Telecom Customer Churn Analysis

(Using SQL Server, PowerBI and Machine Learning)

Project Overview

In today's highly competitive telecom landscape, retaining customers is paramount to business success. To address this challenge, I undertook a **Telecom Customer Churn Analysis** project, leveraging **SQL Server, Microsoft Power BI,** and **Machine Learning** to gain valuable insights and predict customer behavior. This end-to-end project covered everything from data extraction and transformation (ETL) to insightful visualization and predictive modeling, ultimately offering actionable strategies for enhancing customer retention.

By:-

Mohd Parvez Khan

Identifying Missing Values using CASE Statement

We uses **CASE statements** combined with the **SUM** function to count **NULL** values across various columns in this dataset, such as **Customer_ID, Gender, Age, Married, State**, and others related to customer service usage and subscriptions. **Each CASE** statement checks whether the column value is **NULL**. If **NULL**, it counts it as **1**; otherwise, it counts **as 0**. The total number of NULL values for each column is summed and aliased with a corresponding name, such as **Customer_ID_Null_Count, Gender_Null_Count, etc.**

The output shows how many **NULL values**(missing values) are present for each column, allowing you to identify missing data and assess data quality. **This process is crucial for data cleaning and ensuring the dataset is ready for further analysis.**

```
SUM(CASE WHEN Age IS NULL THEN 1 ELSE 0 END) AS Age_Null_Count,
SUM(CASE WHEN Married IS NULL THEN 1 ELSE 0 END) AS Married_Null_Count,
SUM(CASE WHEN State IS NULL THEN 1 ELSE 0 END) AS State_Null_Count,
SUM(CASE WHEN Number_of_Referrals IS NULL THEN 1 ELSE 0 END) AS Number_of_Referrals_Null_Count,
SUM(CASE WHEN Tenure_in_Months IS NULL THEN 1 ELSE 0 END) AS Tenure_in_Months_Null_Count,
SUM(CASE WHEN Value_Deal IS NULL THEN 1 ELSE 0 END) AS Value_Deal_Null_Count,
SUM(CASE WHEN Phone_Service IS NULL THEN 1 ELSE 0 END) AS Phone_Service_Null_Count,
SUM(CASE WHEN Multiple_Lines IS NULL THEN 1 ELSE 0 END) AS Multiple_Lines_Null_Count,
SUM(CASE WHEN Internet_Service IS NULL THEN 1 ELSE 0 END) AS Internet_Service_Null_Count,
SUM(CASE WHEN Internet_Type IS NULL THEN 1 ELSE 0 END) AS Internet_Type_Null_Count,
SUM(CASE WHEN Online_Secur column Internet_Type(nvarchar, null) END) AS Online_Security_Null_Count,
SUM(CASE WHEN Online_Backup IS NULL THEN 1 ELSE 0 END) AS Online_Backup_Null_Count,
SUM(CASE WHEN Device_Protection_Plan IS NULL THEN 1 ELSE 0 END) AS Device_Protection_Plan_Null_Count,
SUM(CASE WHEN Premium_Support IS NULL THEN 1 ELSE 0 END) AS Premium_Support_Null_Count,
SUM(CASE WHEN Streaming_TV IS NULL THEN 1 ELSE 0 END) AS Streaming_TV_Null_Count,
SUM(CASE WHEN Streaming_Movies IS NULL THEN 1 ELSE 0 END) AS Streaming_Movies_Null_Count,
SUM(CASE WHEN Streaming_Music IS NULL THEN 1 ELSE 0 END) AS Streaming_Music_Null_Count,
SUM(CASE WHEN Unlimited_Data IS NULL THEN 1 ELSE 0 END) AS Unlimited_Data_Null_Count,
SUM(CASE WHEN Contract IS NULL THEN 1 ELSE 0 END) AS Contract_Null_Count,
SUM(CASE WHEN Paperless_Billing IS NULL THEN 1 ELSE 0 END) AS Paperless_Billing_Null_Count,
SUM(CASE WHEN Payment_Method IS NULL THEN 1 ELSE 0 END) AS Payment_Method_Null_Count,
SUM(CASE WHEN Monthly_Charge IS NULL THEN 1 ELSE 0 END) AS Monthly_Charge_Null_Count,
SUM(CASE WHEN Total_Charges IS NULL THEN 1 ELSE 0 END) AS Total_Charges_Null_Count,
SUM(CASE WHEN Total_Refunds IS NULL THEN 1 ELSE 0 END) AS Total_Refunds_Null_Count,
SUM(CASE WHEN Total_Extra_Data_Charges IS NULL THEN 1 ELSE 0 END) AS Total_Extra_Data_Charges_Null_Count,
SUM(CASE WHEN Total_Long_Distance_Charges IS NULL THEN 1 ELSE 0 END) AS Total_Long_Distance_Charges_Null_Count,
SUM(CASE WHEN Total_Revenue IS NULL THEN 1 ELSE 0 END) AS Total_Revenue_Null_Count,
SUM(CASE WHEN Customer_Status IS NULL THEN 1 ELSE 0 END) AS Customer_Status_Null_Count,
SUM(CASE WHEN Churn_Category IS NULL THEN 1 ELSE 0 END) AS Churn_Category_Null_Count,
SUM(CASE WHEN Churn_Reason IS NULL THEN 1 ELSE 0 END) AS Churn_Reason_Null_Count
FROM Customer_Data;
```

