

End To End Churn Analysis Project

Agenda

ETL process
Challenges
SQL

Advanced data analytics

Advance visualization

Storytelling and insight

Executive summary dashboard

My recommendation

Introduction to Churn Analysis

In today's competitive business environment, retaining customers is crucial for long-term success. Churn analysis is a key technique used to understand and reduce this customer attrition. It involves examining customer data to identify patterns and reasons behind customer departures. By using advanced data analytics and machine learning, businesses can predict which customers are at risk of leaving and understand the factors driving their decisions. This knowledge allows companies to take proactive steps to improve customer satisfaction and loyalty.

Although **this project focuses on churn analysis for a telecom firm**, the techniques and insights are applicable across various industries. From retail and finance to healthcare and beyond, any business that values customer retention can benefit from churn analysis. We will explore the methods, tools, and best practices for reducing churn and improving customer loyalty, transforming data into actionable insights for sustained success.

Project Target

Create an 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

- **Metrics Required**
- Total Customers
- Total Churn & Churn Rate
- New Joiners

ETL process

(extract, transform, and load)

ETL Process in SQL Server

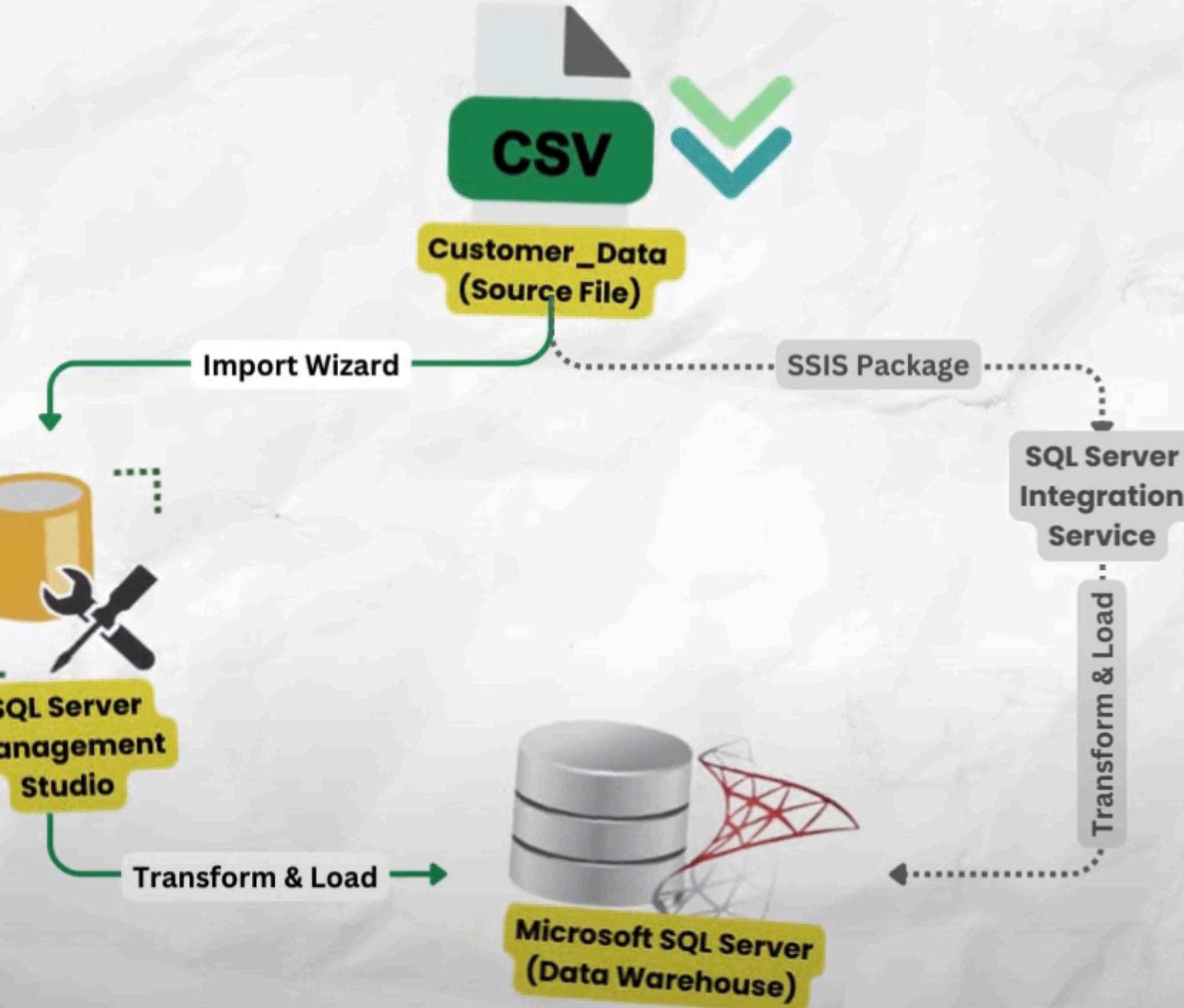
So the first step in churn analysis is to load the data from our source file. For this purpose I will be using Microsoft SQL server because it is a widely used solution across the industry and also because a full-fledged Database System is better at handling recurring data loads and maintaining data integrity compared to an excel file.

