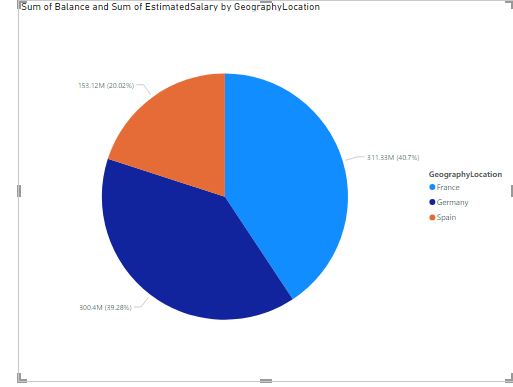
**Learners have to come up with a Report to support the answers to the following questions and suggestions**

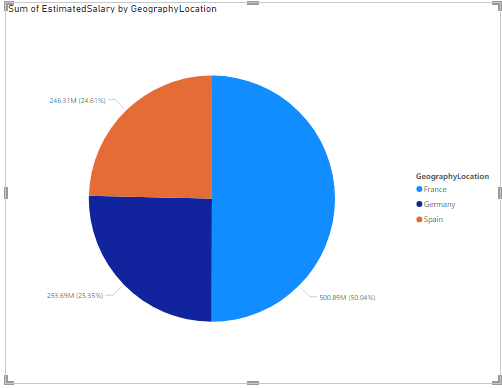
**Objective Questions:**

1. What is the distribution of account balances across different regions?

As we have the data for the balance in the credit card.we have used the pie chart as there are only 3 geographical location and below is the share for each location .



Also we have the sum of estimated salary in the customer Info we have used that as well. Below is the result



1. Identify the top 5 customers with the highest Estimated Salary in the last quarter of the year. (SQL)

**Approach :**

As we are provided with the data of customerInfo table , where Bank DOJ column is in text format , we earlier tried it to change to datetime but it’s not getting imported .   
In order to fix that , we have created another column named “DateOfJoining” with DATE format in the customerInfo table and set the values from the existing column Bank DOJ .

Created new column with date format

alter table customerinfo

add DateOfJoining DATE

Copied values from the existing Bank DOJ column to the newly created DateOfJoining Column

Update customerInfo

Set DateOfJoining = str\_to\_date(`Bank DOJ`, ‘%d’-‘%m’-‘%Y’);

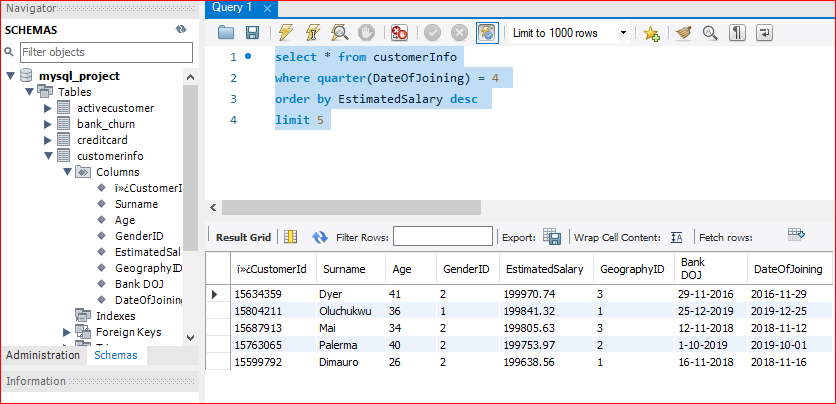
**Conclusion :**

select \* from customerInfo

where quarter(DateOfJoining) = 4

order by EstimatedSalary desc

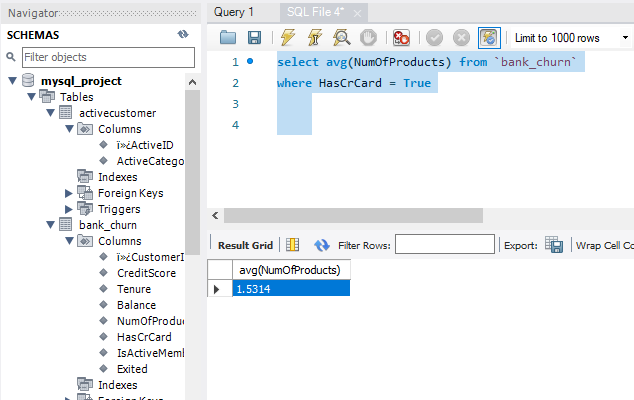
limit 5



1. Calculate the average number of products used by customers who have a credit card. (SQL)

select avg(NumOfProducts) from `bank\_churn`

where HasCrCard = True



1. Determine the churn rate by gender for the most recent year in the dataset.

Churn is basically the sum of total exited customer by total number of customer and then we can group it on the basis of the gender .

