 **Customer Churn Analysis**

**for**

**a Telecom Company**

**AIM**

Customer acquisition costs are high in the telecom industry, making it far more cost-effective to retain existing customers than to attract new ones. Churn not only results in direct revenue loss but also affects long-term brand reputation and market competitiveness. By understanding the reasons why customers leave, telecom companies can take targeted actions—such as personalized offers, service improvements, or proactive support—to retain valuable customers.

This study is essential because:

* It helps **identify patterns** and risk factors associated with churn.
* It supports **data-driven decision-making** for customer retention strategies.
* It enables the **prediction of future churn**, allowing companies to act before a customer actually leaves.
* It improves overall **customer experience and satisfaction**, leading to long-term profitability.

The aim of this project is to analyze customer data from a telecom company to identify the key factors contributing to customer churn, and to develop a predictive model that can accurately classify customers at risk of leaving. The insights gained will support strategic decision-making to enhance customer retention and reduce churn rates.

**OBJECTIVE**

1. **To explore and understand customer demographics and service usage patterns** using exploratory data analysis (EDA).
2. **To calculate churn rates** across different customer segments such as contract types, payment methods, and internet services.
3. **To identify key features influencing customer churn** through statistical analysis and visualization techniques.
4. **To build predictive machine learning models** that classify whether a customer is likely to churn or not.
5. **To evaluate model performance** using appropriate metrics like accuracy, precision, recall, and F1-score.
6. **To provide actionable insights and recommendations** that telecom companies can use to reduce churn and improve customer satisfaction.

**DATASET**

The dataset used for this project is based on customer information from a telecom company, providing data on customer demographics, account details, services subscribed, and whether or not the customer churned (left the company).

Dataset Details:

* Source: Telco Customer Churn dataset (often publicly available).
* Number of Records: 7,043 customers.
* Features: 21 columns in total.

Key Columns:

1. customerID: Unique identifier for each customer.
2. gender: Gender of the customer (Male/Female).
3. SeniorCitizen: Whether the customer is a senior citizen (Yes/No).
4. Partner: Whether the customer has a partner (Yes/No).
5. Dependents: Whether the customer has dependents (Yes/No).
6. tenure: Number of months the customer has been with the company.
7. PhoneService: Whether the customer has phone service (Yes/No).
8. MultipleLines: Whether the customer has multiple lines (Yes/No).
9. InternetService: Type of internet service the customer uses (DSL/Fiber optic/No).
10. OnlineSecurity: Whether the customer has online security (Yes/No).
11. OnlineBackup: Whether the customer has online backup (Yes/No).
12. DeviceProtection: Whether the customer has device protection (Yes/No).
13. TechSupport: Whether the customer has tech support (Yes/No).
14. StreamingTV: Whether the customer has streaming TV (Yes/No).
15. StreamingMovies: Whether the customer has streaming movies (Yes/No).
16. Contract: Type of contract the customer has (Month-to-month/One year/Two year).
17. PaperlessBilling: Whether the customer has paperless billing (Yes/No).
18. PaymentMethod: The payment method used by the customer (Electronic check/Bank transfer/Etc.).
19. MonthlyCharges: The monthly charge for the customer’s account.
20. TotalCharges: Total charges billed to the customer.
21. Churn: The target variable indicating whether the customer has churned (Yes) or stayed (No).

Data Type:

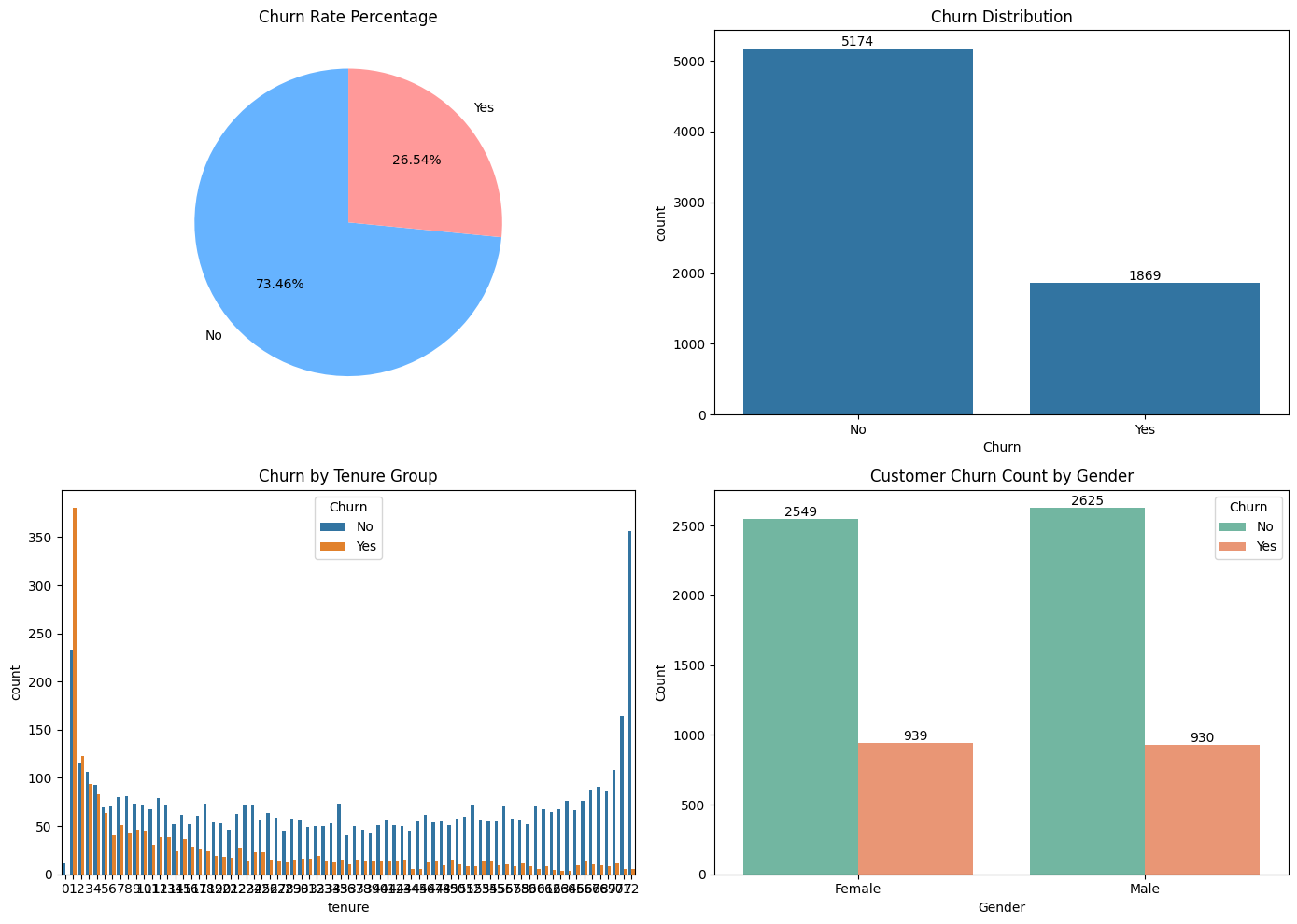
* Categorical Variables: Most variables are categorical (e.g., gender, Contract, Churn).
* Numerical Variables: tenure, MonthlyCharges, TotalCharges.

**CHURN: Churn** refers to the phenomenon where a customer discontinues their relationship with a company or service, essentially "leaving" or canceling their subscription. In the context of the telecom industry (and other service industries), **churn** typically represents customers who stop using the company's services.

**Process & analyze data using Pandas in Python.**

Data preprocessing is an essential step to prepare the data for modeling. It includes handling missing values, encoding categorical variables, scaling numerical features, and splitting the dataset for training and testing. Here's a detailed guide on how to preprocess your **Customer Churn Analysis** data.

Link of python data analysis : [https://colab.research.google.com/drive/1Y86bLC68c3erSjbqWxR18hcLoGms-XHt#scrollTo=-MrhUtmc7aP9](https://colab.research.google.com/drive/1Y86bLC68c3erSjbqWxR18hcLoGms-XHt%23scrollTo=-MrhUtmc7aP9)

****

### 1. ****Churn Rate Percentage (Pie Chart - Top Left)****

* This pie chart shows the **proportion** of customers who **churned** (left the company) vs those who **did not churn**.
* **73.46%** of customers **stayed** with the company (No), while **26.54%** of customers **churned** (Yes).
* **Insight:**  
  The churn rate is **moderate**. About **one-fourth** of customers are leaving, which is a **serious business concern** but not extremely critical yet. Strategies are needed to reduce this.

### 2. ****Churn Distribution (Bar Chart - Top Right)****

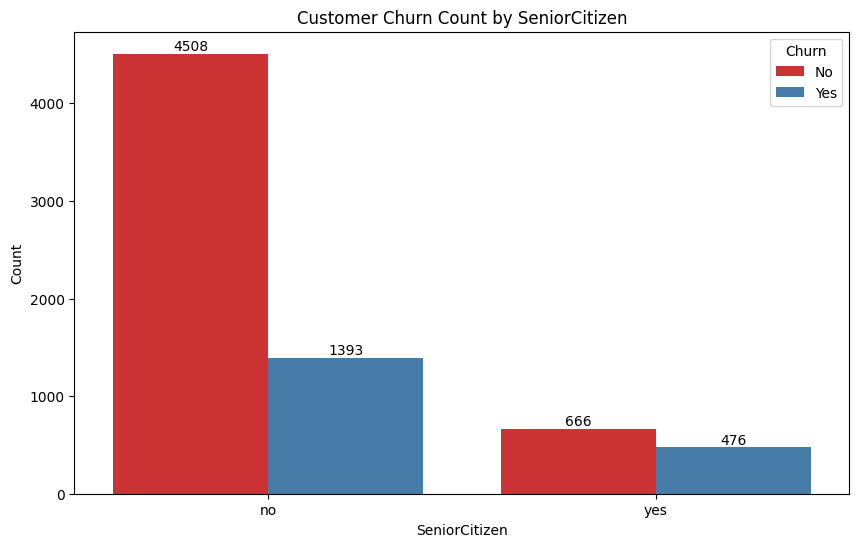
* This bar graph shows the **number of customers** who churned vs who didn't.
* Around **5174 customers did not churn**, while **1869 customers churned**.
* **Insight:**  
  It confirms the churn rate seen in the pie chart. There's a much larger number of loyal customers, but **almost 1 in 4 customers is churning**, which must be addressed.

