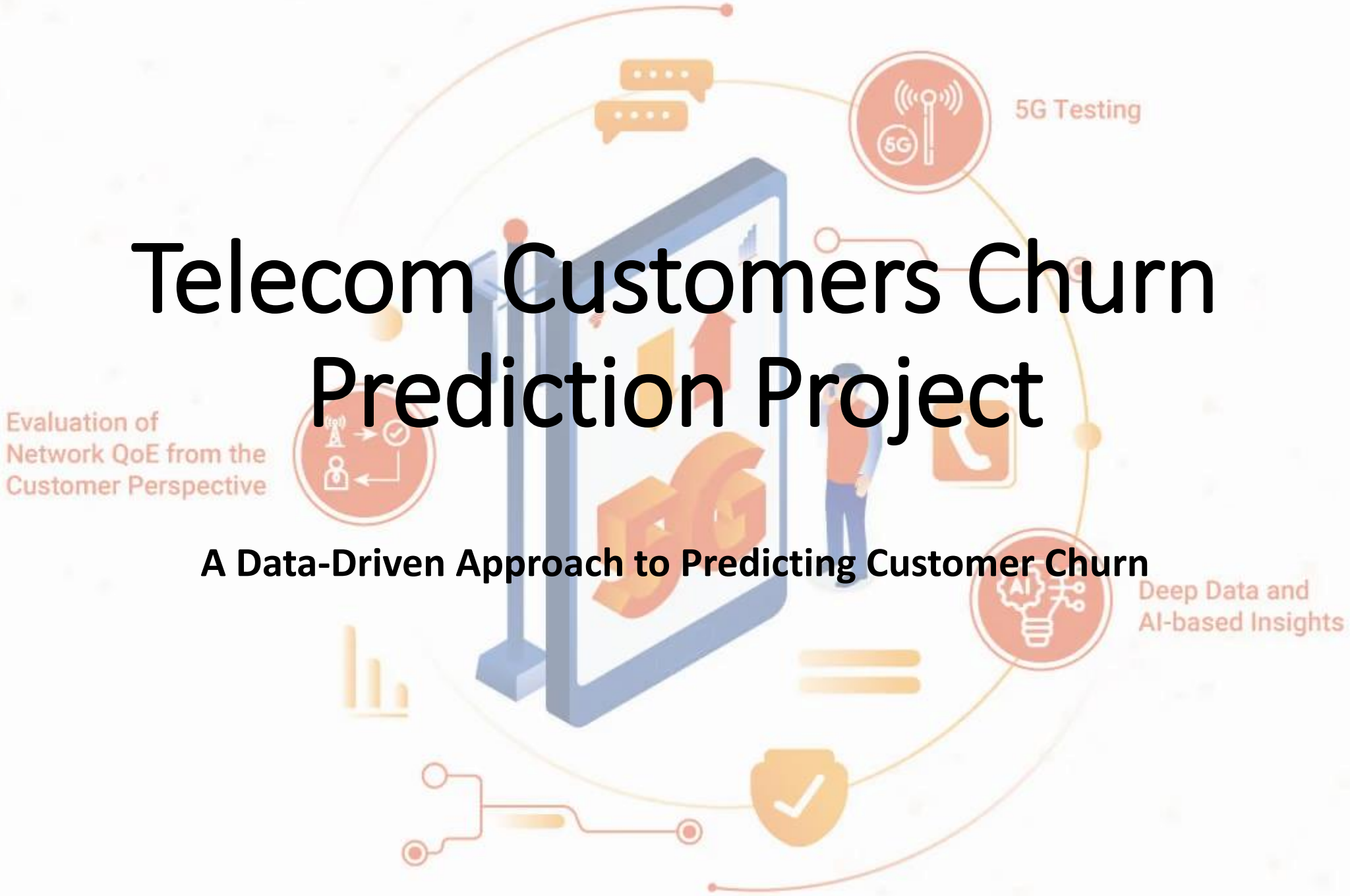


Telecom Customers Churn Prediction Project

A Data-Driven Approach to Predicting Customer Churn

Introduction

The primary objective of the Telecom Customer Churn Prediction Project is to accurately identify customers who are likely to discontinue their services (churn) with the company. By leveraging advanced data analysis and predictive modeling techniques, the project aims to uncover key factors contributing to customer churn, allowing the company to implement targeted retention strategies.

Importance

- Customer retention is crucial for sustaining profitability in the highly competitive telecom industry, where acquiring new customers is significantly more expensive than retaining existing ones. By proactively identifying at-risk customers, the company can take preemptive measures to improve customer satisfaction, reduce churn rates, and ultimately increase long-term revenue. This project not only helps in enhancing customer loyalty but also provides valuable insights into service improvements and marketing strategies, ensuring a stronger market position and better customer relationships.

Dataset Description

- Detailed customer information encompassing demographics, service usage, and account details.
- Includes demographic attributes such as age, gender, and marital status.
- Captures service-related data, including internet service type, phone service and streaming services.
- Financial metrics provided, including monthly charges, total charges and refunds.
- Comprehensive dataset with 2,40,700 records, each representing an individual customer.
- Numerous columns capturing various aspects of customer interactions with the telecom service provider.
- Instrumental in understanding customer behavior and predicting churn.

Key Features

- **Demographic Details**: Provides insights into customer profiles, such as age, gender and marital status.
- **Account Information**: Reflects customer loyalty, contract types and payment methods.
- **Service Usage Patterns**: Helps understand customer interactions with telecom services, including internet and phone services.
- **Financial Metrics**: Offers insights into customer spending behavior through monthly charges, total charges and refunds.
- **Churn-Related Features**: Directly informs about the reasons behind customer attrition.
- **Comprehensive Analysis**: Enables identification of factors most predictive of churn.
- **Targeted Retention Strategies**: Helps the company implement strategies to retain customers effectively.

