TELECOM CUSTOMER CHURN PREDICTION SYSTEM

Project Introduction:

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.

Data Collection and Scope:

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.

Structure of the Customer Dataset:

1. Customer Identification:

These attributes uniquely represent customers in the dataset and help track individual records.

- Customer ID

2 Demographic Attributes:

These features describe the personal and household characteristics of customers and help analyse churn behaviour across different population segments.

- Gender
- Senior Citizen Status
- Partner
- Dependents

3. Service Subscription Attributes:

These attributes indicate the telecom services subscribed by customers and their level of service engagement.

- SIM / Network Provider
- Phone Service
- Multiple Lines
- Internet Service Type (DSL, Fiber Optic, No Internet)
- Streaming TV
- Streaming Movies

4. Value-Added Service Attributes:

These features capture optional services that enhance customer experience and influence retention.

- Online Security
- Online Backup
- Device Protection
- Technical Support

5. Contract and Tenure Attributes:

These attributes describe the nature and duration of the customer's relationship with the service provider.

- Contract Type (Month-to-Month, One-Year, Two-Year)
- Tenure (Number of Months with the Service)

6. Billing and Payment Attributes:

These features represent customer billing behaviour and payment preferences, which are strong indicators of churn.

- Monthly Charges
- Total Charges
- Payment Method
- Paperless Billing Status

7. Time-Based Attributes:

These attributes help analyse customer acquisition trends and cohort-based churn behaviour.

- Customer Joining Year

8. Target Variable:

This attribute represents the outcome variable used for churn analysis and prediction.

- Churn Status (Yes / No)

Analytical Approach and Visualization Strategy:

This project uses exploratory data analysis (EDA) to examine customer behaviour and identify key factors influencing churn. After basic data cleaning and validation, customer attributes such as demographics, service usage, contract type, tenure, and billing preferences are analysed to understand their impact on retention.

Matplotlib visualizations are used to represent patterns clearly and effectively. Bar charts, histograms, and pie charts help compare customer segments, observe tenure trends, and analyse service adoption. These visual insights support the identification of high-risk churn groups and assist in forming data-driven customer retention strategies.

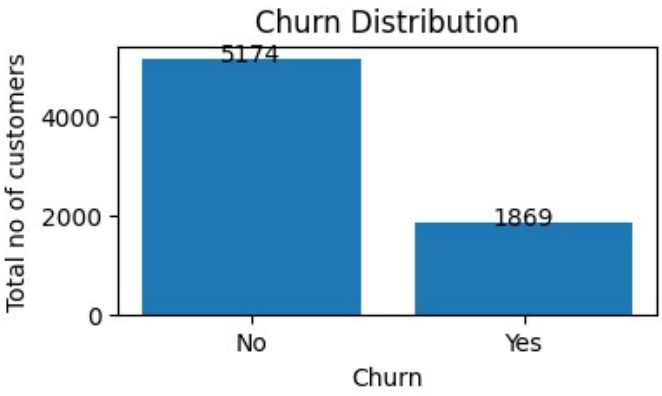
1. **Partner Distribution:**
   This visualization represents the distribution of customers based on their partner status, categorizing them into:

- Customers with a partner
- Customers without a partner

The graph helps analyze how relationship status correlates with customer presence in the telecom service.

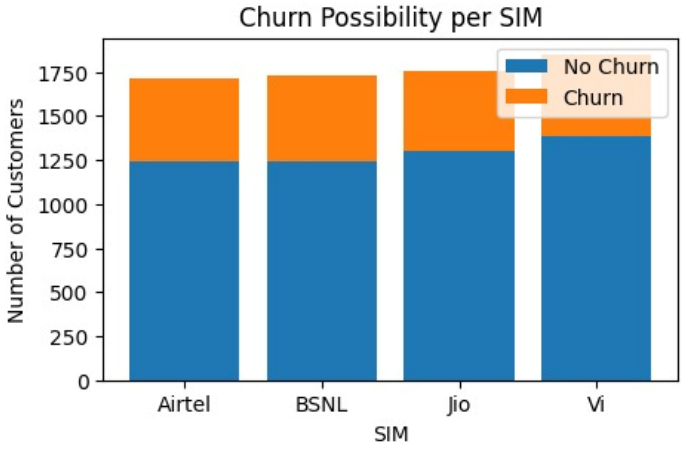
2. **Churn Distribution:**

- This visualization illustrates the overall distribution of customers based on their churn status, categorizing them into:
  1.Customers who continued using the telecom service
  2.Customers who discontinued the service (churned)
- The graph provides a clear overview of customer retention versus customer loss.
- Each bar represents a churn category (Yes / No).
- Reducing churn directly improves customer lifetime value and revenue stability.



