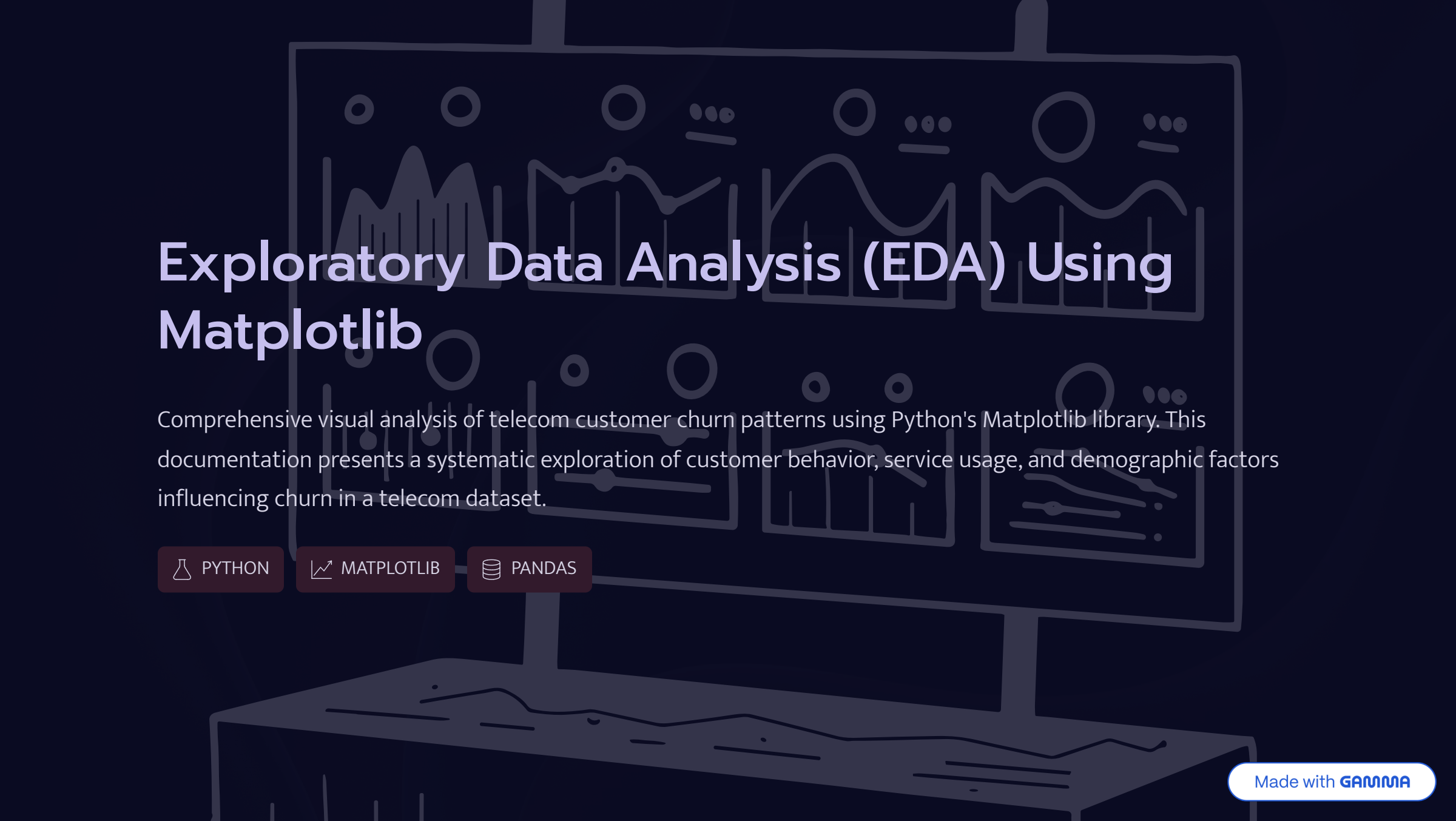


Exploratory Data Analysis (EDA) Using Matplotlib

Comprehensive visual analysis of telecom customer churn patterns using Python's Matplotlib library. This documentation presents a systematic exploration of customer behavior, service usage, and demographic factors influencing churn in a telecom dataset.

PYTHON MATPLOTLIB PANDAS

Project Objective & Technical Foundation

Primary Goal

Understand customer churn behavior by analyzing demographic factors, services, contracts, billing patterns, and usage trends to identify key churn drivers and derive actionable retention strategies.

Technical Stack

Python-based analysis using Pandas for data manipulation, NumPy for numerical operations, and Matplotlib for comprehensive visualization in Google Colab/Jupyter environment.

Dataset Scope

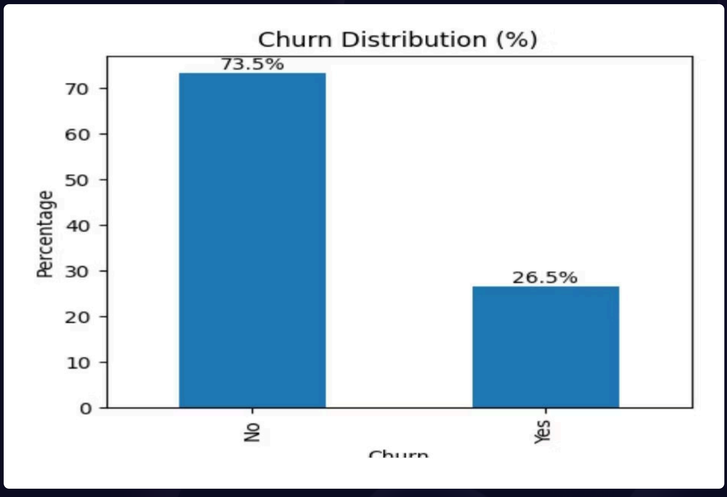
Customer-level telecom data including churn status, demographics, service subscriptions, contract types, billing methods, tenure, charges, and regional information.

The analysis encompasses multiple dimensions of customer behavior, from basic demographic segmentation to complex multi-factor interactions across SIM operators, service bundles, and payment preferences. Each visualization serves to uncover patterns that traditional statistical summaries might miss, providing visual evidence for business decision-making. The comprehensive approach ensures that both obvious and subtle churn indicators are identified and documented for subsequent predictive modeling phases.