### 3. ****Churn by Tenure Group (Bar Chart - Bottom Left)****

* This graph shows the **count of churned and non-churned customers** based on their **tenure** (how many months they've been a customer).
* **Tenure** values range from 0 to 72 months.
* **Observation:**
  + Customers with **very low tenure (0–5 months)** have a **high churn rate**.
  + As tenure increases (customers who have stayed longer), churn decreases.
* **Insight:**  
  New customers are **more likely to leave early** if they are dissatisfied initially.  
  Longer-tenured customers are **more loyal**.  
  So, **early-stage customer experience** is crucial.

### 4. ****Customer Churn Count by Gender (Bar Chart - Bottom Right)****

* This bar chart compares **churn counts across genders** (Male and Female).
* **Observation:**
  + Female: 2549 stayed, 939 churned.
  + Male: 2625 stayed, 930 churned.
* **Insight:**  
  **Churn rate is quite similar between genders**.  
  This suggests **gender is not a strong differentiating factor** in churn behavior for this company.

****

### ****Customer Churn Count by SeniorCitizen****

* The graph shows **churn distribution based on whether the customer is a senior citizen** or not.
* Two groups: **No** → Not a senior citizen **Yes** → Senior citizen

**Observation:**

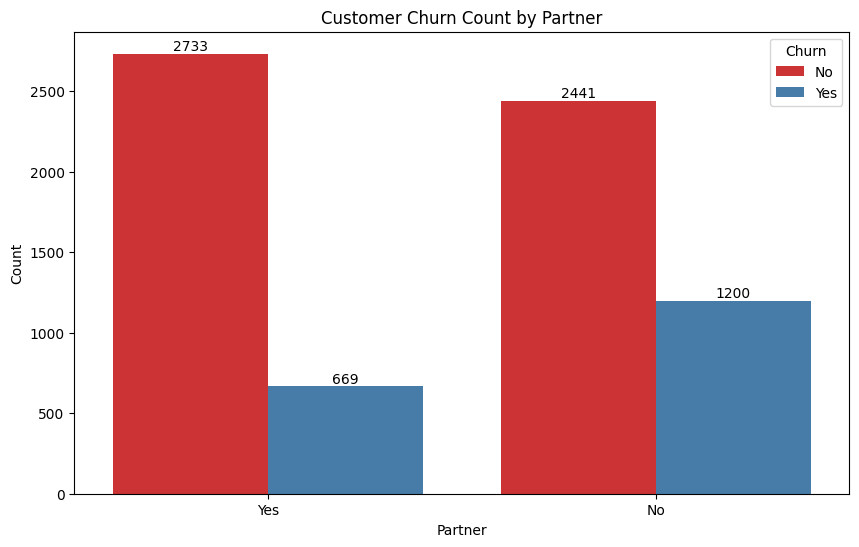
* Among **non-senior citizens**:
  + **4508** customers stayed.
  + **1393** customers churned.
* Among **senior citizens**:
  + **666** customers stayed.
  + **476** customers churned.

**Insights:**

* **Senior citizens have a higher churn rate** compared to non-senior citizens.
* Although the **total number** of senior citizens is smaller, their **churn percentage is higher**.
* **Non-senior citizens** are **more likely to stay**.

🔵 **Simple interpretation:**

Senior citizens seem **less loyal** and **more likely to leave** the telecom company than younger customers.

****

### ****Customer Churn Count by Partner****

* This bar chart shows **churn behavior based on whether the customer has a partner** (spouse) or not.
* Two groups: **Yes** → Has a partner **No** → Does not have a partner

**Observation:**

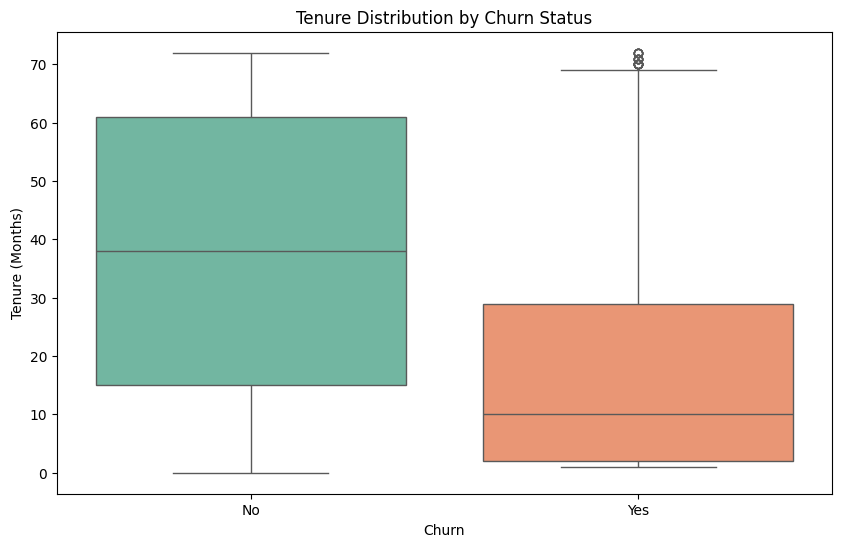
* Customers **with a partner**:
  + **2733** stayed.
  + **669** churned.
* Customers **without a partner**:
  + **2441** stayed.
  + **1200** churned.