Data Quality

- **Total Records**: 7,043 customer entries in the dataset.
- **Duplicates**: No duplicate records found, ensuring data uniqueness.
- **Missing Values**: 5,174 missing values in the 'Churn Category' and 'Churn Reason' columns. The dataset contains significant missing data, including 3,877 missing values in the 'Offer' column and approximately 1,526 missing values each in 'Internet Type' and 'Avg Monthly GB Download', along with several other service-related columns.
- **Working Dataset**: After accounting for missing values, only about 20% of the data is available for robust analysis.
- **Data Reduction**: Unnecessary columns like 'Customer ID', 'City', 'Zip Code', 'Latitude', 'Longitude', and 'Churn Category' were dropped to streamline the dataset for analysis.
- **Data Integrity**: Essential columns such as 'Gender', 'Age', 'Phone Service’, 'Contract' and 'Monthly Charges' have complete data, maintaining the integrity of critical features for churn prediction.

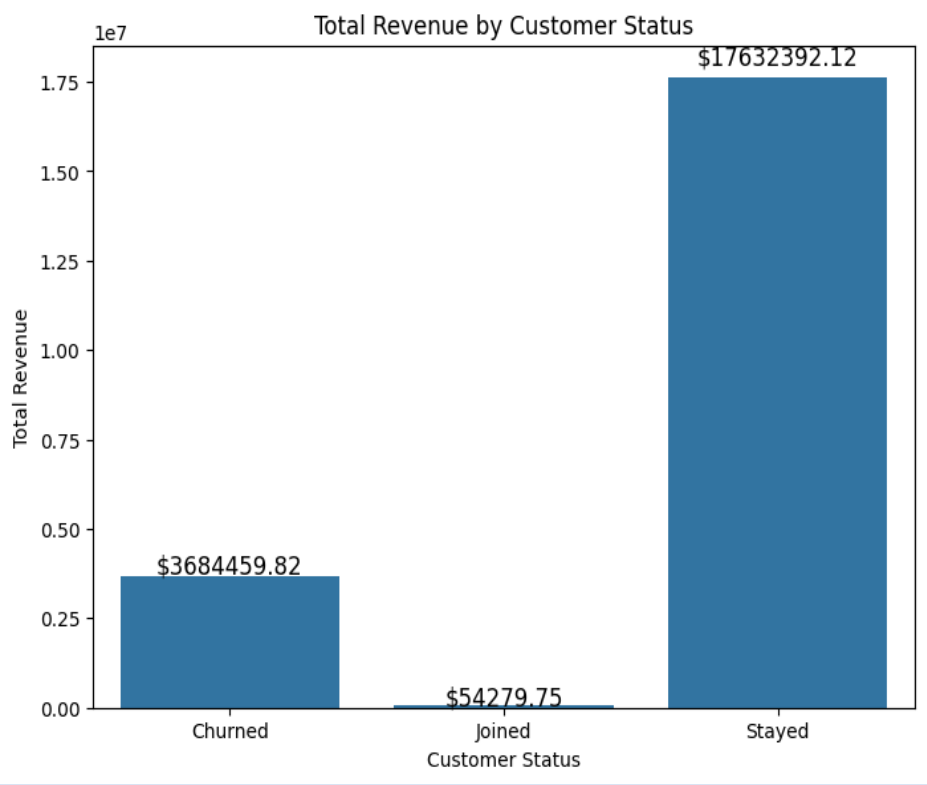
Exploratory Data Analysis (EDA)

Insights from Customer Status Distribution

The bar chart displays the distribution of customers across three categories: "Stayed," "Churned," and "Joined.“

- **Key Insights**:

1. **Stayed**: The majority of customers have remained with the company, with a count exceeding 4,000. This indicates strong customer retention, totalling approximately **$17.63 million**.
2. **Churned**: A significant portion of customers, approximately 1,500, have left the company, still generated a significant revenue amounting to **$3.68 million**. Understanding the reasons for churn is crucial for improving retention strategies.
3. **Joined**: A smaller group of customers has recently joined, with a count around 500, showing potential for growth in the customer base, with total revenue of approximately **$54,279.75**.