Figure 1-3: Foundational Churn & Gender Analysis

Figure 1: Churn Distribution

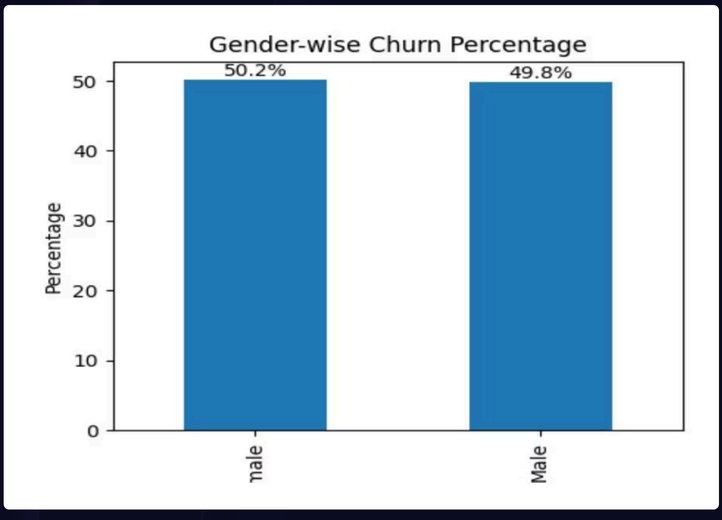
Visualization Type: Bar chart showing binary churn classification (Yes/No). This fundamental plot establishes the baseline churn rate across the entire customer base, revealing the class distribution that will influence all subsequent analyses.



- **Key Observation:** Majority of customers remain with the service, but a significant minority churns
- Clear class imbalance exists, with retained customers outnumbering churned customers
- Churn represents a measurable business problem requiring intervention
- **Business Impact:** The visible churn proportion indicates revenue loss and customer acquisition cost waste

Figure 2: Gender-wise Churn Distribution

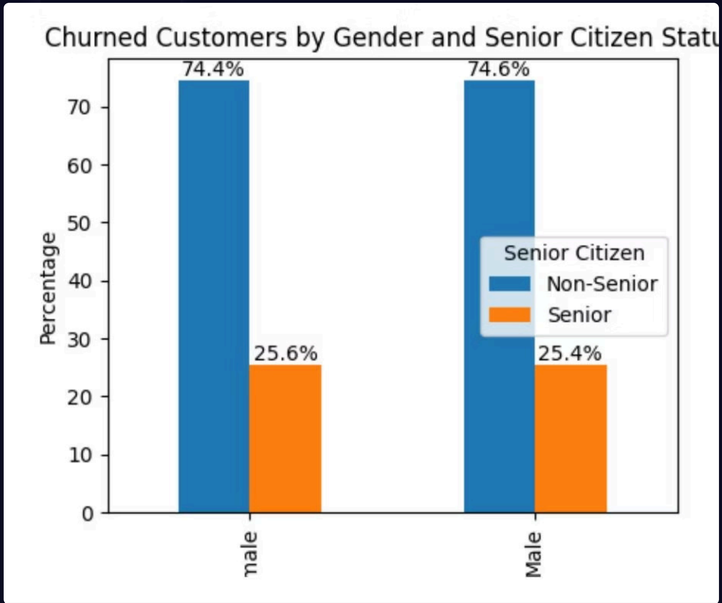
Visualization Type: Bar chart comparing churn across male and female customers. This demographic breakdown tests whether gender serves as a primary churn predictor.



- Both genders contribute to churn with relatively balanced proportions
- Gender difference is not extreme, suggesting gender alone is weak predictor
- **Telecom Interpretation:** Service quality and pricing affect both genders similarly
- Gender may gain predictive power when combined with other features

Figure 3: Churned Customers by Gender & Senior Citizen

Visualization Type: Grouped bar chart examining churn across gender and senior citizen status. This multi-dimensional view reveals age-related churn patterns within gender segments.

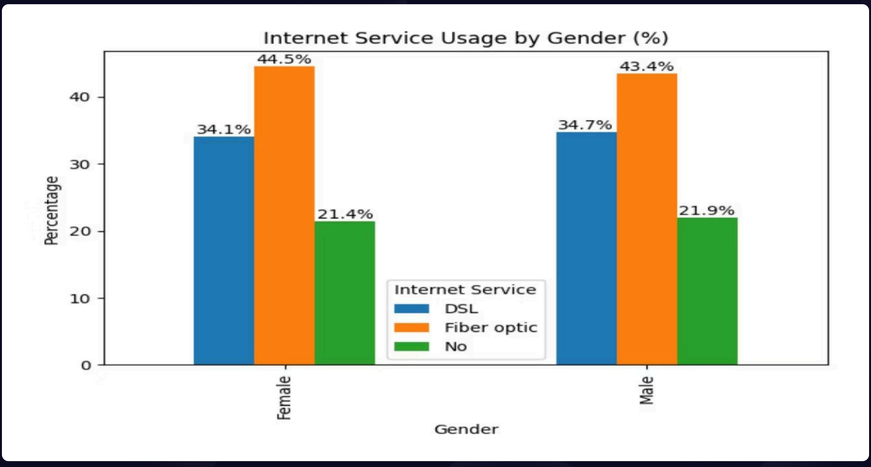


- **Critical Finding:** Senior citizens within both genders exhibit higher churn rates
- Age emerges as more significant factor than gender for churn prediction
- **Customer Behavior:** Senior citizens may have different service expectations or technical support needs
- Suggests targeted retention programs for elderly customers could reduce churn

Figure 4-6: Service Usage Patterns Analysis

Figure 4: Internet Service Usage by Gender

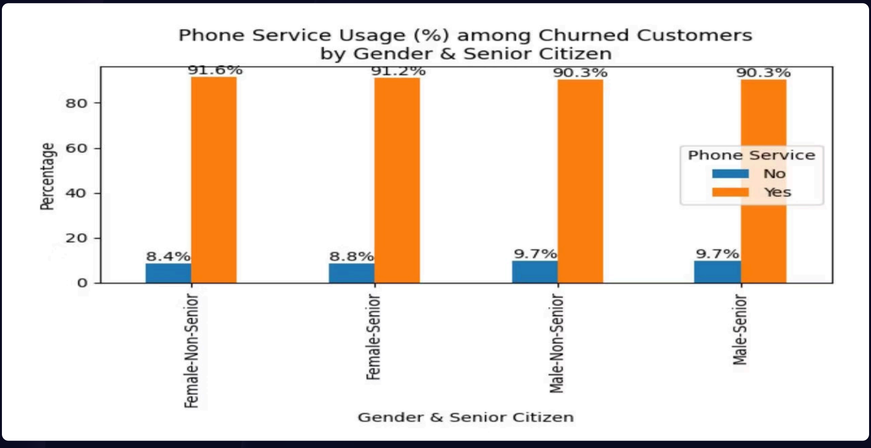
Visualization Type: Percentage bar chart showing internet service type distribution (DSL, Fiber optic, No service) across genders. Reveals service preference patterns and technology adoption rates.