**Insights:**

* Customers **without a partner** have a **higher churn rate** compared to those **with a partner**.
* Having a partner seems to be **associated with higher loyalty** and **lower likelihood of churn**.

**Simple interpretation:** Customers who are **single or without a partner** are **more likely to churn** than those who have a partner.

**Tenure Distribution by Churn Status:**

****

* This boxplot shows the distribution of **tenure (months with the company)** for customers who **churned** (Yes) and **did not churn** (No).

**Observation:**

* **Non-churned customers (No)**:
  + Median tenure is around **38 months**.
  + Most customers stayed between **15 to 60+ months**.
* **Churned customers (Yes)**:
  + Median tenure is around **10 months**.
  + Most churned within the **first 0–30 months**.
  + A few customers with longer tenure (>60 months) still churned (seen as outliers).

**Insights:**

* Customers who have **shorter tenures** are **much more likely to churn**.
* Longer-tenured customers are **more loyal** and **less likely to leave**.
* Early-stage customers (new users) are a **critical group** for churn prevention strategies.

🔵 **Simple Interpretation:** The **longer** a customer stays, the **less likely** they are to churn.  
**Newer customers** need **more attention** to reduce early churn.

**Churn by Tenure Categories: Correlation Heatmap:**



* The **x-axis** represents the **Churn status** (No or Yes).
* The **y-axis** represents **tenure buckets** (in months): 0-12, 12-24, 24-36, 36-48, 48-60, and 60+.
* The **numbers** inside the cells represent the **count of customers** in each group.

**Observation:**  In the **0–12 months** group:

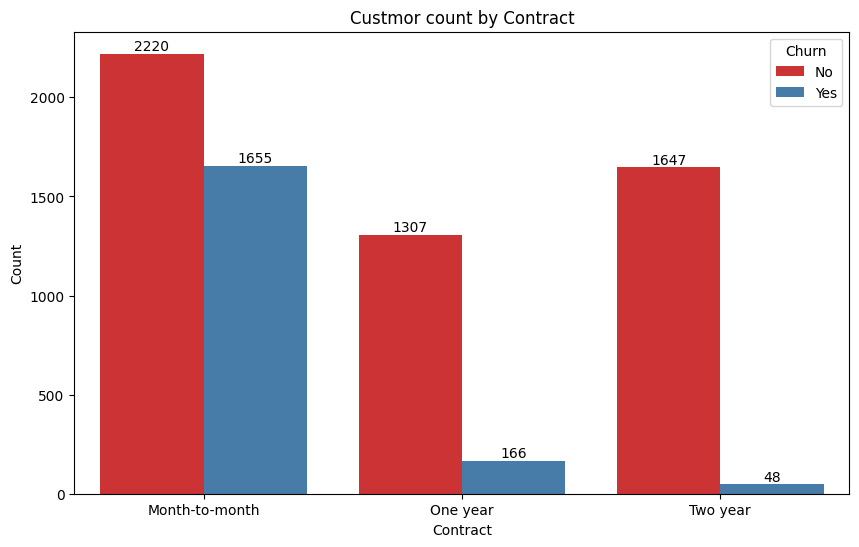
* + Churners (Yes) are **1037**, almost **as high** as non-churners (No) **1138**.
  + ➔ **Very high churn risk** among newer customers.
* As tenure increases:
  + The number of churners **steadily decreases**.
  + For example, in 60+ months, **only 93** customers churned compared to **1314** non-churners.
* From **24 months onwards**, **churners** become significantly **fewer** than non-churners.

**Insights:**

* **Majority of churn** happens **early**, especially within the **first 12 months**.
* Customers staying longer than **2 years** have a **much lower churn rate**.

**Simple Interpretation:** New customers are **much more vulnerable** to churn.  
Retention efforts should **focus heavily** on customers in their **first year**.