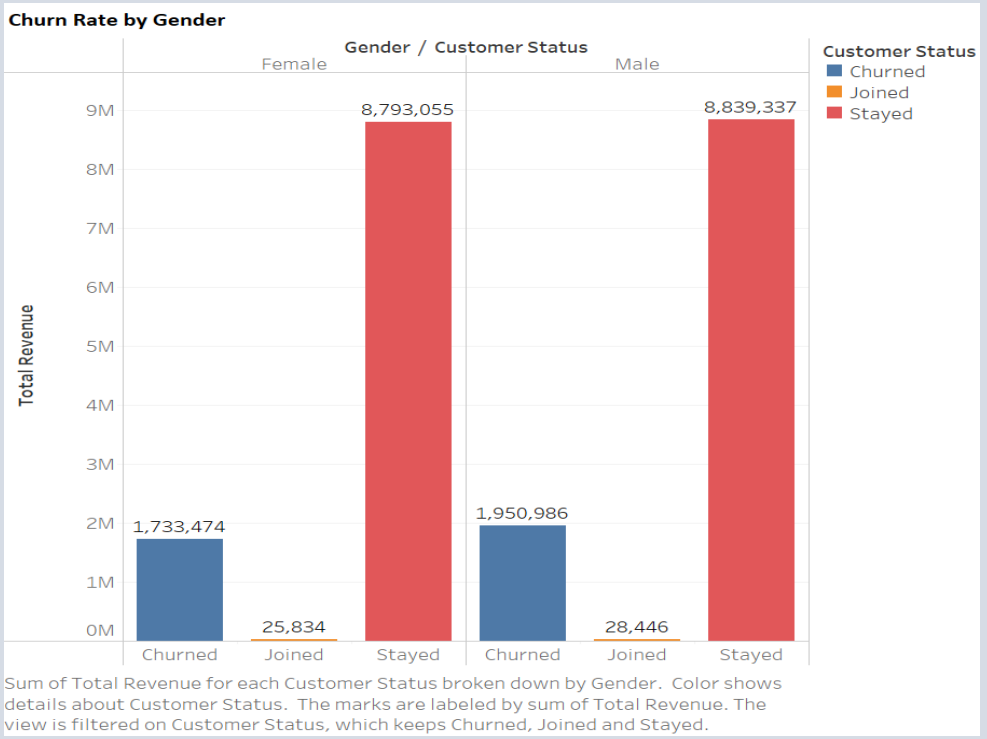


- **Strategic Importance:**

1. **High Revenue from Retained Customers**: The overwhelming majority of revenue comes from customers who have stayed, emphasizing the importance of customer retention strategies.
2. **Revenue Loss Due to Churn**: The revenue from churned customers is substantial, highlighting the potential financial impact of customer churn on the business.
3. **Low Revenue from New Customers**: The relatively low revenue from new customers suggests that while acquiring new customers is important, retaining existing ones is more critical for sustained revenue growth.

Key Insights for Churn Rate by Gender

- **Gender Disparity in Churn:** The visualization reveals a notable difference in churn rates between male and female customers. Male customers have a slightly higher churn rate amounting to **$1.95 million** than female customers of **$1.73 million** .
- **Revenue Impact:** While male customers have a slightly higher churn rate, they also generate more total revenue of **$8.83 million**. This suggests that the company is losing high-value customers at a higher rate.
- **Joined vs. Stayed Revenue:** The revenue generated by customers who have joined and stayed is significantly higher than that of churned customers. This highlights the importance of customer retention and satisfaction.



Recommendations based on above Insights

- **Targeted Retention Efforts:**

1. **Gender-specific campaigns:** Develop tailored retention strategies that address the specific needs and concerns of male customers, who are more likely to churn.
2. **Personalized offers:** Offer personalized promotions, discounts, or upgrades to customers who are at risk of churning.

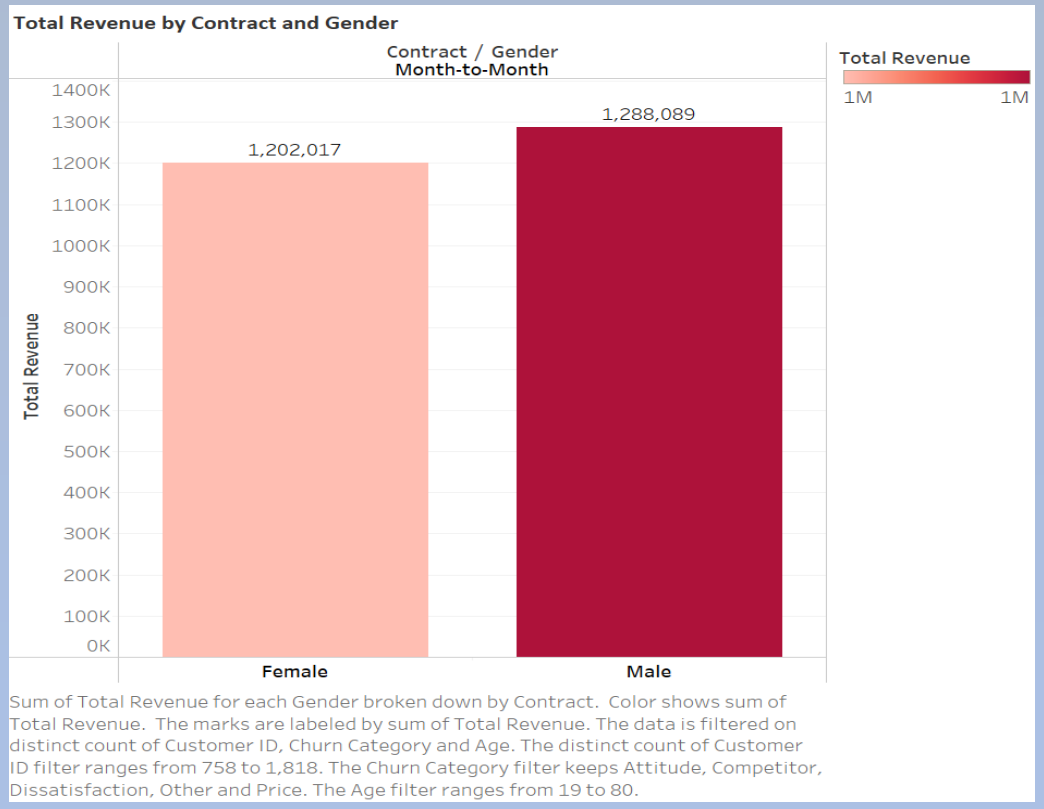
- **Enhanced Customer Experience:**

1. **Improved customer service:** Invest in training and tools to enhance customer service quality, especially for male customers.
2. **Customer satisfaction surveys:** Regularly conduct surveys to gather feedback from customers and identify areas for improvement.