- **Dominant Pattern:** Fiber optic service is most commonly used across both genders
- High-speed internet preference indicates customer expectations for premium service
- **Churn Risk:** Fiber optic users may be more price-sensitive or quality-conscious
- Gender shows minimal influence on internet service choice

Figure 5: Phone Service Usage by Gender & Senior Citizen (Churned Customers)

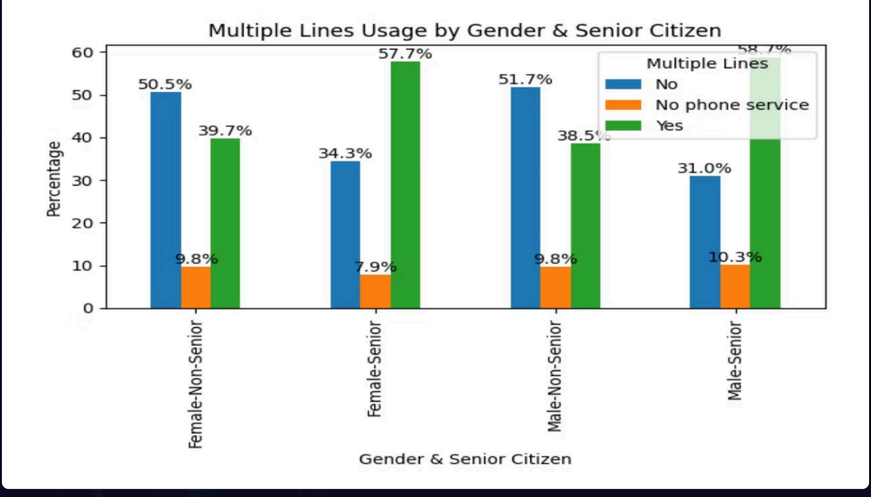
Visualization Type: Bar chart focusing exclusively on churned customers' phone service adoption across demographic segments. Tests whether lack of basic services drives churn.



- Most churned customers still subscribed to phone services before leaving
- **Critical Insight:** Churn is not due to lack of basic service adoption
- **Business Implication:** Issues likely stem from pricing, support quality, or service reliability rather than service availability
- Senior citizens with phone service still churn, indicating unmet needs beyond connectivity

Figure 6: Multiple Lines Usage by Gender & Senior Citizen

Visualization Type: Bar chart examining multiple phone line adoption across demographic segments. Reveals service complexity preferences and household usage patterns.

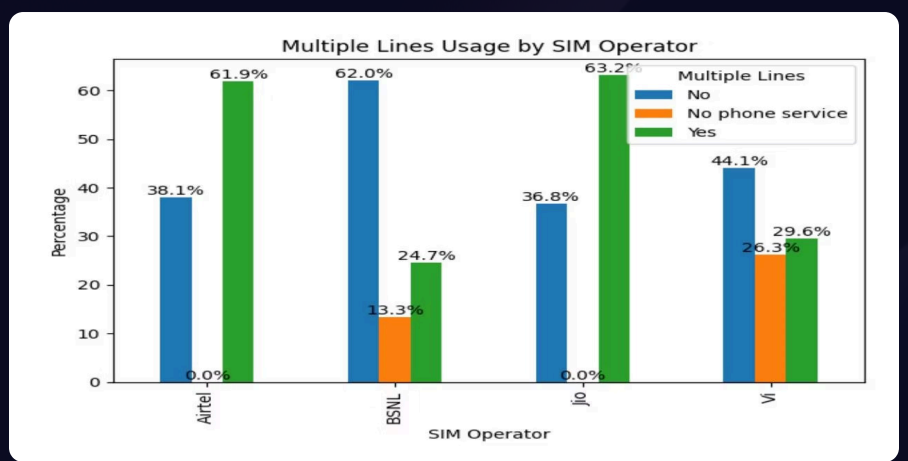


- **Age Pattern:** Senior citizens adopt multiple lines less frequently than younger customers
- Indicates simpler service needs and potentially single-person households among elderly
- Multiple line users may have stronger service dependency, potentially reducing churn
- **Telecom Strategy:** Bundle offerings should consider demographic-specific needs

Figure 7-9: SIM Operator & Multiple Lines Deep Dive

Figure 7: Multiple Lines by SIM Operator

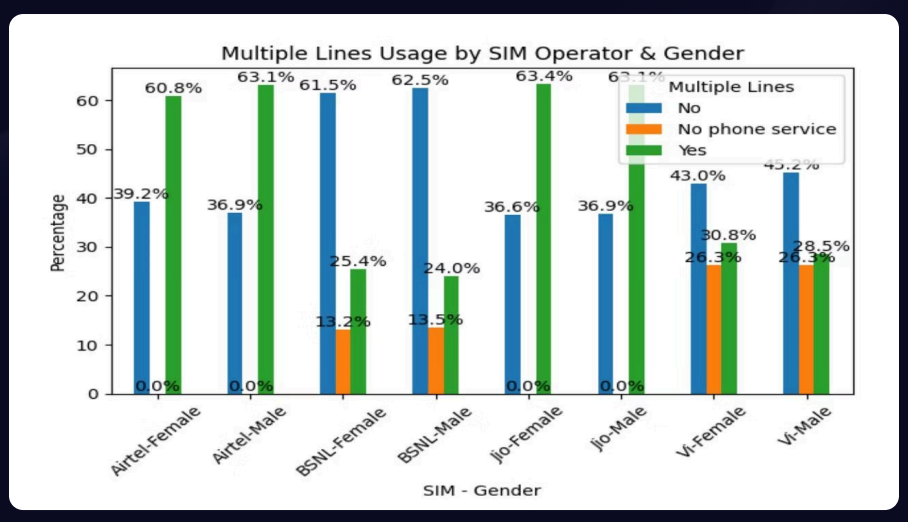
Visualization Type: Bar chart comparing multiple line adoption across different SIM operators (BSNL, Jio, Vi). Reveals operator-specific service offerings and customer preferences.



