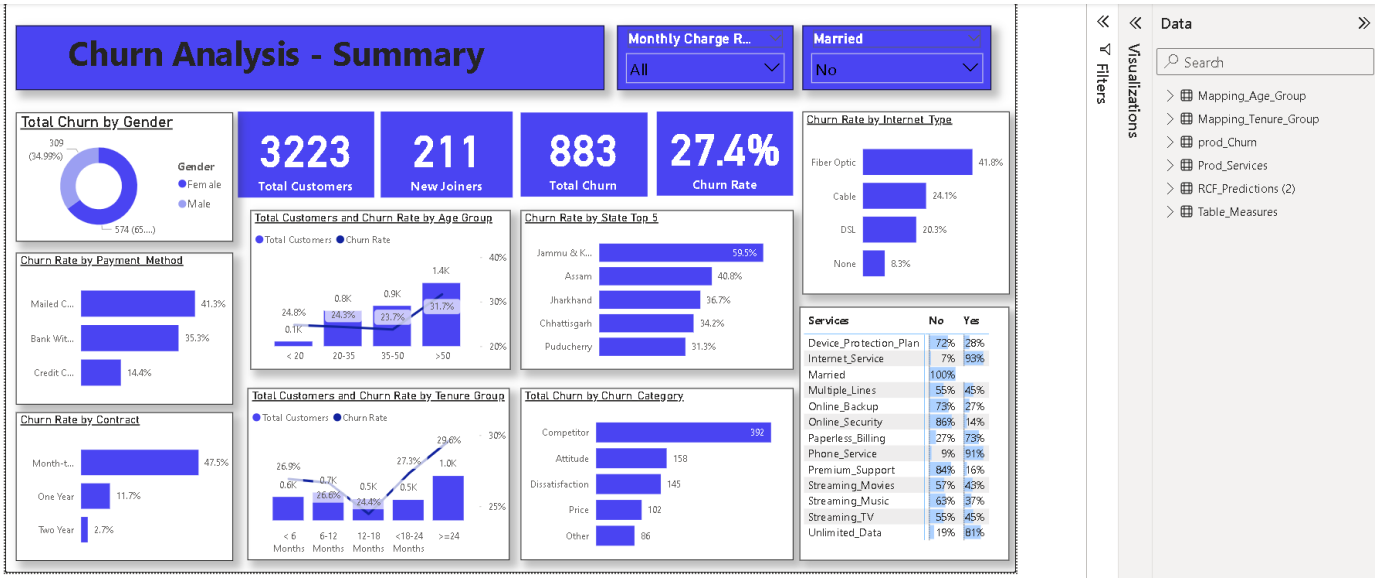
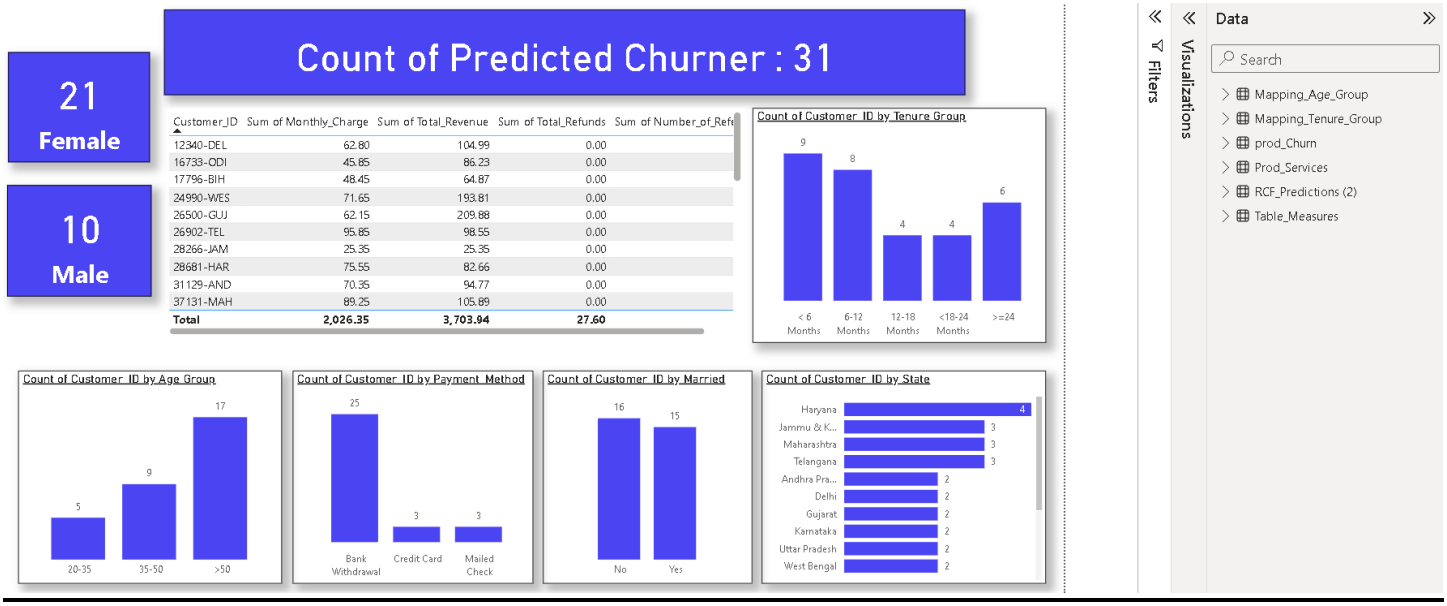
Project Title

