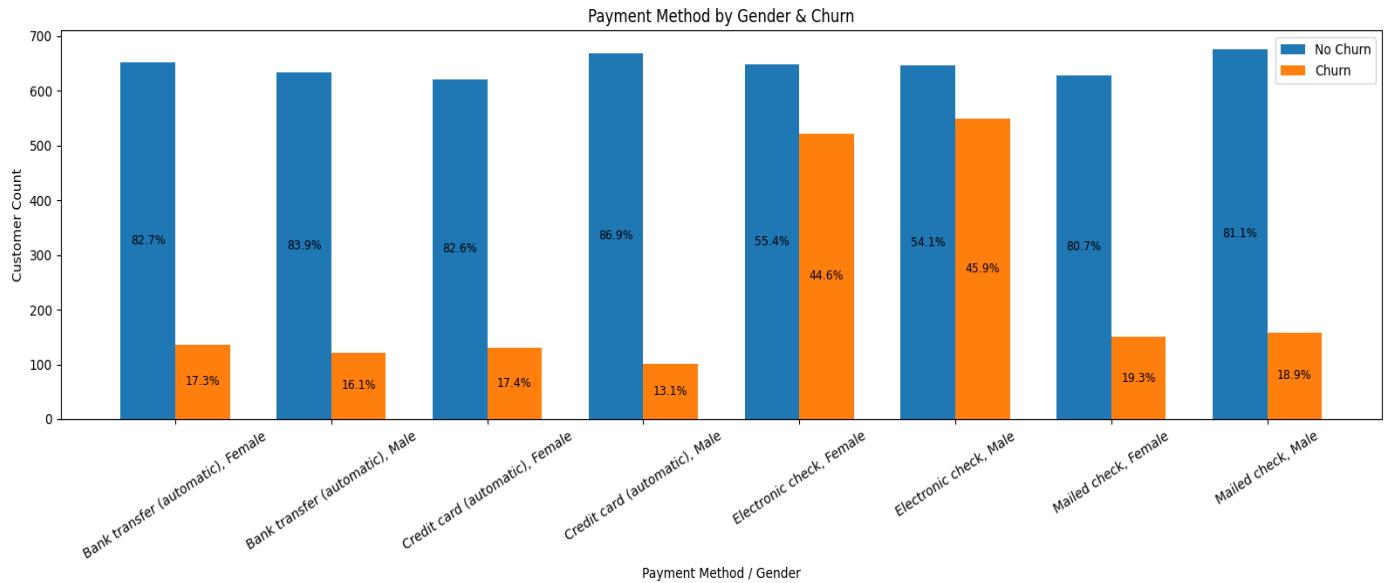
## 18. Paperless Billing by Gender, SIM Operator & Churn

- Graph Used:** Grouped bar chart
- Understanding:** Customers using paperless billing churn more, indicating digitally active but less loyal users.



