

BUSINESS PROCESS ANALYSIS FOR CUSTOMER RETENTION

ETL Process on SQL Server

- a. Database Creation
- b. Table Creation
- c. View Creation

Power BI - Churn Summary

- a. Data Transformation
- b. Custom Measures
- c. Basic Visualization
- d. Advanced Visualization

Data Analysis Approach

ML Model - Random Forest

- a. Data Preparation
- b. Processing
- c. Model Building & Evaluation
- d. Prediction on Joiner Data

Power BI - Churn Prediction

- a. Churner Profile
- b. Customer at Risk



Tech Stack: SQL Server , Power BI , Python (Random Forest) , ETL , KPI Analytics

The Problem:

A telecom company is facing a high, undefined customer churn rate, leading to significant revenue loss and unstable growth.

Why?

"The business had customer data, but no process to analyze it. This created a critical information gap."

- No Root Cause:** They didn't know the primary drivers of churn. (Was it price? Poor service? Contract terms?)
- No Clear Profile:** They couldn't identify who was leaving. (Was it new customers? Old customers? Customers in a specific state?)
- Reactive vs. Proactive:** All retention efforts were *reactive* (e.g., trying to win back someone who already left). There was no *proactive* system to stop it from happening.



The Problem

To analyze the failing 'Customer Retention' process from end-to-end.

ETL IN SQL

This is where I created the database then I loaded the csv file . Then I would be checking values in few of the columns and understand the data distribution.

Data Exploration – Check Distinct Values

Data Exploration – Check Nulls

```
SQLQuery1.requery(true);  
1 Select Gender, Count(Gender) as TotalCount,  
2 Count(Gender) * 100.0 / (Select Count(*) from stg_Churn ) as Percentage  
3 From stg_Churn  
4 Group By Gender  
5  
6 SELECT Contract, Count(Contract) as TotalCount,  
7 Count(Contract) * 1.0 / (Select Count(*) from stg_Churn ) as Percentage  
8 From stg_Churn  
9 Group by Contract  
10  
11 SELECT Customer_Status, Count(Customer_Status) as TotalCount, Sum(Total_Revenue) as TotalRev,  
12 Sum(Total_Revenue) / (Select sum(Total_Revenue) from stg_Churn ) * 100 as RevPercentage  
13 From stg_Churn  
14 Group By Customer_Status  
15
```

```
1    SELECT
2        SUM(CASE WHEN Customer_ID IS NULL THEN 1 ELSE 0 END) AS Customer_ID_Null_Count,
3        SUM(CASE WHEN Gender IS NULL THEN 1 ELSE 0 END) AS Gender_Null_Count,
4        SUM(CASE WHEN Age IS NULL THEN 1 ELSE 0 END) AS Age_Null_Count,
5        SUM(CASE WHEN Married IS NULL THEN 1 ELSE 0 END) AS Married_Null_Count,
6        SUM(CASE WHEN State IS NULL THEN 1 ELSE 0 END) AS State_Null_Count,
7        SUM(CASE WHEN Number_of_Referrals IS NULL THEN 1 ELSE 0 END) AS Number_of_Referrals_Null_Count,
8        SUM(CASE WHEN Tenure_in_Months IS NULL THEN 1 ELSE 0 END) AS Tenure_in_Months_Null_Count,
9        SUM(CASE WHEN Value_Deal IS NULL THEN 1 ELSE 0 END) AS Value_Deal_Null_Count,
10    FROM [AdventureworksLT].[dbo].[Customer]
11
```

```

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, 'N/A') AS Premium_Support

```

The screenshot shows a SQL Server Management Studio window with a query editor containing the provided T-SQL code. Below the editor is a results grid displaying 30 rows of customer data. The columns include Customer_ID, Gender, Age, Married, State, Number_of_Referrals, Tenure_in_Months, Value_Deal, Phone_Service, Multiple_Lines, Internet_Service, Internet_Type, Online_Security, Online_Backup, Device_Protection_Plan, and Premium_Support.

Remove null and insert the new data into Prod table

Checking to see we have all the data or not? – by select statement

```

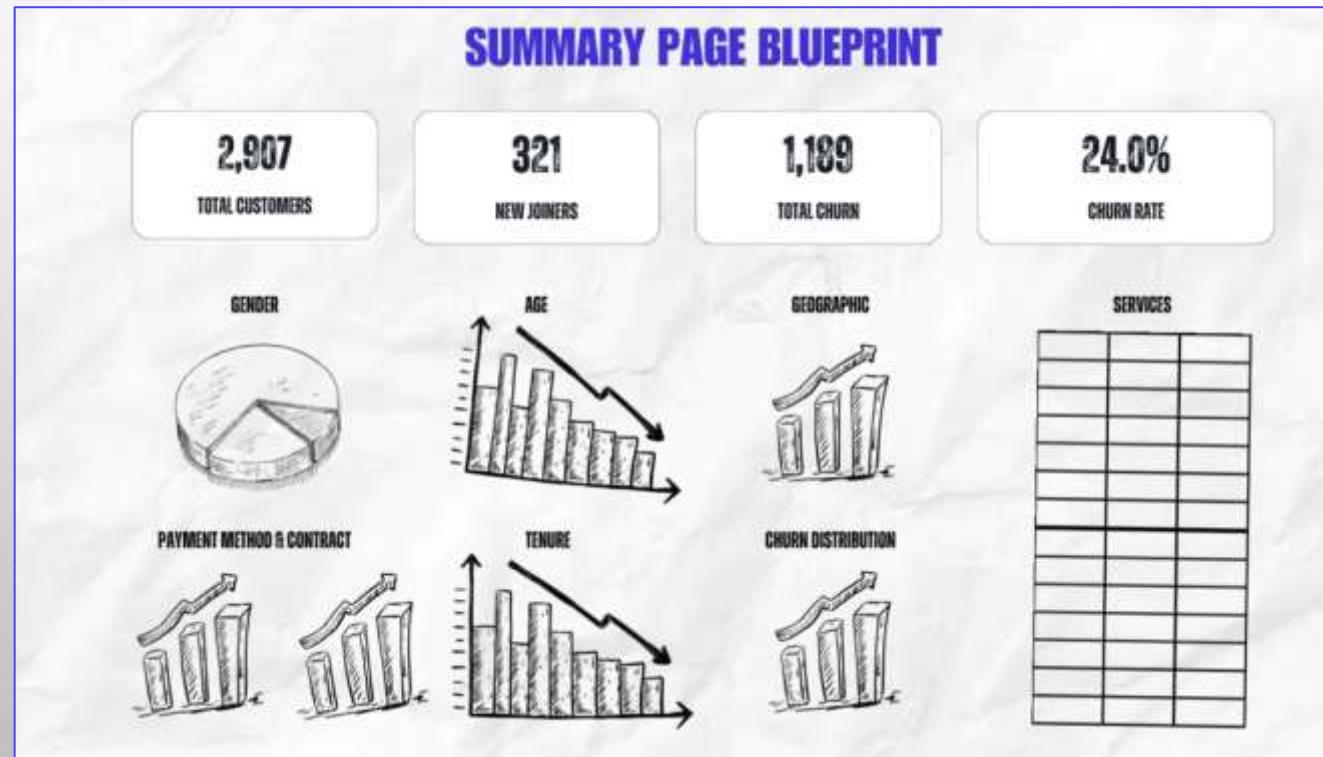
SELECT * FROM stg_Churn

```

The screenshot shows a SQL Server Management Studio window with a query editor containing the provided T-SQL code. Below the editor is a results grid displaying 44 rows of data from the stg_Churn table. The columns listed in the results grid are identical to those in the previous screenshot, including Customer_ID, Gender, Age, Married, State, Number_of_Referrals, Tenure_in_Months, Value_Deal, Phone_Service, Multiple_Lines, Internet_Service, Internet_Type, Online_Security, Online_Backup, Device_Protection_Plan, and Premium_Support.