Key Insights for Total Revenue by Contract and Gender

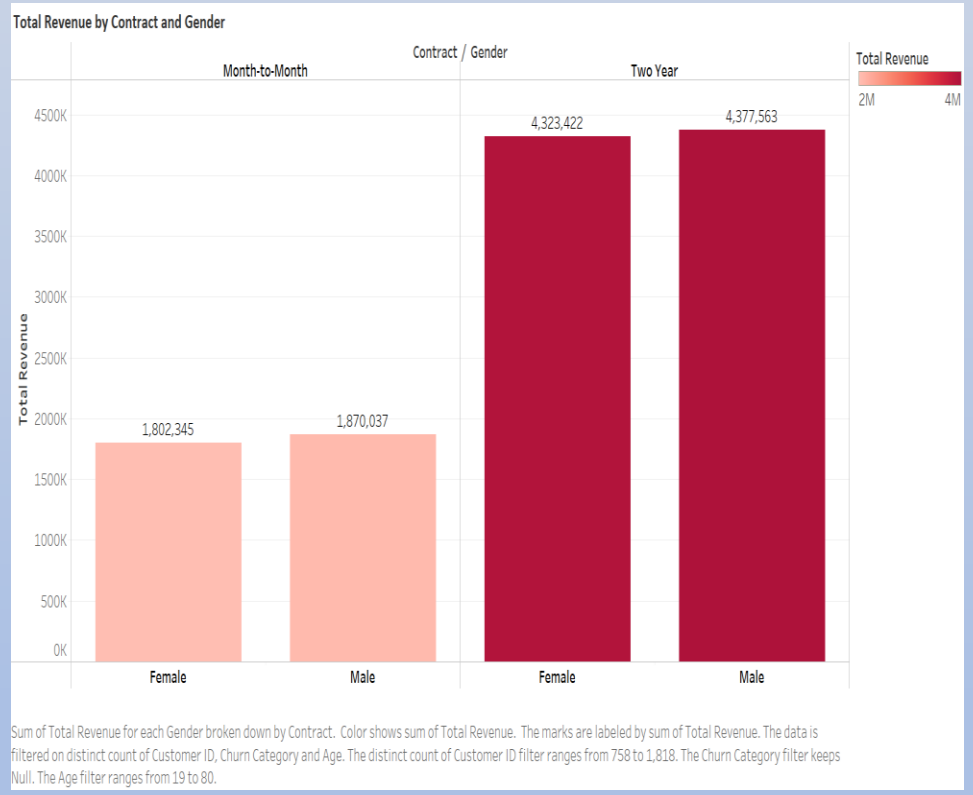
- **Gender Disparity in Total Revenue:** Male customers generate significantly more churn rate compared to female customers, especially for month-to-month contracts.
- **Contract Type Impact:** Month-to-month customers, regardless of gender, generate higher churn rate than those with One Year and Two Year contract.

Churn Category	(Multiple Items)	
Age	(All)	
Row Labels	Count of Customer ID	Sum of Total Revenue
Month-to-Month	1655	2490105.85
Female	842	1202017.3
Male	813	1288088.55
One Year	166	858489.8
Female	75	392619.99
Male	91	465869.81
Two Year	48	335864.17
Female	22	138836.46
Male	26	197027.71
Grand Total	1869	3684459.82



Key Insights for Total Revenue by Contract and Gender (Filter by Null Churn Category)

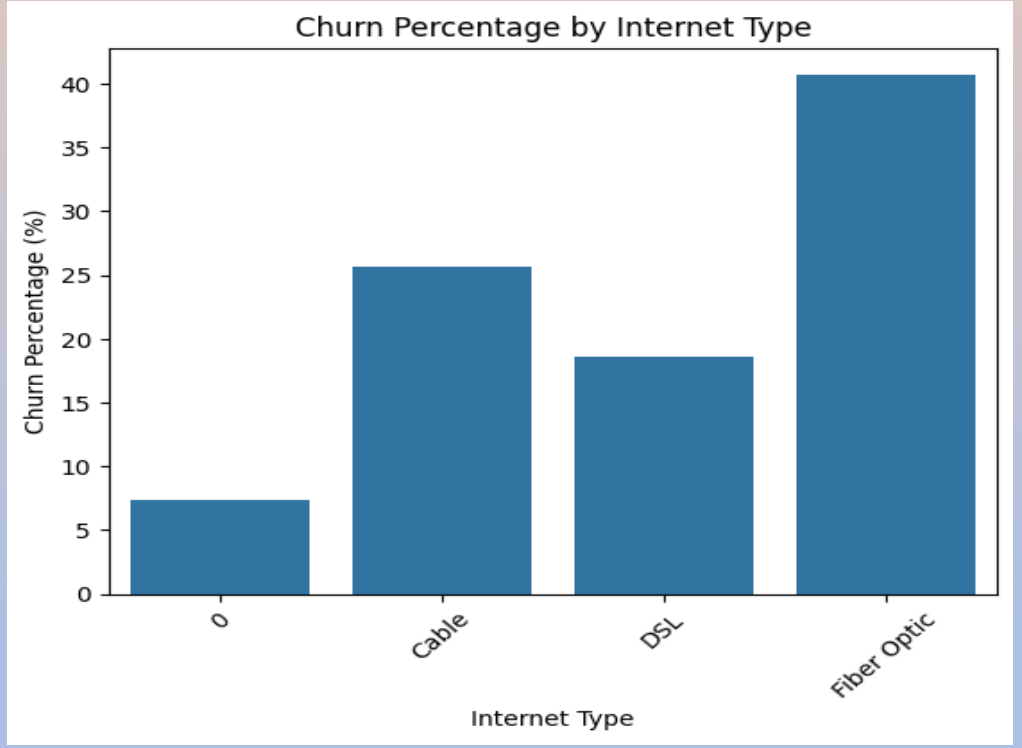
- **Gender and Contract Type Influence Total Revenue:** Both male and female customers generate significant total revenue, especially when they have two-year contracts.
- **High Revenue Earnings:** The data indicates high revenue earnings, suggesting that the customers are generating substantial revenue for the company, especially when we filter Churn Category by null and not considering any churn categories, which shows the companies total revenue is not bad for Two Year contract.



Key Insights for Churn rate by Internet Type and Contract

- **Total Revenue by Internet Type:**

1. **Fiber Optic:** Fiber Optic customers generated the highest total revenue among those who churned, with over $3 million. The majority of this revenue came from customers on month-to-month contracts ($2.1 million), followed by one-year contracts ($648,193.73) and two-year contracts ($254,832.92).
2. **DSL:** DSL customers also contributed significantly, with total churned revenue amounting to approximately $312,450. The highest revenue in this category came from month-to-month contracts ($151,804.32), followed by one-year ($104,085.53) and two-year contracts ($56,560.54).