- Significant variation in multiple line adoption across operators
- **Operator Influence:** Plan structures and pricing strategies differ substantially
- Some operators successfully promote multi-line packages

Figure 8: Multiple Lines by SIM & Gender

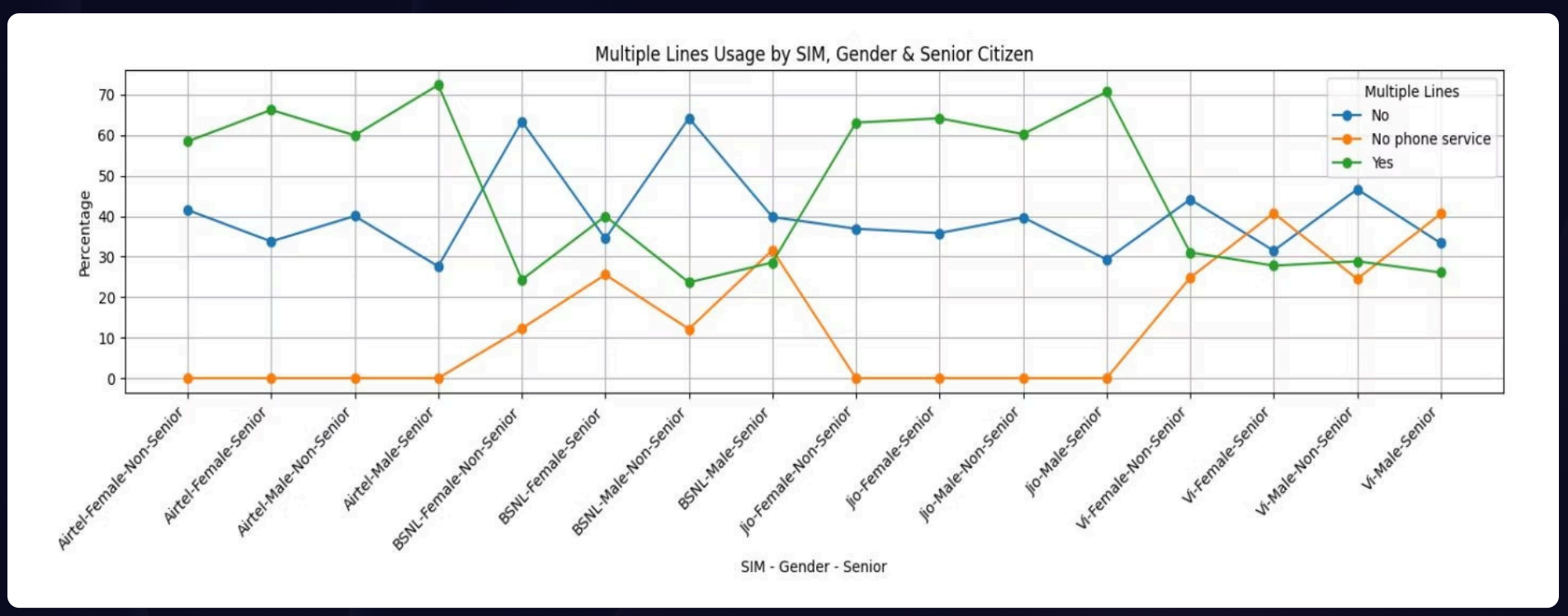
Visualization Type: Grouped bar chart adding gender dimension to operator-level multiple line analysis. Tests interaction between operator choice and gender.



- Gender differences in multiple line usage remain minor across all operators
- **Key Finding:** SIM operator policies dominate over gender preferences
- Marketing strategies should focus on operator-level differentiation

Figure 9: Multiple Lines by SIM, Gender & Senior Citizen

Visualization Type: Line chart presenting three-way interaction between SIM operator, gender, and senior citizen status for multiple line usage. Most complex demographic segmentation in the analysis.



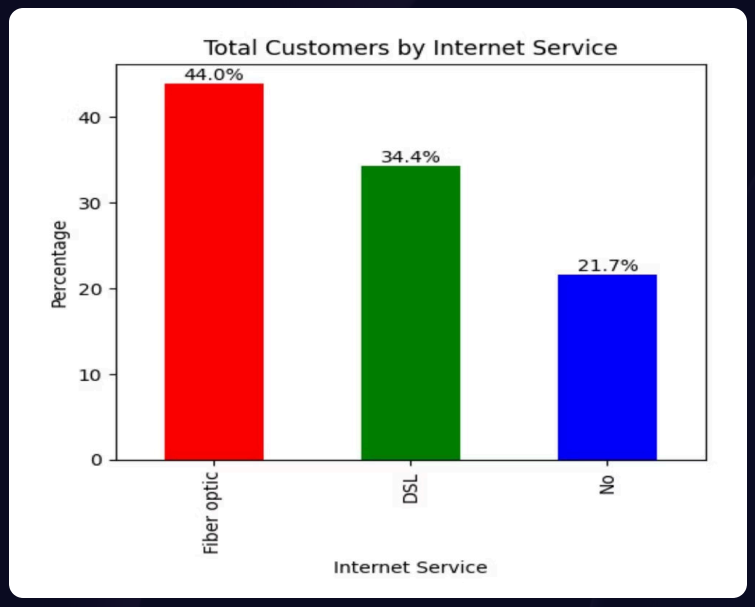
- **Segmentation Value:** Identifies specific demographic-operator combinations with high/low multiple line adoption
- Senior citizens consistently show lower multiple line usage across all operators and genders
- **Targeted Marketing:** Enables precise customer segment identification for promotional campaigns
- Operator-specific patterns emerge more clearly when controlling for demographics

Figure 10-13: Internet Service Comprehensive Analysis

1

Figure 10: Total Customers by Internet Service

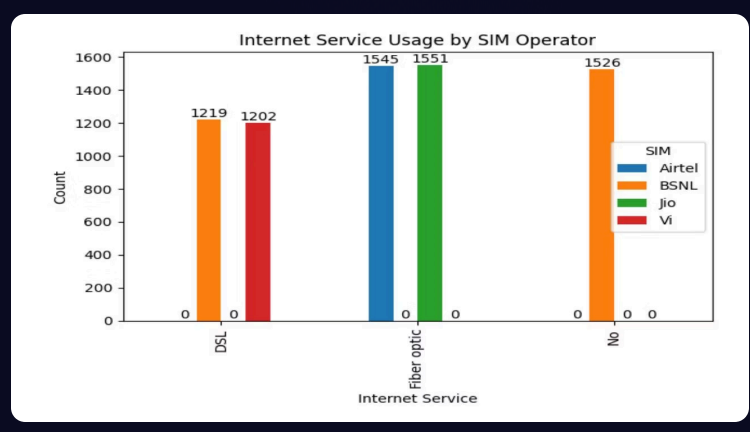
Bar chart showing overall distribution across DSL, Fiber optic, and No internet service. Fiber optic dominates subscriptions, indicating strong demand for high-speed connectivity and modern service expectations.



