Customer Churn Analysis

Technical Requirement Document

**Content**

|  |
| --- |
| 1. [Introduction 03](#_heading=h.gjdgxs) |
| [1.1Purpose of this document 03](#_heading=h.30j0zll) |
| 1. [Purpose of this Project 03](#_heading=h.1fob9te) |
| 1. [Scope 03](#_heading=h.3znysh7) |
| 1. [Data Preparation 03](#_heading=h.2et92p0) |
| 1. [Project Execution Procedure 11](#_heading=h.tyjcwt) |
| 1. [Flow Chart of Preparation of Data 12](#_heading=h.3dy6vkm) |
| 1. [Data Uploading Issue 13](#_heading=h.1t3h5sf) |
| 1. [Data Validation 13](#_heading=h.2s8eyo1) |
| 1. [Power BI Desktop 13](#_heading=h.17dp8vu) |
| 1. [Data preparation and Modeling 15](#_heading=h.3rdcrjn) |
| 1. [Business Requirements 18](#_heading=h.26in1rg) |
| 1. [Required Calculation 20](#_heading=h.lnxbz9) |
| 1. [Report Building 22](#_heading=h.35nkun2) |

## Introduction

## Purpose of this document

The objective of this documents is to describe the Customer churn, or attrition, is the rate at which clients opt out of purchasing more of a company’s products or services. Customer churn analysis is a method of measuring this rate.

## Purpose of this Project

Churn analysis simply tells you what percentage of your customers don’t return compared with the percentage who conduct repeat business. By digging deeper into these numbers, you may be able to identify trends that can avert failure or take an already successful product or service to the next level.

## Scope

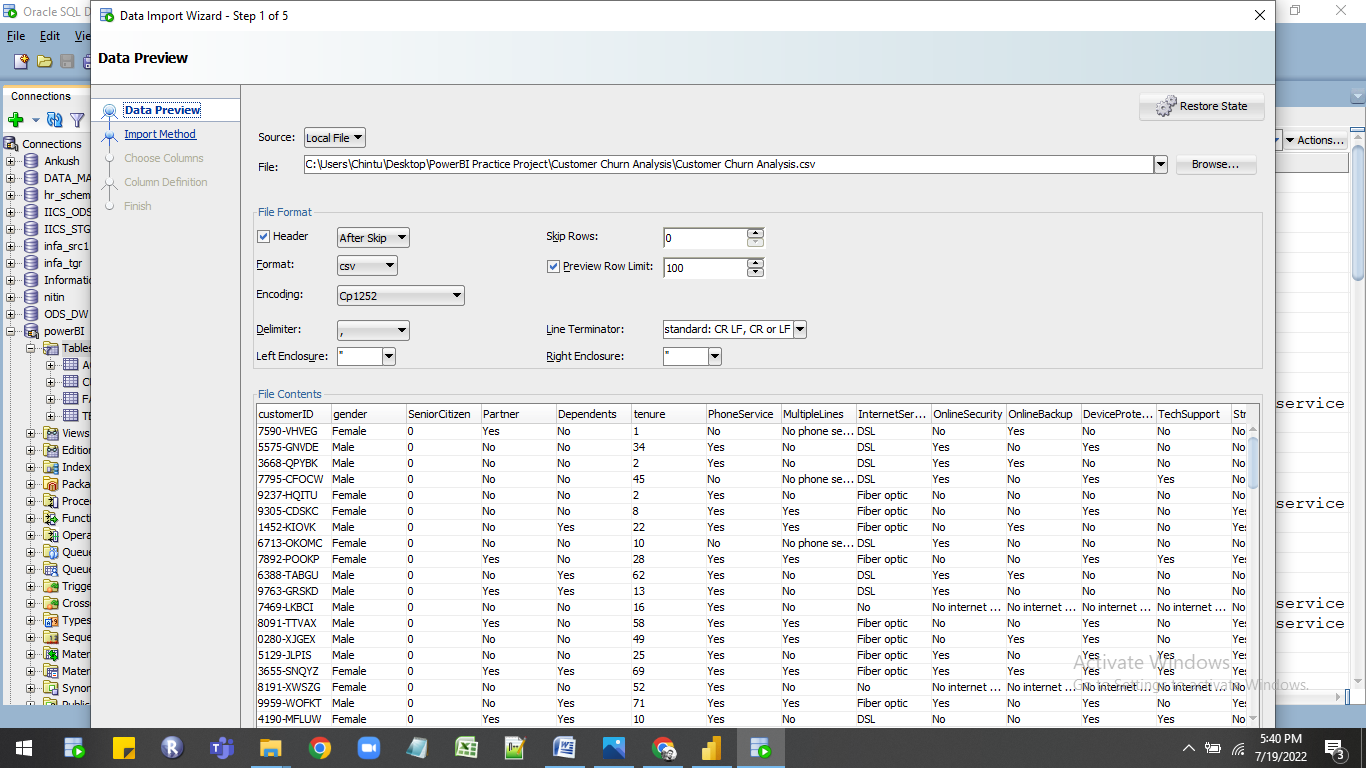
* Churn analysis simply tells you what percentage of your customers don’t return compared with the percentage who conduct repeat business. By digging deeper into these numbers, you may be able to identify trends that can avert failure or take an already successful product or service to the next level.
* Methods to measure customer churn include calculating this KPI over various timeframes and trending those results; high-performing firms also measure the financial results of customers leaving and then benchmark those numbers against key performance indicators (KPIs) critical to the business’s profitability.

## Data Preparation

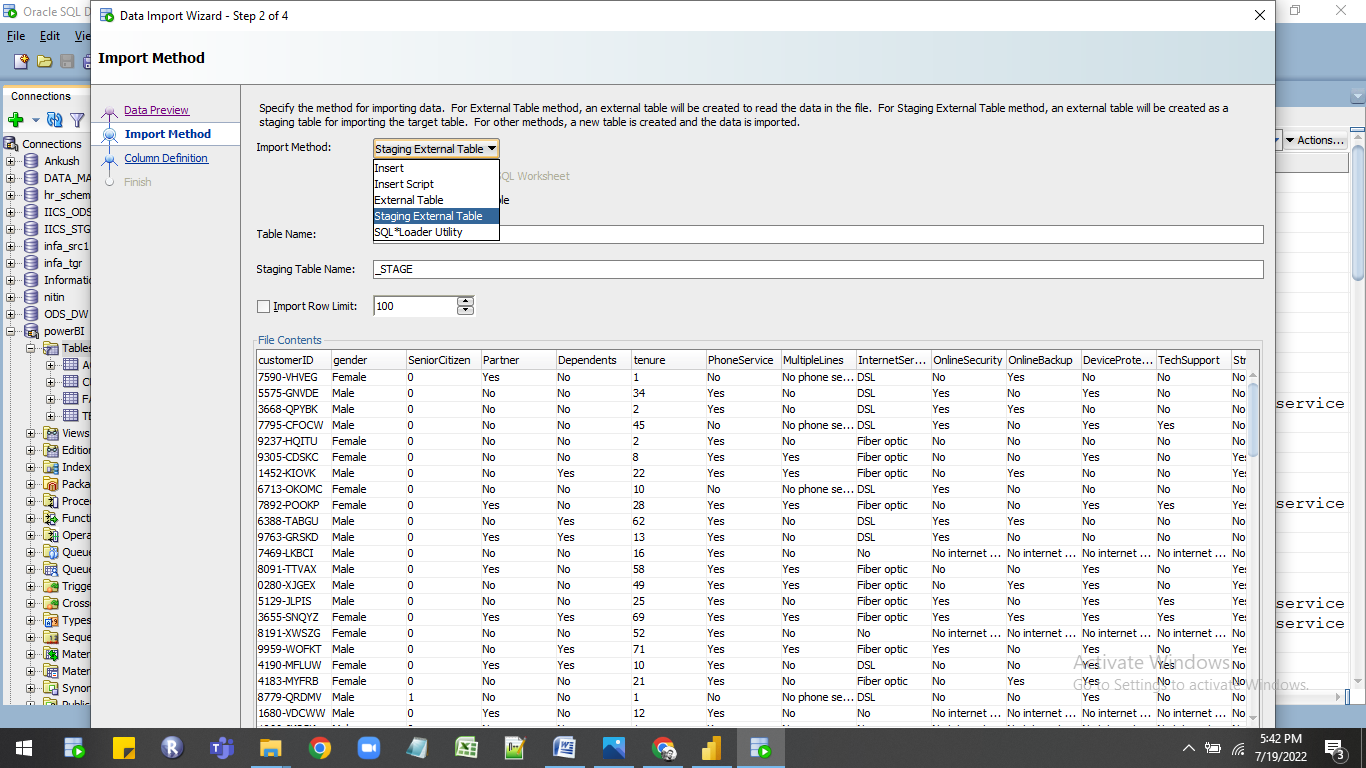
Data given by client that contains following information Data fact table: (excel file)

* Customer id
* gender
* senior citizen
* partner
* dependents
* tenure
* phone service
* multiple lines
* internet service
* online security
* online backup
* device protection
* tech support
* streaming tv
* streaming movies
* contract
* paperless billing
* payment method
* monthly charge
* total charge
* no of admin tickets
* no of tech tickets
* churn
* Load the shared  data files into oracle database as it is.
* Decide the data type for SQL table creation from the shared file which data type should be there to load data into final table with appropriate data type.I have imported all the tables in oracle 11g database.

1. Create schema connection in oracle for the project.
2. Convert excel file type to .Csv format.
3. Load the shared data file into oracle database with proper datatype for each column.