Customer Churn Analysis and Prediction

[Churn Analysis] Dashboard Creation before Prediction :



[Churn Analysis] Dashboard Creation after Prediction :



Insights:

The dashboard shows that out of 31 predicted churners, most are female (21) and over 50 years old (17). Customers with tenure of less than 12 months have the highest churn rates, especially those with bank withdrawals as their payment method. Churners are fairly balanced across marital status and are distributed across several states, with Haryana having the most. This insight can guide targeted retention efforts.

Introduction: In this project, I predicted customer churn for a telecom company using data sourced from Kaggle. SQL was employed for ETL processes, cleaning and preprocessing the data to ensure it was analysis-ready. Power BI was utilized to create an interactive dashboard, visualizing key metrics and providing valuable insights into customer behaviour. The Random Forest Classifier was then applied to predict churners, offering the company actionable insights to improve customer retention strategies.

Data Exploration – Using SQL

SQLQuery1.sql - D...0TOV2K0\DELL (80))*

```sql
SELECT Gender, Count(Gender) as TotalCount,
Count(Gender) * 1.0 / (Select Count(*) from stg_Churn)  as Percentage
from stg_Churn
Group by Gender

SELECT Contract, Count(Contract) as TotalCount,
Count(Contract) * 1.0 / (Select Count(*) from stg_Churn)  as Percentage
from stg_Churn
Group by Contract

SELECT Customer_Status, Count(Customer_Status) as TotalCount, Sum(Total_Revenue) as TotalRev,
Sum(Total_Revenue) / (Select sum(Total_Revenue) from stg_Churn) * 100  as RevPercentage
from stg_Churn
Group by Customer_Status

SELECT State, Count(State) as TotalCount,
Count(State) * 1.0 / (Select Count(*) from stg_Churn)  as Percentage
from stg_Churn
Group by State
Order by Percentage desc
```

Results | Messages

	Gender	TotalCount	Percentage
1	Male	2370	0.369273917108
2	Female	4048	0.630726082891

	Contract	TotalCount	Percentage
1	Month-to-Month	3286	0.511997507011
2	One Year	1413	0.220162044250
3	Two Year	1719	0.267840448737