### ****Customer Count by Contract Type:****



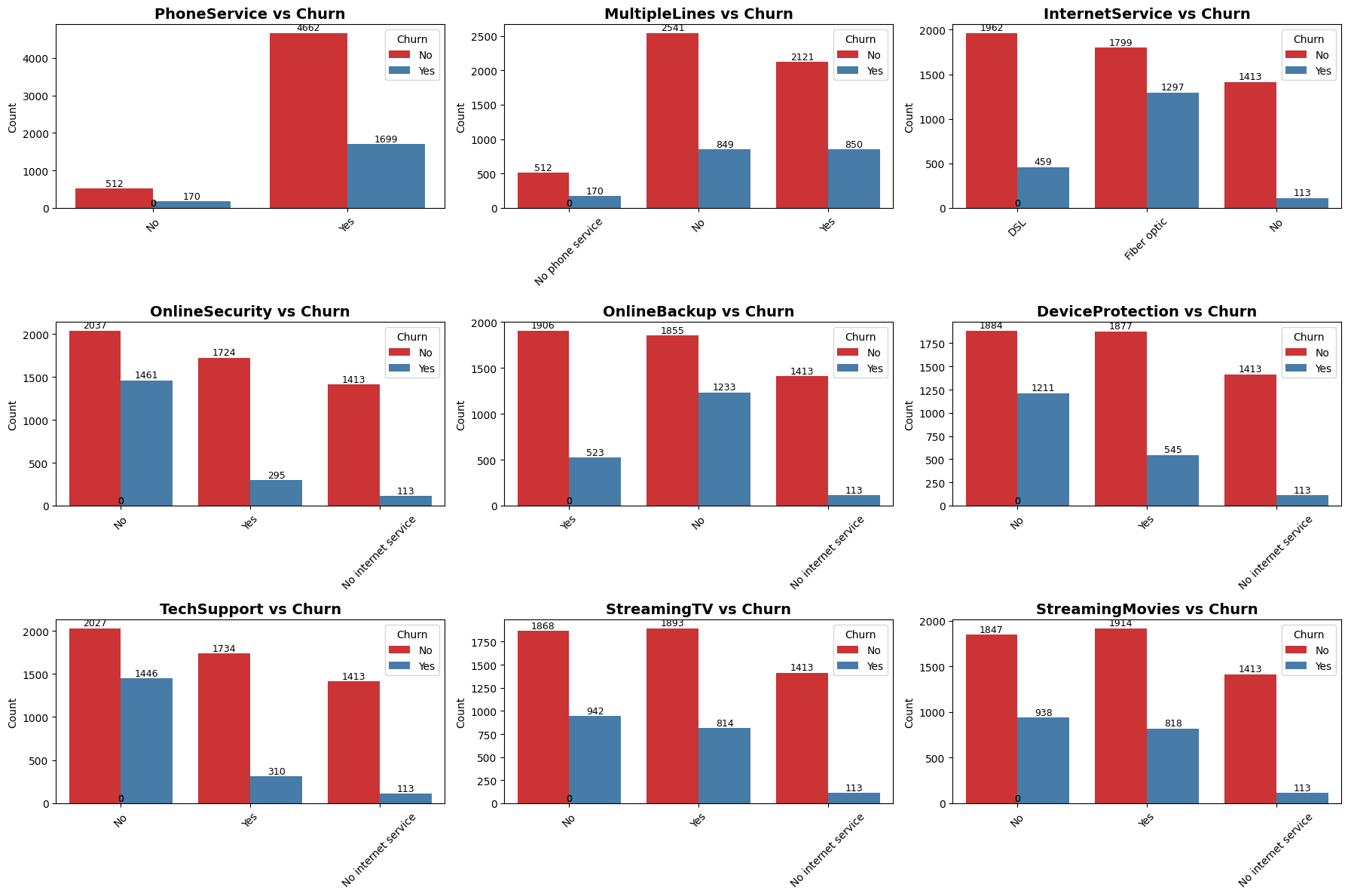
* The **bars** are split by **Churn status** (No in red, yes in blue).
* The **numbers** above the bars indicate **count of customers** in each group.

**Key Observations:**

* **Month-to-Month contracts**:
  + Very high churn (Yes: **1655** customers) compared to churn in **One year** or **Two year** contracts.
  + Also, non-churners are many (**2220**), but **churn rate** is **higher** here.
* **One-year contracts**:
  + Much fewer churners (**166**) compared to non-churners (**1307**).
  + Churn rate **significantly drops** compared to month-to-month.
* **Two-year contracts**:
  + **Lowest churn** (**only 48** customers churned vs **1647** who stayed).
  + Customers with two-year contracts are **most loyal**.

**Simple Interpretation:** Customers with **short-term (month-to-month) contracts** churn a lot more.  
Longer commitments like **one-year or two-year contracts greatly reduce churn**.

**Countplots :**



Each plot is a **countplot** that compares a **service feature** (like PhoneService, OnlineBackup, etc.) with **Customer Churn** (Yes or No).

* **X-axis**: Shows the **category** of the service feature (example: Yes/No for Phone Service).
* **Y-axis**: Shows the **number of customers** (count) in that category.
* **Hue (color)**: Splits the customers based on **whether they churned or not**:
  + Red color → Customers who **did not churn** (Churn = No)
  + Blue color → Customers who **churned** (Churn = Yes)

**SQL: Find churn rate by region, analyze high-churn service plans.**

**SQL QUERY:**

**USE project;**

**Select \* from cleaned\_churn\_analysis;**

**# SQL: Find churn rate by region, analyze high-churn service plans.**

**SELECT**

**Contract,**

**COUNT(\*) AS total\_customers,**

**SUM(CASE WHEN Churn = 'Yes' THEN 1 ELSE 0 END) AS churned\_customers,**

**ROUND(100.0 \* SUM(CASE WHEN Churn = 'Yes' THEN 1 ELSE 0 END) / COUNT(\*), 2) AS churn\_rate\_percentage**

**FROM**

**cleaned\_churn\_analysis**

**GROUP BY**

**Contract**

**ORDER BY**

**churn\_rate\_percentage DESC;**

**OUTPUT :**

**Churn rate by Contract type:**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Contract |  | total\_customers | churned\_customers |  | churn\_rate\_percentage | |
| Month-to-month |  | 3875 | 1655 |  | 42.71 |  |
| One year |  | 1473 | 166 |  | 11.27 |  |
| Two year |  | 1695 | 48 |  | 2.83 |  |