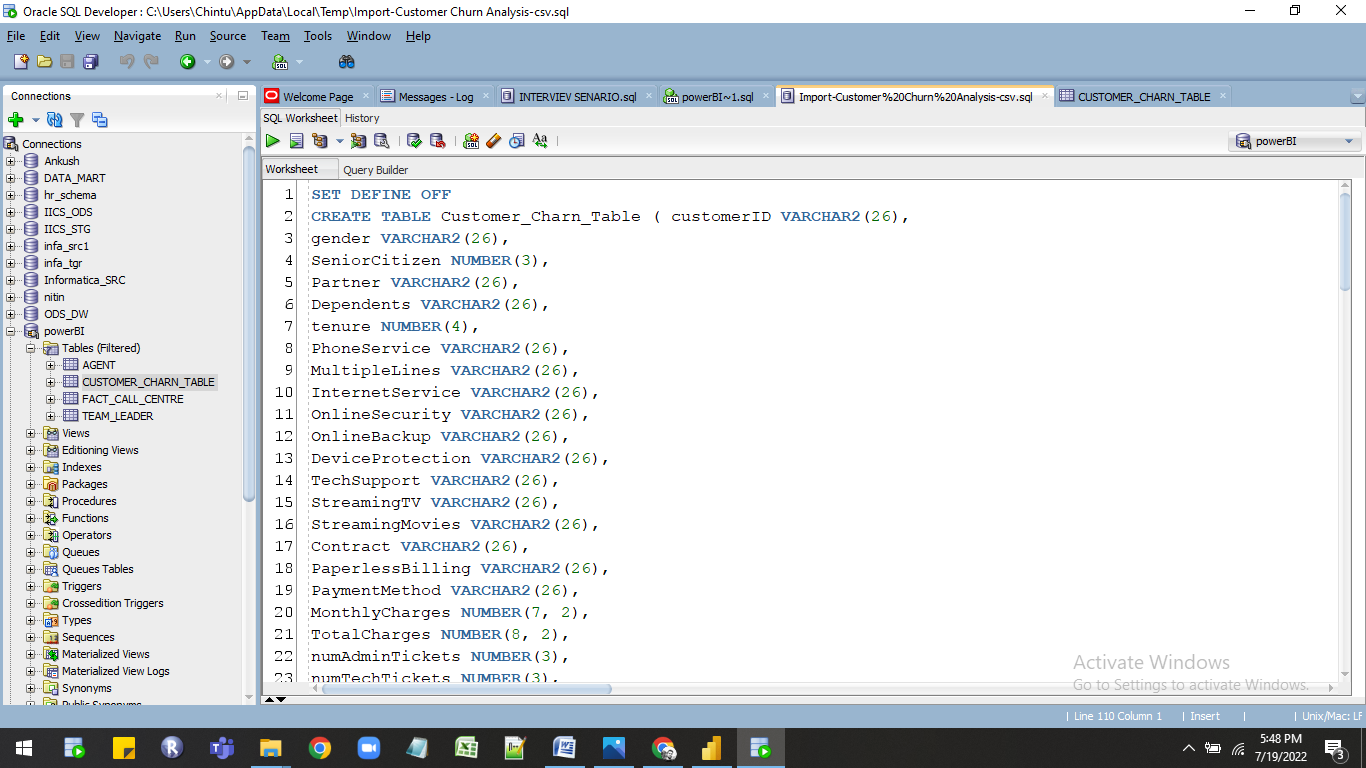
1. First load data into staging table as it is and then load it into final table with right data type.



1. At the data table importing time check all the column datatypes properly.
2. In data there is decimal numbers in that column so for that column use the number datatype

with proper scale and precision.

1. Save the data in oracle which used in power bi desktop for creating  interactive report.
2. Query used in oracle.



SET DEFINE OFF

CREATE TABLE Customer\_Charn\_Table ( customerID VARCHAR2(26),

gender VARCHAR2(26),

SeniorCitizen NUMBER(3),

Partner VARCHAR2(26),

Dependents VARCHAR2(26),

tenure NUMBER(4),

PhoneService VARCHAR2(26),

MultipleLines VARCHAR2(26),

InternetService VARCHAR2(26),

OnlineSecurity VARCHAR2(26),

OnlineBackup VARCHAR2(26),

DeviceProtection VARCHAR2(26),

TechSupport VARCHAR2(26),

StreamingTV VARCHAR2(26),

StreamingMovies VARCHAR2(26),

Contract VARCHAR2(26),

PaperlessBilling VARCHAR2(26),

PaymentMethod VARCHAR2(26),

MonthlyCharges NUMBER(7, 2),

TotalCharges NUMBER(8, 2),

numAdminTickets NUMBER(3),

numTechTickets NUMBER(3),

Churn VARCHAR2(26));

--CREATE OR REPLACE DIRECTORY ORACLECLRDIR AS 'C:\Users\Chintu\Desktop\PowerBI Practice Project\Customer Churn Analysis';

--GRANT READ ON DIRECTORY ORACLECLRDIR TO USER;

--GRANT WRITE ON DIRECTORY ORACLECLRDIR TO USER;

--drop table Customer\_Charn\_Table\_STAGE;

CREATE TABLE Customer\_Charn\_Table\_STAGE

( customerID VARCHAR2(26),

  gender VARCHAR2(26),

  SeniorCitizen NUMBER(3),

  Partner VARCHAR2(26),

  Dependents VARCHAR2(26),

  tenure NUMBER(4),

  PhoneService VARCHAR2(26),

  MultipleLines VARCHAR2(26),

  InternetService VARCHAR2(26),

  OnlineSecurity VARCHAR2(26),

  OnlineBackup VARCHAR2(26),

  DeviceProtection VARCHAR2(26),

  TechSupport VARCHAR2(26),

  StreamingTV VARCHAR2(26),

  StreamingMovies VARCHAR2(26),

  Contract VARCHAR2(26),

  PaperlessBilling VARCHAR2(26),

  PaymentMethod VARCHAR2(26),

  MonthlyCharges NUMBER(7, 2),

  TotalCharges NUMBER(8, 2),

  numAdminTickets NUMBER(3),

  numTechTickets NUMBER(3),

  Churn VARCHAR2(26))

ORGANIZATION EXTERNAL

  (  TYPE ORACLE\_LOADER

     DEFAULT DIRECTORY ORACLECLRDIR

     ACCESS PARAMETERS

       (records delimited BY '\r\n'

           NOBADFILE

           NODISCARDFILE

           NOLOGFILE

           skip 1

           fields terminated BY ','

           OPTIONALLY ENCLOSED BY '"' AND '"'

           lrtrim

           missing field VALUES are NULL

           ( customerID CHAR(4000),

             gender CHAR(4000),

             SeniorCitizen CHAR(4000),

             Partner CHAR(4000),

             Dependents CHAR(4000),

             tenure CHAR(4000),

             PhoneService CHAR(4000),

             MultipleLines CHAR(4000),

             InternetService CHAR(4000),

             OnlineSecurity CHAR(4000),

             OnlineBackup CHAR(4000),

             DeviceProtection CHAR(4000),

             TechSupport CHAR(4000),

             StreamingTV CHAR(4000),

             StreamingMovies CHAR(4000),

             Contract CHAR(4000),

             PaperlessBilling CHAR(4000),

             PaymentMethod CHAR(4000),

             MonthlyCharges CHAR(4000),

             TotalCharges CHAR(4000),

             numAdminTickets CHAR(4000),

             numTechTickets CHAR(4000),

             Churn CHAR(4000)

           )

       )

     LOCATION ('Customer Churn Analysis.csv')

  )

  REJECT LIMIT UNLIMITED;

select \* from Customer\_Charn\_Table\_STAGE WHERE ROWNUM <= 100;

whenever sqlerror exit rollback;

begin

  INSERT INTO Customer\_Charn\_Table (customerID, gender, SeniorCitizen, Partner, Dependents, tenure, PhoneService, MultipleLines, InternetService, OnlineSecurity, OnlineBackup, DeviceProtection, TechSupport, StreamingTV, StreamingMovies, Contract, PaperlessBilling, PaymentMethod, MonthlyCharges, TotalCharges, numAdminTickets, numTechTickets, Churn)

  SELECT customerID, gender, SeniorCitizen, Partner, Dependents, tenure, PhoneService, MultipleLines, InternetService, OnlineSecurity, OnlineBackup, DeviceProtection, TechSupport, StreamingTV, StreamingMovies, Contract, PaperlessBilling, PaymentMethod, MonthlyCharges, TotalCharges, numAdminTickets, numTechTickets, Churn FROM Customer\_Charn\_Table\_STAGE ;

  COMMIT;

  EXECUTE IMMEDIATE 'DROP TABLE Customer\_Charn\_Table\_STAGE';

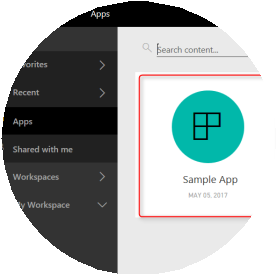
end;