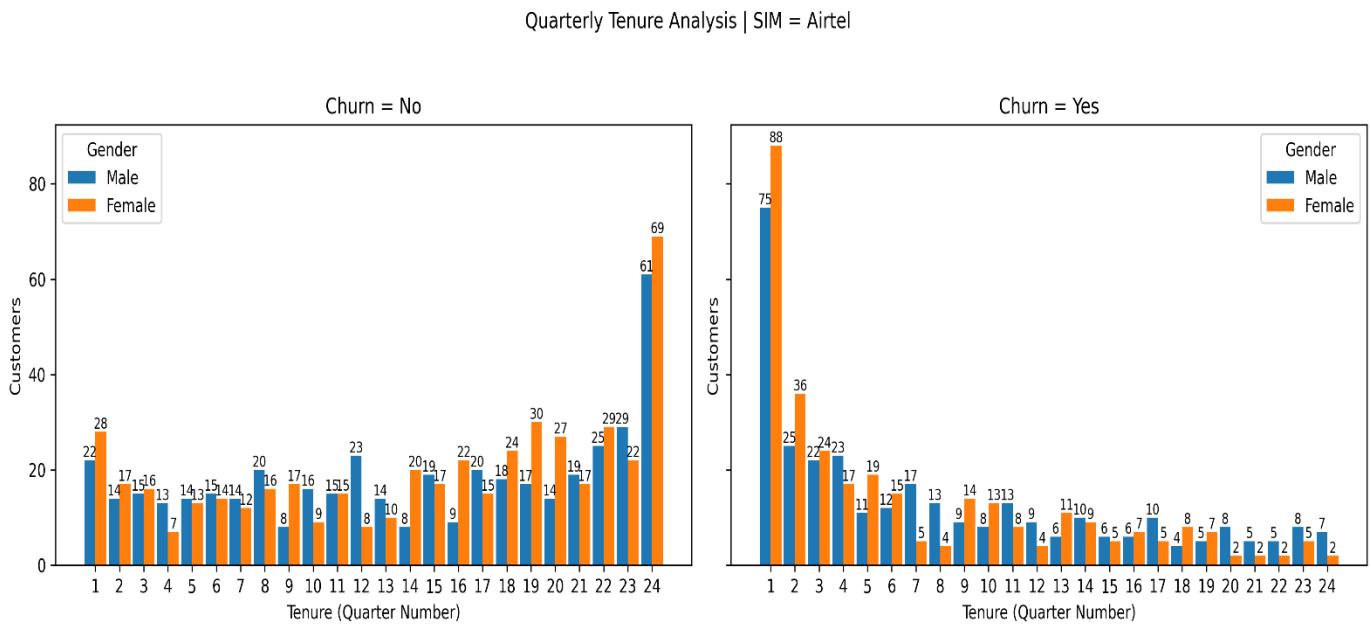
## 19. Payment Method by Gender & Churn

- Graph Used:** Grouped bar chart
- Understanding:** Electronic payment methods show higher churn, whereas automatic payments improve retention.