3. **Churn Possibility Per Sim Distribution:**

- This visualization represents the churn possibility across different SIM providers, showing how customer churn varies by SIM type.
- The SIM providers analyzed include:
  1. Jio
  2. Airtel
  3. VI
  4. BSNL
- The graph helps identify which SIM providers experience higher customer attrition.
- The height of the bar indicates the number or proportion of customers who have churned for that SIM.
- This visualization enables a direct comparison of churn behavior across providers.
- Certain SIM providers exhibit a higher churn tendency, indicating potential issues related to pricing, network quality, or customer satisfaction.
- SIM providers with lower churn demonstrate stronger customer retention and service reliability.



4. **Year-wise Distribution by Customers:**

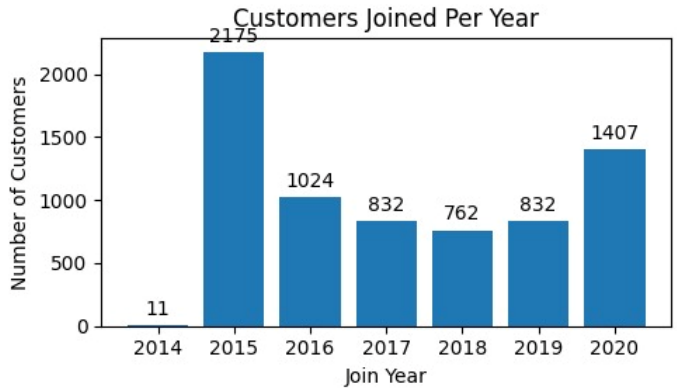
- This visualization summarizes the year-wise distribution of SIMs taken by customers, highlighting overall SIM acquisition trends across different years.

1.Each bar represents a specific year.

2.The height of the bar indicates the number of SIMs activated in that year.

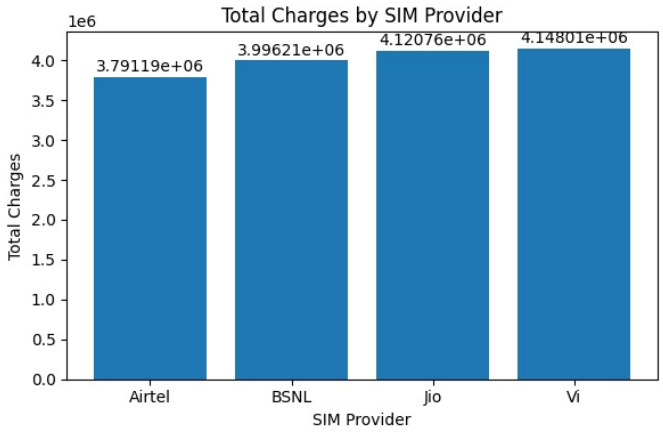
3.The graph provides a clear comparison of SIM adoption over time.

- Recent trends help determine whether SIM adoption is growing, stabilizing, or declining.
- This distribution complements churn and tenure analysis by showing when customers entered the system.



5. **Total Charges Provided by Sim:**

- This visualization represents the distribution of total charges accumulated by customers across different SIM providers.
- Total charges reflect the cumulative amount paid by customers over their entire tenure with a provider.
- The SIM providers analyzed include:
  1. Jio
  2. Airtel
  3. VI
  4. BSNL
- SIM providers with a larger customer base or longer average tenure tend to show higher total charges.
- Providers with lower total charges may have higher churn or shorter customer tenure.
- High total charges indicate strong customer loyalty and long-term engagement.



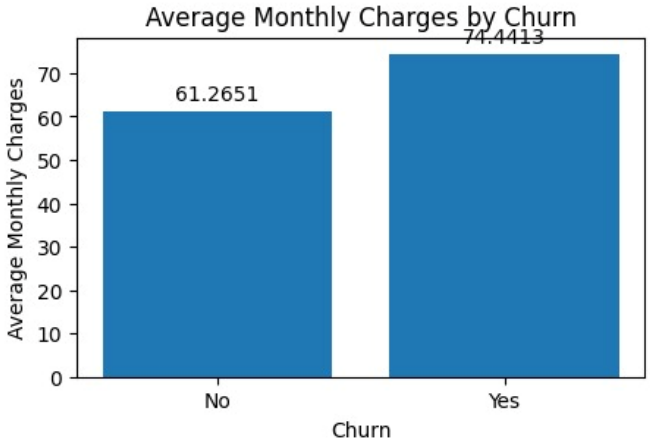
6. **Average Monthly Charges by Churn:**

- This visualization represents the average monthly charges of customers segmented by their churn status:

1. Customers who churned

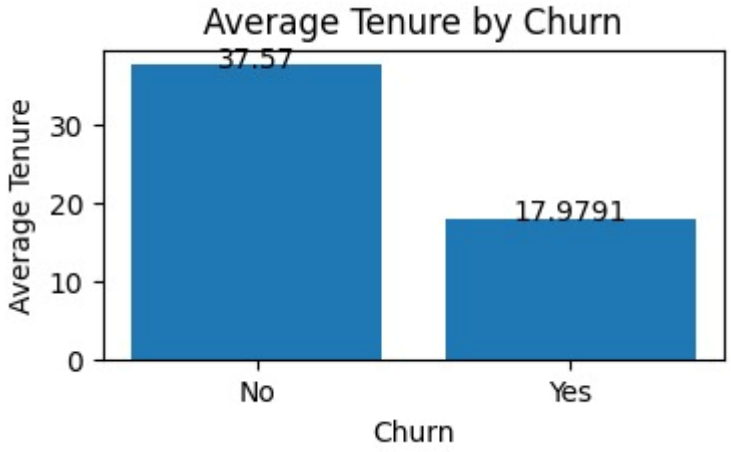
2. Customers who did not churn

- The objective is to understand whether billing amounts influence customer churn behavior.
- Each bar corresponds to a churn category (Yes / No).
- The height of the bar indicates the average monthly charge for customers in that category.
- The visualization enables a direct comparison of billing levels between churned and retained customers.
- Customers who churned generally have higher average monthly charges compared to those who stayed.
- Customers with lower average monthly charges tend to show better retention.
- Pricing plays a significant role in influencing customer loyalty.



7. Average Tenure by Churn:

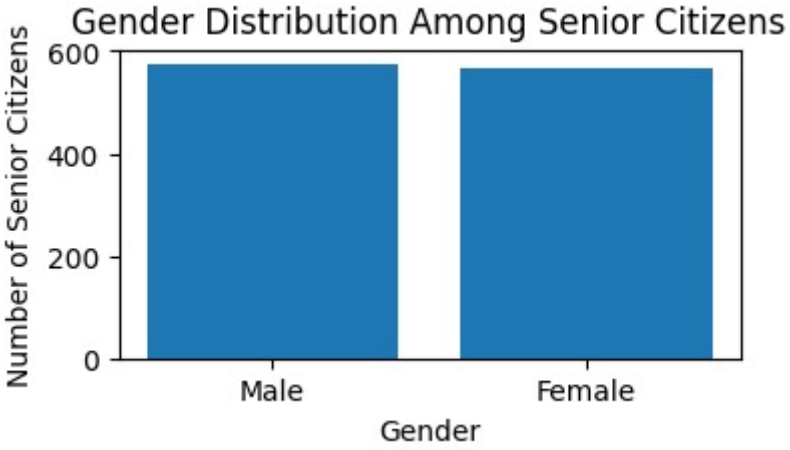
- This visualization represents the average customer tenure (in months) segmented by churn status:
  1. Customers who churned
  2. Customers who did not churn
- The objective is to understand how the length of customer relationship impacts churn behavior.
- Each bar corresponds to a churn category (Yes / No).
- The height of the bar indicates the average tenure of customers in that category.
- The visualization provides a clear comparison of customer longevity between churned and retained users.
- Customers who did not churn have a significantly higher average tenure.
- The likelihood of churn decreases as tenure increases.



8. Gender Distribution Among Senior Citizens:

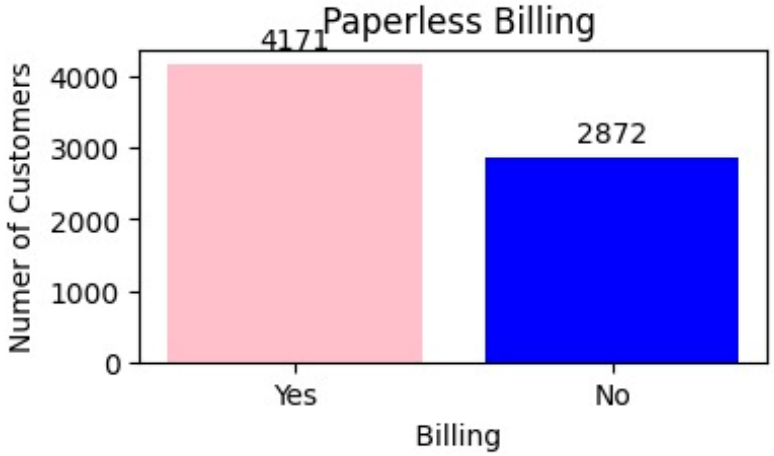
- This visualization represents the distribution of senior citizen customers based on gender, categorizing them into:
  1. Male senior citizens
  2. Female senior citizens
- Each bar corresponds to a gender category among senior citizens.
- The height of the bar indicates the number of senior customers in each gender group.
- The visualization highlights gender-based participation within the senior citizen population.
- Senior male customers often constitute a slightly higher proportion compared to female senior customers.
- Senior citizens, regardless of gender, typically show higher churn sensitivity due to pricing, service complexity, or support needs.