1. **Cable:** Cable customers, though fewer in number, generated about $307,794.16 in churned revenue, with month-to-month contracts ($196,602.45) being the primary contributor, followed by one-year ($96,941.71) and two-year contracts ($14,250).
2. **Unspecified Type:** A small portion of the churned revenue ($43,107.19) came from customers with an unspecified internet type, primarily on month-to-month contracts.

- **Strategic Recommendation:**

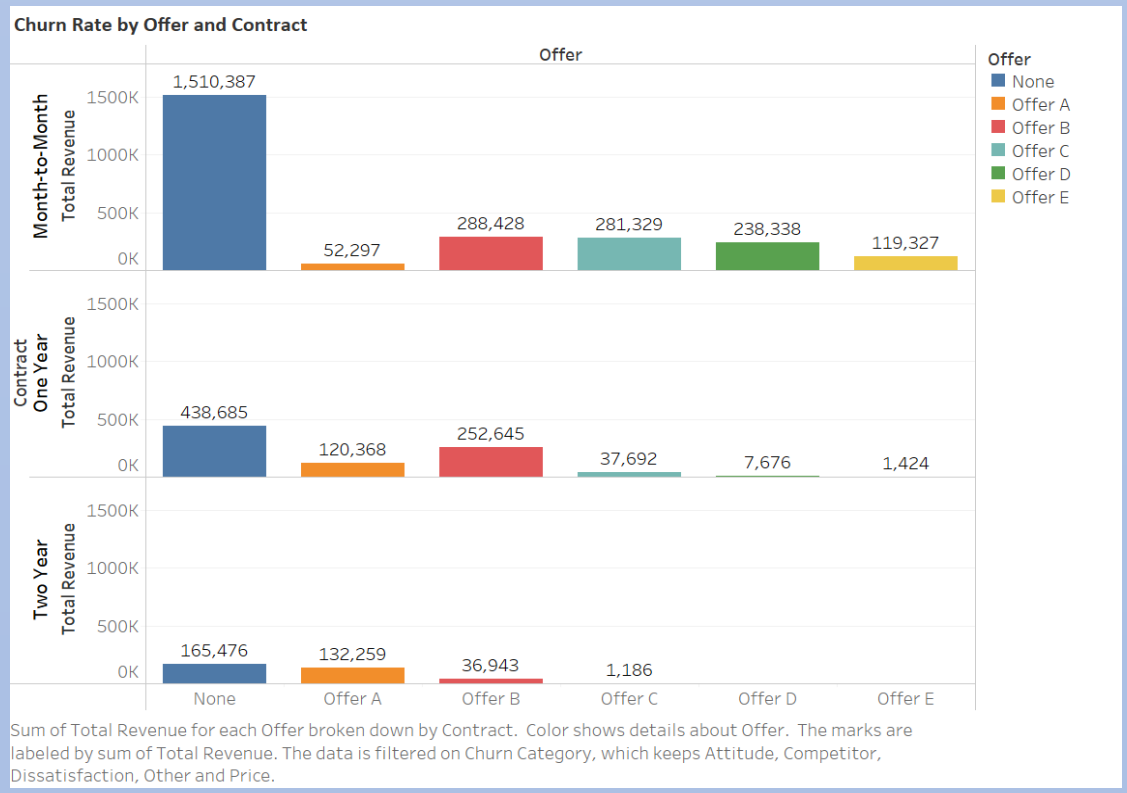
1. To reduce churn, the company should focus on improving the value proposition and customer experience for Fiber Optic customers on month-to-month contracts, as they represent the highest revenue loss in churn.

Customer Status	(All)
Churn Category	(Multiple Items)
Gender	(All)
Row Labels	Sum of Total Revenue
Cable	$307,794.16
Month-to-Month	$196,602.45
One Year	$96,941.71
Two Year	$14,250.00
DSL	$312,450.39
Month-to-Month	$151,804.32
One Year	$104,085.53
Two Year	$56,560.54
Fiber Optic	$3,021,108.08
Month-to-Month	$2,118,081.43
One Year	$648,193.73
Two Year	$254,832.92
(blank)	$43,107.19
Month-to-Month	$23,617.65
One Year	$9,268.83
Two Year	$10,220.71
Grand Total	$3,684,459.82

Key Insights for Churn rate by Offer and Contract

- **Overall Churn Revenue by Offer:**

1. **Month-to-Month Contracts:** The majority of churned revenue ($2,490,105.85) and customers (1,655) came from month-to-month contracts. Within this category, the highest churned revenue was generated by customers without any specific offer ($1,510,386.72 from 936 customers), followed by those who accepted Offer B ($288,427.53 from 53 customers) and Offer C ($281,329.40 from 83 customers).
2. **One-Year Contracts:** The total churned revenue from one-year contracts amounted to $858,489.80, with 166 customers. The most churned revenue in this category was also from customers without any specific offer ($438,684.74 from 90 customers), followed by Offer B ($252,644.87 from 41 customers) and Offer A ($120,368.07 from 14 customers).



3. Two-Year Contracts: The least amount of churned revenue came from two-year contracts, totaling $335,864.17 from 48 customers. Here too, the highest churned revenue was from customers without any specific offer ($165,476.37 from 25 customers), followed by Offer A ($132,258.73 from 15 customers).

- **Offer Performance:**

1. **Offer E:** Despite being offered to 426 customers, Offer E generated the least churned revenue ($120,751.68), indicating it might have been more effective in reducing churn compared to other offers.
2. **Offer B and Offer C:** These offers generated relatively high churned revenue across all contract types, suggesting that they may not be as effective in retaining customers.