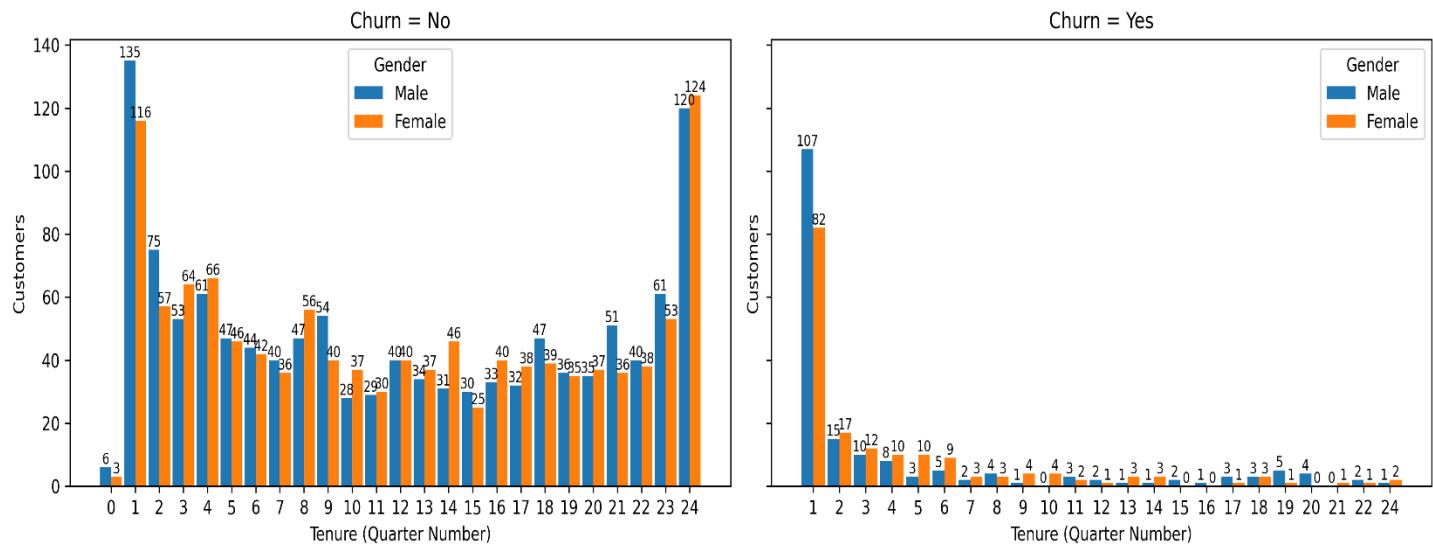


## 20. Quarterly Tenure Analysis with SIM

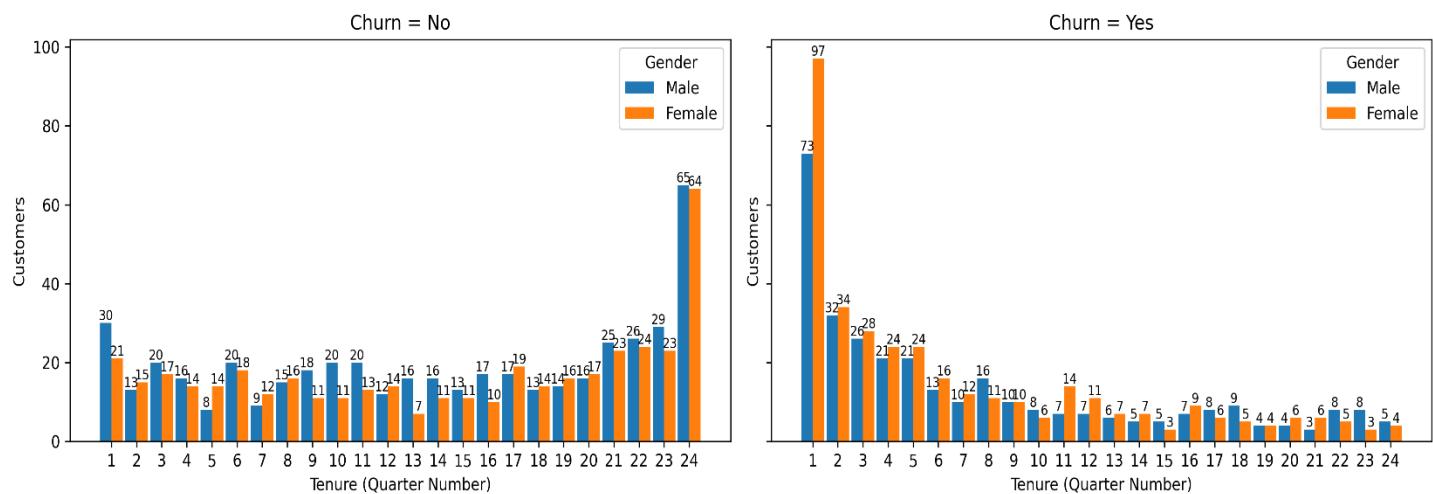
- Graph Used:** Grouped bar charts
- Understanding:** Most churn occurs in early tenure periods, emphasizing the importance of onboarding experience.



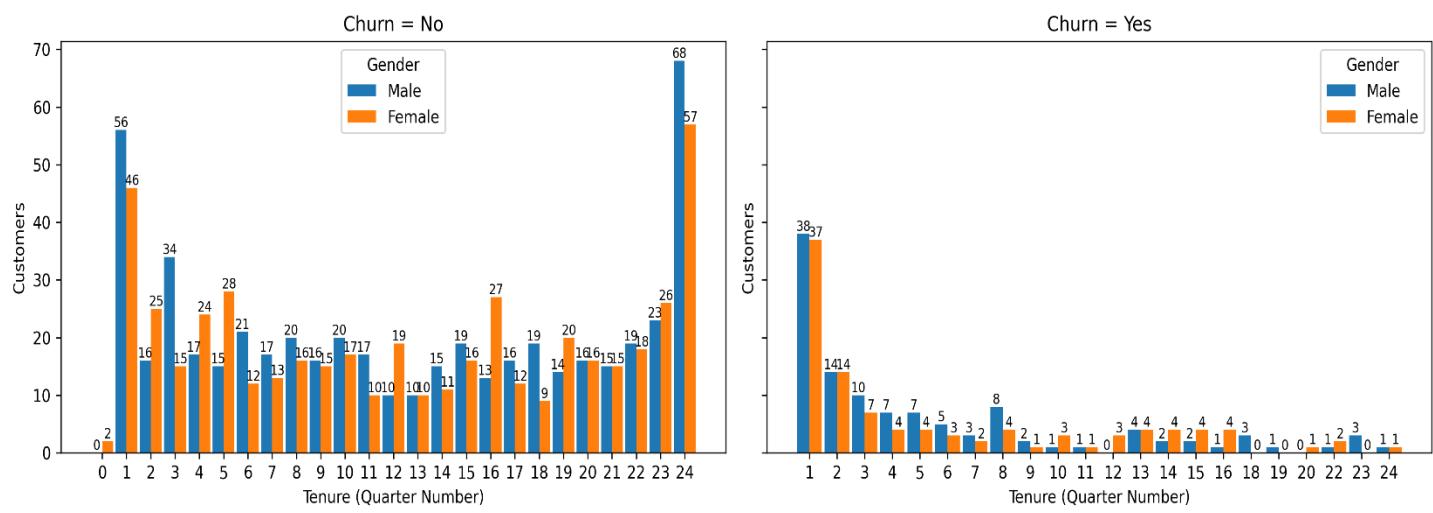
### Quarterly Tenure Analysis | SIM = BSNL