- Gender differences among senior citizens can influence service usage patterns and support requirements.



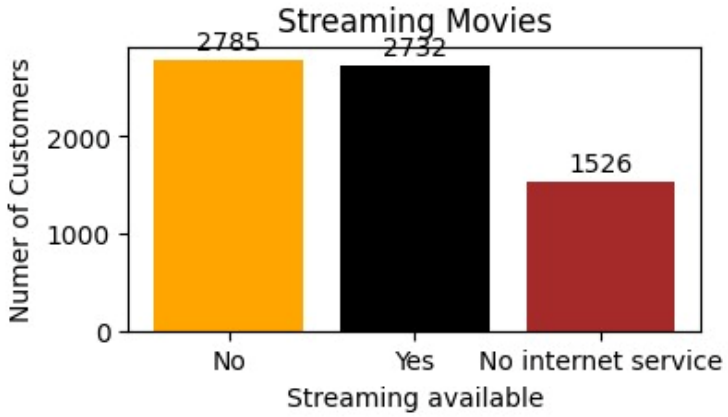
9. Paperless Billing:

- This visualization represents the distribution of customers based on their paperless billing preference.
- Customers are categorized into:
  1. Customers who have enabled paperless billing
  2. Customers who have not enabled paperless billing
- This analysis helps understand customer adoption of digital billing practices.
- Each bar corresponds to a paperless billing category (Yes / No).
- The height of the bar indicates the number of customers in each category.
- The visualization provides a clear comparison between digital and traditional billing preferences.
- A significant portion of customers have opted for paperless billing, indicating growing acceptance of digital services.
- Customers who use paperless billing often show higher engagement with online platforms.
- Non-paperless billing customers may prefer traditional communication or have lower digital adoption.
- Paperless billing users are frequently associated with lower churn risk when combined with auto-payment options.



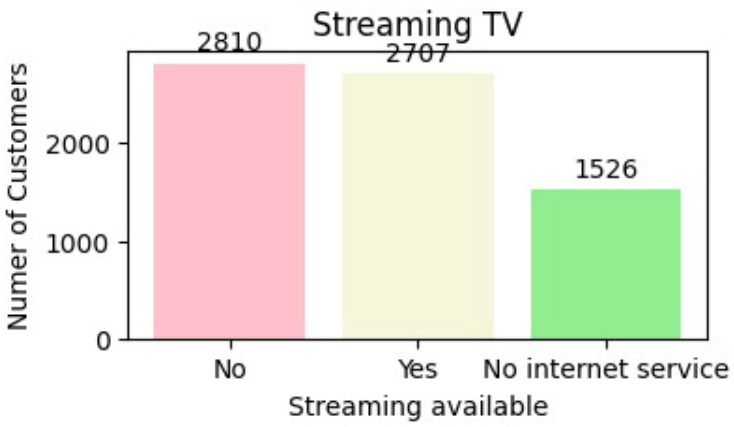
10. Streaming Movies:

- This visualization represents the distribution of customers based on their subscription to the Streaming Movies service.
- Customers are categorized into:
  1. Customers who have Streaming Movies service
  2. Customers who do not have Streaming Movies service
  3. Customers with no internet service, where streaming is not applicable
- This analysis helps understand the adoption level of entertainment-related value-added services.
- Each bar corresponds to a Streaming Movies service category.
- The height of the bar indicates the number of customers in each category.
- The visualization clearly differentiates between streaming subscribers and non-subscribers.
- A substantial number of customers have not subscribed to the Streaming Movies service.
- Customers with internet service are more likely to adopt streaming services.



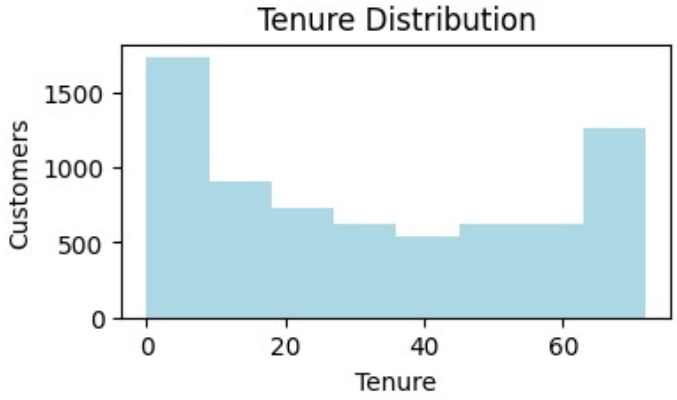
11. Streaming TV:

- This visualization represents the distribution of customers based on their subscription to the Streaming TV service.
- Customers are categorized into:
  1. Customers who have Streaming TV service
  2. Customers who do not have Streaming TV service
  3. Customers with no internet service, where streaming services are not applicable
- Each bar corresponds to a Streaming TV service category.
- The height of the bar indicates the number of customers in each category.
- The visualization highlights how widely Streaming TV is adopted among customers.
- Streaming TV adoption is higher among customers with reliable internet services.
- Customers subscribed to Streaming TV generally demonstrate higher engagement and loyalty.



