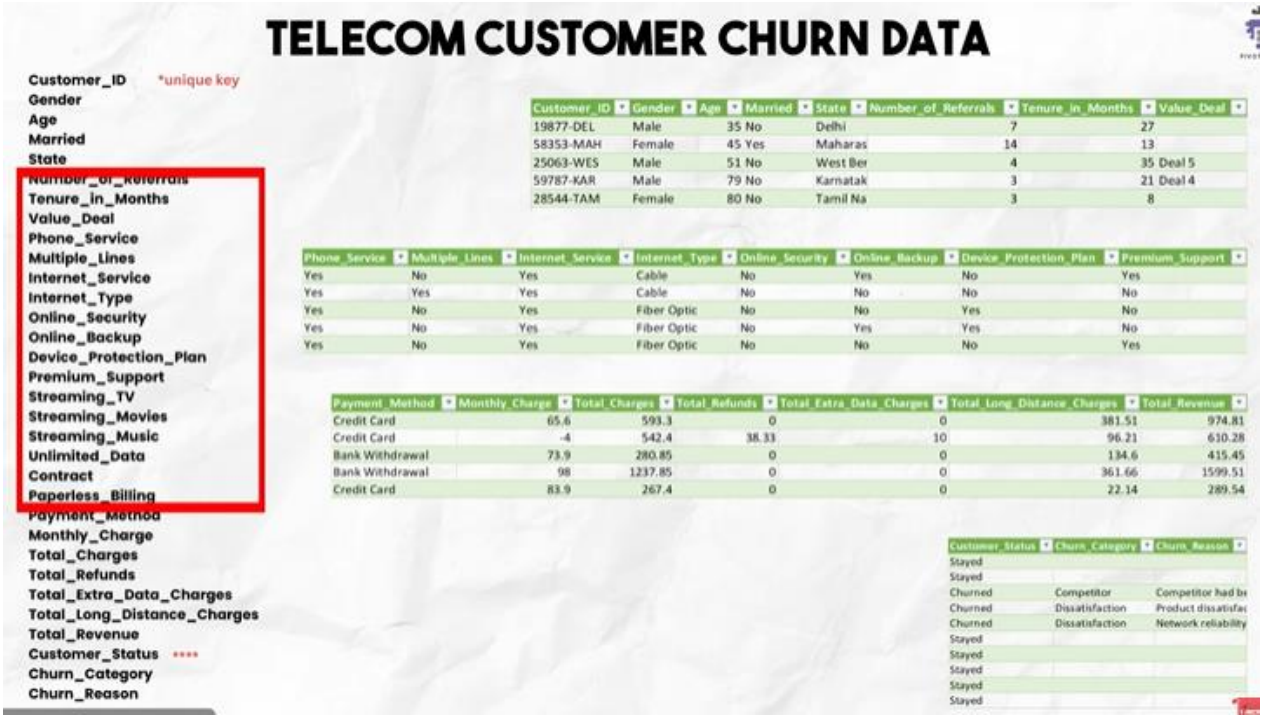
Telecom Churn analysis in customer data

Churn analysis is an important skill in any kind of customer data.

- Churn analysis dashboard
- Extract, Transform and Load Data viz ETL process in SQL database.
- Build a ML model to predict the customer churn on new data and then use that final prediction to build another page in that dashboard to create the churn profile.



Project Target

Create an entire ETL process in a database & a Power BI dashboard to utilize the Customer Data and achieve below goals:

- Visualize & Analyse Customer Data at below levels
- Demographic
- Geographic
- Payment & Account Info
- Services
- Study Churner Profile & Identify Areas for Implementing Marketing Campaigns
- Identify a Method to Predict Future Churners

Metrics Required

- Total Customers
- Total Churn & Churn Rate
- New Joiners

To load the data we use the microsoft SQL server because it is the widely used solution across the industry also because the full fledged database system is better at handling the recurring data loads and maintaining the data integrity as compared to an excel.