Handling Missing Values using ISNULL Function

We uses the **ISNULL** function to handle missing data. The query selects various customer attributes such as **Customer_ID, Gender, Age, Married,** and more, while replacing **NULL** values with specified default values. For instance, in fields like Value_Deal, Internet_Type, and Streaming_TV, if a **NULL** value is encountered, it is replaced with **'None' or 'No',** indicating the absence of a particular service or attribute.

This approach improves the reliability of summary statistics and other data processing tasks, as **NULL values can skew results if not handled properly**.

```sql
SELECT
    Customer_ID,Gender,Age,Married,State,Number_of_Referrals,Tenure_in_Months,
    ISNULL(Value_Deal, 'None') AS Value_Deal,
    Phone_Service,
    ISNULL(Multiple_Lines, 'No') As Multiple_Lines,
    Internet_Service,
    ISNULL(Internet_Type, 'None') AS Internet_Type,
    ISNULL(Online_Security, 'No') AS Online_Security,
    ISNULL(Online_Backup, 'No') AS Online_Backup,
    ISNULL(Device_Protection_Plan, 'No') AS Device_Protection_Plan,
    ISNULL(Premium_Support, 'No') AS Premium_Support,
    ISNULL(Streaming_TV, 'No') AS Streaming_TV,
    ISNULL(Streaming_Movies, 'No') AS Streaming_Movies,
    ISNULL(Streaming_Music, 'No') AS Streaming_Music,
    ISNULL(Unlimited_Data, 'No') AS Unlimited_Data,
    Contract,
    Paperless_Billing,
    Payment_Method,
    Monthly_Charge,
    Total_Charges,
    Total_Refunds,
    Total_Extra_Data_Charges,
    Total_Long_Distance_Charges,
    Total_Revenue,
    Customer_Status,
    ISNULL(Churn_Category, 'Others') AS Churn_Category,
    ISNULL(Churn_Reason , 'Others') AS Churn_Reason

INTO [Db_customer_churn].[dbo].[prod_Churn]
FROM [Db_customer_churn].[dbo].[Customer_Data];
```

Creating Views for easier Data analysis

We have created Views using the **CREATE VIEW** command, which is used to create **virtual tables (views)** based on the prod_Churn dataset. In this case, two views are being created: **vw_ChurnData** filters data for customers whose status is either **'Churned' or 'Stayed'**, and **vw_JoinData** focuses on customers with the status **'Joined'.** The purpose of creating these views is to simplify complex queries and organize the data into relevant customer segments for easier analysis, improving data management and query efficiency.

```sql
Create View vw_ChurnData as
	select * from prod_Churn where Customer_Status In ('Churned', 'Stayed')


Create View vw_JoinData as

	select * from prod_Churn where Customer_Status = 'Joined'
```