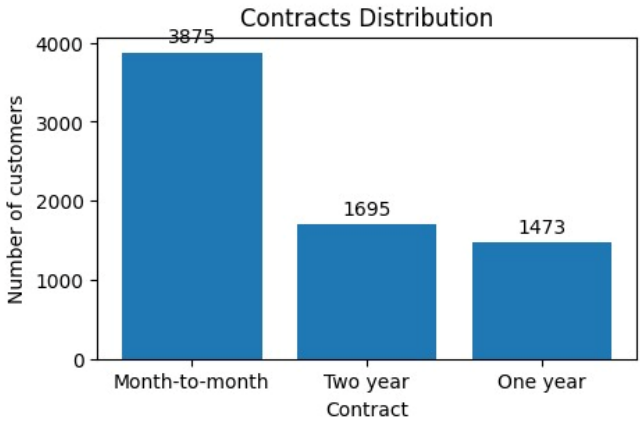
12. Tenure Distribution:

- This visualization represents the distribution of customers based on their tenure, i.e., the length of time (in months) they have been associated with the telecom service provider.
- Tenure is a key indicator of customer loyalty and service engagement.
- Each bar corresponds to a tenure range (in months).
- The height of the bar indicates the number of customers within that tenure range.
- A large proportion of customers fall within the lower tenure range, indicating many customers are relatively new.
- Customer count gradually decreases as tenure increases, suggesting churn is more likely during the early stages.
- Customers with longer tenure represent a smaller but more stable and loyal segment.



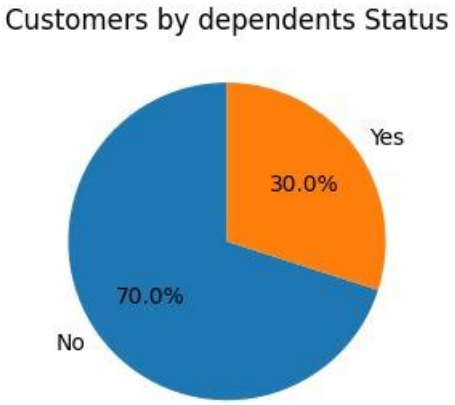
13. Contract Distribution:

- This visualization represents the distribution of customers based on the type of contract they have subscribed to.
- The contract types typically include:
  1. Month-to-Month contracts
  2. One Year contracts
  3. Two-Year contracts
- This graph helps in understanding customer commitment levels and service stability.
- The height of the bar indicates the number of customers subscribed to that contract.
- The visualization clearly differentiates between short-term and long-term customer commitments.A large proportion of customers are subscribed to month-to-month contracts, indicating a preference for flexibility.
- Month-to-month contracts are typically associated with higher churn risk.
- Long-term contracts reflect stronger customer trust and satisfaction.



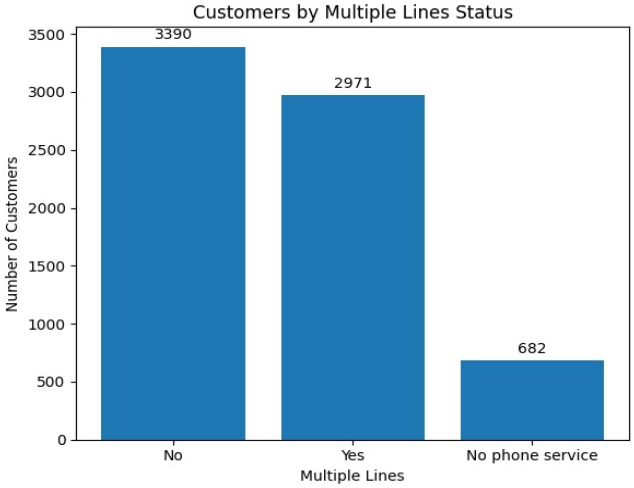
14. Customers by Dependents:

- This visualization represents the distribution of customers based on their dependent status, categorizing them into:
  1. Customers with dependents
  2. Customers without dependents
- Dependent status is an important demographic factor that influences customer needs, service preferences, and retention behavior.
- Each bar corresponds to a dependent status category (Yes / No).
- The height of the bar indicates the number of customers in each category.
- A larger proportion of customers do not have dependents, indicating a significant segment of individual or single users.
- Customers without dependents tend to be more price-sensitive and flexible, increasing their churn risk.