2

Figure 11: Internet Service by SIM Operator

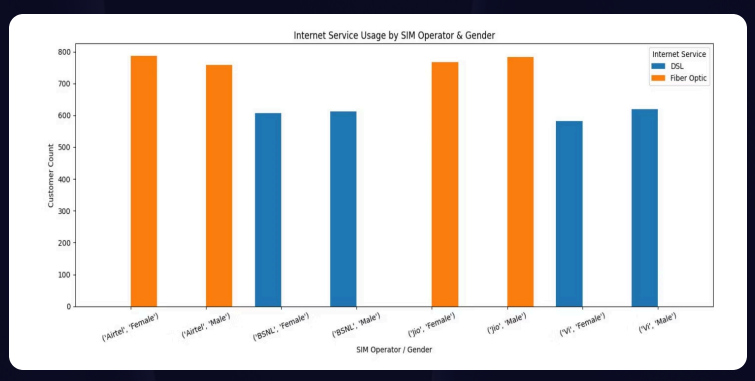
Count-based bar chart revealing operator specialization in internet service types. Different operators show distinct service portfolios, affecting customer experience and satisfaction levels.



3

Figure 12: Internet Service by SIM & Gender

Grouped bar chart adding gender dimension to operator-level internet service analysis. Service preferences vary slightly by gender, but operator influence remains dominant factor.



4

Figure 13: Internet Service by SIM, Gender & Churn

Multi-dimensional grouped bar chart incorporating churn status. Certain SIM-gender combinations exhibit higher churn, suggesting service quality or pricing disparities across segments.

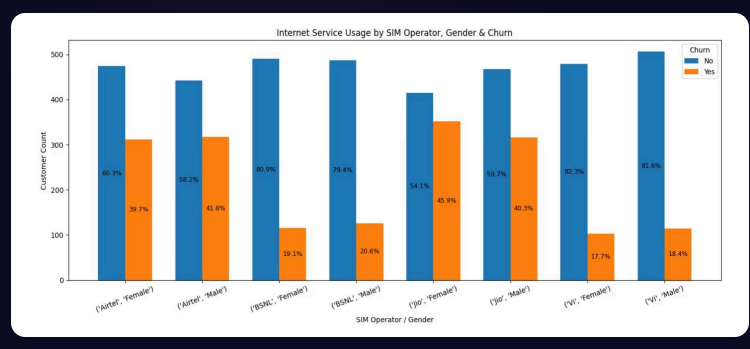
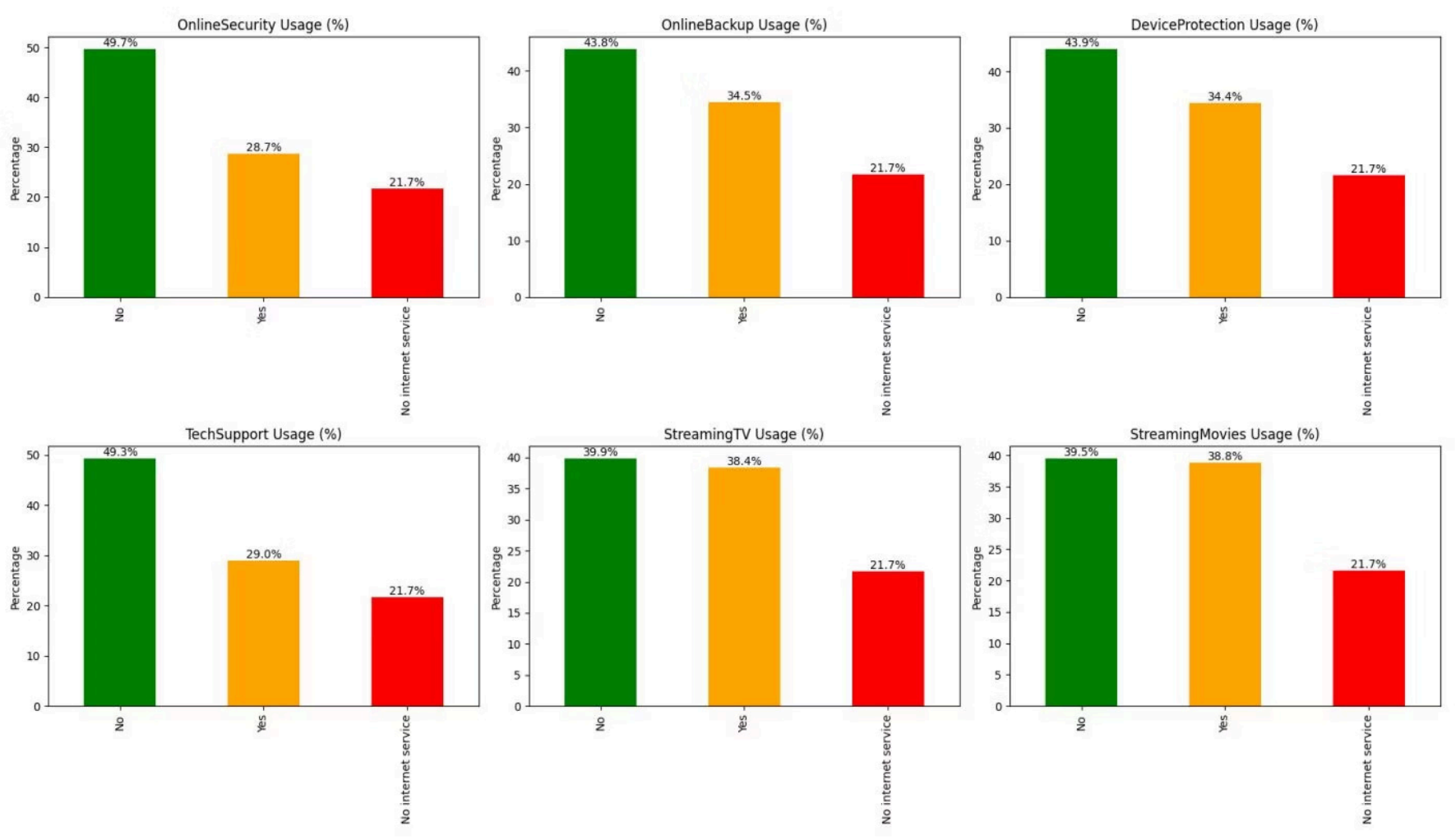