Microsoft provides us with the GUI interface known SQL server management database.
Download SQL Server Management Studio (SSMS)

SQL server name used while using the Power BI:
DESKTOP-7VTC0GS\SQLEXPRESS

Open the SQL server management.
We will be creating the database and that database will host our entire data.
3 major steps to be done:

- Create the database.
- Create the table
- Load the data

Now open up new query editor create the “db_Churn” that is the database

Data exploration has been done for checking :

- Number of people in different genders
- Contract and their percentage
- Customer_Status that whether they are stayed , joined or churned
- Number of services in different states
- Number of distinct internet type (i.e, Fibre Optics, Cable, NULL, DSL)

Then the number of the NULL values in each columns and they are changes to 1 or 0 , so that we can get the count of the NULL values.

Remove null and insert the new data into Prod table
Those NULL values are changed to NONE or No and a production table in the same database is created known as “prod_Churn” which will actually be used from now on .

Create View for Power BI:
A view is a virtual table in a database that displays data from one or more tables through a predefined query. It does not store the data itself but provides a way to look at and interact with data from the underlying tables as if it were a regular table. It has its own advantages and disadvantages.

STEP 2

Now we will jump into powerbi and import the data from there.

We have to get the data from the SQL server , also we can write the sql query in the power bi itself and fetch the data directly but we will create the tables and views in SQL itself and importing it directly.

Now we have to do some transformation before visualizing the data instead of loading it. So we click on the transform data.

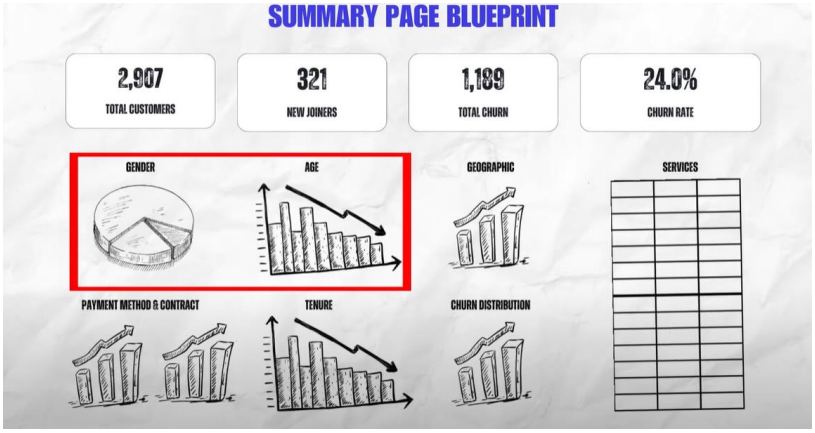
Add a new column in prod_Churn based on certain condition.

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"

Now after this we go to HOME > close and apply, this includes all the added columns also in the table.

Now we save the dashboard



demographic related visuals - Age and gender
Account related visuals - Payment method & Contract and Tenure
Geographic - State related data
Churn distribution - we will look at churn category and churn reasons

Power BI Measure

So we will keep the measures separate from the main table because it is easy to handle i..e, find the measures and edit it in future if required.
So we will make the dummy table. Enter data name it “tbl_measures”and name the columns as “Z_Dummy” by clicking on “New data”.

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]

Create a New Table Reference for mapping_AgeGrp:
The reference table is used to create the new table which will fetch the data which is present in the prod_Churn table and it doesn't hit the SQL server again and again to fetch the data, which reduces the time for processing. If we would have created the duplicate table than it would have hot the SQL server again and again which would have taken time

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 to Numbers

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 to Numbers

Create a new table reference for prod_Services

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

For editing the interaction between the charts you have to go to Format -> Edit interaction

This has two options, one is only filter and one is highlight.
Highlight is default but filter is good for the analysis.

STEP 4 - Power BI Visualization

Summary Page

1. Top Card
   a. Total Customers
   b. New Joiners
   c. Total Churn
   d. Churn Rate%
2. Demographic
   a. Gender – Churn Rate
   b. Age Group – Total Customer & Churn Rate
3. Account Info
   a. Payment Method – Churn Rate
   b. Contract – Churn Rate
   c. Tenure Group - Total Customer & Churn Rate
4. Geographic
   a. Top 5 State – Churn Rate
5. Churn Distribution

a. Churn Category – Total Churn
b. Tooltip : Churn Reason – Total Churn
6. Service Used
a. Internet Type – Churn Rate
b. prod_Service >> Services – Status – % RT Sum of Churn Status

Churn Reason Page (Tooltip)
1. Churn Reason – Total Churn

Now for the analysis of the chart we will take help of the AI features in the power BI.
The AI visuals give you a good starting point before you dig deeper into the data.
For doing that we check from the chart **narrative** tools, start with the insight it is giving and then you can start drilling deep into the insight you get.