POWER BI

- To import the data of what we have done in the SQL Server and present it in Power BI perspective



Now, for the visualization, I have created a blueprint / agenda to create the first page. So, we have to cover a couple of metrics:

- I. Overall level of Numbers.
- II. Demographic
 - a. Gender related visuals
 - b. Account related visuals
 - c. Geographic
 - d. Churn Distribution
- III. Services used by the Customers

THE DASHBOARD

CHURN ANALYSIS - SUMMARY

6,418
Total Customers

411
New Joiners

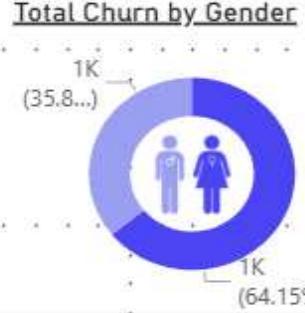
1,732
Total Churn

27.0%
Churn Rate

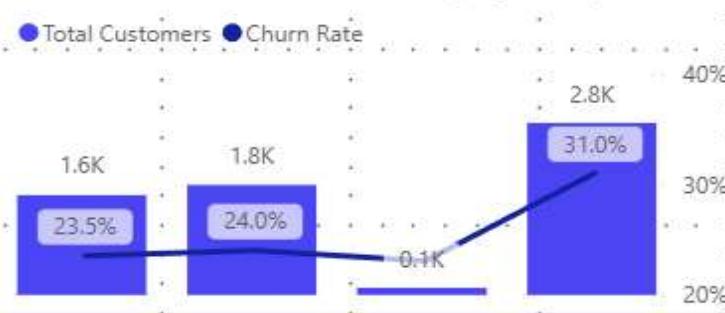
Churn Prediction



DEMOGRAPHIC



Total Customers and Churn Rate by Age Group



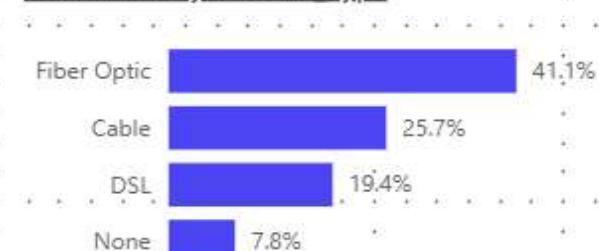
GEOGRAPHIC

Churn Rate by State



SERVICES USED

Churn Rate by Internet Type



ACCOUNT INFO

Churn Rate by Contract

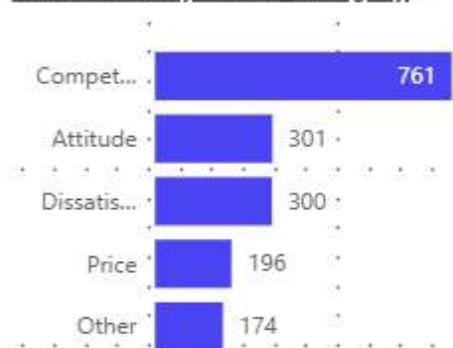


Total Customers and Churn Rate by Tenure Group



CHURN DISTRIBUTION

Total Churn by Churn Category



Services

No

Yes

Services	No	Yes
Device_Protection_Plan	71.02%	28.98%
Internet_Service	6.29%	93.71%
Multiple_Lines	54.79%	45.21%
Online_Backup	71.88%	28.12%
Online_Security	84.64%	15.36%
Paperless_Billing	25.40%	74.60%
Phone_Service	9.41%	90.59%
Premium_Support	83.49%	16.51%
Streaming_Movies	56.00%	44.00%
Streaming_Music	61.14%	38.86%
Streaming_TV	56.76%	43.24%
Unlimited_Data	19.92%	80.08%

CHURN ANALYSIS - SUMMARY

Monthly Charge Range

All

Married

All

Churn Prediction

6,418

411

1,732

27.0%

Total Customers

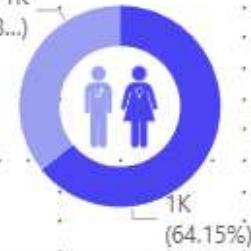
New Joiners

Total Churn

Churn Rate



Total Churn by Gender



Total Cust...

Total Custo...

1.6K

23.5%

Churn_Reason

Total Churn

Competitor had better devices	289
Competitor made better offer	274
Competitor offered more data	106
Competitor offered higher download speeds	92
Total	761

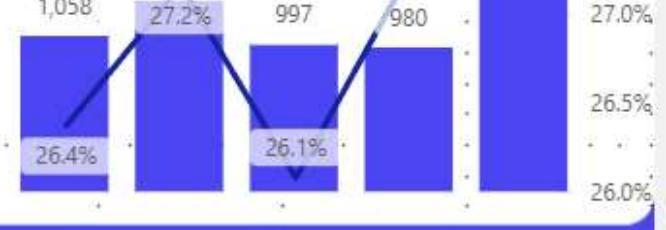
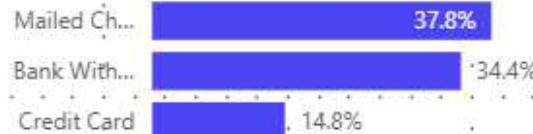
Churn Rate by Contract



Total C...

Total

Churn Rate by Payment Method



Demographic

State



Distribution

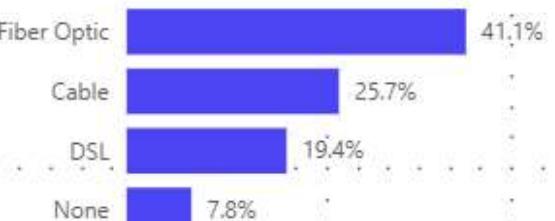
Churn Category

761



Services Used

Churn Rate by Internet Type



Services

No

Yes

Service	No	Yes
Device_Protection_Plan	71.02%	28.98%
Internet_Service	6.29%	93.71%
Multiple_Lines	54.79%	45.21%
Online_Backup	71.88%	28.12%
Online_Security	84.64%	15.36%
Paperless_Billing	25.40%	74.60%
Phone_Service	9.41%	90.59%
Premium_Support	83.49%	16.51%
Streaming_Movies	56.00%	44.00%
Streaming_Music	61.14%	38.86%
Streaming_TV	56.76%	43.24%
Unlimited_Data	19.92%	80.08%