Figure 14-15: Value-Added Services Adoption

Figure 14: Service Columns Distribution

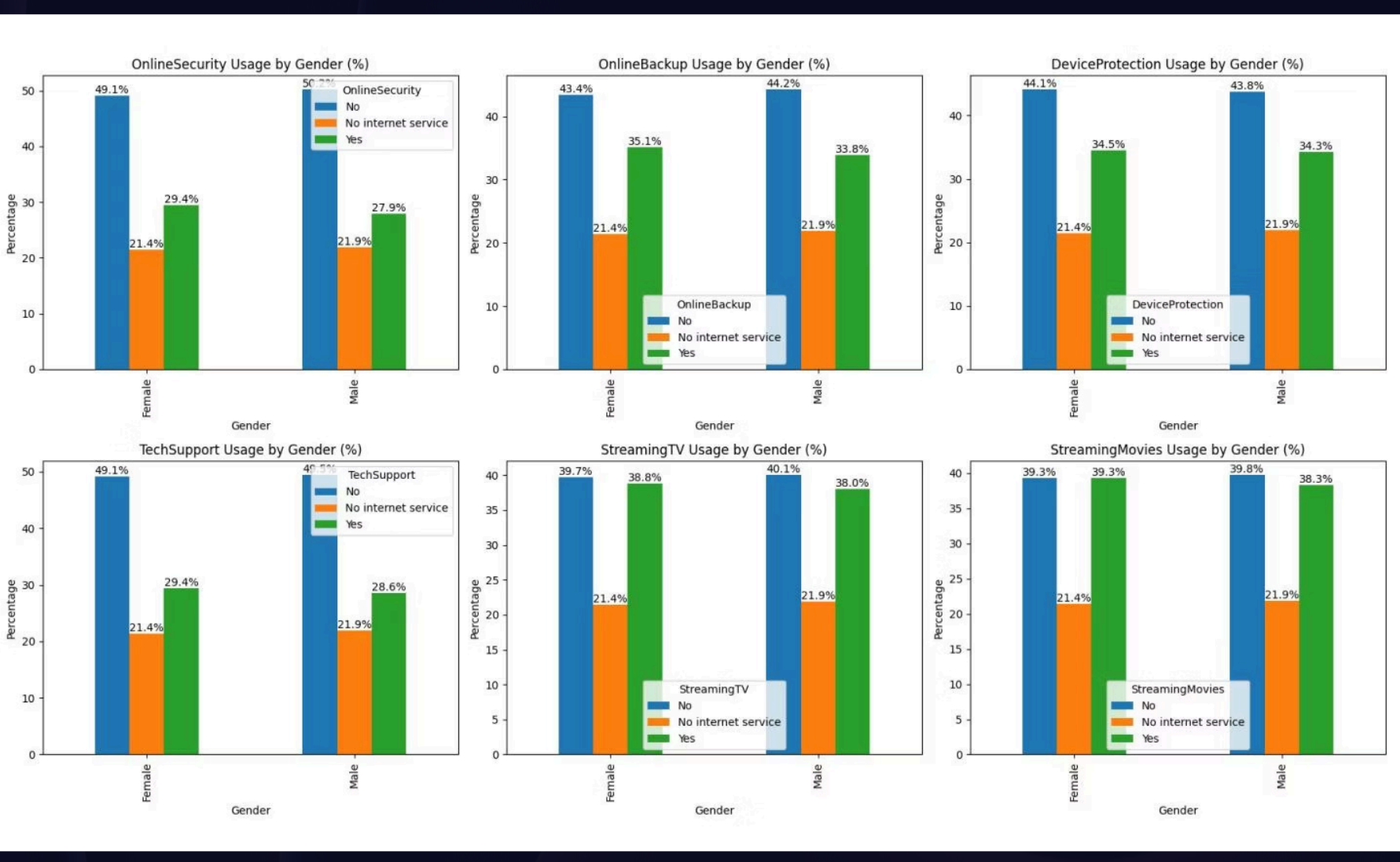
Visualization Type: Subplot bar charts displaying adoption rates for multiple value-added services (OnlineSecurity, OnlineBackup, DeviceProtection, TechSupport, StreamingTV, StreamingMovies). Comprehensive view of service bundle penetration.



- **Critical Finding:** Value-added services like OnlineSecurity and TechSupport show low adoption rates
- Customers without these protective services face higher churn risk
- **Revenue Opportunity:** Significant untapped potential for service bundling and upselling
- Streaming services show better adoption than security/support services
- **Business Implication:** Customers may not perceive value in security offerings or lack awareness

Figure 15: Service Usage by Gender

Visualization Type: Subplot grouped bar charts comparing value-added service adoption across genders. Tests whether service preferences differ by gender.

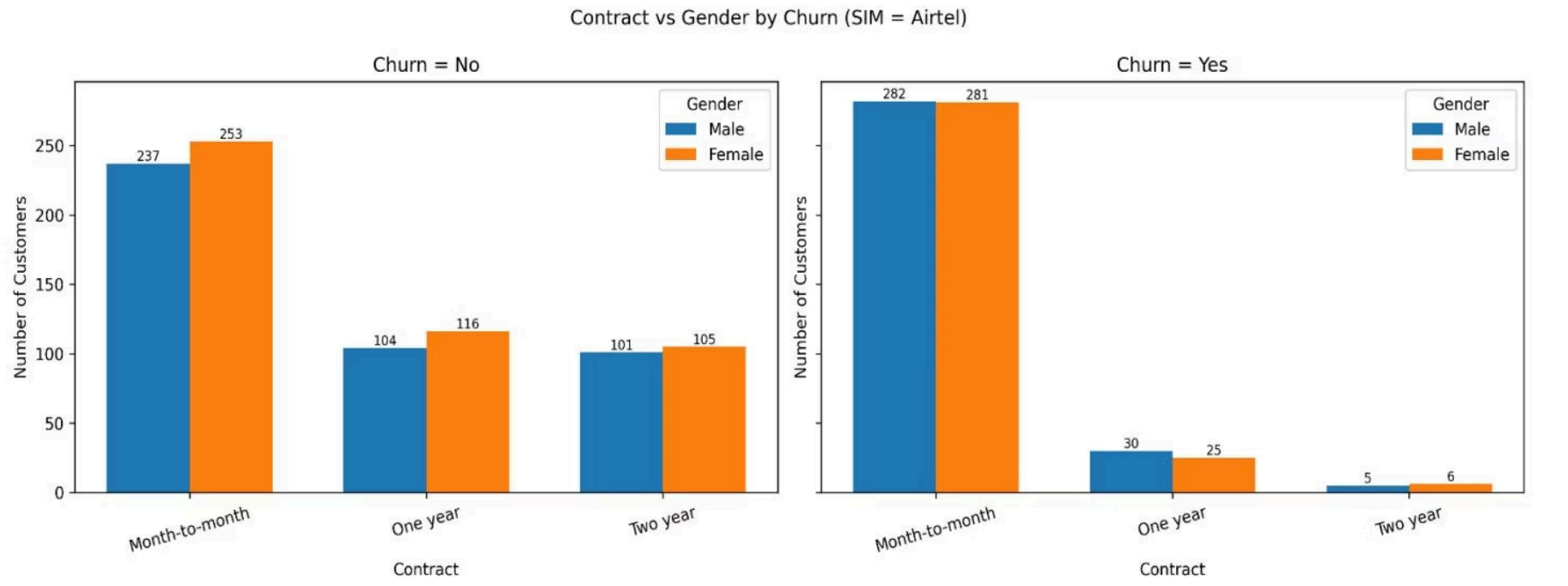


- Service adoption patterns remain similar across male and female customers
- **Key Insight:** Gender is not a differentiating factor for value-added service uptake
- **Marketing Focus:** Awareness campaigns and pricing strategies matter more than gender-targeted messaging
- Both genders show similar gaps in security and support service adoption
- Suggests universal barriers to adoption (price, awareness, perceived value) rather than demographic preferences