### Quarterly Tenure Analysis | SIM = Jio

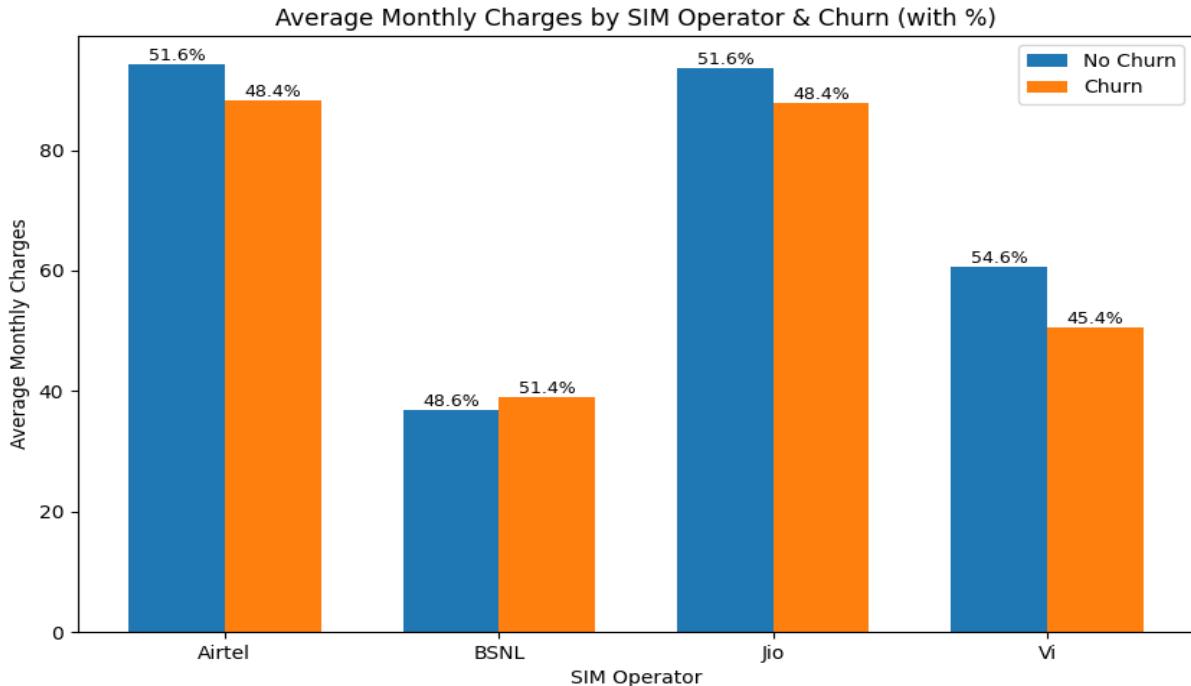


### Quarterly Tenure Analysis | SIM = Vi



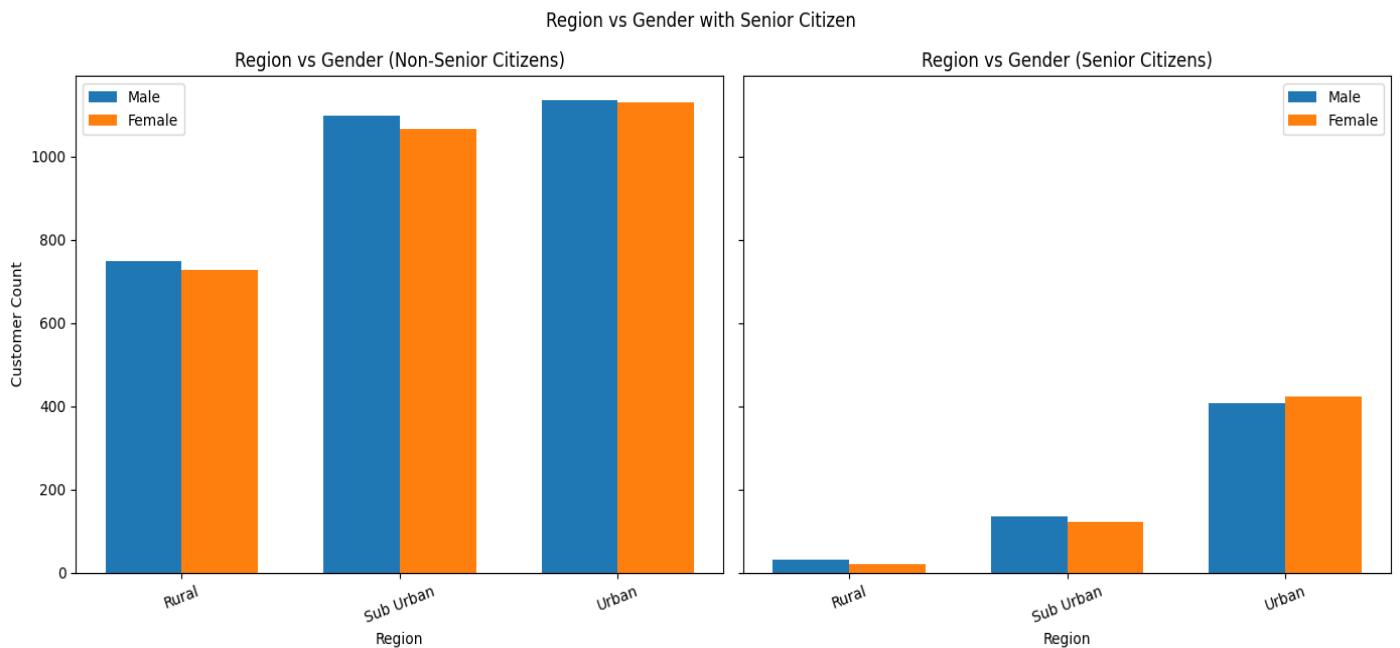
## 21. Monthly Charges vs SIM Operator by Churn