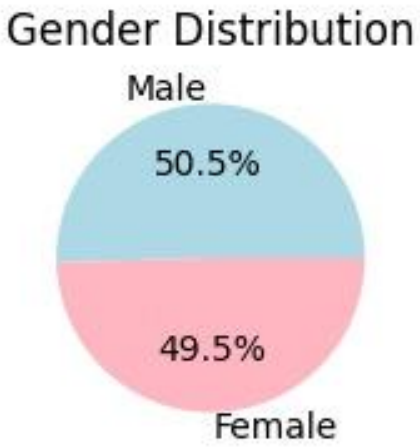
15. Customers by Multiple Lines:

- This visualization represents the distribution of customers based on whether they have multiple phone lines subscribed under their account.
- Customers are categorized into:
  1. Customers with multiple lines
  2. Customers with a single line
- 3.Customers with no phone service
- This analysis helps understand customer usage patterns and service complexity.
- Each bar corresponds to a multiple-lines category.
- The height of each bar indicates the number of customers in that category.
- Customers with multiple lines generally show higher engagement and lower churn risk



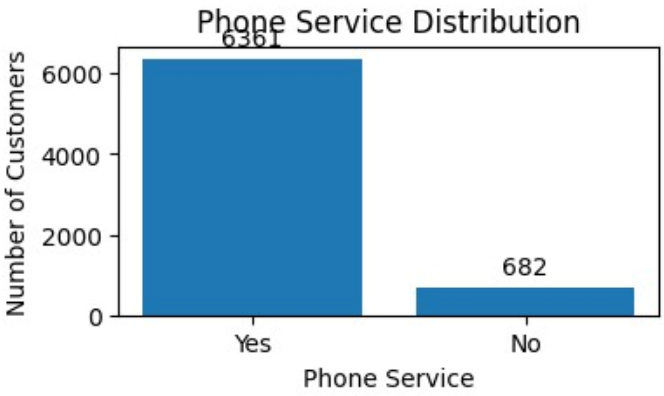
16. Gender Distribution:

- This visualization represents the distribution of customers based on gender, categorizing them into:
  1. Male customers
  2. Female customers
- Each bar corresponds to a gender category.
- The height of the bar indicates the total number of customers belonging to that gender.
- The visualization offers a clear comparison of male and female representation within the telecom customer base.
- Gender distribution helps in understanding overall customer composition.



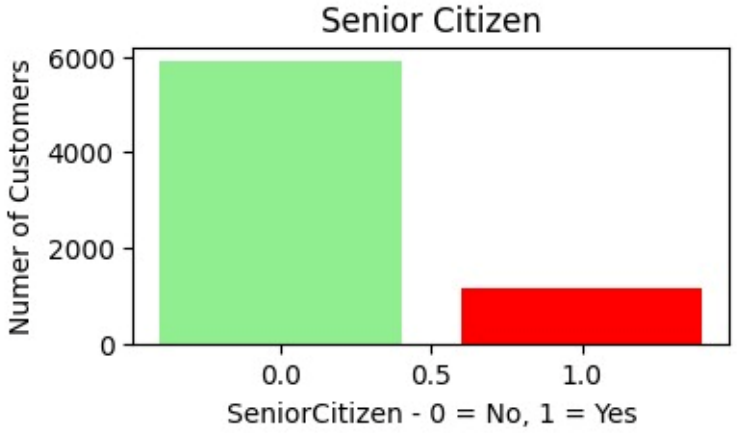
17. Phone Service Distribution:

- This visualization represents the distribution of customers based on whether they have subscribed to the phone service offered by the telecom provider.
- Customers are categorized into:
  1. Customers with phone service
  2. Customers without phone service
- This graph helps assess the adoption level of basic telecom services among customers.
- Each bar corresponds to a phone service category (Yes / No).
- The height of each bar indicates the number of customers in that category.
- The visualization provides a clear view of how widely phone services are used within the customer base.
- The majority of customers have subscribed to phone services, indicating it is a core offering.
- A smaller segment of customers does not use phone services, possibly relying on alternative communication platforms.
- Lack of phone service may indicate customers who are more likely to churn or switch providers.