Figure 16-17: Contract Analysis & Retention Patterns

Figure 16: Contract Distribution by SIM, Gender & Churn

Visualization Type: Multiple grouped bar charts showing contract type distribution (Month-to-month, One year, Two year) across SIM operators, segmented by gender and churn status. Most comprehensive contract analysis in the dataset.



Month-to-Month Contracts

Strongest association with churn across all operators and demographics. Flexibility comes at cost of customer commitment and loyalty.

Long-Term Contracts

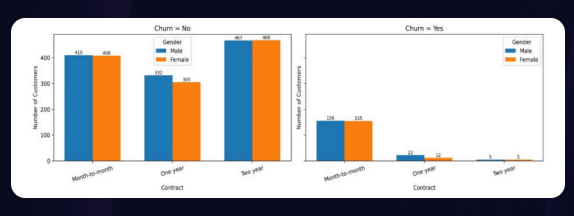
One-year and two-year contracts demonstrate significantly better retention rates, providing stability and reducing churn risk substantially.

Senior Citizen Preference

Elderly customers favor shorter contracts, partially explaining their higher churn rates and need for flexible arrangements.

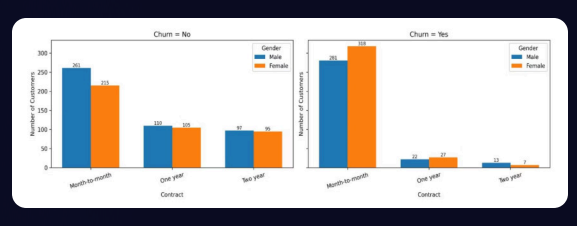
Operator-Specific Contract Patterns

BSNL Contracts



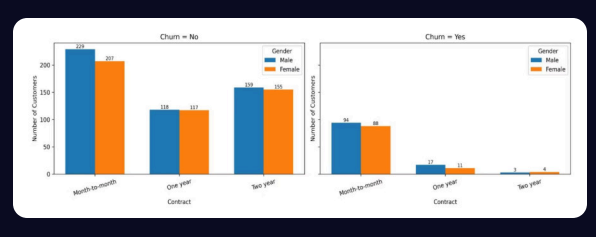
Contract vs Gender by Churn analysis for BSNL operator showing distinct patterns in customer commitment levels.

Jio Contracts



Jio's contract distribution reveals operator-specific retention strategies and customer base characteristics.

Vi Contracts



Vi operator contract patterns demonstrate unique market positioning and customer relationship approaches.

Figure 17: Contract Distribution by Gender, Senior Citizen & SIM

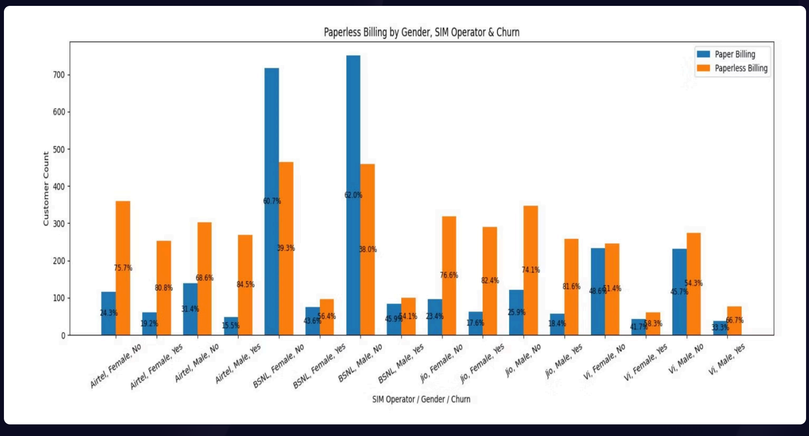
Visualization Type: Percentage bar chart examining contract preferences across demographic segments and operators. Reveals age-related contract behavior patterns.

- **Age Factor:** Senior citizens consistently prefer shorter contract durations across all operators
- Aligns with higher churn levels observed in elderly customer segments
- **Retention Strategy:** Incentivizing long-term contracts for seniors could significantly reduce churn
- Operator-specific patterns suggest different success rates in converting customers to long-term commitments

Figure 18-20: Billing, Payment & Tenure Analysis

Figure 18: Paperless Billing by Gender, SIM Operator & Churn

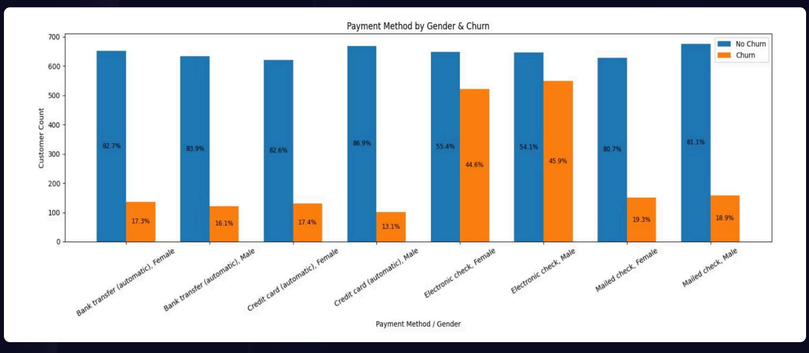
Visualization Type: Grouped bar chart analyzing paperless billing adoption across demographics and churn status. Tests relationship between digital billing preferences and customer loyalty.