Churn Category	(Multiple Items)													
Age	(All)													
Married	(All)													
	Column Labels													
	None		Offer A		Offer B		Offer C		Offer D		Offer E		Total Sum of Total Revenue	Total Count of Customer ID
Row Labels	Sum of Total Revenue	Count of Customer ID	Sum of Total Revenue	Count of Customer ID	Sum of Total Revenue	Count of Customer ID	Sum of Total Revenue	Count of Customer ID	Sum of Total Revenue	Count of Customer ID	Sum of Total Revenue	Count of Customer ID		
Month-to-Month	$1,510,386.72	936	$52,297.25	6	$288,427.53	53	$281,329.40	83	$238,337.57	155	$119,327.38	422	$2,490,105.85	1655
One Year	$438,684.74	90	$120,368.07	14	$252,644.87	41	$37,691.63	11	$7,676.19	6	$1,424.30	4	$858,489.80	166
Two Year	$165,476.37	25	$132,258.73	15	$36,943.22	7	$1,185.85	1					$335,864.17	48
Grand Total	$2,114,547.83	1051	$304,924.05	35	$578,015.62	101	$320,206.88	95	$246,013.76	161	$120,751.68	426	$3,684,459.82	1869

- **Strategic Recommendation:**

1. To reduce churn, the company should re-evaluate the effectiveness of its current offers, particularly focusing on why Offers B and C are leading to higher churn rates and considering enhancements or alternatives to make these offers more appealing and effective in customer retention.

Model Performance Validation

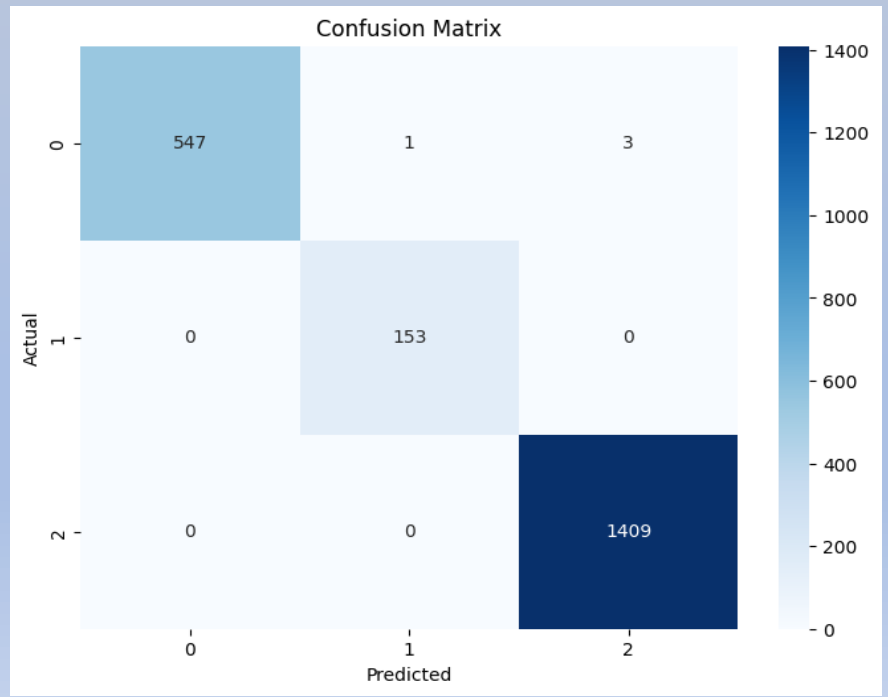
- **Accuracy and Precision:**

1. **Perfect Accuracy:** The model achieved 100% accuracy, meaning it correctly predicted every instance in the test set. This reflects the model's high reliability in classifying customer churn.
2. **Precision, Recall, and F1-Score:** All metrics (precision, recall, and F1-score) are 1.00 for each class (0, 1, and 2). This indicates that the model not only identifies churners accurately but also minimizes false positives and false negatives, ensuring reliable and actionable predictions.

```
Classification Report:
              precision    recall  f1-score   support

           0       1.00      0.99      1.00       551
           1       0.99      1.00      1.00       153
           2       1.00      1.00      1.00      1409

    accuracy                           1.00      2113
   macro avg       1.00      1.00      1.00      2113
weighted avg       1.00      1.00      1.00      2113

Confusion Matrix:
[[ 547    1    3]
 [   0  153    0]
 [   0    0 1409]]
```