CHURN ANALYSIS - SUMMARY

Monthly Charge Range

All

Married

All

Churn Prediction

4,048

Total Customers

269

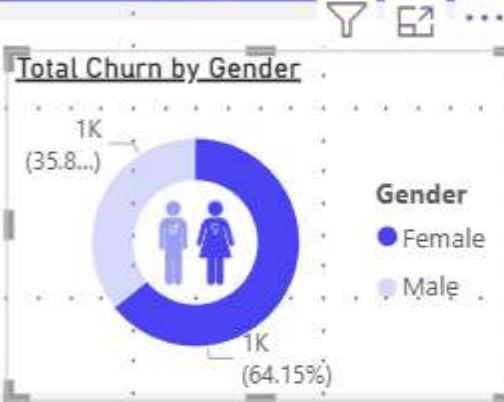
New Joiners

1,111

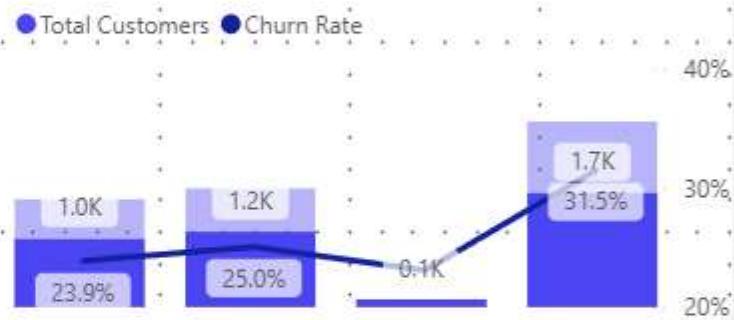
Total Churn

27.4%

Churn Rate

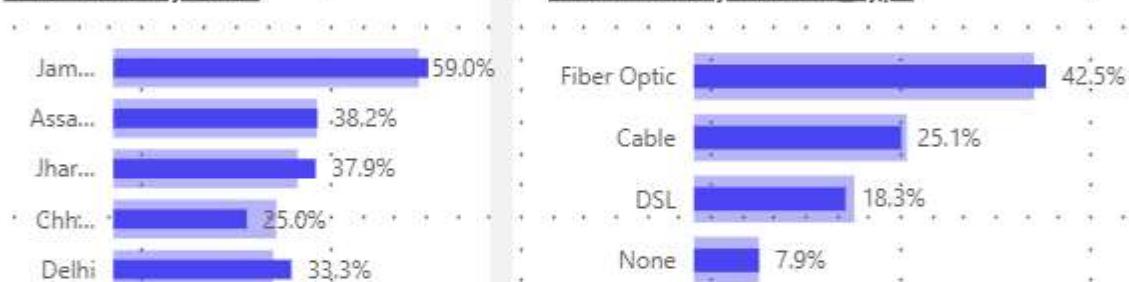


Total Customers and Churn Rate by Age Group



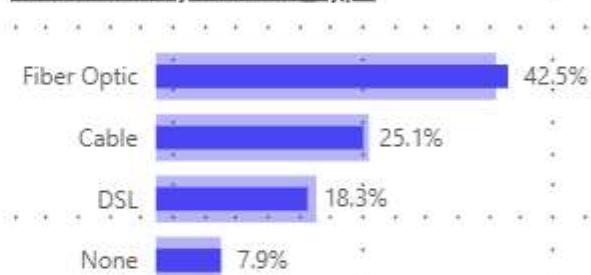
GEOGRAPHIC

Churn Rate by State



SERVICES USED

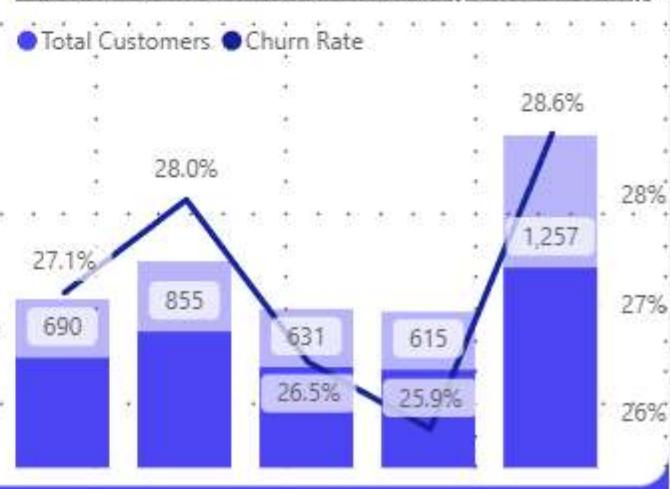
Churn Rate by Internet Type



Churn Rate by Contract

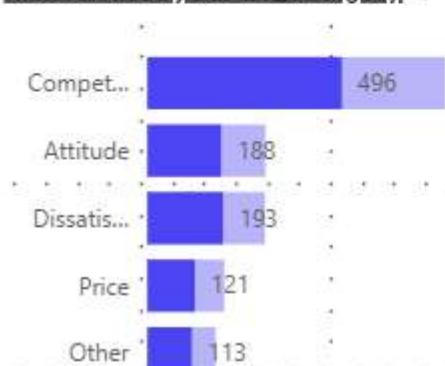


Total Customers and Churn Rate by Tenure Group

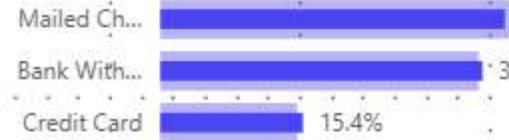


CHURN DISTRIBUTION

Total Churn by Churn Category



Churn Rate by Payment Method



Services	No	Yes
Device_Protection_Plan	69.49%	30.51%
Internet_Service	6.39%	93.61%
Multiple_Lines	52.30%	47.70%
Online_Backup	70.66%	29.34%
Online_Security	83.26%	16.74%
Paperless_Billing	25.02%	74.98%
Phone_Service	8.19%	91.81%
Premium_Support	82.63%	17.37%
Streaming_Movies	54.46%	45.54%
Streaming_Music	59.05%	40.95%
Streaming_TV	55.63%	44.37%
Unlimited_Data	19.44%	80.56%

CHURN ANALYSIS - SUMMARY

Monthly Charge Range

All

Married

All

Churn Prediction

2,370

Total Customers

142

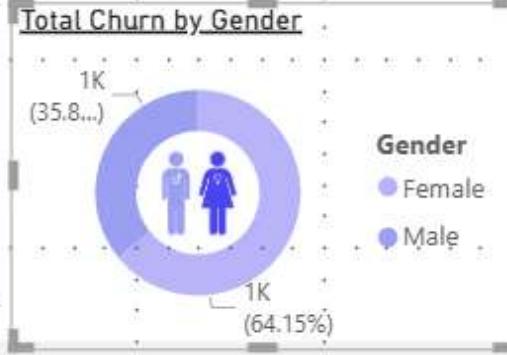
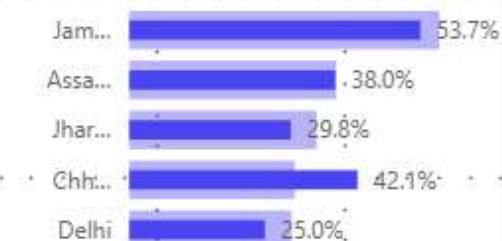
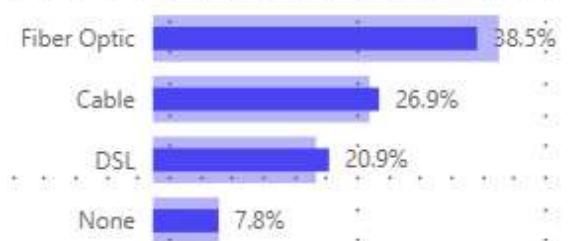
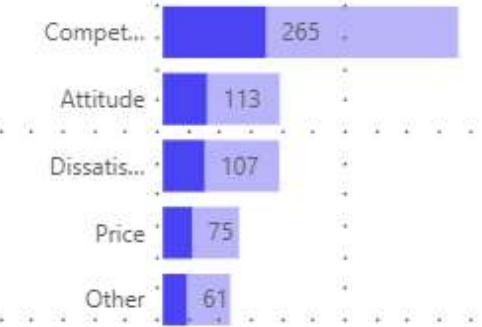
New Joiners

621

Total Churn

26.2%

Churn Rate

Total Customers and Churn Rate by Age Group**GEOGRAPHIC**Churn Rate by State**SERVICES USED**Churn Rate by Internet Type**ACCOUNT INFO**Churn Rate by ContractTotal Customers and Churn Rate by Tenure Group**CHURN DISTRIBUTION**Total Churn by Churn Category

Services	No	Yes
Device_Protection_Plan	73.75%	26.25%
Internet_Service	6.12%	93.88%
Multiple_Lines	59.26%	40.74%
Online_Backup	74.07%	25.93%
Online_Security	87.12%	12.88%
Paperless_Billing	26.09%	73.91%
Phone_Service	11.59%	88.41%
Premium_Support	85.02%	14.98%
Streaming_Movies	58.78%	41.22%
Streaming_Music	64.90%	35.10%
Streaming_TV	58.78%	41.22%
Unlimited_Data	20.77%	79.23%

CHURN ANALYSIS - SUMMARY

Monthly Charge Range

50-100

- < 20
- > 100
- 20-50
- 50-100

Married

All

Churn Prediction

3,428

Total Customers

143

New Joiners

1,

Total

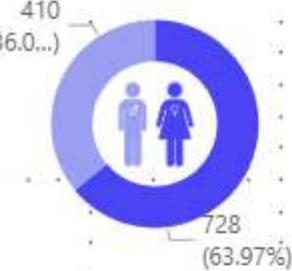
33.2%

Churn Rate

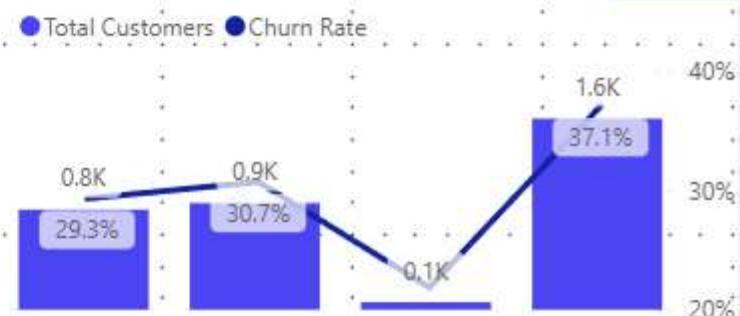


DEMOGRAPHIC

Total Churn by Gender



Total Customers and Churn Rate by Age Group



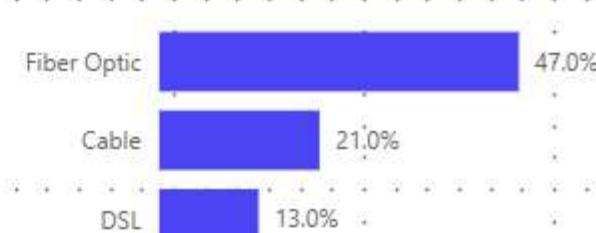
EOGRAPHIC

Churn Rate by State



SERVICES USED

Churn Rate by Internet Type



ACCOUNT INFO

Churn Rate by Contract



Total Customers and Churn Rate by Tenure Group



CHURN DISTRIBUTION

Total Churn by Churn Category



Services	No	Yes
Device_Protection_Plan	74.96%	25.04%
Internet_Service	100.00%	
Multiple_Lines	51.41%	48.59%
Online_Backup	74.34%	25.66%
Online_Security	85.50%	14.50%
Paperless_Billing	22.14%	77.86%
Phone_Service	2.02%	97.98%
Premium_Support	85.15%	14.85%
Streaming_Movies	58.00%	42.00%
Streaming_Music	61.78%	38.22%
Streaming_TV	58.26%	41.74%
Unlimited_Data	14.32%	85.68%

CHURN ANALYSIS - SUMMARY

1,726

Total Customers

74

New Joiners

554

Total Churn

3

Married

Monthly Charge Ran...