Download SSMS

In order for us to run our sql queries Microsoft provides us with GUI interface which is known as **SQL Server Management Studio**. You can download the latest version from the link provided below.

<https://learn.microsoft.com/en-us/sql/ssms/download-sql-server-management-studio-ssms?view=sql-server-ver16>

ETL FRAMEWORK



Our framework uses below components :

- **CSV file - This is our source file**
- **SQL Server Management Studio - We will use its import wizard to transform & load the data**
- **SQL Server Database - This is where our final data will be loaded and host our data warehouse, temporary tables & views for final usage.**

Alternative Framework :

- **CSV file - This is our source file**
- **SQL Server Integration Services - We can use SSIS to make SSIS package to transform & load data. It is usually used when you have recurring data loads and you want to automate the process**
- **SQL Server Database - Package loads our final data into the data warehouse in SQL Server**

Challenges

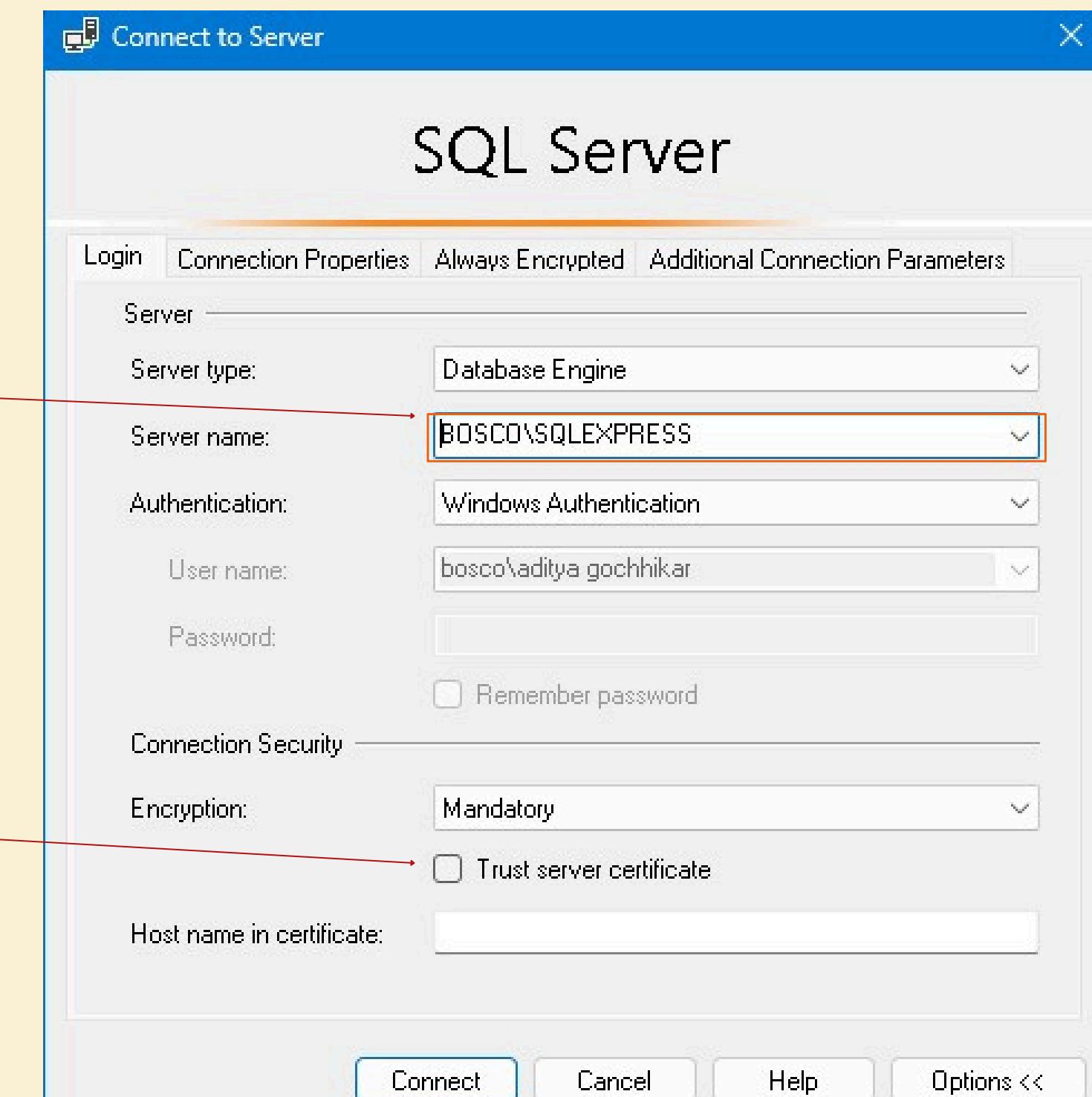
- After I download my SQL management studio (2019) I will face create a local server in my device so to tackle this problem follow this You Tube video