### Insights:

* **Month-to-month contracts** have the highest churn rate (42.71%), indicating that customers without long-term commitments are more likely to leave.
* **One year contracts** reduce churn significantly (11.27%), suggesting some commitment improves retention.
* **Two year contracts** have the **lowest churn rate (2.83%)**, showing that longer commitments strongly correlate with customer retention.

**SELECT**

**InternetService,**

**COUNT(CASE WHEN Churn = 'Yes' THEN 1 END) AS Churned\_Customers,**

**COUNT(\*) AS Total\_Customers,**

**ROUND(**

**(COUNT(CASE WHEN Churn = 'Yes' THEN 1 END) \* 100.0) / COUNT(\*),**

**2**

**) AS Churn\_Rate\_Percentage**

**FROM**

**cleaned\_churn\_analysis**

**GROUP BY**

**InternetService**

**ORDER BY**

**Churn\_Rate\_Percentage DESC;**

**OUTPUT:**

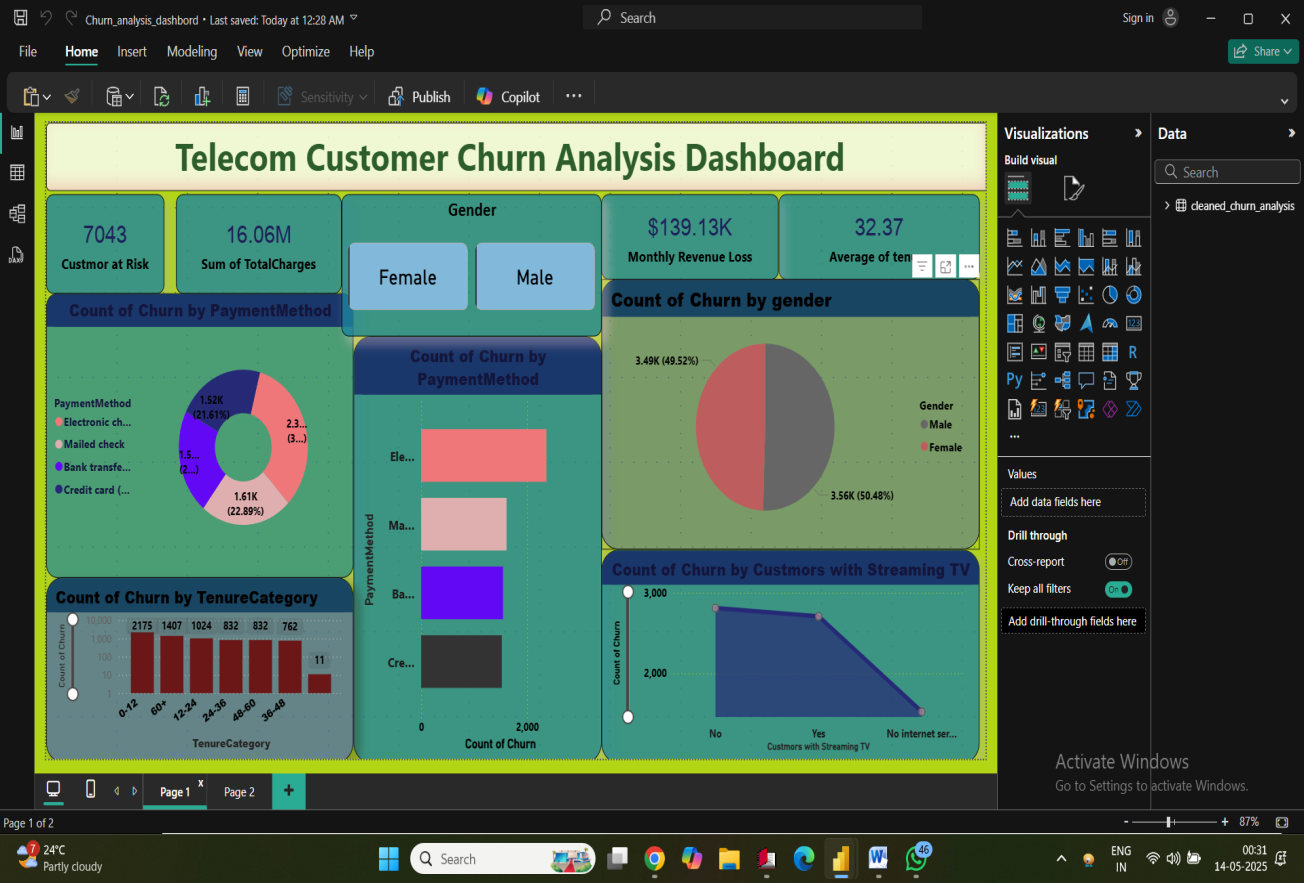
**churn data by Internet service type**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | InternetService | Churned\_Customers |  | Total\_Customers |  | Churn\_Rate\_Percentage | | Fiber optic | 1297 |  | 3096 |  | 41.89 | | DSL | 459 |  | 2421 |  | 18.96 | | No | 113 |  | 1526 |  | 7.4 | |

### Insights:

* **Fiber optic users** have the highest churn rate (41.89%). This may indicate dissatisfaction with the service, pricing issues, or more competition in that segment.
* **DSL customers** show a moderate churn rate (18.96%), significantly lower than fiber optic.
* Customers with **no internet service** churn the least (7.40%), possibly due to less usage dependency or lower service expectations.