50-100

Married

Yes

No

Yes

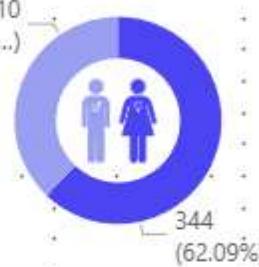
Churn Prediction

Churn Prediction



DEMOGRAPHIC

Total Churn by Gender



Gender

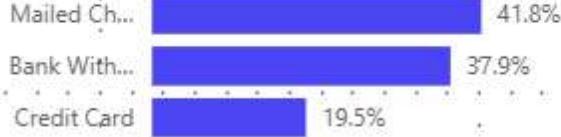
- Female
- Male

ACCOUNT INFO

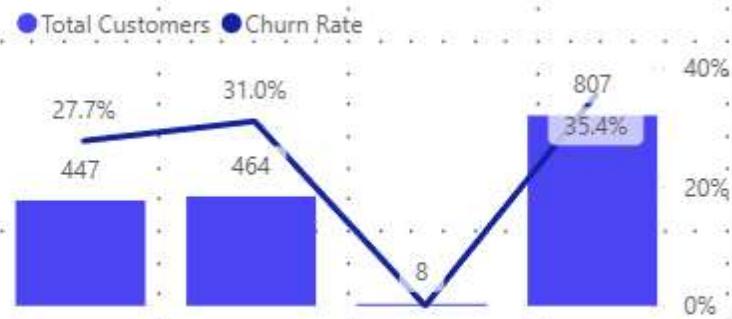
Churn Rate by Contract



Churn Rate by Payment Method



Total Customers and Churn Rate by Age Group

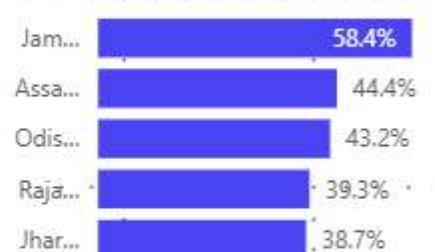


Total Customers and Churn Rate by Tenure Group



GEOGRAPHY

Churn Rate by State



Churn Rate by Internet Type



CHURN DISTRIBUTION

Total Churn by Churn Category



Services

Services	No	Yes
Device_Protection_Plan	74.55%	25.45%
Internet_Service		100.00%
Multiple_Lines	51.44%	48.56%
Online_Backup	72.38%	27.62%
Online_Security	84.66%	15.34%
Paperless_Billing	20.40%	79.60%
Phone_Service	2.53%	97.47%
Premium_Support	85.92%	14.08%
Streaming_Movies	57.22%	42.78%
Streaming_Music	60.47%	39.53%
Streaming_TV	61.19%	38.81%
Unlimited_Data	15.70%	84.30%

CHURN ANALYSIS - SUMMARY

1,726

Total Customers

74

New Joiners

554

Total Churn

3

Monthly Charge Range

50-100

Married

Yes

No

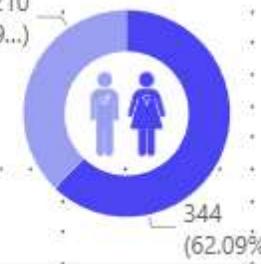
Yes

Churn Prediction



DEMOGRAPHIC

Total Churn by Gender



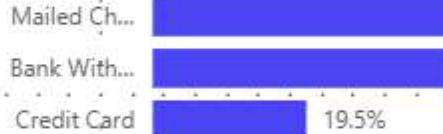
Gender

- Female
- Male

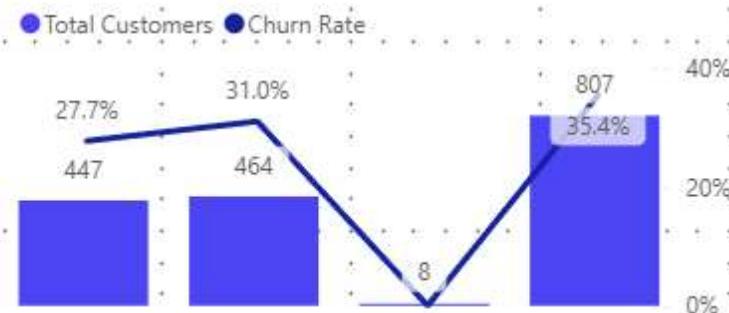
Churn Rate by Contract



Churn Rate by Payment Method



Total Customers and Churn Rate by Age Group

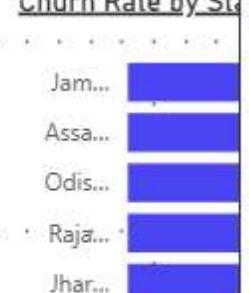


Total Customers and Churn Rate by Tenure Group



GEOGRAPHY

Churn Rate by State



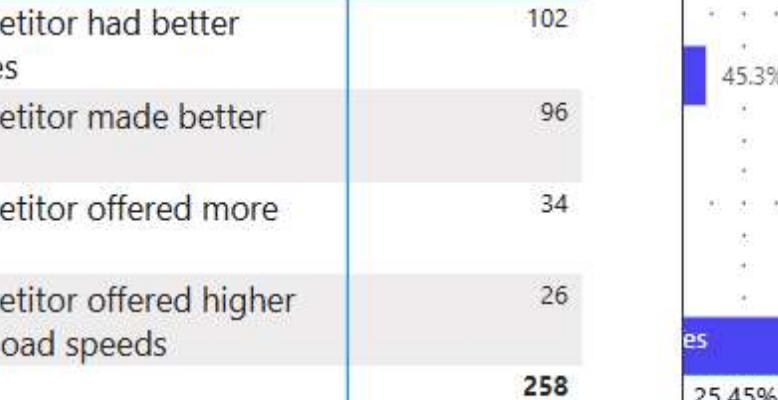
CHURN DIS

Total Churn by Churn Reason

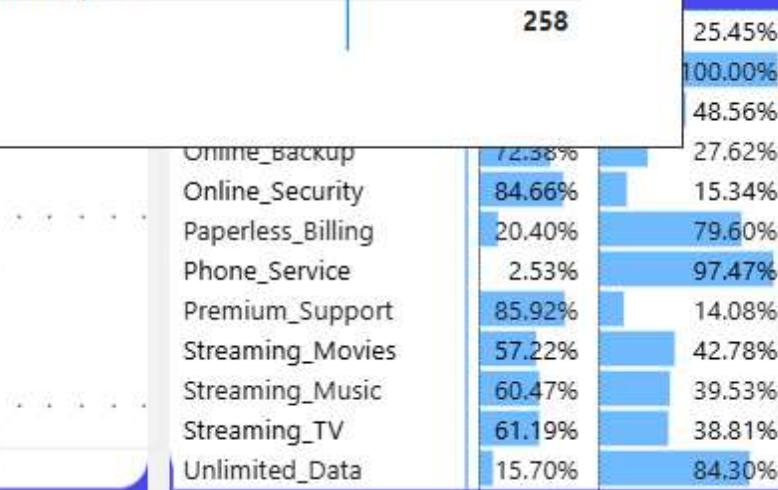


Churn_Reason

Total Churn



Total



ACCOUNT INFO

WHAT'S DONE TILL NOW?

- I. ETL DESIGN
- II. DATA MODEL
- III. CUSTOM MEASURES AND METRICS
- IV. VISUALIZATION DASHBOARD
- V. EXPLORATORY DATA ANALYSIS
- VI. DESCRIPTIVE ANALYSIS

WHAT'S NEXT ?

MACHINE LEARNING ALGORITHM

Now, I will use **ML Algorithm** which we gonna use with this data and use that final machine learning output which will be a predictive output to predict future churners. I'm going to use **Random Forest** (one of the ML Algorithm)

PLATFORM AND PROGRAMMING LANGUAGE NEEDED



THE EXCEL FILE WITH ALL THE CATEGORICAL DATA

Prediction.xlsx - Excel

Debina Mandal

Share

File Home Insert Page Layout Formulas Data Review View Help PDFelement Tell me what you want to do