You can check and after seeing the chart you can find out the various analyses and find out where the campaign should be focused such that the churn rate reduces. Don't depend on only the AI tools entirely

Conclusion:

-
-
-

Now we will introduce something new in this project.

Which will introduce you to the ML and build predictive models which will say which are more likely to churn in coming months. Or to Predict the future churners.

Here we will be using the Random forest.
For churn analysis, Random forest is mostly used.

STEP 5 – Predict Customer Churn
For predicting customer churn, we will be using a widely used Machine Learning algorithm called RANDOM FOREST.

What is Random Forest?A random forest is a machine learning algorithm that consists of multiple decision trees. Each decision tree is trained on a random subset of the data and features. The final prediction is made by averaging the predictions (in regression tasks) or taking the majority vote (in classification tasks) from all the trees in the forest. This ensemble approach improves the accuracy and robustness of the model by reducing the risk of overfitting compared to using a single decision tree.

Data Preparation for ML model
Let us first import views in an Excel file.

- Go to Data >> Get Data >> SQL Server Database
- Enter the Server Name & Database name to connect to SQL Server
- Import both vw_ChurnData & vw_JoinData
- Save the file as Prediction_Data

Note : But here it is important to improve the accuracy by using SMOTE **(synthetic minority oversampling technique)** to balance that data and also see which features are important in prediction and only **keeping those features in the prediction process**.

STEP 6 – Power BI Visualization of Predicted Data
Import CSV Data or Load Predicted data in SQL server & connect to server
Create Measures

Count Predicted Churner = COUNT(Predictions[Customer_ID]) + 0
Title Predicted Churners = "COUNT OF PREDICTED CHURNERS : " &
COUNT(Predictions[Customer_ID])

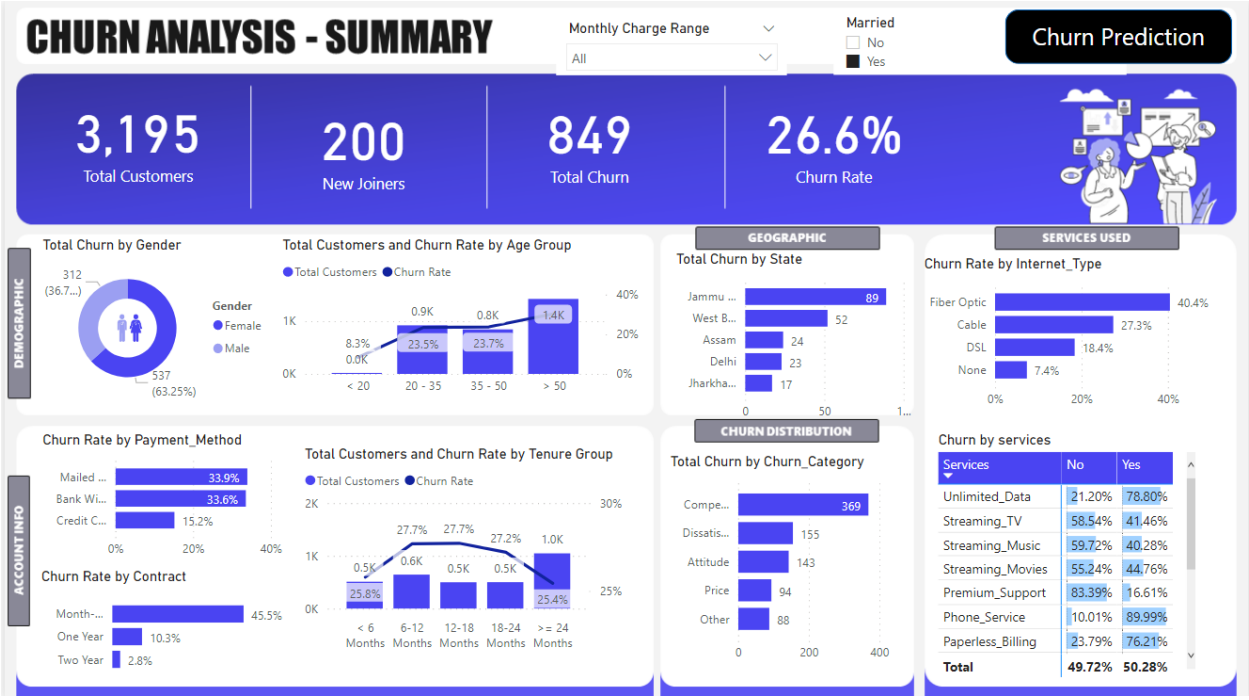
Churn Prediction Page (Using New Predicted Data)