) %

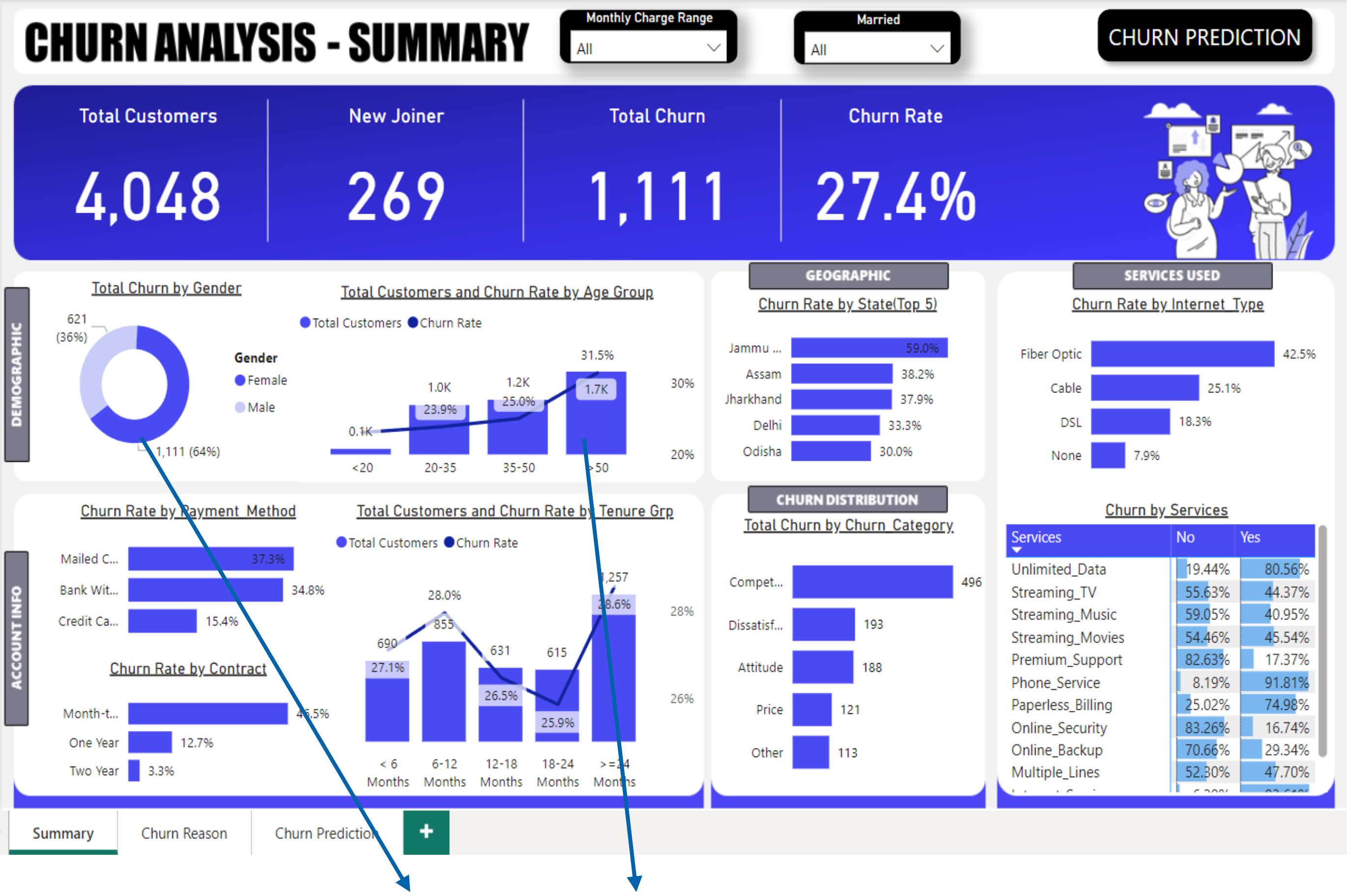
Messages

```
Commands completed successfully.

Completion time: 2024-09-26T02:17:22.9563268+05:30
```