**Power BI: Create visuals for churn rate trends, customer segmentation, and revenue loss.**

****

### ****Top Summary KPIs (Cards at the top):****

1. **Customer at Risk: 7043**
   * Number of customers identified as likely to churn.
2. **Sum of TotalCharges: 16.06M**
   * Total revenue from all customers (or churned customers depending on the filter).
3. **Gender Filter (Female / Male):**
   * Interactive filter to view gender-based churn metrics.
4. **Monthly Revenue Loss: $139.13K**
   * Estimated revenue lost per month due to churned customers.
5. **Average of Tenure: 32.37**
   * Average tenure (in months) of customers before churning.

### ****Visualizations & Charts:****

#### 1. ****Count of Churn by Payment Method (Donut Chart & Bar Chart):****

* **Payment Methods:**
  + Electronic Check (most churn: 2.3K)
  + Mailed Check
  + Bank Transfer
  + Credit Card
* **Insight:** Customers using electronic checks have the highest churn rate.

#### 2. ****Count of Churn by Gender (Pie Chart):****

* Female: 50.48% (3.56K)
* Male: 49.52% (3.49K)
* **Insight:** Churn is almost equally distributed between genders.

#### 3. ****Count of Churn by Tenure Category (Bar Chart):****

* Categories like 0-12, 60+, etc.
* Highest churn is in 0–12-month category (2,175 customers), indicating new customers leave quickly.
* **Insight:** Retention issues are high in early months.

#### 4. ****Count of Churn by Customers with Streaming TV (Line Chart):****

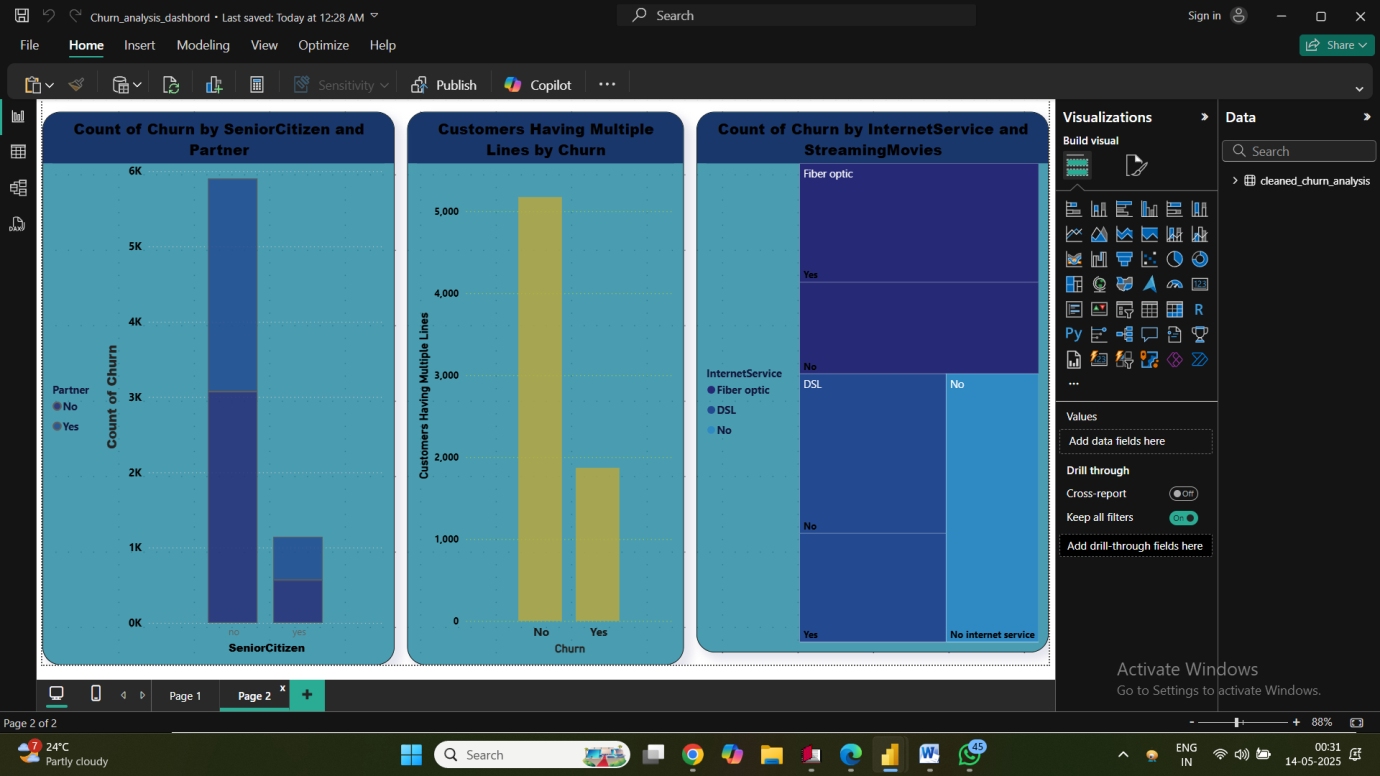
* Comparison between:
  + Customers without streaming TV
  + With streaming TV
  + Without internet service
* **Insight:** Customers with no streaming services seem to churn more.