1. **Right Side Grid**
   - Customer ID
   - Monthly Charge
   - Total Revenue
   - Total Refunds
   - Number of Referrals

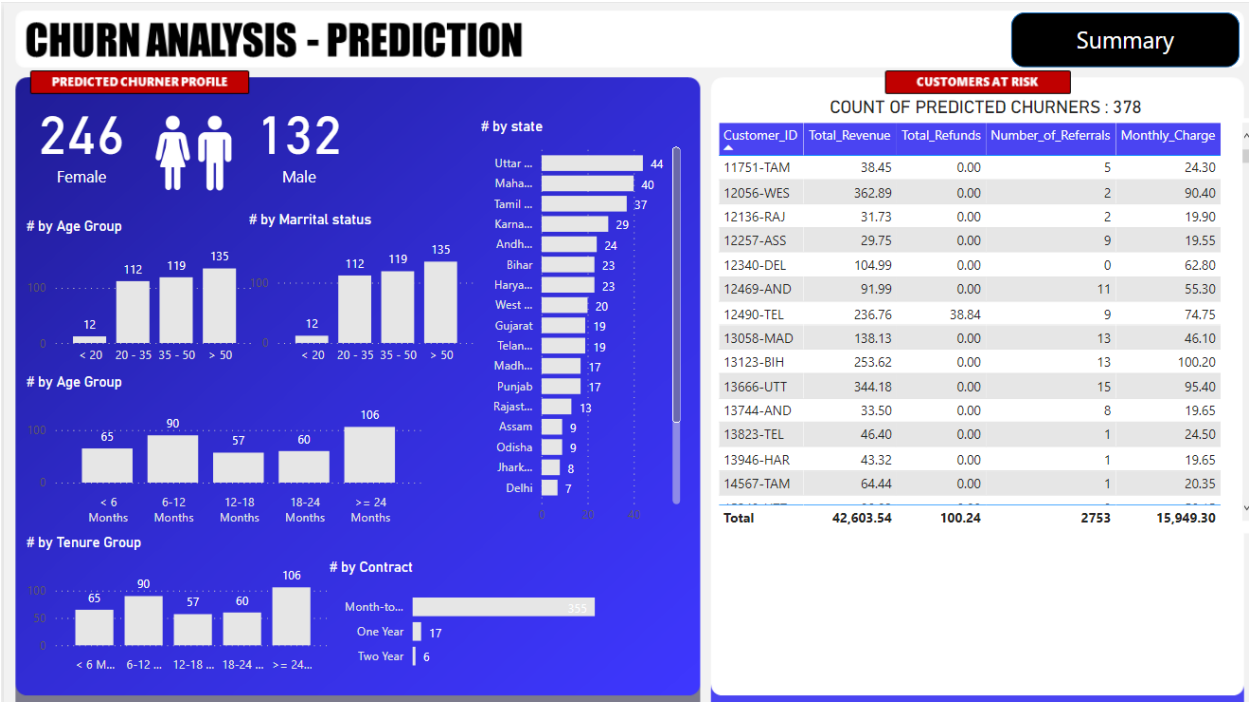
2. **Demographic**
   - Gender – Churn Count
   - Age Group – Churn Count
   - Marital Status – Churn Count
3. **Account Info**
   - Payment Method – Churn Count
   - Contract – Churn Count
   - Tenure Group - Churn Count
4. **Geographic**
   - State – Churn Count

Result

- Overall comprehensive Power BI dashboard with Executive Summary to analyze historical data.



- Churn Prediction page to predict future churners. Hope I was able to provide value with this content.



Reference:

https://youtu.be/QFDslca5AX8?si=KUg784U--IQT5LDI

https://www.pivotalstats.com/post/end-end-churn-analysis-portfolio-project