Cut Copy Format Painter Paste

Font: Calibri 11pt

Wrap Text

General

Conditional Formatting

Format as Table

Cell Styles

Insert Delete Format

AutoSum

Fill

Sort & Find & Filter

Select

Add-ins

Clipboard

Font Alignment Number Styles Cells Editing Add-ins

Customer_Status_Predicted

	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF
1	Number_cTenure_in_Value_Deal	Phone_Se_Multiple_	Internet_SInternet_TOnline_Se	Online_Be_Device_Pr_Premium_	Streaming_Streaming_Streaming_Streaming_	Unlimited_Contract_Paperless_Payment_Monthly_(Total_Cha	Total_Ref_Total_Exi_Total_Lon_Total_Rev_Customer_Churn_Ca	Churn_Reason																			
2	5	7 Deal 5	No No Yes	DSL No No No	No Yes No No Yes	Month-to-No Mailed Ch	24.3 38.45	0 0 0 0	0 0 0 0	38.45	Joined Others																
3	2	20 Yes	No Yes	Fiber Opti No Yes	No No No No Yes	Month-to-No Bank With	90.4 268.45	0 0 0 0	0 0 0 0	94.44	362.89	Joined Others															
4	2	35 Yes	No No	No No No No	No No No No No	Month-to-No Bank With	19.9 19.9	0 0 0 0	0 0 0 0	11.83	31.73	Joined Others															
5	9	1 Yes	No No	No No No No	No No No No No	Month-to-Yes Credit Car	19.55 19.55	0 0 0 0	0 0 0 0	10.2	29.75	Joined Others															
6	0	10 Yes	No Yes	Cable No No No	No Yes Yes Yes	Month-to-Yes Credit Car	62.8 62.8	0 0 0 0	0 0 0 0	42.19	104.99	Joined Others															
7	11	4 Yes	No Yes	DSL No No No	No Yes No Yes	Month-to-Yes Credit Car	55.3 55.3	0 0 0 0	0 0 0 0	36.69	91.99	Joined Others															
8	9	6 Yes	No Yes	Fiber Opti No Yes	No No No No No	Month-to-Yes Mailed Ch	74.75 238.1	38.84	0 0 0 0	37.5	236.76	Joined Others															
9	13	22 Yes	No Yes	DSL No No No	No No No No Yes	Month-to-Yes Mailed Ch	46.1 130.15	0 0 0 0	0 0 0 0	7.98	138.13	Joined Others															
10	13	18 Deal 5 Yes	No Yes	Fiber Opti Yes	Yes No No Yes	Month-to-No Bank With	100.2 198.5	0 0 0 0	10 0 0 0	45.12	253.62	Joined Others															
11	15	16 Yes	No Yes	Fiber Opti No Yes	No No Yes Yes	Month-to-No Bank With	95.4 293.15	0 0 0 0	0 0 0 0	51.03	344.18	Joined Others															
12	8	3 Deal 5 Yes	No No	No No No No	No No No No Yes	Month-to-Yes Bank With	19.65 19.65	0 0 0 0	0 0 0 0	13.85	33.5	Joined Others															
13	1	28 Deal 5 No	No No	Yes DSL No No	No No No No No	Month-to-No Bank With	24.5 46.4	0 0 0 0	0 0 0 0	46.4	100.2	Joined Others															
14	1	3 Deal 5 Yes	No No	No No No No	No No No No No	Month-to-No Credit Car	19.65 19.65	0 0 0 0	0 0 0 0	23.67	43.32	Joined Others															
15	1	9 Yes	No No	No No No No	No No No No No	Month-to-No Credit Car	20.35 20.35	0 0 0 0	0 0 0 0	44.09	64.44	Joined Others															
16	9	12 Deal 5 Yes	Yes Yes	DSL No No No	No No No No No	Month-to-Yes Bank With	50.15 50.15	0 0 0 0	0 0 0 0	39.87	90.02	Joined Others															
17	3	31 Yes	No No	No No No No	No No No No No	Month-to-No Credit Car	20.4 20.4	0 0 0 0	0 0 0 0	45.91	66.31	Joined Others															
18	6	31 Yes	No No	No No No No	No No No No No	Month-to-No Credit Car	19.15 19.15	0 0 0 0	0 0 0 0	22.37	41.52	Joined Others															
19	8	18 Yes	No Yes	DSL No No No	No No No No No	Month-to-Yes Bank With	46.6 87.9	0 0 0 0	0 0 0 0	3.74	91.64	Joined Others															
20	4	25 Deal 5 No	No No	Yes DSL No No	No No No No No	Month-to-No Bank With	25.25 25.25	0 0 0 0	10 0 0 0	35.25	55.25	Joined Others															
21	5	7 Yes	No No	No No No No	No No No No No	Month-to-No Credit Car	19.55 19.55	0 0 0 0	0 0 0 0	5.1	24.65	Joined Others															
22	10	1 Yes	No Yes	DSL No No No	No No No No Yes	Month-to-Yes Credit Car	45.85 45.85	0 0 0 0	0 0 0 0	40.38	86.23	Joined Others															
23	15	13 Deal 5 Yes	No No	No No No No	No No No No No	Month-to-No Credit Car	20.8 20.8	0 0 0 0	0 0 0 0	48.14	68.94	Joined Others															
24	8	23 Deal 5 Yes	No No	No No No No	No No No No No	Month-to-Yes Credit Car	19.3 19.3	0 0 0 0	0 0 0 0	8.66	27.96	Joined Others															
25	13	4 Yes	No No	No No No No	No No No No No	One Year No Mailed Ch	18.8 18.8	0 0 0 0	0 0 0 0	43.57	62.37	Joined Others															
26	12	2 Yes	No Yes	DSL No Yes	No No No No No	Month-to-No Credit Car	49.25 97	0 0 0 0	0 0 0 0	81.5	178.5	Joined Others															
27	2	6 Deal 5 No	No No	Yes DSL No Yes	No No No No No	Month-to-Yes Bank With	29.85 29.85	0 0 0 0	0 0 0 0	29.85	55.25	Joined Others															
28	5	1 Yes	No Yes	Cable Yes No	No No No No No	Month-to-Yes Bank With	48.45 48.45	0 0 0 0	0 0 0 0	16.42	64.87	Joined Others															
29	15	9 Yes	No Yes	DSL No No	Yes No No No Yes	Month-to-Yes Bank With	49.9 49.9	0 0 0 0	0 0 0 0	14.93	64.83	Joined Others															
30	15	30 Deal 5 Yes	No No	No No No No	No No No No No	Month-to-Yes Credit Car	19.9 33.7	0 0 0 0	0 0 0 0	7.64	41.34	Joined Others															
31	7	36 Yes	No No	No No No No	No No No No No	Month-to-No Mailed Ch	19.1 19.1	0 0 0 0	0 0 0 0	5.02	24.12	Joined Others															
32	4	31 Deal 5 Yes	No No	No No No No	No No No No No	Month-to-No Credit Car	18.8 56	0 0 0 0	0 0 0 0	19.83	75.83	Joined Others															
33	6	21 Deal 5 No	No No	Yes Cable Yes	No No Yes No No	Month-to-No Credit Car	35.1 101.1	0 0 0 0	0 0 0 0	101.1	152.2	Joined Others															
34	14	8 Deal 5 Yes	No No	Yes DSL No No	No No No No No	Month-to-Yes Credit Car	45.4 45.4	0 0 0 0	0 0 0 0	49.33	94.73	Joined Others															
35	16	44 Deal 5 Yes	No No	Yes Cable No No	No No No No No	Month-to-Yes Credit Car	45.25 74.1	0 0 0 0	0 0 0 0	70	152.2	Joined Others															

HERE, WE HAVE IMPORTED BOTH THE CHURN_DATA AND THE JOIN DATA INTO 1 FILE. THEN WE WILL IMPORT THIS FILE FOR ML ALGORITHM.

File Edit View Run Kernel Settings Help

Trustee

JupyterLab Python 3 (ipykernel)

```
[1]: pip install pandas numpy matplotlib seaborn scikit-learn joblib
Defaulting to user installation because normal site-packages is not writable
Requirement already satisfied: pandas in d:\deblina\anaconda\lib\site-packages (2.2.2)
Requirement already satisfied: numpy in d:\deblina\anaconda\lib\site-packages (1.26.4)
Requirement already satisfied: matplotlib in d:\deblina\anaconda\lib\site-packages (3.9.2)
Requirement already satisfied: seaborn in d:\deblina\anaconda\lib\site-packages (0.13.2)
Requirement already satisfied: scikit-learn in d:\deblina\anaconda\lib\site-packages (1.5.1)
Requirement already satisfied: joblib in d:\deblina\anaconda\lib\site-packages (1.4.2)
Requirement already satisfied: python-dateutil<=2.8.2 in d:\deblina\anaconda\lib\site-packages (from pandas) (2.9.0.post0)
Requirement already satisfied: pytz>=2028.1 in d:\deblina\anaconda\lib\site-packages (from pandas) (2024.1)
Requirement already satisfied: tzdata>=2022.7 in d:\deblina\anaconda\lib\site-packages (from pandas) (2023.3)
Requirement already satisfied: contourpy<1.8.1 in d:\deblina\anaconda\lib\site-packages (from matplotlib) (1.1.0)
Requirement already satisfied: cycler>=0.10 in d:\deblina\anaconda\lib\site-packages (from matplotlib) (0.11.0)
Requirement already satisfied: fonttools<=4.22.8 in d:\deblina\anaconda\lib\site-packages (from matplotlib) (4.51.0)
Requirement already satisfied: kiwisolver<1.3.1 in d:\deblina\anaconda\lib\site-packages (from matplotlib) (1.4.4)
Requirement already satisfied: packaging<20 in d:\deblina\anaconda\lib\site-packages (from matplotlib) (24.1)
Requirement already satisfied: pillow<9.0 in d:\deblina\anaconda\lib\site-packages (from matplotlib) (10.4.0)
Requirement already satisfied: pyparsing<2.3.1 in d:\deblina\anaconda\lib\site-packages (from scikit-learn) (3.1.2)
Requirement already satisfied: scipy<1.6.0 in d:\deblina\anaconda\lib\site-packages (from scikit-learn) (1.13.1)
Requirement already satisfied: threadpoolctl<=3.1.0 in d:\deblina\anaconda\lib\site-packages (from scikit-learn) (3.5.0)
Requirement already satisfied: six<1.5 in d:\deblina\anaconda\lib\site-packages (from python-dateutil>=2.8.2->pandas) (1.16.0)
Note: you may need to restart the kernel to use updated packages.
```

```
[3]: import pandas as pd
import numpy as np

import matplotlib.pyplot as plt

import seaborn as sns

from sklearn.model_selection import train_test_split

from sklearn.ensemble import RandomForestClassifier

from sklearn.metrics import classification_report, confusion_matrix

from sklearn.preprocessing import LabelEncoder

import joblib
```

Define the path to the Excel file

Importing certain libraries and loading the data in the python environment

```
confusion_matrix([[True, False], [False, True]])
classification_report()
precision, recall, f1-score, support
```