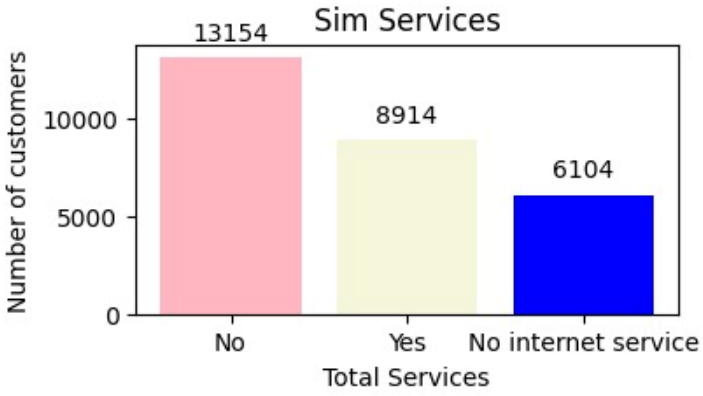
18. Senior Citizens:

- This visualization represents the distribution of customers based on senior citizen status, categorizing them into:
  1. Senior citizens
  2. Non-senior citizens
- This analysis helps in understanding the proportion of elderly customers within the overall telecom customer base.
- Each bar corresponds to a senior citizen category (Yes / No)
- The height of the bar indicates the number of customers in each category.
- The visualization provides a clear demographic split between senior and non-senior customers.
- The majority of customers belong to the non-senior citizen category.
- This group often exhibits higher churn sensitivity due to cost, technology usage, or service complexity.



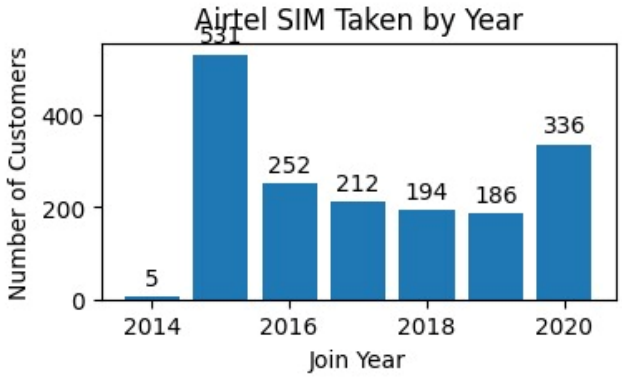
19. Sim Services:

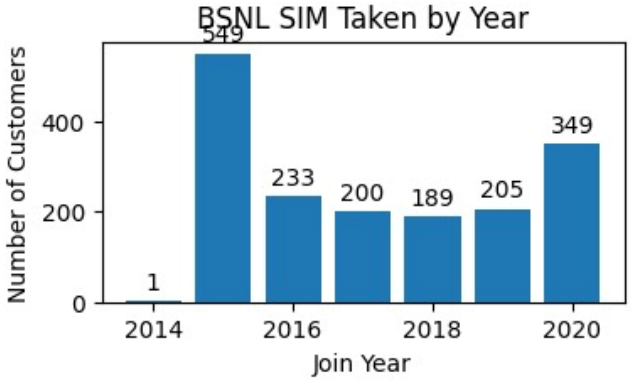
- This visualization represents the distribution of customers based on the services subscribed under their SIM connection.
- SIM services typically include:
  1. Phone Service
  2. Internet Service
  3. Streaming Services
- The graph helps understand how customers utilize different services provided through their SIM.
- Each bar corresponds to a specific SIM service category.
- The height of the bar indicates the number of customers subscribed to that service.
- Value-added services have comparatively lower adoption, highlighting potential upselling opportunities.
- Customers subscribed to multiple SIM services tend to have higher engagement and lower churn probability.

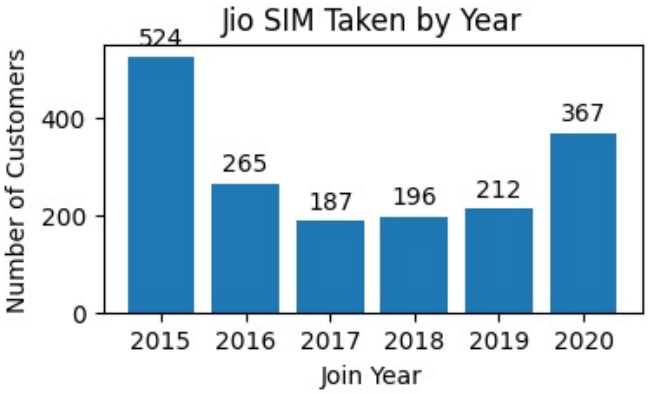


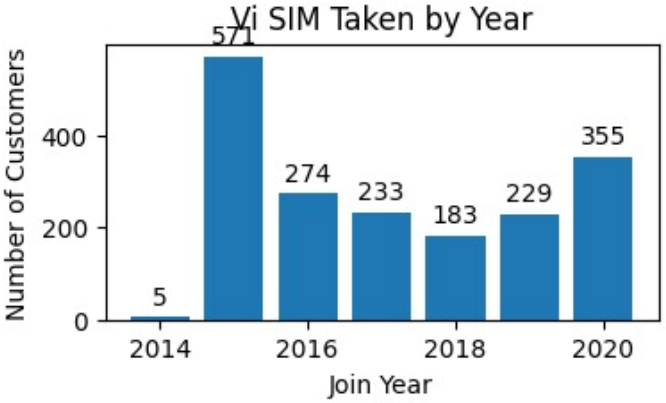
20. Sim's Taken by Year:

- This visualization summarizes the year-wise distribution of SIMs taken by customers
- Each bar represents a specific year.
- The height of the bar indicates the number of SIMs activated in that year.
- The graph provides a clear comparison of SIM adoption over time.
- Years with higher bars indicate strong SIM acquisition, often driven by promotions, competitive pricing, or market expansion.





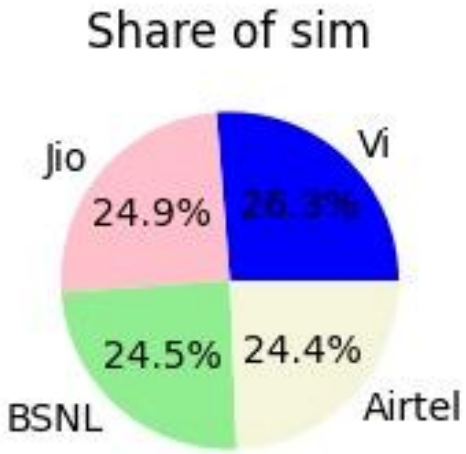




21. Share of Sim:

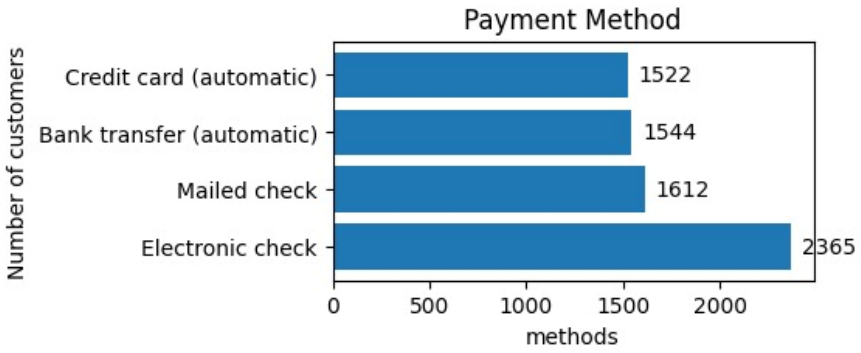
- This visualization represents the distribution of customers based on the SIM provider they are using.
- The SIM providers included in the analysis are:
  1. Jio
  2. Airtel
  3. VI
  4. BSNL

- Each bar corresponds to a SIM provider.
- The height of the bar indicates the number of customers using that particular SIM.
- One or two SIM providers dominate the customer base, indicating stronger brand presence and service reach.
- SIM providers with lower customer counts may face competitive challenges such as pricing or network coverage.
- SIM distribution provides a foundational understanding for further churn and performance analysis by provider.



22. Payment Method:

- This visualization represents the distribution of customers based on the payment methods they use to pay their telecom bills.
- Common payment methods include:
  1. Electronic Check
  2. Credit Card (Automatic)
  3. Bank Transfer (Automatic
  4. Mailed Check
- This graph helps understand customer payment preferences and billing behavior.
- Each bar corresponds to a specific payment method.
- The height of the bar indicates the number of customers using that payment method.
- The visualization provides insight into how customers prefer to manage their payments.
- A significant number of customers use electronic check, indicating a preference for manual or flexible payment options.
- Customers using automatic payment methods (credit card or bank transfer) generally show more stable behavior.
- Manual payment methods are often associated with higher churn risk.



Conclusion:

The exploratory analysis conducted in this project highlights the key factors that significantly influence customer churn in the telecom domain. The results indicate that customer tenure, contract type, payment preferences, service usage, and billing characteristics play a crucial role in determining retention behaviour. Customers with shorter tenure periods, flexible month-to-month contracts, manual payment methods, and limited-service subscriptions are more prone to churn, whereas customers committed to long-term contracts, using multiple services, and opting for digital billing show stronger retention.

The insights derived from visual analysis provide valuable business intelligence and reinforce the importance of targeted customer engagement. By focusing on early-stage customers, promoting bundled services, encouraging automated payment methods, and designing customized plans for high-risk segments, organizations can effectively reduce churn. Overall, the findings support data-driven decision-making and lay a strong foundation for building predictive churn models and implementing strategic retention initiatives.