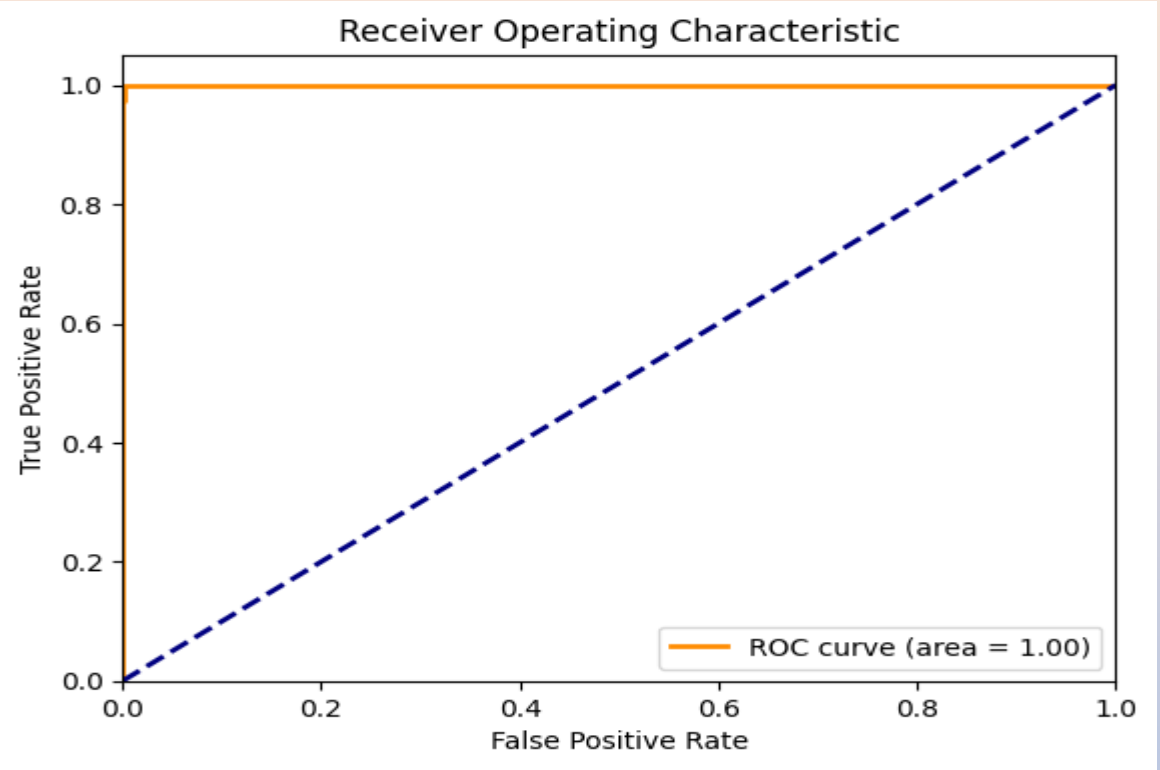
3. Confusion Matrix Insights:

- **Class-Specific Performance:** The confusion matrix shows that:
  1. **Class 0:** The model correctly identified 547 churners with only 4 misclassifications, demonstrating high precision and recall.
  2. **Class 1:** The model perfectly predicted all 153 instances without any errors.
  3. **Class 2:** The model accurately predicted all 1409 non-churners, showcasing its effectiveness in distinguishing this class.
- This detailed performance across classes underscores the model's robust ability to handle multiple churn categories effectively.



- **ROC Curve and AUC Score:**

1. **AUC Score:** The ROC curve achieves a perfect AUC score of 1.00. This means the model has an ideal capability to distinguish between churn and non-churn instances across different thresholds.
2. **Visual Representation:** The ROC curve's optimal performance further validates the model's exceptional discrimination ability and overall predictive power.



- **Implications for the Company:**

1. **Strategic Decisions:** With such high performance, the model can be confidently used to predict customer churn accurately, enabling proactive strategies for retention and targeted interventions.
2. **Operational Efficiency:** The reliability of this model ensures minimal false predictions, leading to more efficient resource allocation and better customer relationship management.