FIX: SSMS Server Name Not Showing [video link](#)

**After resolve all of the challenge here
I find my server name**

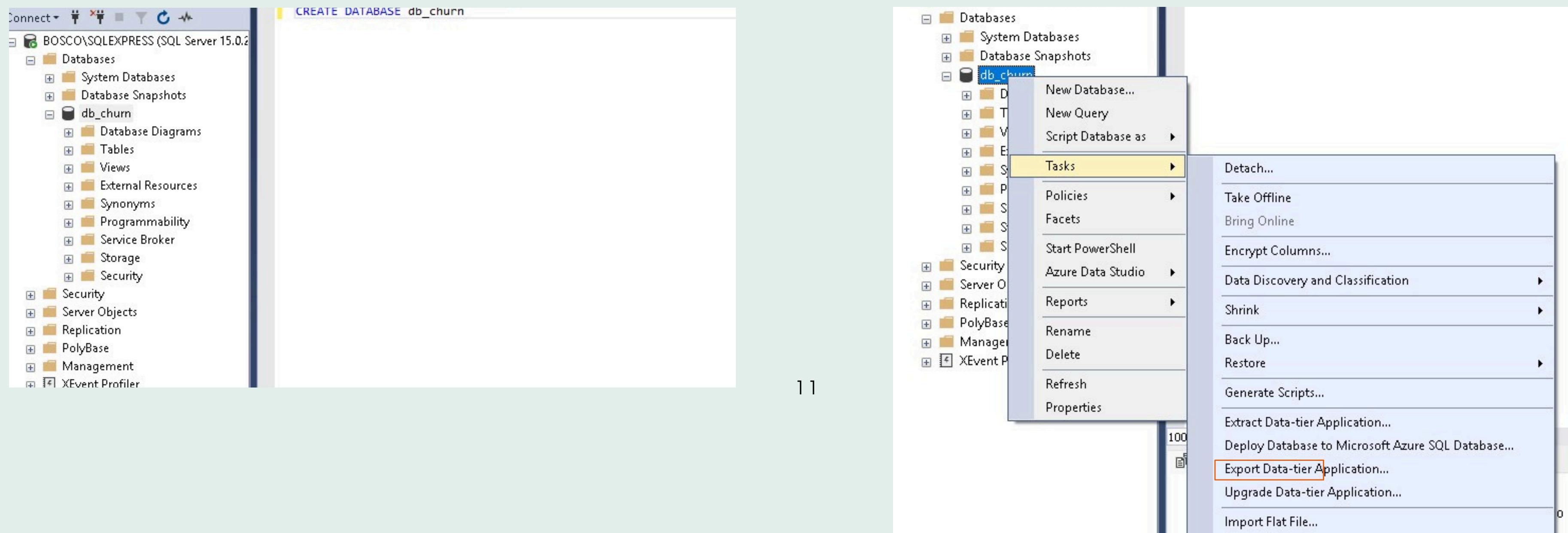
Allow this

10



Creating Database

After installation, I will land on the following screen. Then to copy paste the **server name** somewhere because we will need this at a later stage. Also enable the checkbox which says “**Trust Server Certificate**” and then click on Connect Once connected, click on NEW QUERY button at the top ribbon and then write below query. This will create a new Database named db_Churn



The screenshot shows two windows side-by-side. The left window is titled 'Import Flat File' and has a blue header bar with 'Introduction', 'Specify Input File', 'Preview Data', 'Modify Columns', 'Summary', and 'Results'. It displays a list of steps: 'Specify the input file containing the data.' and 'Preview the automatically generated table schema and optionally modify columns.' Below this is a diagram showing a document icon, a gear icon, and an SQL database icon connected by a flow. A red arrow points from the 'Specify Input File' step in the first window to the second window. The right window is titled 'Specify Input File' and has a blue header bar with 'Introduction', 'Specify Input File', 'Preview Data', 'Modify Columns', 'Summary', and 'Results'. It contains fields for 'Location of file to be imported' (C:\Users\aditya gochhikar\OneDrive\Desktop\Customer_Data.csv), 'New table name' (Customer_Data), and 'Table schema' (dbo). A red box highlights the 'Browse...' button next to the file path field.