/

## Project Execution Procedure

## Flow Chart of Preparation of Data





Importing Data from data source

Creation of Fact Table,Dimension Table,

Performing joins/Union all/create new table if required

Data cleaning ,Creation of Model,Dax

Calculation , Visualization like charts

,graph etc,

Publishing Report in to Power Bi Service

Creation of Apps and share to end user

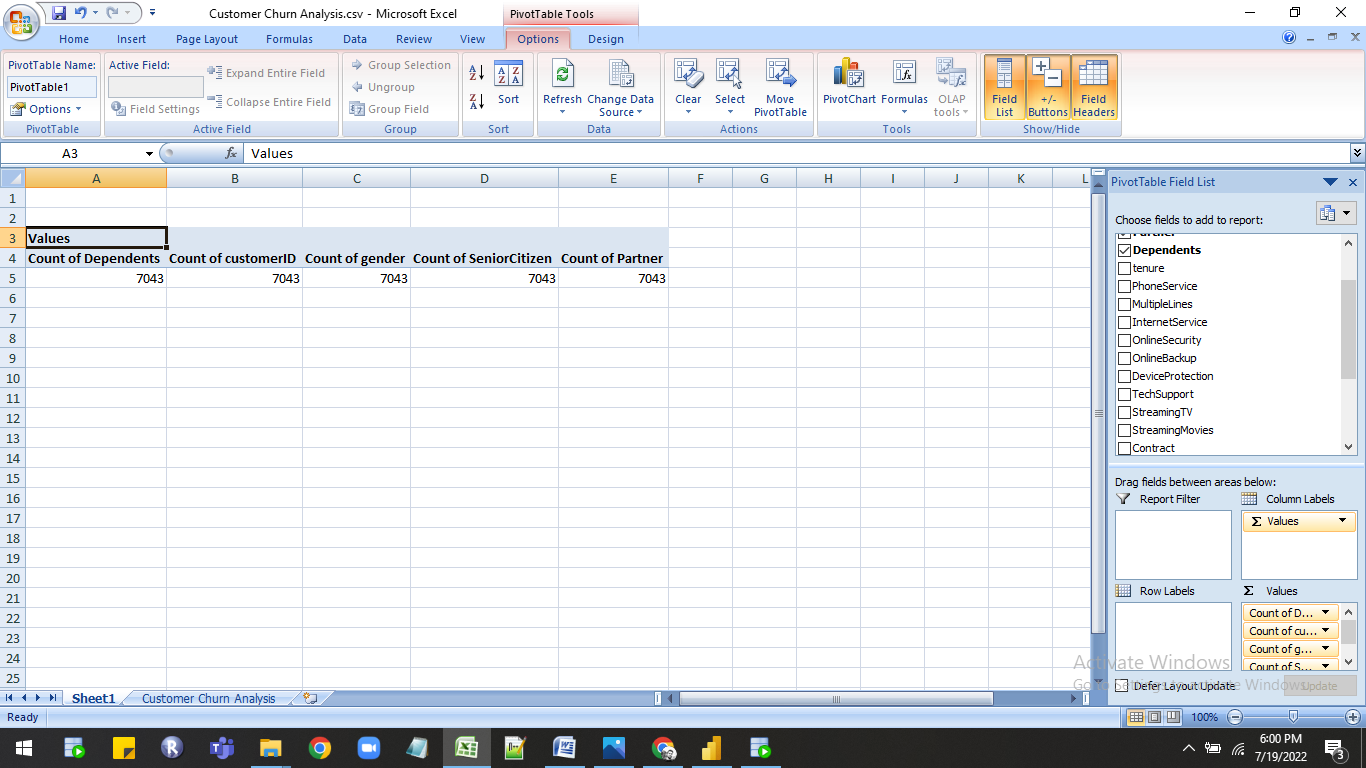
## Data Uploading Issue

Convert excel file type to .Csv format.

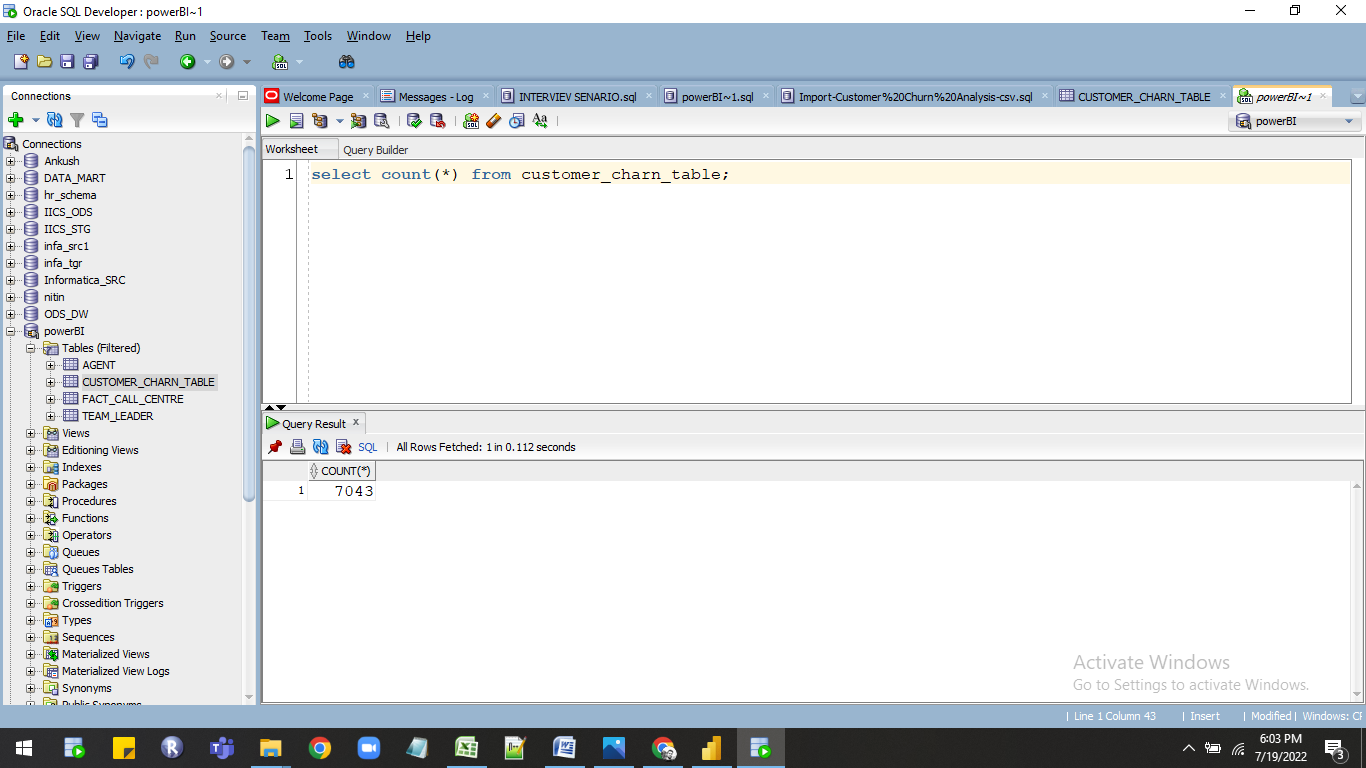
Load the shared data file into oracle database with proper datatype for each column.

## Data Validation

* Data validation is **the practice of checking the integrity, accuracy and structure of data before it is used for a business operation**. Data validation operation results can provide data used for data analytics, business intelligence or training a machine learning model.
* Validate the data in excel in pivot table and in oracle using aggregate functions.
* Count of records in excel pivot table is 7043.



* Count of records in oracle is 7043 used aggregate function count.



* Table showing fields which are validated in both excel and oracle.

|  |  |  |  |
| --- | --- | --- | --- |
| sr.no | Fields | Excel records | Oracle records |
| 1. | Count CustomerID | 7043 | 7043 |
| 2. | Count Gender | 7043 | 7043 |
| 3. | Count Dependant | 7043 | 7043 |
| 4. | Count Partner | 7043 | 7043 |

* Quary used in oracle for data validation.

select \* from customer\_charn\_table;

select count(CUSTOMERID) from customer\_charn\_table;

select count(Gender) from customer\_charn\_table;

select count(PARTNER) from customer\_charn\_table;

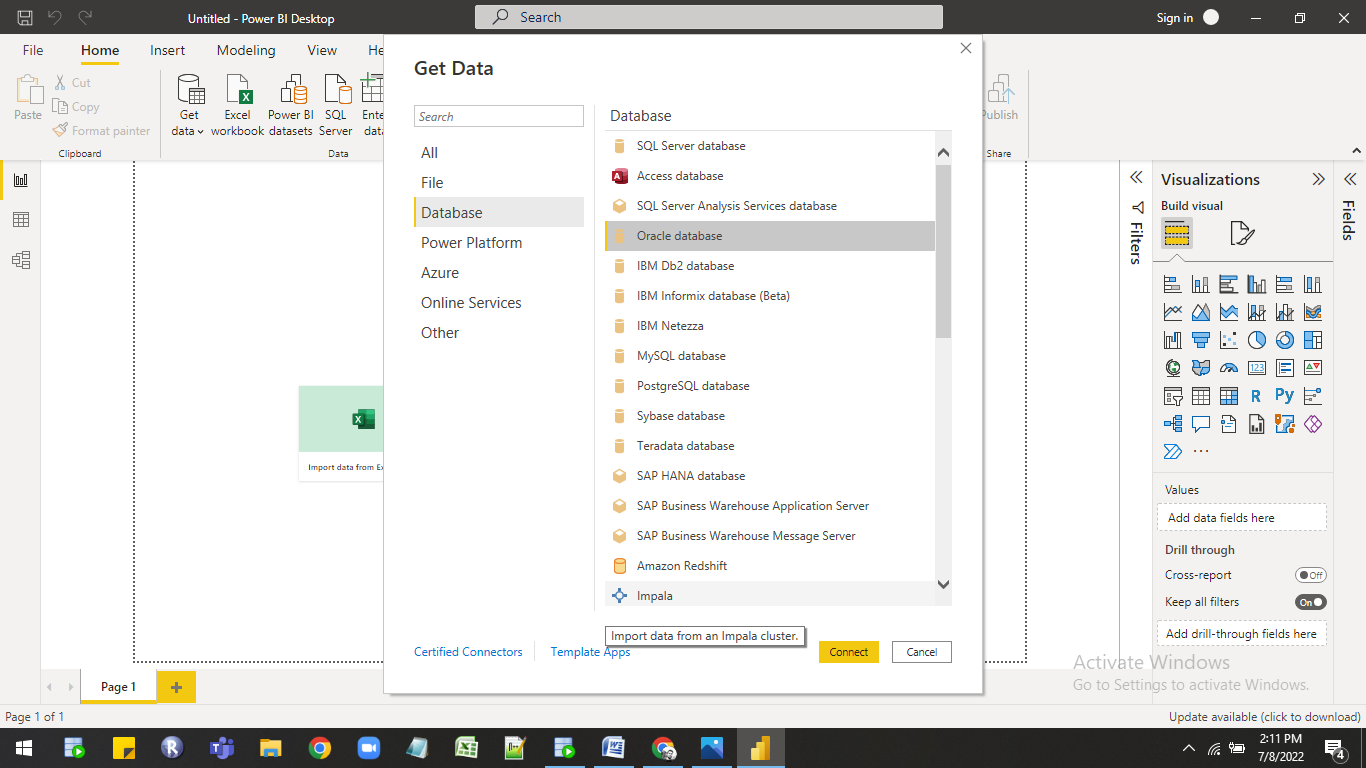
select count(DEPENDENTS) from customer\_charn\_table;