- **Counterintuitive Finding:** Customers using paperless billing exhibit higher churn rates
- Indicates digitally active but potentially less loyal customer segment
- **Customer Profile:** Paperless users may be more price-conscious and willing to switch providers
- Digital convenience does not translate to emotional loyalty or commitment

Figure 19: Payment Method by Gender & Churn

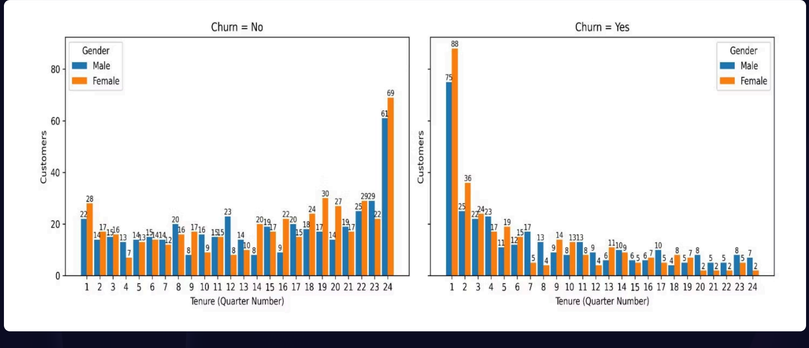
Visualization Type: Grouped bar chart comparing payment methods (Electronic check, Mailed check, Bank transfer, Credit card) across gender and churn status. Reveals payment behavior's impact on retention.



- **High-Risk Method:** Electronic payment methods show elevated churn rates
- **Retention Advantage:** Automatic payment methods (bank transfer, credit card auto-pay) improve customer retention
- **Behavioral Insight:** Payment friction and manual processes increase likelihood of service cancellation
- Automated payments create habit formation and reduce conscious decision-making about service continuation

Figure 20: Quarterly Tenure Analysis with SIM

Visualization Type: Multiple grouped bar charts showing customer distribution across tenure quarters (0-18 months, 18-36 months, 36-54 months, 54-72 months) for each SIM operator. Critical for understanding churn timing patterns.



Early Tenure (0-18 months)

Highest churn concentration occurs in initial quarters. Critical onboarding period determines long-term retention success.

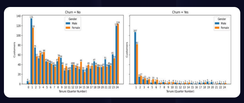
Stabilization (18-36 months)

Customers surviving initial period show improved retention. Relationship strengthens with continued service satisfaction.

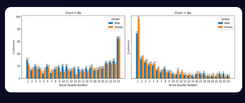
Long-Term (36+ months)

Established customers demonstrate strong loyalty. Churn risk decreases significantly after three-year milestone.

BSNL Tenure



Jio Tenure



Vi Tenure

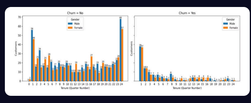
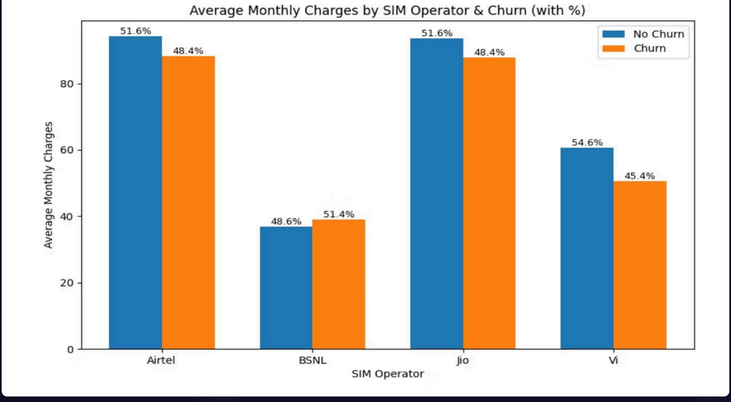


Figure 21-22 & Overall EDA Insights

Figure 21: Monthly Charges vs SIM Operator by Churn

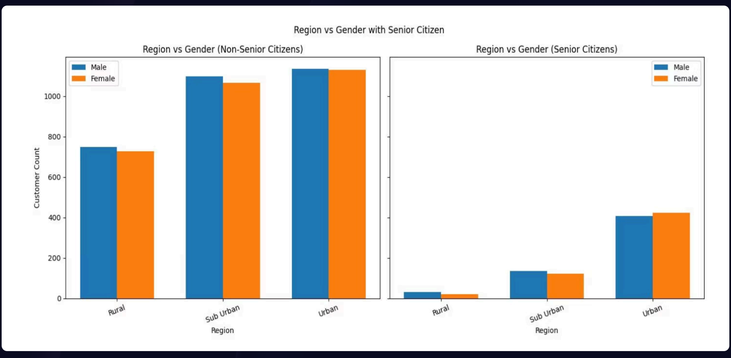
Visualization Type: Bar chart displaying average monthly charges across SIM operators, segmented by churn status. Tests price sensitivity's role in customer attrition.



- **Critical Price-Churn Relationship:** Higher monthly charges strongly correlate with increased churn rates
- Customers perceive insufficient value for premium pricing tiers
- **Operator Variation:** Different operators show varying price points and corresponding churn levels
- **Strategic Implication:** Price optimization and value communication are essential retention levers

Figure 22: Region vs Gender with Senior Citizen

Visualization Type: Subplot grouped bar charts examining regional distribution across gender and senior citizen status. Tests geographic influence on customer composition.



- Regional differences exist in customer distribution patterns
- **Key Finding:** Churn is more strongly influenced by service and contract factors than geographic region
- Regional strategies should focus on service quality rather than location-specific marketing

---

Overall Insights from EDA – Step 1

Top Churn Drivers Identified

Month-to-month contracts and high monthly charges emerge as strongest churn predictors. Senior citizens and early-tenure customers require focused retention strategies. Payment method automation significantly impacts retention rates.

Customer Segments at Risk

Senior citizens across all demographics show elevated churn. Customers in first 18 months of tenure are highly vulnerable. Paperless billing users demonstrate lower loyalty despite digital engagement. High-charge customers without value-added services face retention challenges.

Service Combinations & Churn

Fiber optic users show higher churn despite premium service. Low adoption of OnlineSecurity, TechSupport, and DeviceProtection correlates with increased churn. Multiple line users demonstrate stronger service dependency and lower churn risk.

Data Quality & ML Readiness

Dataset exhibits class imbalance requiring sampling techniques. Categorical features dominate, necessitating encoding strategies. Monthly charges show skewness requiring transformation. Early tenure concentration suggests feature engineering opportunities. Dataset is ready for preprocessing with identified transformation requirements.

> **Business Impact Summary:** This comprehensive EDA provides actionable insights for churn prediction modeling. The analysis reveals that contract type, pricing strategy, early customer experience, and value-added service bundling represent the most promising intervention points for reducing customer attrition and improving long-term retention rates.