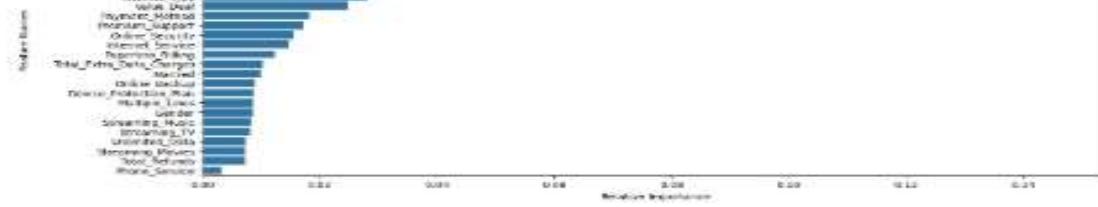
	0	1
0	0.82	0.80
1	0.78	0.73

```
precision
recall
f1-score
support
```

	0	1
0	0.82	0.80
1	0.78	0.73

```
weighted avg
0.82
0.79
0.80
0.82
```

F1 Score



File Edit View Run Kernel Settings Help

jupyter Customer Retention Last Checkpoint: 21 hours ago

JupyterLab Python 3 (ipykernel)

```
[1]: # Define the path to the Excel file
file_path = "C:\Users\DEBLINA\OneDrive\Documents\GitHub\CustomerRetention\customer_churn.csv"

# Define the sheet name to read data from
sheet_name = 'Customer'

# Read the data from the specified sheet into a pandas DataFrame
data = pd.read_excel(file_path, sheet_name=sheet_name)

# Adapt the first few rows of the fetched data
print(data.head(10))

customer_id service age married state number_of_referrals
0 11100-0001 Female 28 Yes Madison Wisconsin 0
1 11111-0001 Male 35 Yes New York New York 0
2 11112-0001 Female 35 Yes West Bengal India 0
3 11113-0001 Male 35 Yes Maharashtra India 0
4 11130-1001 Male 74 Yes Las Vegas Nevada 0

Inference_id Month value_deal phone_service multiple_lines ...
0 12 2023-01-01 1000 Yes No ----
1 12 2023-01-01 1000 Yes Yes ----
2 12 2023-01-01 1000 Yes Yes Yes ----
3 12 2023-01-01 1000 Yes Yes No ----
4 12 2023-01-01 1000 Yes Yes No ----

Payment_Method monthly_charge total_charges total_revenue
0 Bank Withdrawal 0.000000 0.000000 0.00
1 Bank Withdrawal 0.000000 100.000000 100.00
2 Bank Withdrawal 0.000000 100.000000 100.00
3 Credit Card 0.000000 100.000000 100.00
4 Credit Card 0.000000 100.000000 100.00

total_bills_data_charge total_long_distance_charge total_revenue
0 0 0.000000 0.000000
1 0 100.000000 100.000000
2 0 100.000000 100.000000
3 0 100.000000 100.000000
4 0 100.000000 100.000000

customer_status churn_category churn_reason
0 Unsubscribed OTHERS OTHERS
1 Churned COMPETITOR COMPETITOR bad internet service OTHERS
2 Unsubscribed OTHERS OTHERS
3 Unsubscribed OTHERS OTHERS
4 Unsubscribed OTHERS OTHERS

In [1]: 42 columns
```

```
[1]: customer_id service age married state number_of_referrals
0 11111-0001 Male 35 Yes New York New York 0
1 11112-0001 Female 35 Yes West Bengal India 0
2 11113-0001 Male 35 Yes Maharashtra India 0
3 11130-1001 Male 74 Yes Las Vegas Nevada 0

Inference_id Month value_deal phone_service multiple_lines ...
0 12 2023-01-01 1000 Yes No ----
1 12 2023-01-01 1000 Yes Yes ----
2 12 2023-01-01 1000 Yes Yes Yes ----
3 12 2023-01-01 1000 Yes Yes No ----
4 12 2023-01-01 1000 Yes Yes No ----

Payment_Method monthly_charge total_charges total_revenue
0 Mailed Check 0.000000 0.000000 0.00
1 Bank Withdrawal 0.000000 100.000000 100.00
2 Bank Withdrawal 0.000000 100.000000 100.00
3 Credit Card 0.000000 100.000000 100.00
4 Credit Card 0.000000 100.000000 100.00

total_bills_data_charge total_long_distance_charge total_revenue
0 0 0.000000 0.000000
1 0 100.000000 100.000000
2 0 100.000000 100.000000
3 0 100.000000 100.000000
4 0 100.000000 100.000000

customer_status churn_category churn_reason
0 Unsubscribed OTHERS OTHERS
1 Churned OTHERS OTHERS
2 Unsubscribed OTHERS OTHERS
3 Unsubscribed OTHERS OTHERS
4 Unsubscribed OTHERS OTHERS

In [1]: 42 columns
```

Preview of the loaded data and the new cell that was recently added

Finally storing the new prediction model and saving the output with the csv file

The Predicted Column

Deblina Mandal DM

Share

Font: Calibri, Size: 11, Bold, Italic, Underline, Color: Black, Alignment: Wrap Text, General, Merge & Center, Number, Styles, Conditional Formatting, Table, Cell Styles, Insert, Delete, Format, AutoSum, Fill, Sort & Filter, Clear, Add-ins.

Cells, Editing, Add-ins.