select count(SENIORCITIZEN) from customer\_charn\_table;

## Data preparation and Modeling

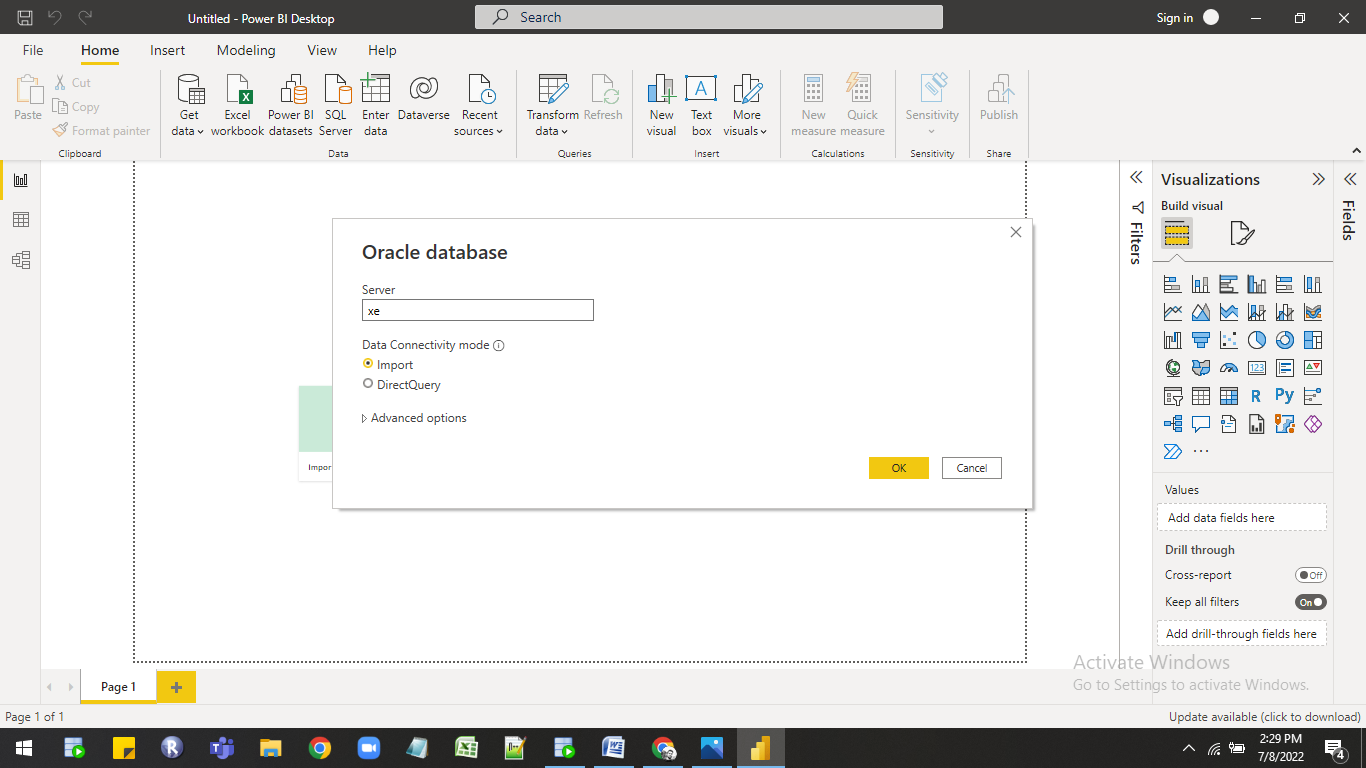
**Steps :-**

1. Import the data in power BI desktop for importing data first you have to connect to oracle database.
2. To connecting data click on GET DATA option then select the DATABASE option after that select oracle database option then click on connect.



1. Then new window is open in that you have to give server name for e.g Localhost and then you have to select data connectivity mode there are two modes

* import :- file size 1gb limit.
* Direct query:- No limit for file size.



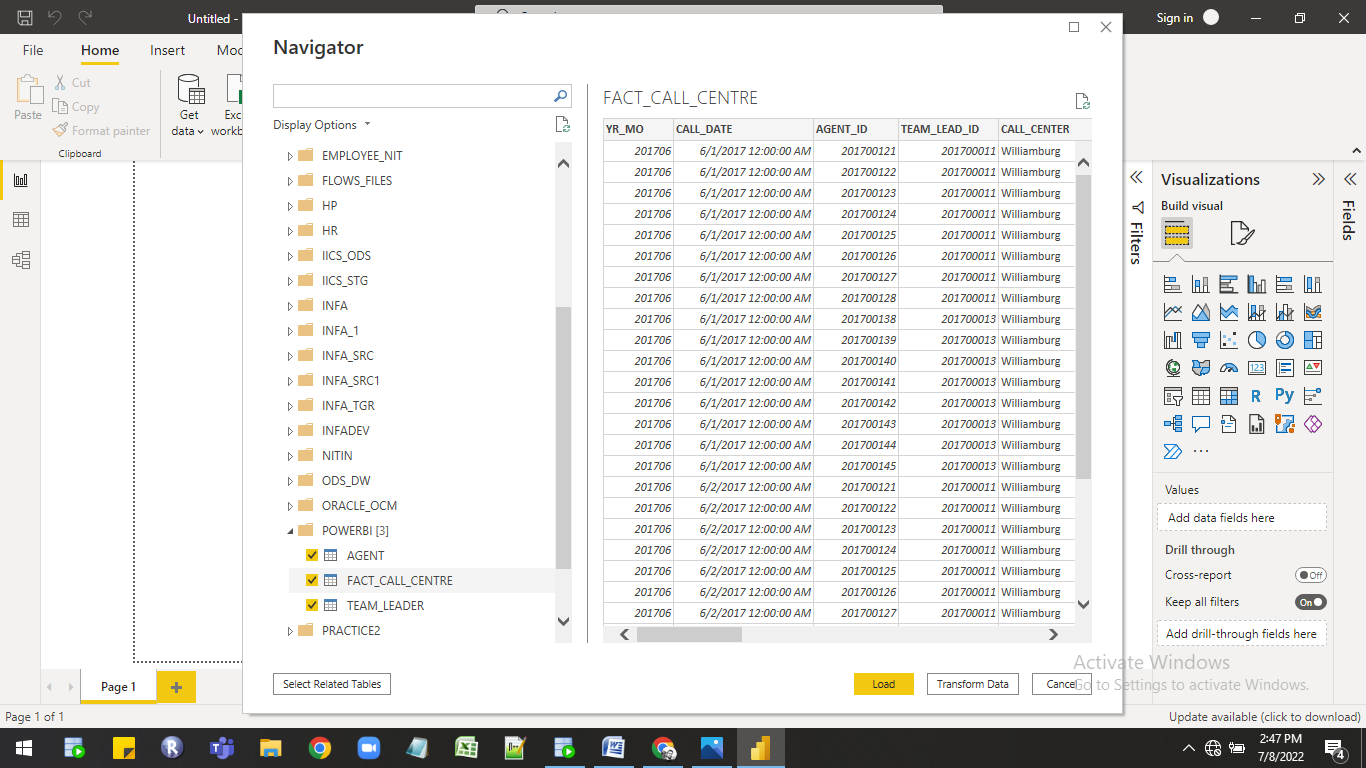
1. Select import option then click on ok then it will connect to oracle new window open in that you can see the oracle schema with table list at left hand side from that list select the connection schema containing created tables then select whatever table you want to import in power bi they are three options
2. 1.Load:-If you do not want any transformation in data then you can use

               Directly use load option.