After clicking next then I browse data set in your system

The screenshot shows two windows of the 'Import Flat File 'db_churn'' wizard. The left window is titled 'Modify Columns' and displays a table schema for 20 columns. The right window is titled 'Summary' and shows import information. A large red arrow points from the 'Modify Columns' window to the 'Summary' window.

Modify Columns

This operation generated the following table schema. Please verify if schema is accurate, and if not, please make any changes.

Column Name	Data Type	Primary Key	Allow Nulls
Device_Protection_Plan	varchar(50)	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Premium_Support	varchar(50)	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Streaming_TV	varchar(50)	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Streaming_Movies	varchar(50)	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Streaming_Music	varchar(50)	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Unlimited_Data	varchar(50)	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Contract	nvarchar(50)	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Paperless_Billing	varchar(50)	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Payment_Method	nvarchar(50)	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Monthly_Charge	float	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Total_Charges	float	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Total_Refunds	float	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Total_Extra_Data_Charges	tinyint	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Total_Long_Distance_Charges	float	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Total_Revenue	float	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Customer_Status	nvarchar(50)	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Churn_Category	nvarchar(50)	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Churn_Reason	nvarchar(50)	<input type="checkbox"/>	<input checked="" type="checkbox"/>

Row granularity of error reporting (performance impact with smaller ranges)

< Previous Cancel

Summary

To complete the operation using the specified inputs, click Finish.

Import Information

- Name: BOSCO\SQLEXPRESS
- Database Name: db_churn
- Table Name: dbo.stg_Churn
- File to be imported: C:\Users\aditya gochhikar\OneDrive\Desktop\Customer_Data.csv

< Previous Cancel

Remember to add customerId as primary key and allow nulls for all remaining columns. This is done to avoid any errors while data load. Also make sure to change the datatype where it say **Bit** to **Varchar(50)**. I doing this because while using import wizard I faced issues with the BIT data type, however Varchar(50) works fine.

SQL

Data Exploration – Check Distinct Values

```
2 Count(Gender) * 100.0 / (select count(*) from stg_Churn) as percentage
3   from stg_Churn
4 group by Gender
5
6
7 8 SELECT Contract, Count(Contract) as TotalCount,
8   Count(Contract) * 1.0 / (Select Count(*) from stg_Churn) as Percentage
9   from stg_Churn
10 Group by Contract
11
12
13 SELECT Customer_Status, Count(Customer_Status) as TotalCount, Sum(Total_Revenue) as TotalRev,
14   Sum(Total_Revenue) / (Select sum(Total_Revenue) from stg_Churn) * 100 as RevPercentage
15   from stg_Churn
16 Group by Customer_Status
17
18
19 SELECT State, Count(State) as TotalCount,
20   Count(State) * 100.0 / (Select Count(*) from stg_Churn) as Percentage
21   from stg_Churn
22 Group by State
23 Order by Percentage desc
24
```

Data Exploration – Check Nulls

```
2     SUM(CASE WHEN Customer_ID IS NULL THEN 1 ELSE 0 END) AS Customer_ID_Null_Count,  
3  
4     SUM(CASE WHEN Gender IS NULL THEN 1 ELSE 0 END) AS Gender_Null_Count,  
5  
6     SUM(CASE WHEN Age IS NULL THEN 1 ELSE 0 END) AS Age_Null_Count,  
7  
8     SUM(CASE WHEN Married IS NULL THEN 1 ELSE 0 END) AS Married_Null_Count,  
9  
10    SUM(CASE WHEN State IS NULL THEN 1 ELSE 0 END) AS State_Null_Count,  
11  
12    SUM(CASE WHEN Number_of_Referrals IS NULL THEN 1 ELSE 0 END) AS Number_of_Referrals_Null_Count,  
13  
14    SUM(CASE WHEN Tenure_in_Months IS NULL THEN 1 ELSE 0 END) AS Tenure_in_Months_Null_Count,  
15  
16    SUM(CASE WHEN Value_Deal IS NULL THEN 1 ELSE 0 END) AS Value_Deal_Null_Count,  
17  
18    SUM(CASE WHEN Phone_Service IS NULL THEN 1 ELSE 0 END) AS Phone_Service_Null_Count,  
19  
20    SUM(CASE WHEN Multiple_Lines IS NULL THEN 1 ELSE 0 END) AS Multiple_Lines_Null_Count,  
21  
22    SUM(CASE WHEN Internet_Service IS NULL THEN 1 ELSE 0 END) AS Internet_Service_Null_Count,  
23  
24    SUM(CASE WHEN Internet_Type IS NULL THEN 1 ELSE 0 END) AS Internet_Type_Null_Count,  
25  
26    SUM(CASE WHEN Online_Security IS NULL THEN 1 ELSE 0 END) AS Online_Security_Null_Count,  
27  
28    SUM(CASE WHEN Online_Backup IS NULL THEN 1 ELSE 0 END) AS Online_Backup_Null_Count,  
29  
30    SUM(CASE WHEN Device_Protection_Plan IS NULL THEN 1 ELSE 0 END) AS Device_Protection_Plan_Null_Co  
31  
32    SUM(CASE WHEN Premium_Support IS NULL THEN 1 ELSE 0 END) AS Premium_Support_Null_Count,  
33  
34    SUM(CASE WHEN Streaming_TV IS NULL THEN 1 ELSE 0 END) AS Streaming_TV_Null_Count,  
35  
36    SUM(CASE WHEN Streaming_Movies IS NULL THEN 1 ELSE 0 END) AS Streaming_Movies_Null_Count,
```