### ****Other Dashboard Features:****

* **Page Navigation (Bottom Tab):**
  + Indicates there's a second page (Page 2) with potentially more insights.
* **Visual Interactivity:**
  + Filters and slicers (like gender selection) allow dynamic exploration of data.

### ****Overall Insights:****

* **High churn risk** among customers with:
  + Electronic check payments
  + Low tenure (especially <12 months)
  + No streaming services
* Gender doesn't have a major impact.
* Revenue loss is significant, underlining the importance of retention strategies.

****

### 1. ****Count of Churn by SeniorCitizen and Partner (Bar Chart - Left Panel):****

* **X-axis:** SeniorCitizen (Yes/No)
* **Bar Color Split:** By Partner status (Yes/No)
* **Insight:**
  + Most churned customers are **not senior citizens** and **do not have a partner**.
  + Churn is relatively lower among senior citizens with a partner.
  + This suggests **married non-senior customers churn the least**, and **single younger customers churn more**.

### 2. ****Customers Having Multiple Lines by Churn (Clustered Column Chart - Middle Panel):****

* **X-axis:** Churn (Yes/No)
* **Y-axis:** Number of customers with multiple lines
* **Insight:**
  + Majority of **non-churned customers** have multiple lines.
  + Fewer customers with multiple lines have churned.
  + **Having multiple lines might be a sign of customer commitment** and could reduce churn likelihood.

### 3. ****Count of Churn by Internet Service and Streaming Movies (Treemap - Right Panel):****

* **Split by:**
  + **InternetService:** Fiber Optic, DSL, No internet
  + **StreamingMovies:** Yes/No
* **Insight:**
  + **Customers using fiber optic with streaming services churn the most**.
  + **Customers with no internet service churn the least**.
  + Indicates dissatisfaction may be higher with premium services like fiber + streaming, potentially due to cost or service issues.

### ✅ ****Overall Insights from Page 2:****

* **High churn risk groups:**
  + Customers **not senior citizens** and **without a partner**
  + **Single-line** customers
  + **Fiber optic users with streaming**
* **Lower churn risk groups:**
  + Customers with **multiple lines**
  + **Senior citizens with a partner**
  + Those with **no internet service**

### ****Key Metrics:****

* **Total Customers at Risk:** 7,043
* **Total Charges (Churned Customers):** $16.06M
* **Monthly Revenue Loss:** $139.13K
* **Average Customer Tenure:** 32.37 months

### ****Key Findings:****

#### 1. ****Demographics & Customer Profile:****

* **Gender:** Churn is almost equally distributed among **male (49.5%)** and **female (50.5%)** customers.
* **Senior Citizens:** Majority of churned customers are **not senior citizens**.
* **Partner Status:** Customers **without partners** are more likely to churn.
* **Tenure:** Most churn occurs in the **first 12 months**, indicating early dissatisfaction or onboarding issues.

#### 2. ****Services & Usage Patterns:****

* **Internet Service:**
  + Customers with **Fiber Optic internet** and **streaming movies** have **the highest churn rates**.
  + Customers with **no internet service** churn the least.
* **Streaming Services:** Lack of streaming services seems correlated with higher churn.
* **Multiple Lines:** Customers with **multiple lines churn less**, possibly due to higher engagement or bundled services.

#### 3. ****Payment Behavior:****

* **Payment Method:**
  + Highest churn among customers using **Electronic Checks**.
  + Customers using **Credit Cards or Bank Transfers** churn less frequently.

**Recommendations:**

1. **Improve Onboarding Experience:**
   * Focus on reducing churn in the first 12 months by improving support, education, and welcome offers.
2. **Incentivize Long-Term Engagement:**
   * Offer loyalty discounts or bundled services to customers with multiple lines.
3. **Review Pricing & Experience for Fiber Optic + Streaming Customers:**
   * These customers are likely premium users; ensure they receive high-quality service and personalized support.
4. **Address Electronic Payment Users:**
   * Consider offering benefits or alternatives to users of electronic checks to reduce dissatisfaction.
5. **Targeted Retention Campaigns:**
   * Focus on **young, single, non-senior** customers who are more prone to churn.

**Conclusion:** The dashboard provides a comprehensive view of churn behavior. By acting on these insights, the telecom company can **reduce revenue loss**, **improve customer retention**, and **enhance overall satisfaction**.

**THANK YOU**