Customer_Status_Predicted

Age	Married	State	Number_cTenure_inValue_DelPhone_SeMultiple_	Internet_3GOnline_5GOnline_8GDevice_PrPremium_StreamingStreamingUnlimitedContract	PaperlessPayment_Monthly_TotalChTotal_RefiTotal_ExtrTotal_ConTotal_RevCustomerChurn_CaChurn_Reason	Customer_Status_Predicted
18 No	Tamil Nat	5 7 Deal 5	No Yes	DSL No No No No No No Yes Month-to-No Mailed Cr 24.3 38.45 0 0 0 38.45 Joined Others Others	1	
27 No	West Benj	2 20	Yes No Yes	Fiber Opt No Yes No No Yes Yes Month-to-No Bank Wifl 90.4 268.45 0 0 0 94.44 Joined Others Others	1	
25 Yes	Rajasthan	2 35	Yes No No	No No No No No No No Month-to-Yes Bank Wifl 19.9 19.95 0 0 0 11.85 81.73 Joined Others Others	1	
39 No	Assam	9 1	Yes No No	No No No No No No No Month-to-Yes Credit Cr 19.53 19.55 0 0 0 10.2 29.75 Joined Others Others	1	
51 Yes	Delhi	0 10	Yes No Yes	Cable No No No No Yes Yes Month-to-Yes Credit Cr 62.8 61.8 0 0 0 42.19 104.99 Joined Others Others	1	
24 Yes	Andhra Pr	11 4	Yes No Yes	DSL No No No Yes No Yes Month-to-Yes Credit Cr 55.3 55.3 0 0 0 36.69 91.99 Joined Others Others	1	
52 No	Telengana	9 6	Yes No Yes	Fiber Opt No Yes No No No Yes Month-to-Yes Mailed Cr 74.75 258.1 38.84 0 0 57.5 256.76 Joined Others Others	1	
58 No	Madhya P	13 22	Yes No Yes	DSL No No No No No No Yes Month-to-Yes Mailed Cr 46.1 130.15 0 0 0 7.98 138.13 Joined Others Others	1	
46 No	Bihar	13 18 Deal 5	Yes No Yes	Fiber Opt Yes Yes No Yes Yes Yes Month-to-No Bank Wifl 100.2 198.5 0 0 10 45.12 253.62 Joined Others Others	1	
57 Yes	Uttarakhu	15 16	Yes No Yes	Fiber Opt No Yes No No Yes Yes Month-to-No Bank Wifl 95.4 283.15 0 0 0 51.03 344.18 Joined Others Others	1	
41 Yes	Andhra Pr	8 5 Deal 5	Yes No No	No No No No No No No Month-to-Yes Bank Wifl 19.65 19.65 0 0 0 13.85 33.5 Joined Others Others	1	
31 Yes	Telengana	1 28 Deal 5	No No Yes	DSL No No No No No No Yes Month-to-No Bank Wifl 24.5 46.4 0 0 0 46.4 Joined Others Others	1	
41 Yes	Haryana	1 5 Deal 5	Yes No No	No No No No No No No Month-to-No Credit Cr 19.65 19.65 0 0 0 23.67 43.32 Joined Others Others	1	
43 No	Tamil Nat	1 9	Yes No	No No No No No No No Month-to-No Credit Cr 20.35 20.35 0 0 0 44.09 64.44 Joined Others Others	1	
47 Yes	Uttar Praj	9 12 Deal 5	Yes Yes	DSL No No No No No No Yes Month-to-Yes Bank Wifl 50.15 50.15 0 0 0 39.87 90.02 Joined Others Others	1	
26 No	Karnataka	5 31	Yes No	No No No No No No No Month-to-No Credit Cr 20.4 20.4 0 0 0 45.91 66.31 Joined Others Others	1	
34 No	Uttar Praj	6 81	Yes No No	No No No No No No No Month-to-No Credit Cr 19.15 19.15 0 0 0 22.97 41.52 Joined Others Others	1	
47 Yes	Andhra Pr	8 18	Yes No Yes	DSL No No No No No No Yes Month-to-Yes Bank Wifl 46.6 87.9 0 0 0 5.74 92.64 Joined Others Others	1	
44 Yes	Bihar	4 25 Deal 5	No No Yes	DSL No No No No No No No Month-to-No Bank Wifl 25.25 25.25 0 0 10 0 35.25 Joined Others Others	1	
19 Yes	Uttar Praj	5 7	Yes No No	No No No No No No No Month-to-No Credit Cr 19.55 19.55 0 0 0 5.1 24.65 Joined Others Others	1	
76 Yes	Odisha	10 1	Yes No Yes	DSL No No No No No No Yes Month-to-Yes Credit Cr 45.85 45.85 0 0 0 40.38 86.23 Joined Others Others	1	
40 No	West Benj	15 13 Deal 5	Yes No No	No No No No No No No Month-to-No Credit Cr 20.8 20.8 0 0 0 48.14 68.94 Joined Others Others	1	
39 Yes	Bihar	8 23 Deal 5	Yes No No	No No No No No No No Month-to-Yes Credit Cr 19.3 19.3 0 0 0 8.66 27.96 Joined Others Others	1	
32 No	Haryana	13 4	Yes No No	No No No No No No No Month-to-No One Year No Mailed Cr 18.8 18.8 0 0 0 43.57 62.37 Joined Others Others	1	
23 Yes	Haryana	12 2	Yes No Yes	DSL No Yes No No No No Yes Month-to-No Credit Cr 49.25 97 0 0 0 81.5 178.5 Joined Others Others	1	
33 Yes	Haryana	2 6 Deal 5	No No Yes	DSL No Yes No No No No Yes Month-to-Yes Bank Wifl 29.85 29.85 0 0 0 29.85 Joined Others Others	1	
65 No	Bihar	5 1	Yes No	Cable Yes No No No No No Yes Month-to-No Bank Wifl 48.45 48.45 0 0 0 16.42 64.87 Joined Others Others	1	
28 Yes	Odisha	15 9	Yes No Yes	DSL No No No Yes No No Yes Month-to-Yes Bank Wifl 49.9 49.9 0 0 0 14.95 64.83 Joined Others Others	1	
58 Yes	Uttar Praj	15 30 Deal 5	Yes No No	No No No No No No No Month-to-Yes Credit Cr 19.9 33.7 0 0 0 7.64 41.34 Joined Others Others	1	
33 Yes	Karnataka	7 36	Yes No No	No No No No No No No Month-to-No Mailed Cr 19.1 19.1 0 0 0 5.02 24.12 Joined Others Others	1	
49 Yes	Tamil Nat	4 31 Deal 5	Yes No No	No No No No No No No Month-to-No Credit Cr 18.8 56 0 0 0 19.83 75.83 Joined Others Others	1	
36 No	Andhra Pr	6 21 Deal 5	No Yes	Cable Yes No No Yes No No Yes Month-to-No Credit Cr 35.1 101.1 0 0 0 101.1 Joined Others Others	1	
37 Yes	Punjab	14 8 Deal 5	Yes No Yes	DSL No No No No No No Yes Month-to-Yes Credit Cr 45.4 45.4 0 0 0 49.33 94.73 Joined Others Others	1	
37 Yes	Odisha	15 14 Deal 5	Yes No Yes	Cable No No No No No No Yes Month-to-Yes Credit Cr 45.25 74.2 0 0 0 79 153.2 Joined Others Others	1	
47 No	West Benj	15 19 Deal 5	Yes No Yes	DSL Yes No Yes No Yes Yes No Month-to-Yes Bank Wifl 64.3 177.4 0 60 118.65 356.05 Joined Others Others	1	
35 Yes	Maharast	10 7 Deal 5	Yes No Yes	Fiber Opt Yes Yes No Yes Yes No Yes Month-to-No Bank Wifl 89.85 248.4 0 0 0 99.9 348.3 Joined Others Others	1	
57 Yes	Tamil Nat	4 6	No No Yes	Cable No Yes No No Yes Yes Yes Month-to-No Bank Wifl 51.4 96.8 0 0 0 96.8 Joined Others Others	1	
31 Yes	Haryana	3 30 Deal 5	Yes No No	No No No No No No No Month-to-No Credit Cr 19.45 19.45 0 0 0 29.15 48.6 Joined Others Others	1	
18 No	Andhra Pr	0 1 Deal 5	Yes No Yes	Fiber Opt No Yes Yes No Yes Yes Yes Month-to-No Credit Cr 85.7 256.75 0 0 0 89.52 348.27 Joined Others Others	1	
27 No	Haryana	5 10	Yes No No	No No No No No No No Month-to-No Credit Cr 20.3 20.3 0 0 0 10.84 81.14 Joined Others Others	1	

Now, All the categorical data are there but at the end we have one more column that's says "Customer Status Predicted"

CHURN PREDICTION PAGE BLUEPRINT

PREDICTED CHURNER PROFILE



CUSTOMER AT RISK

LIST OF PREDICTED CHURNERS

CHURN ANALYSIS - PREDICTION

Summary

PREDICTED CHURNER PROFILE

246



Female

132

Male

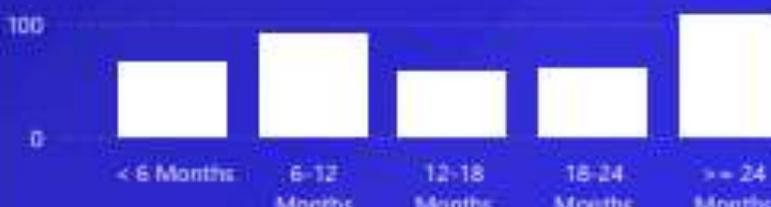
by Age Group



by Marital Status



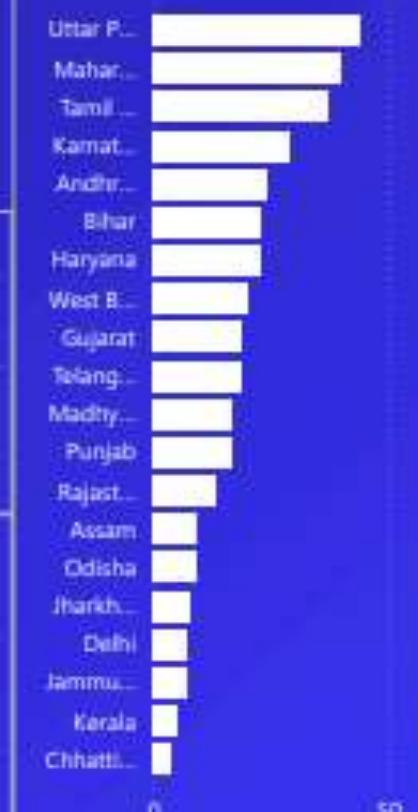
by Tenure Group



by Payment Method



by State



by Contract



CUSTOMERS AT RISK

COUNT OF PREDICTED CHURNERS : 378

Customer_ID	Monthly_Charge	Number_of_Referrals	Total_Refunds	Total_Revenue
11751-TAM	24.30	5	0.00	38.45
12056-WES	90.40	2	0.00	362.85
12136-RAJ	19.90	2	0.00	31.75
12257-ASS	19.55	9	0.00	29.75
12340-DEL	62.80	0	0.00	104.95
12469-AND	55.30	11	0.00	91.95
12490-TEL	74.75	9	38.84	236.76
13058-MAD	46.10	13	0.00	138.15
13123-BIH	100.20	13	0.00	253.62
13666-UTT	95.40	15	0.00	344.15
13744-AND	19.65	8	0.00	33.50
13823-TEL	24.50	1	0.00	46.45
13946-HAR	19.65	1	0.00	43.32
14567-TAM	20.35	1	0.00	64.45
15349-UTT	50.15	9	0.00	90.02
15591-KAR	20.40	3	0.00	66.31
15803-UTT	19.15	6	0.00	41.52
16032-AND	46.60	8	0.00	91.54
16068-BIH	25.25	4	0.00	35.25
16244-UTT	19.55	5	0.00	24.65
16733-ODI	45.85	10	0.00	86.25
16764-WES	20.80	15	0.00	68.95
16931-BIH	19.30	8	0.00	27.95
17444-LAK	19.80	11	0.00	49.95