WHEN Contract IS NULL THEN 1 ELSE 0 END) AS Contract_Null_Count,
WHEN Paperless_Billing IS NULL THEN 1 ELSE 0 END) AS Paperless_Billing_Null_Count,
WHEN Payment_Method IS NULL THEN 1 ELSE 0 END) AS Payment_Method_Null_Count,
WHEN Monthly_Charge IS NULL THEN 1 ELSE 0 END) AS Monthly_Charge_Null_Count,
WHEN Total_Charges IS NULL THEN 1 ELSE 0 END) AS Total_Charges_Null_Count,
WHEN Total_Refunds IS NULL THEN 1 ELSE 0 END) AS Total_Refunds_Null_Count,
WHEN Total_Extra_Data_Charges IS NULL THEN 1 ELSE 0 END) AS Total_Extra_Data_Charges_N
WHEN Total_Long_Distance_Charges IS NULL THEN 1 ELSE 0 END) AS Total_Long_Distance_Ch
WHEN Total_Revenue IS NULL THEN 1 ELSE 0 END) AS Total_Revenue_Null_Count,
WHEN Customer_Status IS NULL THEN 1 ELSE 0 END) AS Customer_Status_Null_Count,
WHEN Churn_Category IS NULL THEN 1 ELSE 0 END) AS Churn_Category_Null_Count,
WHEN Churn_Reason IS NULL THEN 1 ELSE 0 END) AS Churn_Reason_Null_Count
};

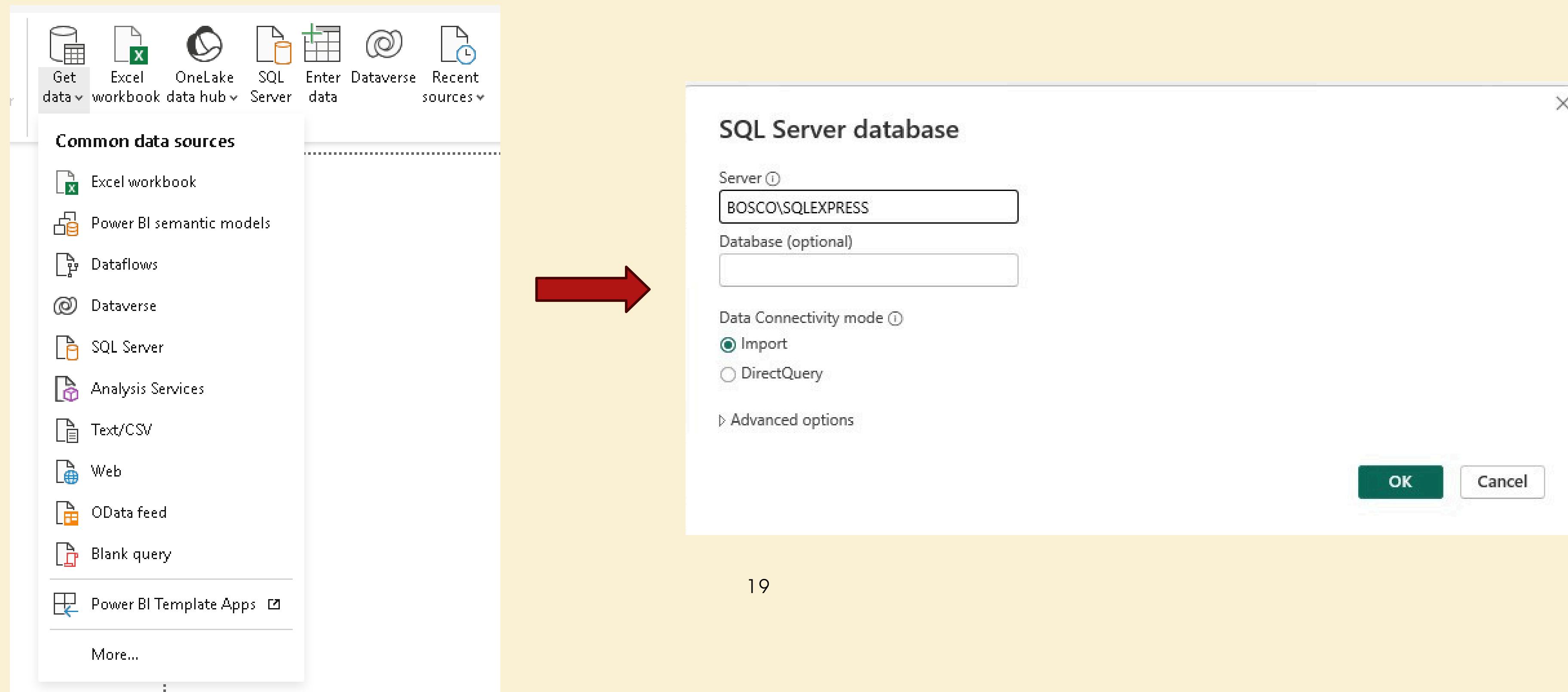
Remove null and insert the new data into Prod table

```
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(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]
```