PowerBI Transformations

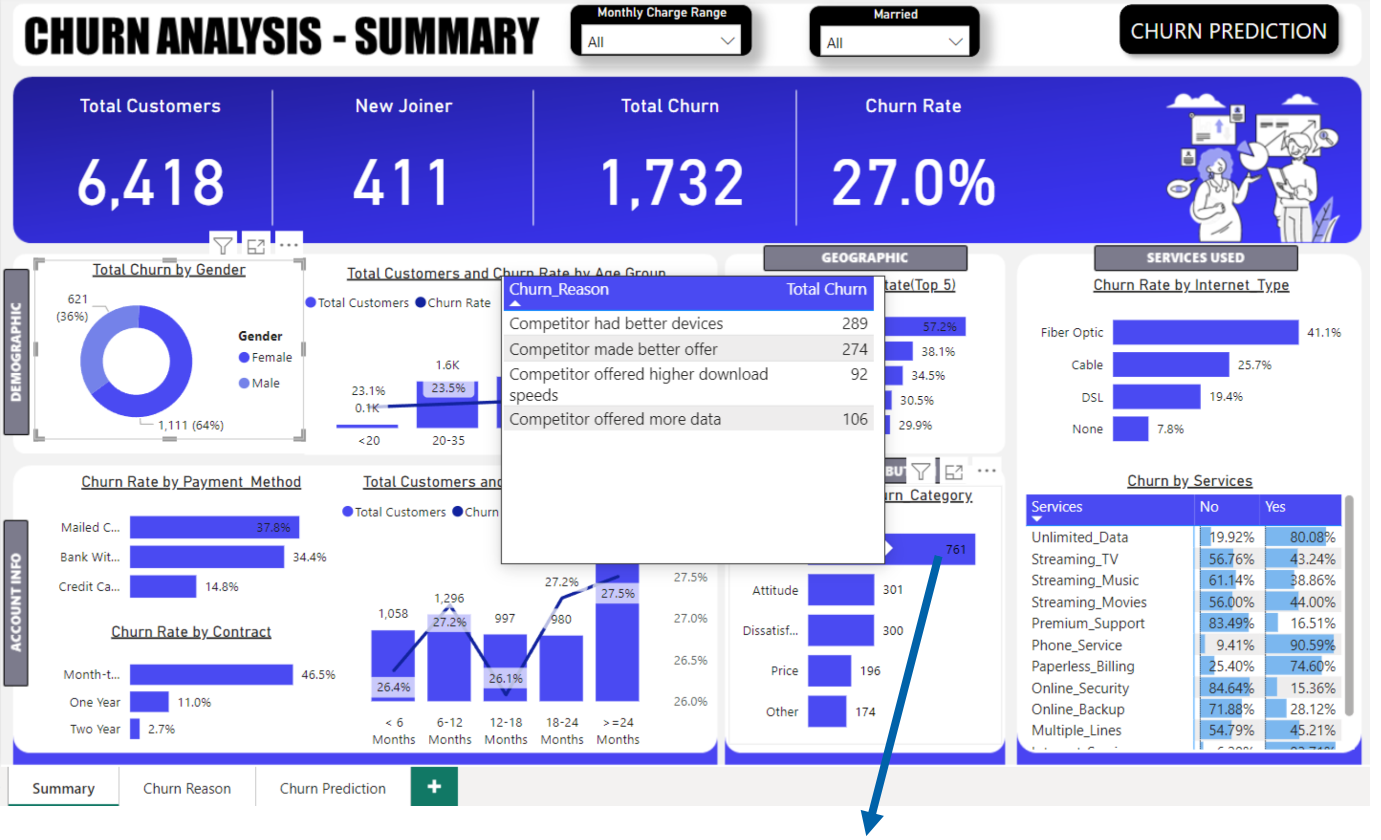
Next, the focus shifted to Power BI, where the cleaned data from SQL Server was imported **(using import method)** and transformed further for better visualization and analysis.