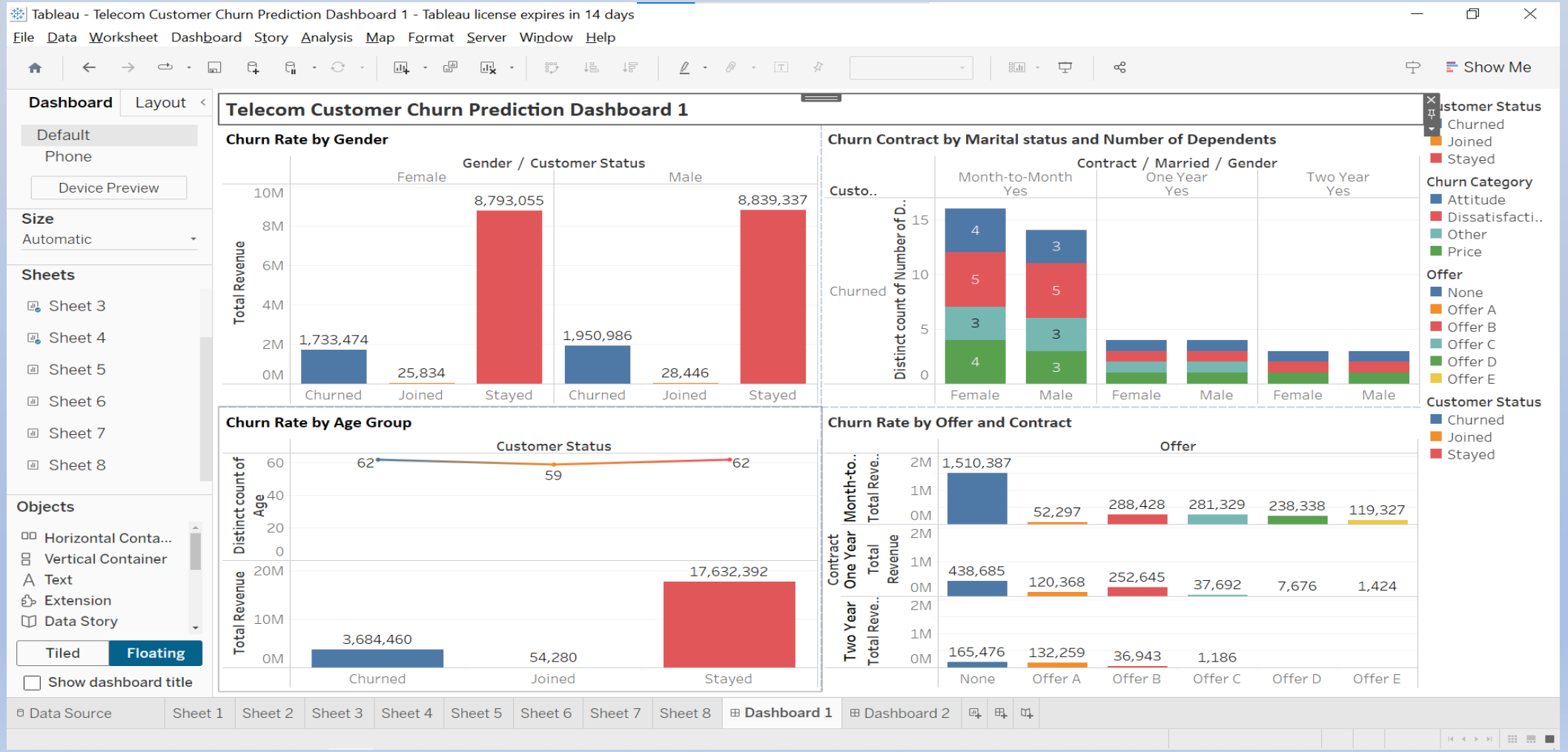
Conclusion

The Telecom Customer Churn Prediction model has achieved remarkable performance with 100% accuracy and flawless precision, recall, and F1-score, underscoring its reliability in identifying customer churn. The ROC curve, with an AUC score of 1.00, highlights the model's exceptional ability to differentiate between churn and non-churn cases. This powerful tool not only provides actionable insights for targeted customer retention strategies but also enhances overall operational efficiency. Looking ahead, the model's robustness and scalability open opportunities for continuous refinement and adaptation to evolving customer behaviors, positioning the company to effectively mitigate churn and drive sustained growth.

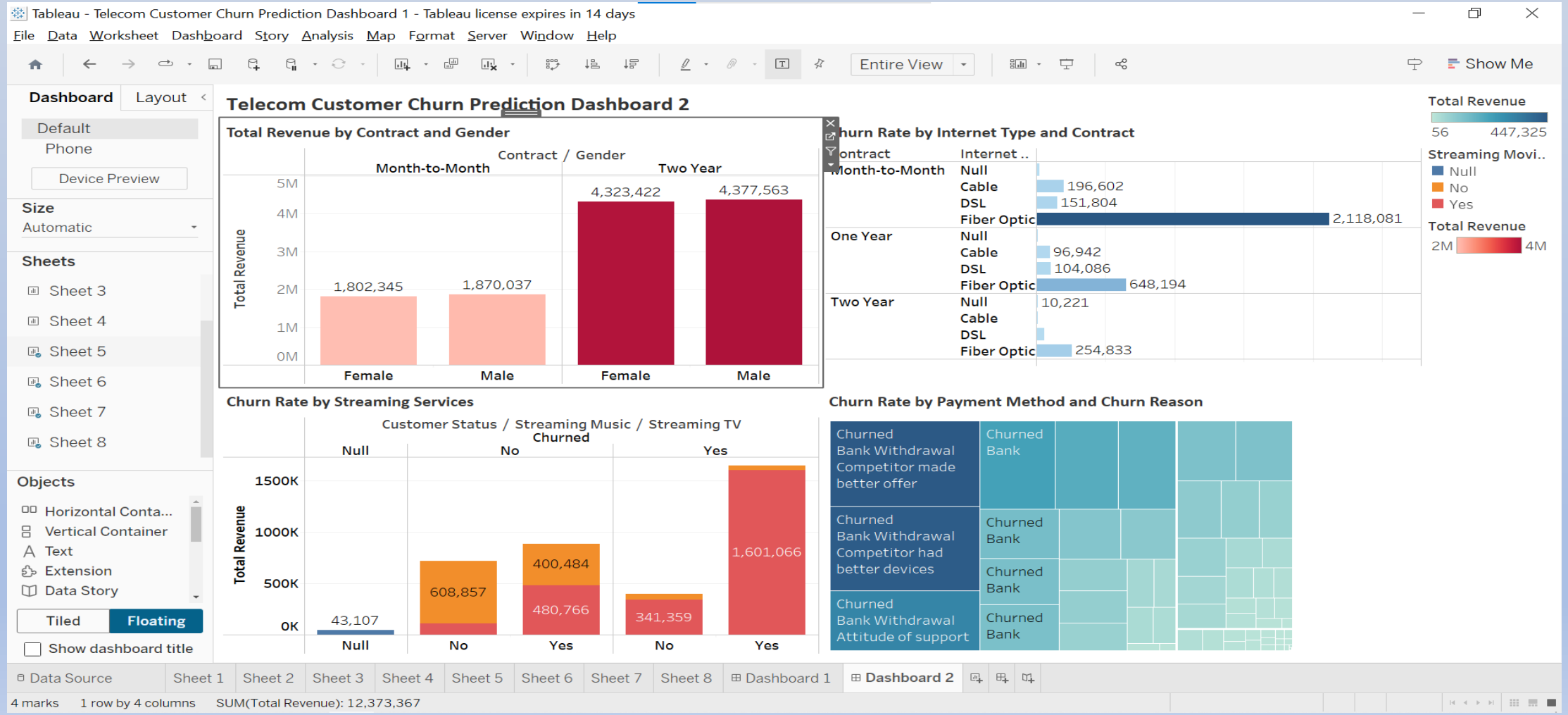
Future Predictions

For future predictions in the Telecom Customer Churn Prediction project, the model can be enhanced by incorporating more real-time customer data, such as usage patterns, customer service interactions, and billing history. By continuously updating the model with fresh data and leveraging advanced techniques like deep learning, the accuracy of churn predictions can be improved. This will enable the company to identify at-risk customers earlier and implement targeted retention strategies, ultimately reducing churn rates and increasing customer loyalty. Additionally, exploring new features, such as customer sentiment analysis, could provide deeper insights into customer behavior and further refine the predictions.

Tableau Dashboards
Telecom Customer Churn Prediction Dashboard 1



Telecom Customer Churn Prediction Dashboard 2



Project Code

I have attached my Telecom Customer Churn Project code link for your reference, the project code is performed in Python language:

https://colab.research.google.com/drive/1A1GMYn86Q1q-JyxZpxCA0qI_goakDbFo?usp=sharing