Power BI Transform

Connect my data into power BI



Add a new column in prod_Churn

DAX

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**

DAX

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

DAX

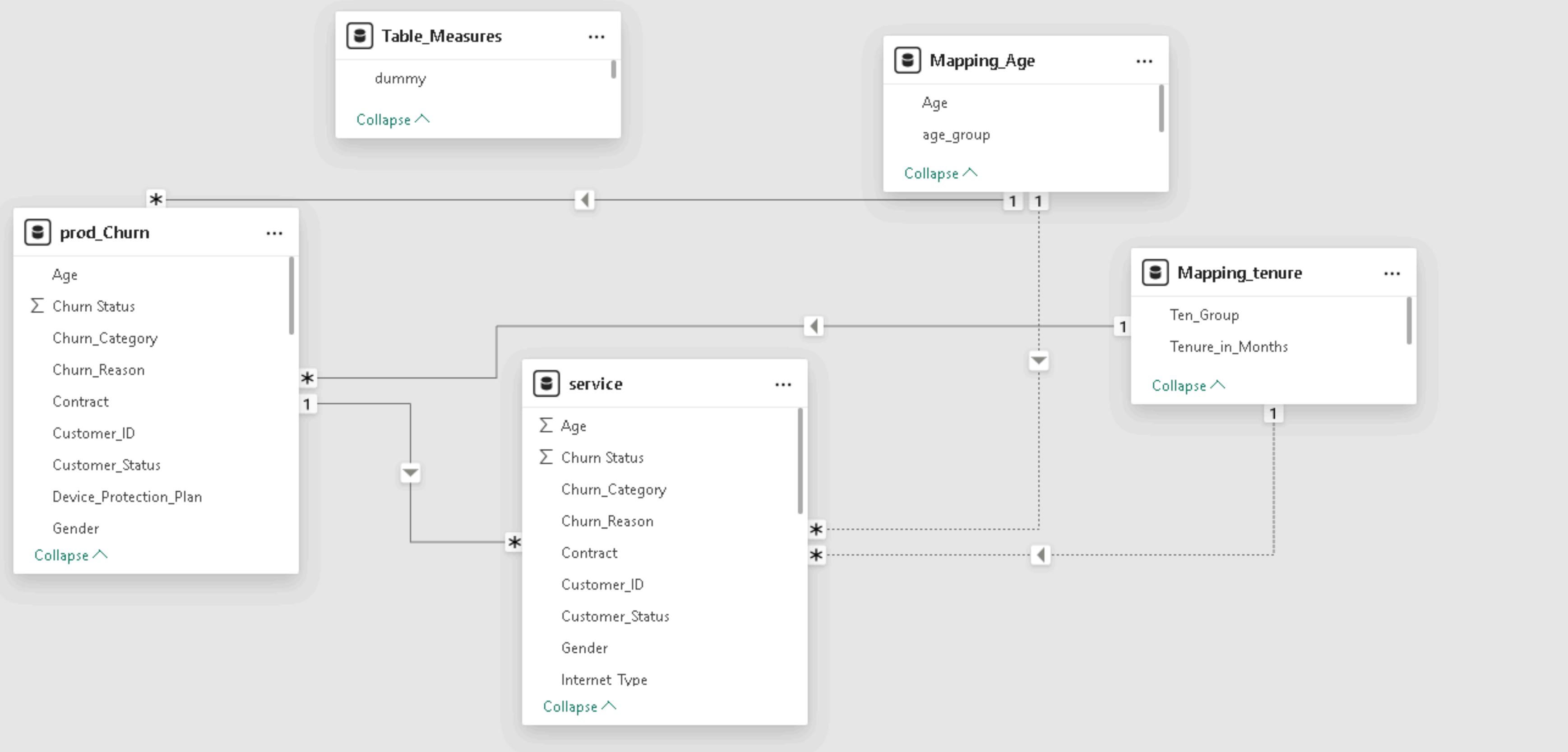
- 1 Keep only Tenure_in_Months and remove duplicates
2. Tenure Group = if [Tenure_in_Months] < 6 then "< 6 Months" else if [Tenure_in_Months] < 12 then "6-12 Months" else if [Tenure_in_Months] < 18 then "12-18 Months" else if [Tenure_in_Months] < 24 then "18-24 Months" else ">= 24 Months"
- 3.TenureGrpSorting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