- **Graph Used:** Bar chart (average values)
- **Understanding:** Higher monthly charges are closely associated with churn, suggesting price sensitivity.



## 22. Region vs Gender with Senior Citizen

- **Graph Used:** Subplot grouped bar charts
- **Understanding:** Regional differences exist, but churn is more influenced by service and contract factors than region.



## 5. Overall Insights

Based on the visual analysis performed, the following key understandings were derived from the charts:

### Churn Behavior

- The churn distribution chart clearly shows that although the majority of customers are retained, a **significant portion is churning**, making churn prediction and prevention important.
- Churn is **not evenly distributed** across demographics and services.

### Demographic Insights (Gender & Senior Citizen)

- Gender-wise churn charts indicate that **both males and females churn**, but the proportion varies slightly depending on services and SIM operators.
- Senior citizens show **higher churn sensitivity**, especially when combined with month-to-month contracts and limited service subscriptions.
- Charts combining gender and senior citizen reveal that **senior citizens within both genders are more likely to churn** compared to non-seniors.

### Service Usage Patterns

- Internet service charts show that **Fiber optic users dominate**, but they also exhibit **higher churn**, suggesting cost or service quality concerns.
- Customers with **no internet service** tend to churn less, indicating lower expectations and simpler usage.
- Value-added services such as **OnlineSecurity, TechSupport, and DeviceProtection** have **lower adoption rates**, and customers without these services churn more frequently.

### Phone & Multiple Lines Analysis

- Multiple line usage varies significantly by SIM operator.
- Customers with **multiple lines generally show lower churn**, suggesting stronger dependency on the service.
- Charts combining SIM, gender, and senior citizen show that **complex customer profiles** help in identifying niche churn-prone segments.

### Contract-Based Insights

- Contract distribution charts strongly indicate that **month-to-month contracts have the highest churn**.
- One-year and two-year contracts show **much better retention**, across all SIM operators and demographics.
- Senior citizens prefer shorter contracts, which partially explains their higher churn rates.

### Billing & Payment Behavior

- Paperless billing charts show that **churned customers are more inclined toward paperless billing**, indicating digital-first but less loyal users.
- Certain payment methods (like electronic check) are associated with **higher churn**, while auto-pay methods show better retention.

## Tenure Analysis

- Quarterly tenure analysis clearly shows that **most churn happens in the early quarters**.
- Customers who stay beyond initial tenure periods are more likely to remain long-term.
- This highlights the importance of **early customer engagement and onboarding strategies**.

## Charges & Revenue Insights

- Monthly charges vs churn charts indicate that **customers with higher monthly charges churn more often**.
- This suggests that customers may not perceive sufficient value for the price they pay.

## Regional & SIM Operator Insights

- Regional charts reveal **variation in customer distribution**, but churn is more influenced by services and contracts than by region.
  - SIM operators show distinct patterns in service usage and churn, meaning **operator-specific retention strategies** are necessary.
- 

## 6. Overall Insights

- Month-to-month contracts and high monthly charges are the strongest churn drivers
- Senior citizens and early-tenure customers require focused retention strategies
- Bundling value-added services can help reduce churn
- Early-stage customer experience plays a crucial role in long-term retention

This EDA provides actionable business insights and forms a strong foundation for churn prediction modeling and strategic decision-making.