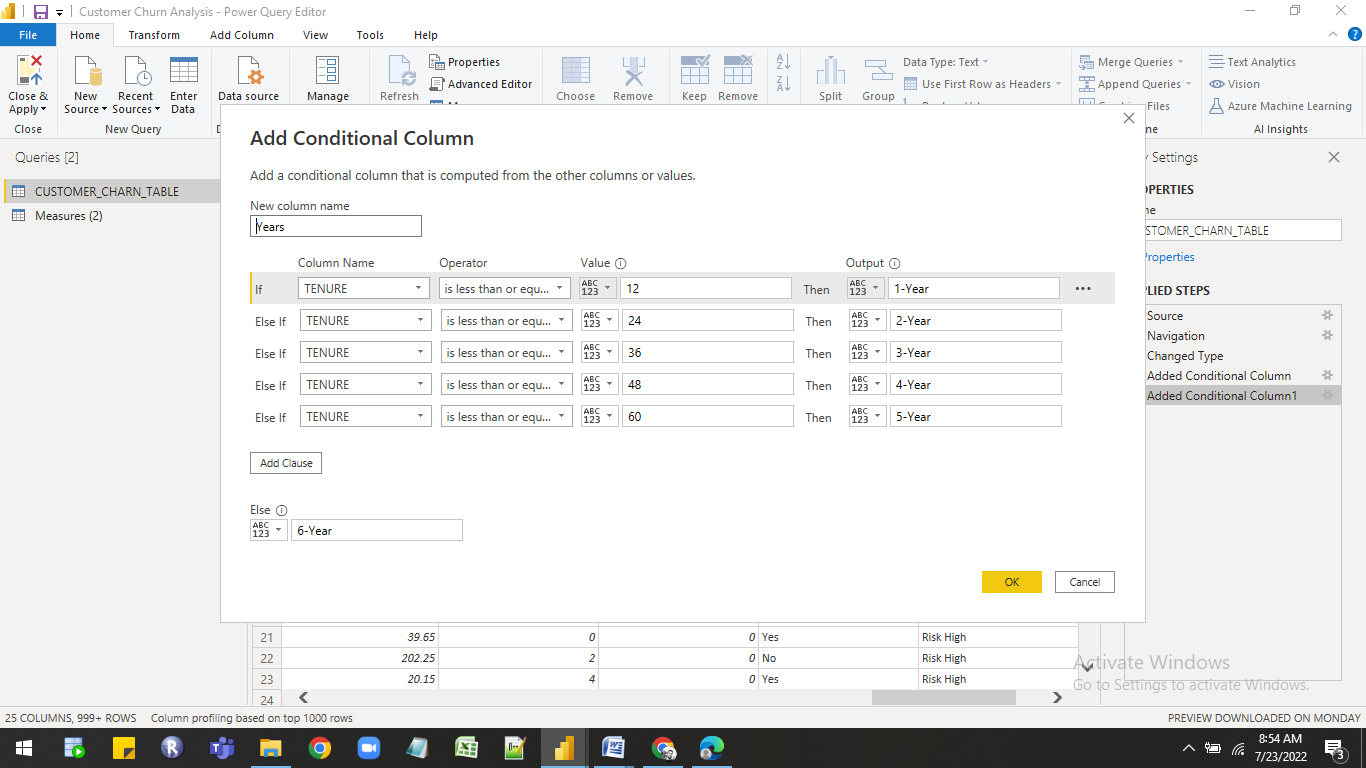
2.Tranform:-Using transform option you can make changes in original data

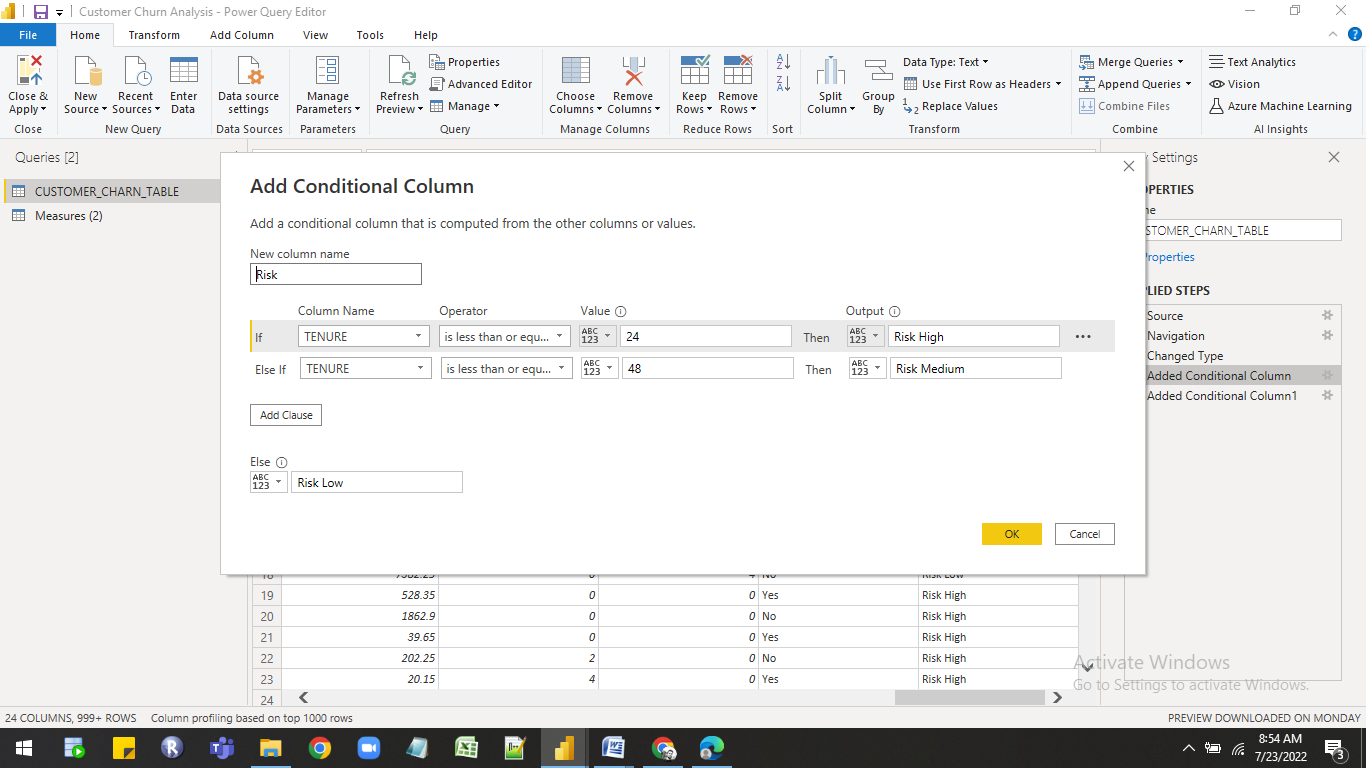
                     In power quary.

3.Cancel:- Using cancel option you can directly cancel the import.

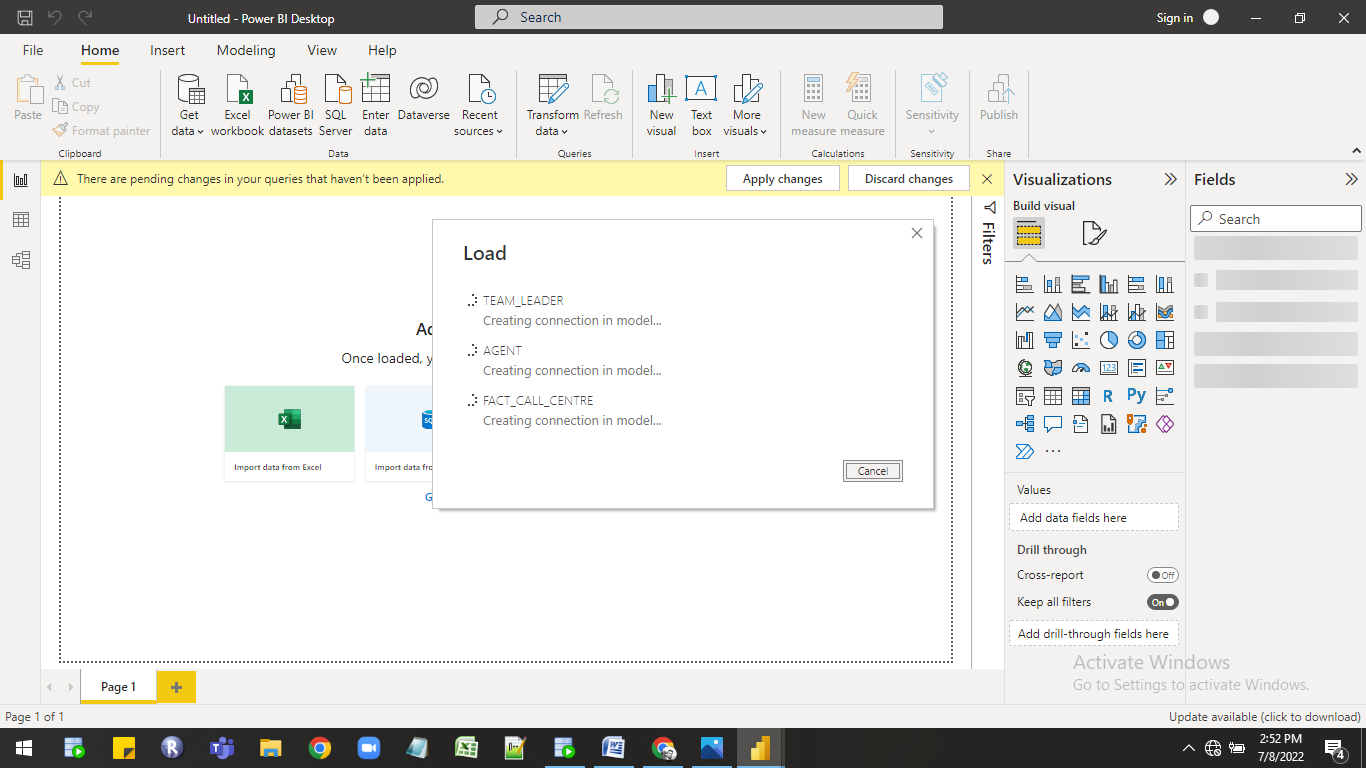


6. Click on Tranform data option to open power  query to clean data i havecreated new conditional column risk  and year use tenure column for that and checkthe datatypes of column.