Unpivot services columns

- . Rename Column – Attribute >> Services & Value >> Status



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

SUMMARY PAGE BLUEPRINT

2,907

TOTAL CUSTOMERS

321

NEW JOINERS

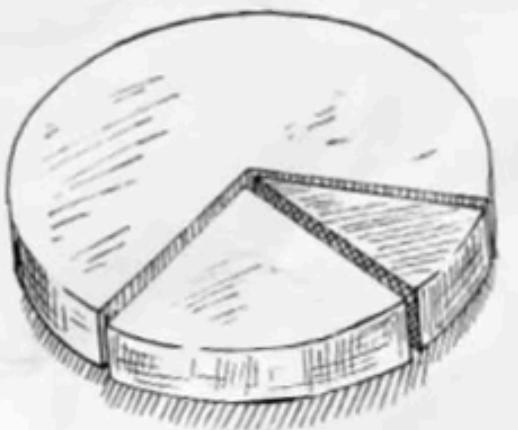
1,189

TOTAL CHURN

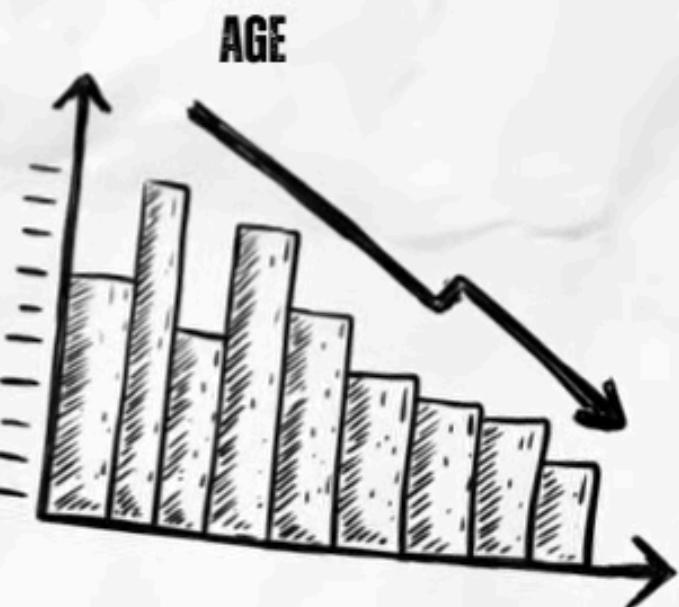
24.0%

CHURN RATE

GENDER



AGE



GEOGRAPHIC



SERVICES



PAYMENT METHOD & CONTRACT



TENURE



CHURN DISTRIBUTION



Churn Analysis Executive summary

4,048

Total customers

269

New Joiners

1,111

total Churn

27.45%

Churn rate

Monthly Charge Status

All