INSIGHTS FROM THIS VISUAL

- As you can see from above visual **(only demographic portion).** There are **64%** of entire Churner list are **female** population.
- And out of these 64% of females **31% were in the greater than 50 years age.**
- So If I have to specifically design a **marketing campaigns** which focuses of females greater than 50 years only then I would be able to save this much customer and **stop them from Churning.**
- The top **5 States** with the most churners does not change with Gender filter. It stills remains the same so for this, states does not matter.

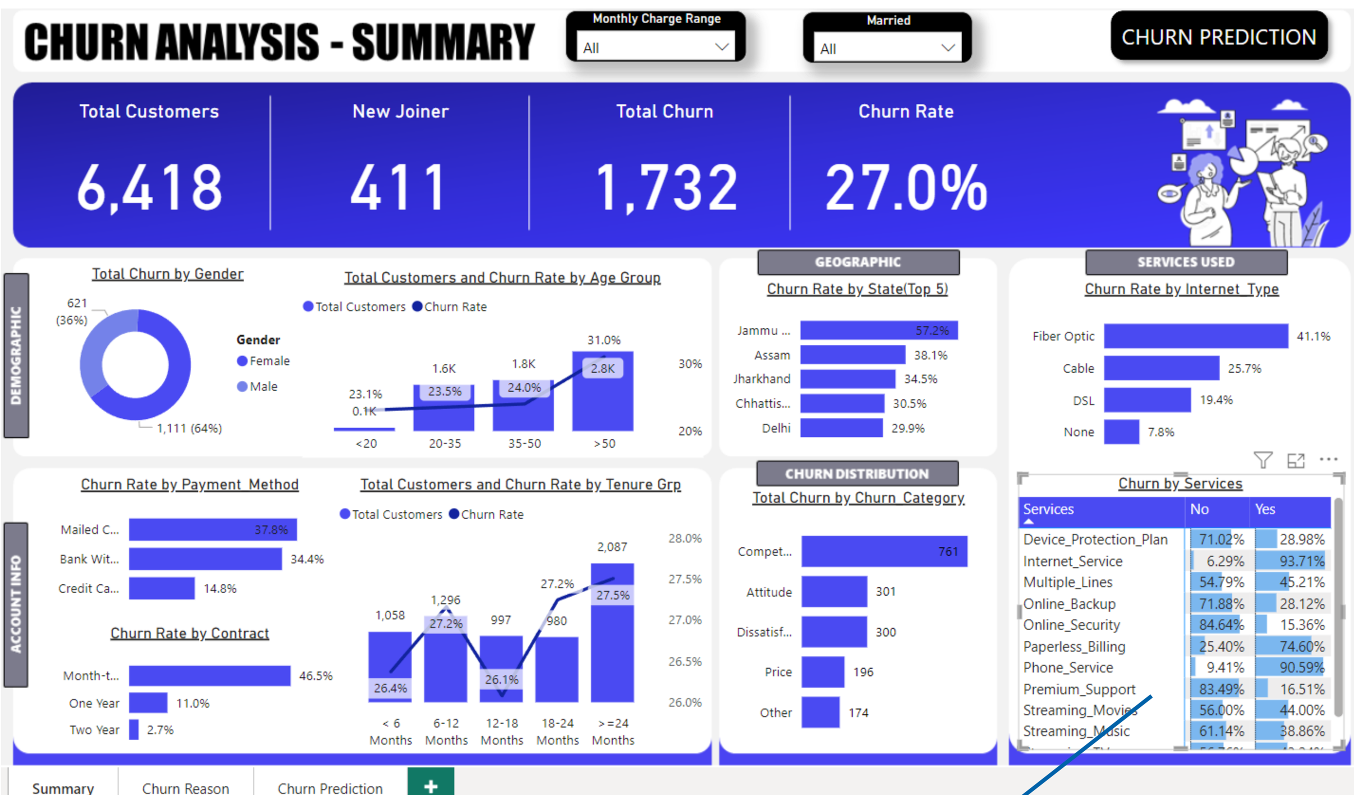
Reason of Churning



INSIGHTS FROM THIS VISUAL

- There are **761** Customers which based on Gender are getting churned because of Competitor.
- In that 761 customers, **289** customers are churned because competitor had **better device** and **274** are because competitor **made better offer.**
- So, Now I know that these are the two Categories we need to work upon to stop those customers from churning.