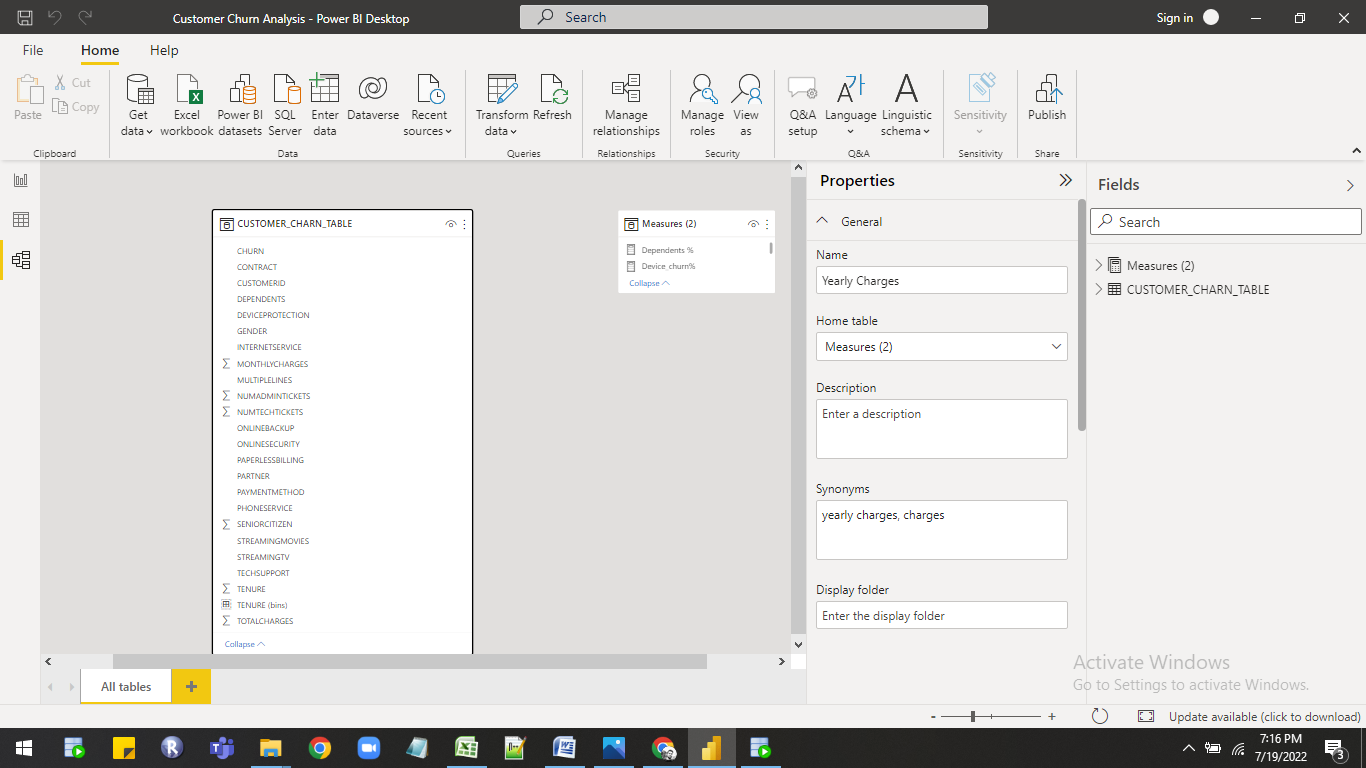




8)Click on Load option then data directly imported into power bi in power pivot             stores in compressed form in columnar database vertically.



9)Tables is loded in power pivot.



## Business Requirements:-

## Following are the requirements

* + Customer Churn Analysis will integrate data from multiple sources and serve as integrated reporting and analytics data repository. Customer Churn Analysis will contain all commercial data:
  + Customer id
  + gender
  + senior citizen
  + partner
  + dependents
  + tenure
  + phone service
  + multiple lines
  + internet service
  + online security
  + online backup
  + device protection
  + tech support
  + streaming tv
  + streaming movies
  + contract
  + paperless billing
  + payment method
  + monthly charge
  + total charge
  + no of admin tickets
  + no of tech tickets
  + churn

**KPI’s**

⮚ No of Customer at risk

⮚ No of Technical tickets

⮚ No of Technical tickets

⮚ Yearly charges

⮚ Monthly charges

⮚ Senior Citizen %

⮚ Partner %

⮚ Dependents %

⮚ Male and Female %

⮚ Subscription time with % for different category

⮚ Customer Account information should have following information,

* Payment method and payment %
* Paperless billing and its Billing %
* Average Monthly and total charges

⮚ Contract type and its % of contract

⮚ Phone Service %

⮚ Streaming TV %

⮚ Streaming Movies %

⮚ Device Protection %

⮚ Online backup %

⮚ Tech Support %

⮚ Online Security %

⮚ Internet Service Provide along with its % of Service providers

⮚ Multiple line yes % with phone services

⮚ Multiple line No % with phone services

⮚ Total Customers

⮚ Churn Rate %

⮚ Total Churn Customers

⮚ Yearly charges

⮚ Technical tickets

⮚ Admin tickets

⮚ Churning by Internet service provider in %

⮚ No of customer Churn by Internet service provider.