CHURN ANALYSIS - PREDICTION

Summary

PREDICTED CHURNER PROFILE

11

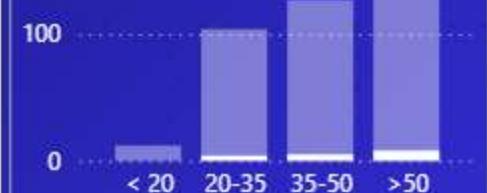


Female

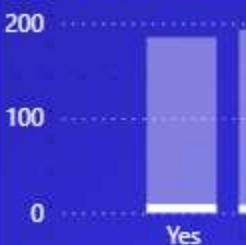
6

Male

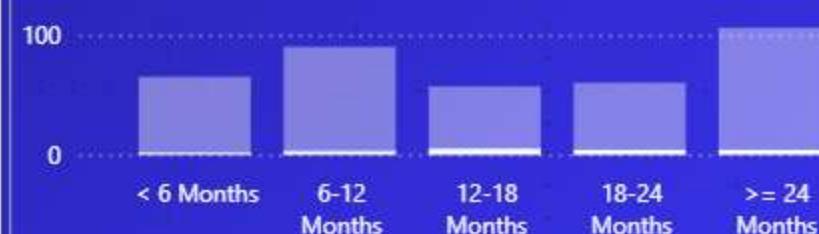
by Age Group



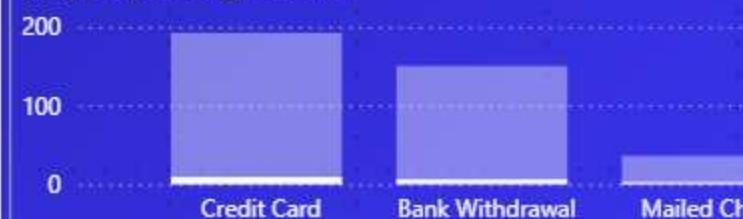
by Marital Status



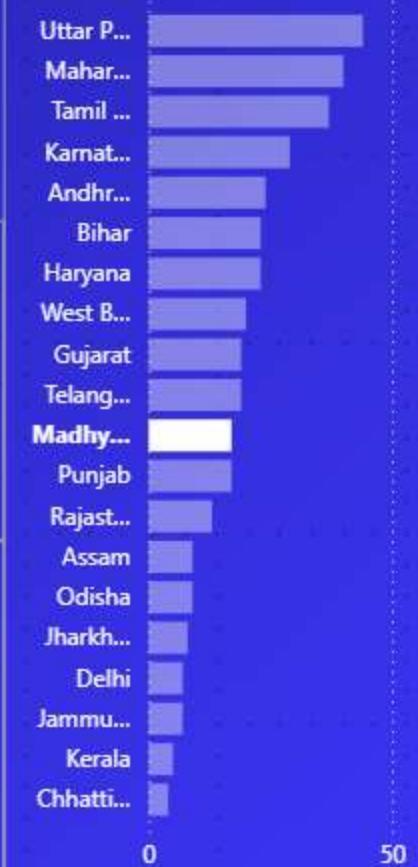
by Tenure Group



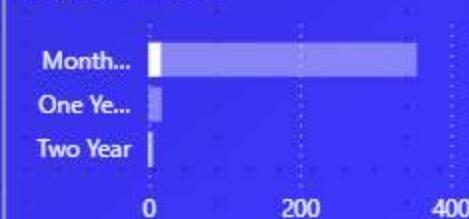
by Payment_Method



by State



by Contract



CUSTOMERS AT RISK

COUNT OF PREDICTED CHURNERS : 17

Customer_ID	Monthly_Charge	Number_of_Referrals	Total_Refunds	Total_Revenue
13058-MAD	46.10	13	0.00	138.13
27461-MAD	44.55	7	0.00	67.59
28378-MAD	19.45	12	0.00	66.04
33621-MAD	70.35	13	0.00	155.65
46792-MAD	35.90	6	0.00	35.90
51597-MAD	19.90	10	0.00	21.40
57457-MAD	20.00	14	0.00	106.76
58974-MAD	20.30	11	0.00	66.56
67994-MAD	46.30	12	0.00	57.81
72357-MAD	44.30	14	0.00	87.25
72469-MAD	55.70	9	0.00	73.38
74979-MAD	29.05	10	0.00	44.75
77311-MAD	20.25	9	0.00	60.71
81118-MAD	45.80	1	0.00	59.41

CHURN ANALYSIS - PREDICTION

Summary

PREDICTED CHURNER PROFILE

15

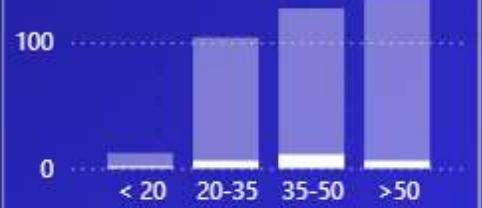


Female

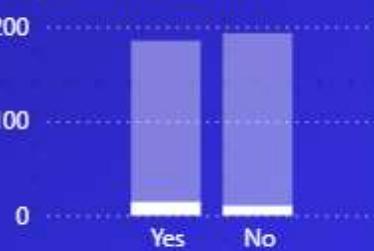
9

Male

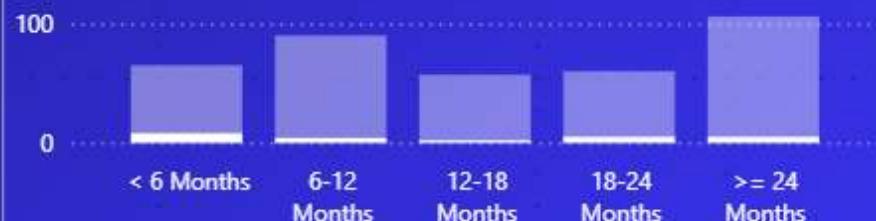
by Age Group



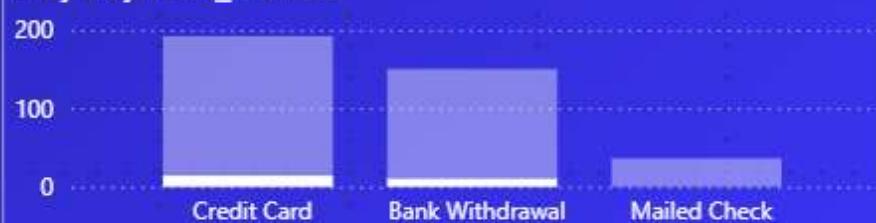
by Marital Status



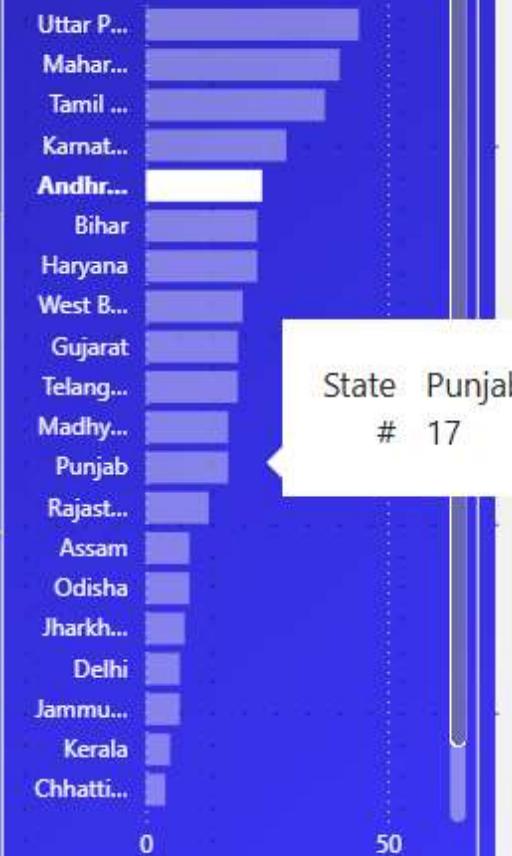
by Tenure Group



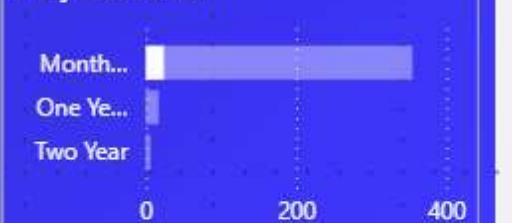
by Payment_Method



by State



by Contract



CUSTOMERS AT RISK

COUNT OF PREDICTED CHURNERS : 24

Customer_ID	Monthly_Charge	Number_of_Referrals	Total_Refunds	Total_Revenue
12469-AND	55.30	11	0.00	91.99
13744-AND	19.65	8	0.00	33.50
16032-AND	46.60	8	0.00	91.64
19041-AND	35.10	6	0.00	101.10
19998-AND	85.70	0	0.00	346.27
24754-AND	69.55	0	33.80	171.41
31129-AND	70.35	15	0.00	94.77
34024-AND	20.05	4	0.00	27.71
38748-AND	19.65	13	0.00	47.53
44208-AND	69.15	2	0.00	307.84
45213-AND	61.20	15	0.00	190.07
47492-AND	20.15	15	0.00	38.78
54564-AND	76.10	11	0.00	300.65
59750-AND	74.90	11	0.00	218.21
62359-AND	33.60	5	0.00	83.60
62535-AND	70.70	8	0.00	300.77
63464-AND	60.65	7	0.00	264.37
64327-AND	20.65	8	0.00	55.89
77658-AND	24.40	0	0.00	24.40
82473-AND	29.80	10	0.00	94.40
83939-AND	20.10	8	0.00	108.87
89303-AND	69.95	6	0.00	202.10
92398-AND	20.00	5	0.00	107.81
95822-AND	19.20	3	0.00	37.09

THANK YOU