Reason of Churning



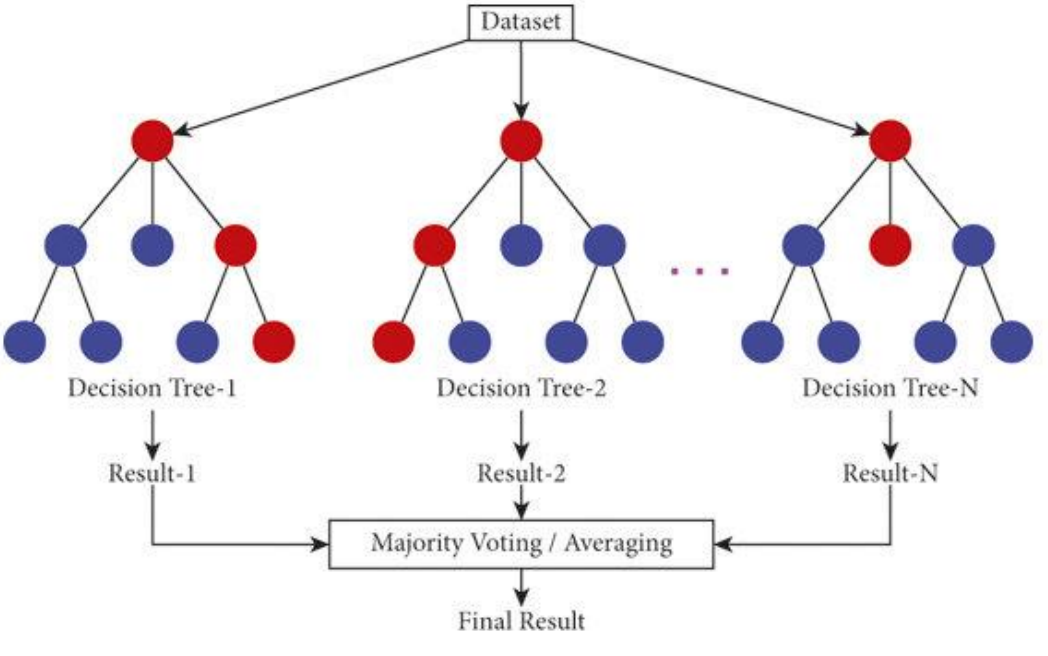
INSIGHTS FROM THIS VISUAL

- In this, I am considering anything services **greater than 60%** (this benchmark can vary organization to organisation) to be **bad for me**. So if any of the categories have more than 60%, I should look at those services.
- **Device Protection Plan, Online Backup, online Security, Premium Support, Streaming Music**, people who have **not subscribed** to these services are more likely to churn than people who have subscribed this.
- Similarly, **Internet Services, Paperless Billing, Phone Service, Unlimited Data** are more likely to churn which means their **services is not upto the mark**.
- So, when we are designing a market campaign we need to look at these services and come out with ways with which we can improve these services.