⮚ Internet service provider wise Monthly charges

⮚ Contract type wise No of customers and its churn %

⮚ Years of contract by its customer and churn %

⮚ Payment method wise churn % and Monthly charges

## Required Calculation

For creating report required DAX calculation are as follow

| Sr.No | Dax Measure |
| --- | --- |
| 1 | Charn rate % = DIVIDE(CUSTOMER\_CHARN\_TABLE[Count of CHURN forYes],Count(CUSTOMER\_CHARN\_TABLE[CUSTOMERID]),0) |
| 2 | Count of Churn = count(CUSTOMER\_CHARN\_TABLE[CHURN]) |
| 3 | Count of DEPENDENTS for Yes =  CALCULATE(      COUNTA('CUSTOMER\_CHARN\_TABLE'[DEPENDENTS]),      'CUSTOMER\_CHARN\_TABLE'[DEPENDENTS] IN { "Yes" }  ) |
| 4 | Count of PHONESERVICE for Yes =  CALCULATE(      COUNTA('CUSTOMER\_CHARN\_TABLE'[PHONESERVICE]),      'CUSTOMER\_CHARN\_TABLE'[PHONESERVICE] IN { "Yes" }  ) |
| 5 | Count of SENIORCITIZEN for 1 =  CALCULATE(      COUNTA('CUSTOMER\_CHARN\_TABLE'[SENIORCITIZEN]),      'CUSTOMER\_CHARN\_TABLE'[SENIORCITIZEN] IN { 1 }  ) |
| 6 | Dependents % = DIVIDE(CALCULATE(COUNT('CUSTOMER\_CHARN\_TABLE'[DEPENDENTS]),'CUSTOMER\_CHARN\_TABLE'[DEPENDENTS]="Yes",'CUSTOMER\_CHARN\_TABLE'[Churn]="Yes"),CALCULATE(COUNT('CUSTOMER\_CHARN\_TABLE'[DEPENDENTS]),'CUSTOMER\_CHARN\_TABLE'[Churn]="Yes"),0) |
| 7 | Device\_churn% = DIVIDE(CALCULATE(COUNT('CUSTOMER\_CHARN\_TABLE'[DeviceProtection]),'CUSTOMER\_CHARN\_TABLE'[DeviceProtection]="Yes",'CUSTOMER\_CHARN\_TABLE'[Churn]="Yes"),CALCULATE(COUNT('CUSTOMER\_CHARN\_TABLE'[DeviceProtection]),'CUSTOMER\_CHARN\_TABLE'[Churn]="Yes"),0) |
| 8 | Multiple\_no\_churn% = DIVIDE(CALCULATE(COUNT('CUSTOMER\_CHARN\_TABLE'[MultipleLines]),'CUSTOMER\_CHARN\_TABLE'[MultipleLines]="No",'CUSTOMER\_CHARN\_TABLE'[Churn]="Yes"),CALCULATE(COUNT('CUSTOMER\_CHARN\_TABLE'[MultipleLines]),'CUSTOMER\_CHARN\_TABLE'[Churn]="Yes",'CUSTOMER\_CHARN\_TABLE'[MultipleLines] <> "No phone service"),0) |
| 9 | Multiple\_yes\_churn% = DIVIDE(CALCULATE(COUNT('CUSTOMER\_CHARN\_TABLE'[MultipleLines]),'CUSTOMER\_CHARN\_TABLE'[MultipleLines]="Yes",'CUSTOMER\_CHARN\_TABLE'[Churn]="Yes"),CALCULATE(COUNT('CUSTOMER\_CHARN\_TABLE'[MultipleLines]),'CUSTOMER\_CHARN\_TABLE'[Churn]="Yes",'CUSTOMER\_CHARN\_TABLE'[MultipleLines] <> "No phone service"),0) |
| 10 | Onlinebackup\_churn% = DIVIDE(CALCULATE(COUNT('CUSTOMER\_CHARN\_TABLE'[OnlineBackup]),'CUSTOMER\_CHARN\_TABLE'[OnlineBackup]="Yes",'CUSTOMER\_CHARN\_TABLE'[Churn]="Yes"),CALCULATE(COUNT('CUSTOMER\_CHARN\_TABLE'[OnlineBackup]),'CUSTOMER\_CHARN\_TABLE'[Churn]="Yes"),0) |
| 11 | Onlinesec\_churn% = DIVIDE(CALCULATE(COUNT('CUSTOMER\_CHARN\_TABLE'[OnlineSecurity]),'CUSTOMER\_CHARN\_TABLE'[OnlineSecurity]="Yes",'CUSTOMER\_CHARN\_TABLE'[Churn]="Yes"),CALCULATE(COUNT('CUSTOMER\_CHARN\_TABLE'[OnlineSecurity]),'CUSTOMER\_CHARN\_TABLE'[Churn]="Yes"),0) |
| 12 | Partner\_churn% = DIVIDE(CALCULATE(COUNT('CUSTOMER\_CHARN\_TABLE'[Partner]),'CUSTOMER\_CHARN\_TABLE'[Partner]="Yes",'CUSTOMER\_CHARN\_TABLE'[Churn]="Yes"),CALCULATE(COUNT('CUSTOMER\_CHARN\_TABLE'[Partner]),'CUSTOMER\_CHARN\_TABLE'[Churn]="Yes"),0) |
| 13 | Phone Service % = [Count of PHONESERVICE for Yes] /7043 |
| 14 | Senior Citizen % = DIVIDE([Senior Citizen for 1],[Senior Citizen for churn yes],0) |
| 15 | Senior Citizen for 1 = CALCULATE(COUNT('CUSTOMER\_CHARN\_TABLE'[SeniorCitizen]),'CUSTOMER\_CHARN\_TABLE'[SeniorCitizen]=1,'CUSTOMER\_CHARN\_TABLE'[Churn]="Yes") |
| 16 | Senior Citizen for churn yes = CALCULATE(COUNT('CUSTOMER\_CHARN\_TABLE'[SeniorCitizen]),'CUSTOMER\_CHARN\_TABLE'[Churn]="Yes") |
| 17 | Streammovie\_churn% = DIVIDE(CALCULATE(COUNT('CUSTOMER\_CHARN\_TABLE'[StreamingMovies]),'CUSTOMER\_CHARN\_TABLE'[StreamingMovies]="Yes",'CUSTOMER\_CHARN\_TABLE'[Churn]="Yes"),CALCULATE(COUNT('CUSTOMER\_CHARN\_TABLE'[StreamingMovies]),'CUSTOMER\_CHARN\_TABLE'[Churn]="Yes"),0) |
| 18 | StreamTV\_churn% = DIVIDE(CALCULATE(COUNT('CUSTOMER\_CHARN\_TABLE'[StreamingTV]),'CUSTOMER\_CHARN\_TABLE'[StreamingTV]="Yes",'CUSTOMER\_CHARN\_TABLE'[Churn]="Yes"),CALCULATE(COUNT('CUSTOMER\_CHARN\_TABLE'[StreamingTV]),'CUSTOMER\_CHARN\_TABLE'[Churn]="Yes"),0) |
| 19 | Tech\_churn% = DIVIDE(CALCULATE(COUNT('CUSTOMER\_CHARN\_TABLE'[TechSupport]),'CUSTOMER\_CHARN\_TABLE'[TechSupport]="Yes",'CUSTOMER\_CHARN\_TABLE'[Churn]="Yes"),CALCULATE(COUNT('CUSTOMER\_CHARN\_TABLE'[TechSupport]),'CUSTOMER\_CHARN\_TABLE'[Churn]="Yes"),0) |
| 20 | Count of CHURN for Yes =  CALCULATE(      COUNTA('CUSTOMER\_CHARN\_TABLE'[CHURN]),      'CUSTOMER\_CHARN\_TABLE'[CHURN] IN { "Yes" }  ) |

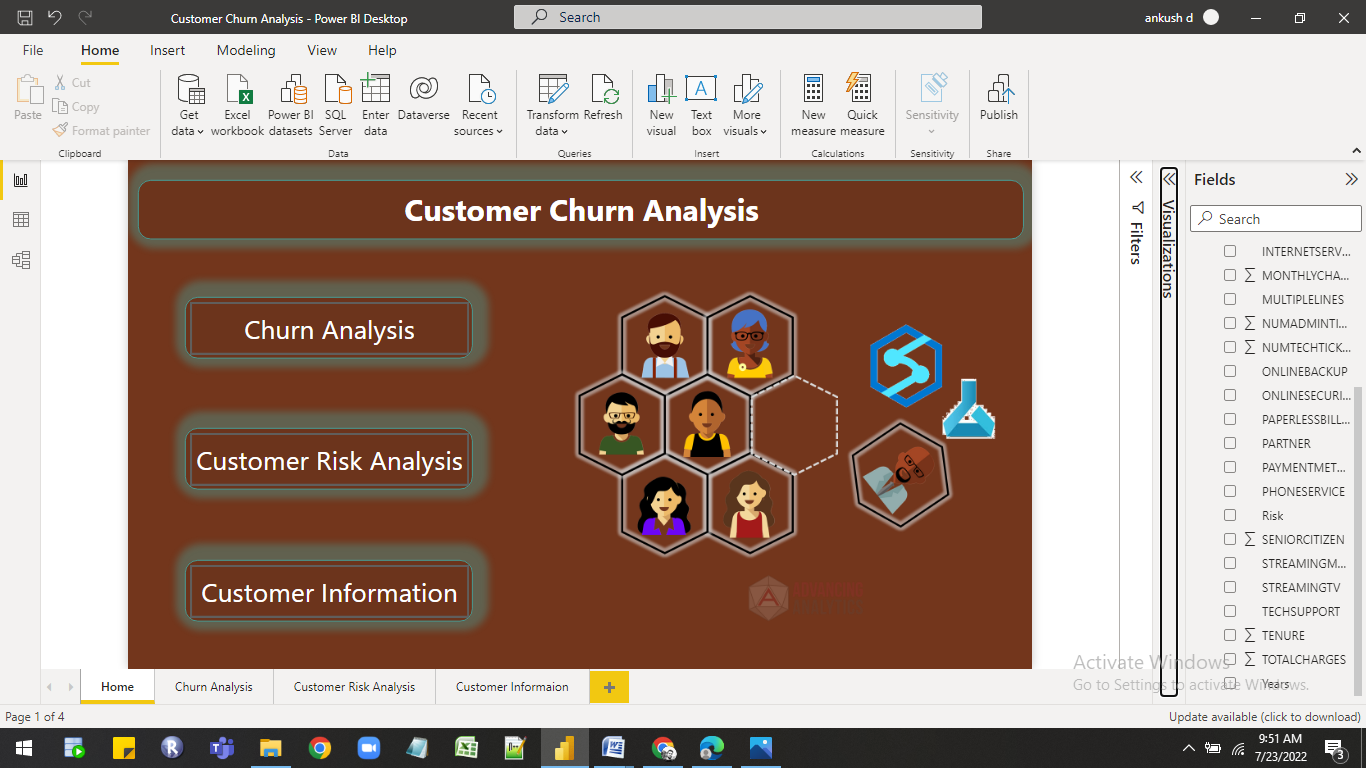
## Report Building :-

As per the requirements next steps is to create report with multiple pages. For creating report the following are the components are used in this report.