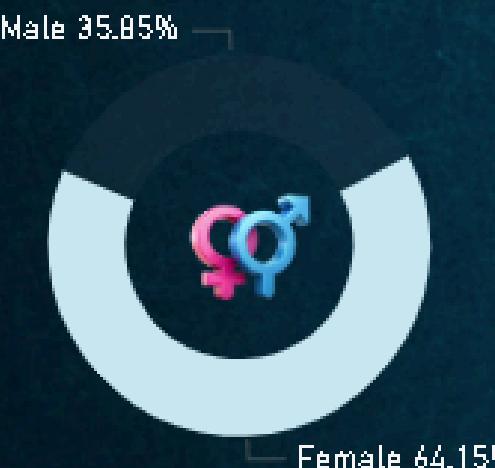
Married

No

Yes

DEMOGRAPHIC

total Churn by Gender



Churn rate by Payment_Method

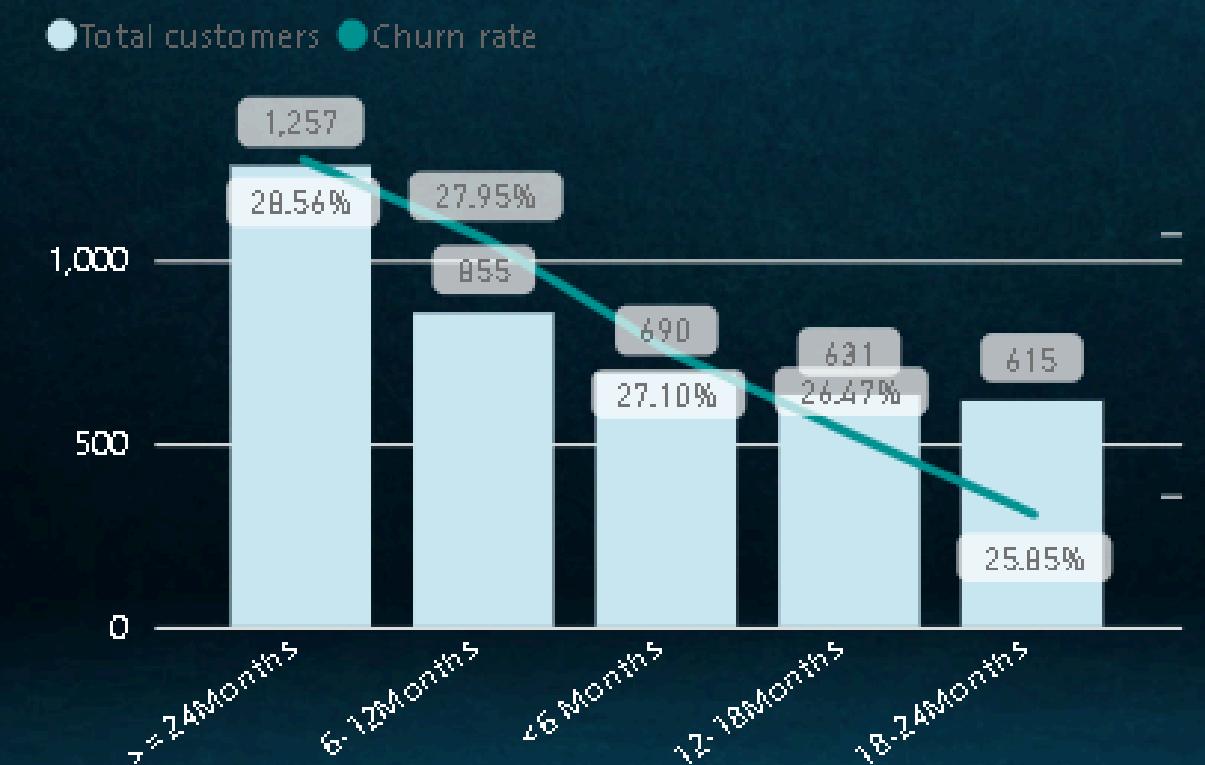


Total customers and Churn rate by age_group



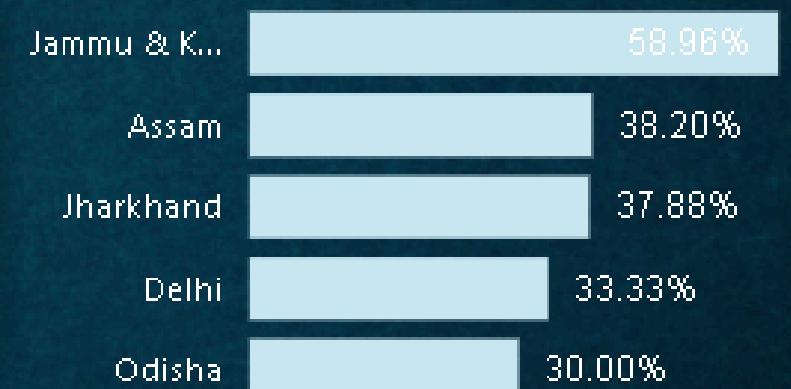
ACCOUNT INFO

Total customers and Churn rate by Ten_Group



GEOGRAPHIC

Churn rate by State(Top 5)



CHURN DISTRIBUTION

total Churn by Churn_Category



Churn by services

service	No	Yes
Device_Protection_Plan	69.49%	30.51%
Internet_Service	6.39%	93.61%
Married	51.67%	48.33%
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%

thank you

Aditya gochhikar