	Customer_Status	TotalCount	TotalRev	RevPercentage
1	Joined	411	49281.5598697662	0.253097281975677
2	Churned	1732	3411960.5796299	17.5229426827105
3	Stayed	4275	16010148.2622757	82.2239600353138

Results Messages

	State	TotalCount	Percentage
1	Uttar Pradesh	629	0.098005609224
2	Tamil Nadu	600	0.093487067622
3	Maharashtra	504	0.078529136802
4	Karnataka	470	0.073231536304
5	Haryana	398	0.062013088189
6	Andhra Pradesh	395	0.061545652851
7	West Bengal	368	0.057338734808
8	Punjab	342	0.053287628544
9	Bihar	336	0.052352757868
10	Gujarat	335	0.052196946089
11	Jammu & Kashmir	320	0.049859769398
12	Madhya Pradesh	288	0.044873792458
13	Telangana	281	0.043783110003
14	Rajasthan	259	0.040355250856
15	Kerala	200	0.031162355874
16	Odisha	152	0.023683390464
17	Assam	139	0.021657837332
18	Delhi	127	0.019788095980
19	Jharkhand	113	0.017606731068
20	Uttarakhand	62	0.009660330320
21	Chhattisgarh	59	0.009192894982
22	Puducherry	41	0.006388282954

Data Exploration -,Remove null and insert the new data into Prod table

```
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_Churn].[dbo].[prod_Churn]
FROM [db_Churn].[dbo].[stg_Churn];
```

Create View for Power BI

```
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'
```

Power Query Transformations

Add a new column in prod_Churn

1. Churn Status = if [Customer_Status] = "Churned" then 1 else 0
2. Change Churn Status data type to numbers
3. Monthly Charge Range = if [Monthly_Charge] < 20 then "< 20" else if [Monthly_Charge] < 50 then "20-50" else if [Monthly_Charge] < 100 then "50-100" else "> 100"

Create a New Table Reference for mapping_AgeGrp

1. Keep only Age column and remove duplicates
2. Age Group = if [Age] < 20 then "< 20" else if [Age] < 36 then "20 - 35" else if [Age] < 51 then "36 - 50" else "> 50"

3. AgeGrpSorting = if [Age Group] = "< 20" then 1 else if [Age Group] = "20 - 35" then 2 else if [Age Group] = "36 - 50" then 3 else 4
4. Change data type of AgeGrpSorting

Create a new table reference for mapping_TenureGrp

1. Keep only Tenure_in_Months and remove duplicates
2. Tenure Group = if [Tenure_in_Months] < 6 then "< 6 Months" else if [Tenure_in_Months] < 12 then "6-12 Months" else if [Tenure_in_Months] < 18 then "12-18 Months" else if [Tenure_in_Months] < 24 then "18-24 Months" else ">= 24 Months"
3. TenureGrpSorting = if [Tenure_in_Months] = "< 6 Months" then 1 else if [Tenure_in_Months] = "6-12 Months" then 2 else if [Tenure_in_Months] = "12-18 Months" then 3 else if [Tenure_in_Months] = "18-24 Months " then 4 else 5
4. Change data type of TenureGrpSorting

Create a new table reference for prod_Services

1. Unpivot services columns
2. Rename Column –
    a. Attribute >> Services
    b. Value >> Status

<u>Summary Page - Measures</u>

Total Customers = Count (prod_Churn[Customer_ID])

New Joiners = CALCULATE(COUNT(prod_Churn[Customer_ID]), prod_Churn[Customer_Status] = "Joined")

Total Churn = SUM(prod_Churn[Churn Status])

Churn Rate = [Total Churn] / [Total Customers]

<u>Churn Prediction Page - Measures</u>

Count Predicted Churner = COUNT(Predictions[Customer_ID]) + 0

Title Predicted Churners = "COUNT OF PREDICTED CHURNERS : " & COUNT(Predictions[Customer]