Pages In the Report:-

1.Home page:-

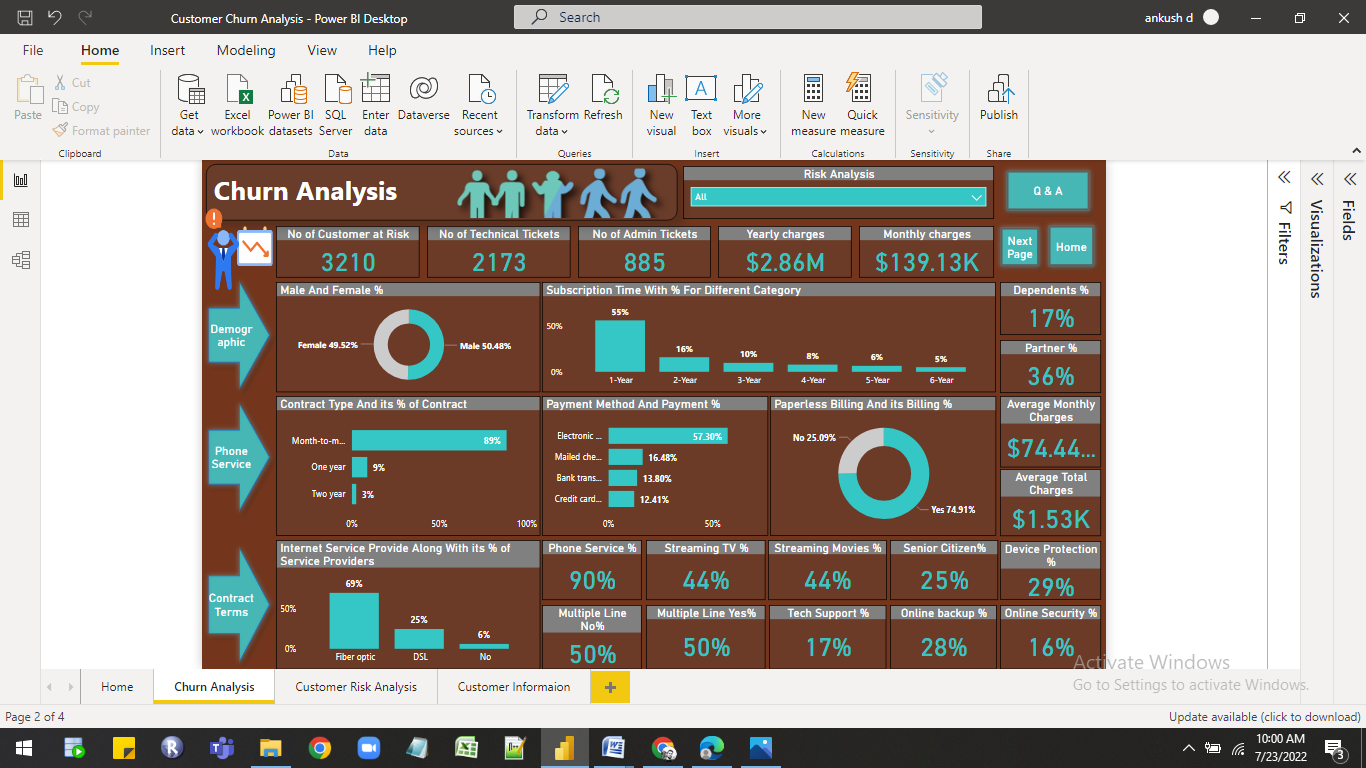
Visualization



Feature Used in report

| 1 | Page Navigation |
| --- | --- |
| 2 | Button |
| 3 | Logo,Image |
| 4 | Tooltip |

2.Churn Analysis:-



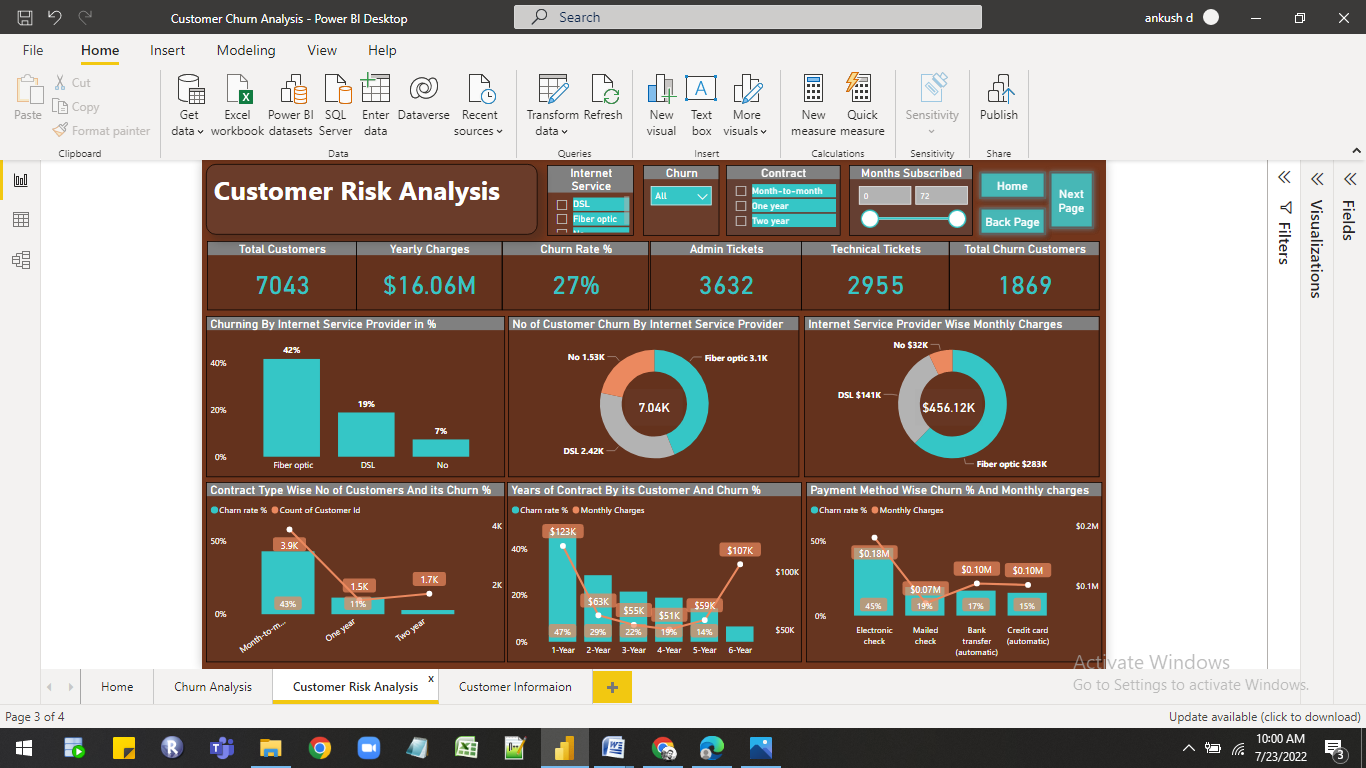
Visualization List

| 1 | Cards |
| --- | --- |
| 2 | Bar Chart |
| 3 | Column Chart |
| 4 | Line Chart |
| 5 | Donut Chart |

Feature Used in report

| 1 | Page Navigation |
| --- | --- |
| 2 | Slicer |
| 3 | Button |
| 4 | Logo,Image,Sticker |

3.Customer Risk Analysis:-



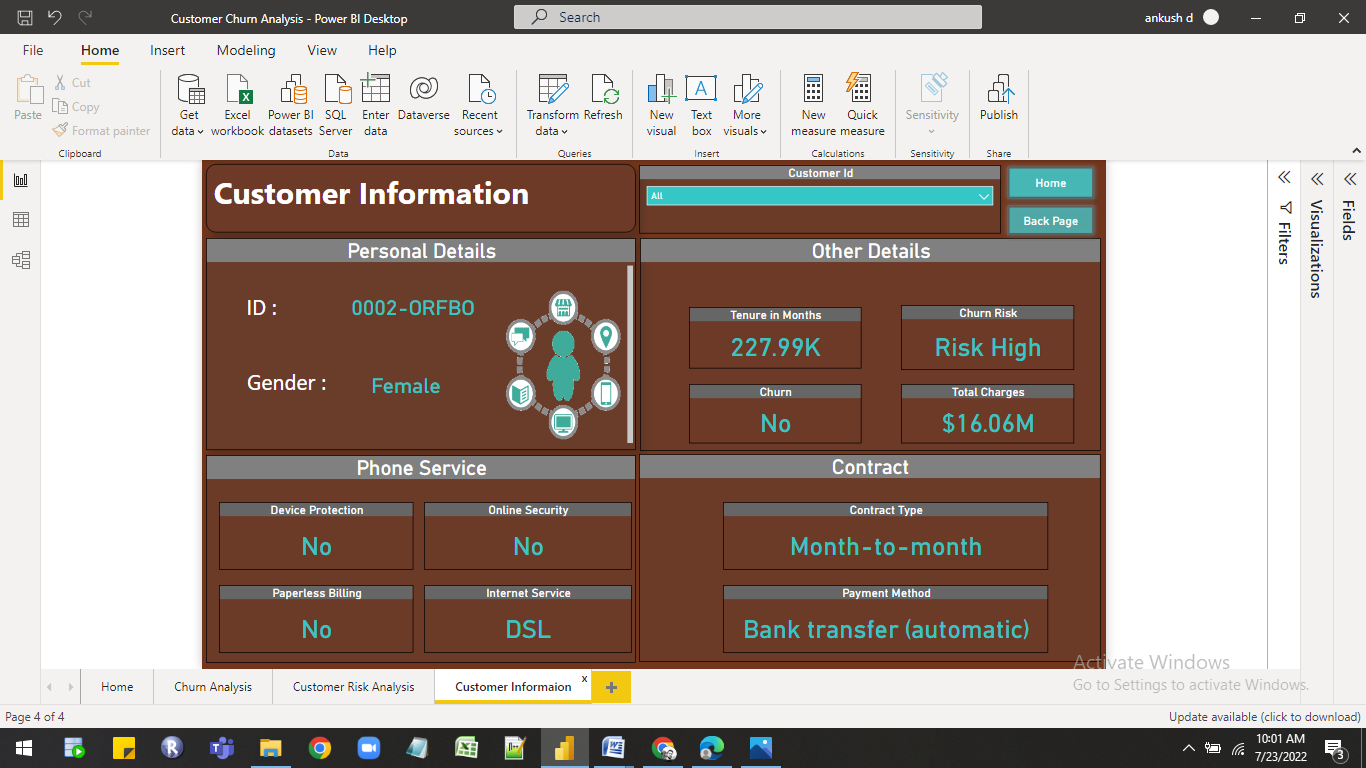
Visualization List

| 1 | Cards |
| --- | --- |
| 2 | Bar Chart |
| 3 | Slicer |
| 4 | Donut Chart |

Feature Used in report

| 1 | Page Navigation |
| --- | --- |
| 2 | Tooltip |
| 3 | Button |

4.Customer Information:-



Visualization List

| 1 | Cards |
| --- | --- |
| 2 | Slicer |

Feature Used in report

| 1 | Page Navigation |
| --- | --- |
| 2 | Tooltip |
| 3 | Button |
| 4 | Sticker |