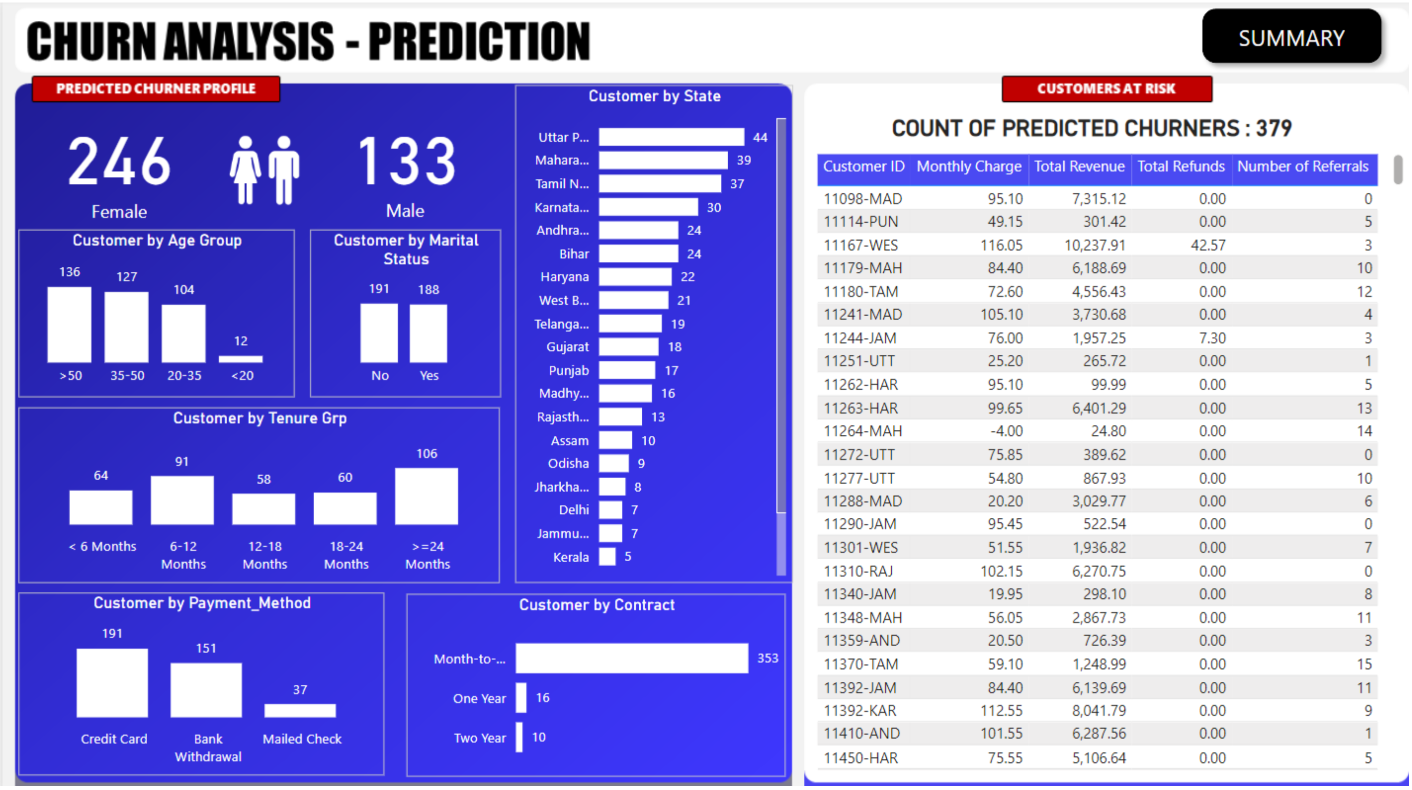
Machine Learning for Predictive Analysis

Predicting Future Churner:

To predict future churner, a **Random Forest model** was developed using Machine Learning techniques. The model was trained on historical customer data and subsequently validated to ensure its reliability and accuracy in predicting future churners. Once the predictive model was finalized, **the forecasted data was imported back into Power BI**, seamlessly integrating with the existing customer data for strategic insights. The final Power BI dashboard provided an executive summary, combining historical analysis with a dedicated prediction page to forecast future churn trends.



Reason for Future Churners



INSIGHTS FROM THIS VISUAL

- *From above Visual (Customer by Age Group),* **Customers aged 35-50** and **above 50** are more likely to churn, possibly because their needs aren't being fully met or because they are finding **better offers from competitors.**
- *From (Customer by Tenure Group),* there **are 106 customers with tenures greater than 24** months are predicted to churn, which suggests they are unhappy because they aren't getting enough rewards or recognition for their loyalty.
- *From (Customer by State),* there are top 3 states with high churn which is **Uttar Pradesh, Maharashtra, and Tamil Nadu**, which means there are issues with our service quality or more competition in these areas.
- *From (Customer by Contract),* there are **353 customers are on month-to-month** contracts, which provides customers with flexibility to leave, indicating they may not be fully committed to the service, or that they are not seeing enough value to commit long-term.

Thank You