select

G.GenderCategory,

G.ï»¿GenderID,

Max(year(CI.`DateOfJoining`)) mostRecentYear,

count(\*) total\_customers,

sum(BC.Exited) total\_exited,

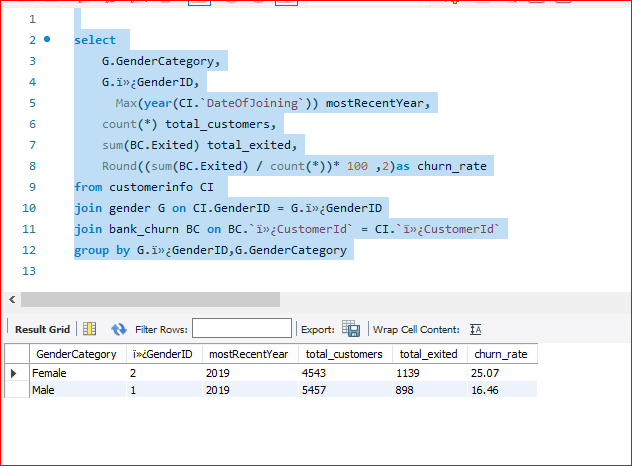
Round((sum(BC.Exited) / count(\*))\* 100 ,2)as churn\_rate

from customerinfo CI

join gender G on CI.GenderID = G.ï»¿GenderID

join bank\_churn BC on BC.`ï»¿CustomerId` = CI.`ï»¿CustomerId`

group by G.ï»¿GenderID,G.GenderCategory



1. Compare the average credit score of customers who have exited and those who remain. (SQL)

select

avg(`CreditScore`),

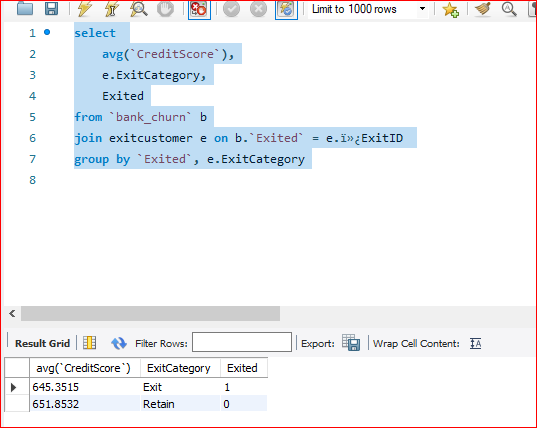
e.ExitCategory,

Exited

from `bank\_churn` b

join exitcustomer e on b.`Exited` = e.ï»¿ExitID

group by `Exited`, e.ExitCategory



1. Which gender has a higher average estimated salary, and how does it relate to the number of active accounts? (SQL)

select

round(avg(c.EstimatedSalary),2) avgEstimateSalary,

g.GenderCategory,

count(case when b.IsActiveMember = True then b.ï»¿CustomerId end) ActiveMembers

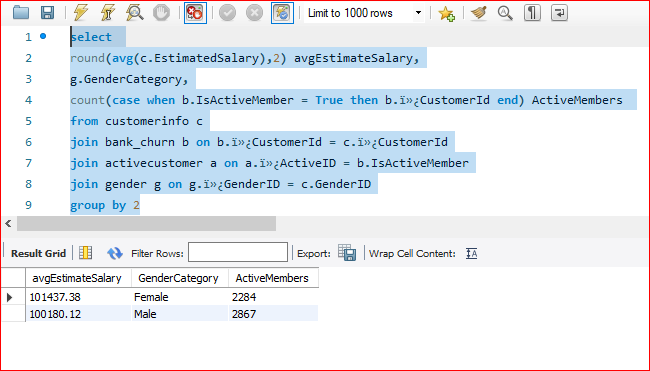
from customerinfo c

join bank\_churn b on b.ï»¿CustomerId = c.ï»¿CustomerId

join activecustomer a on a.ï»¿ActiveID = b.IsActiveMember

join gender g on g.ï»¿GenderID = c.GenderID

group by 2



1. Segment the customers based on their credit score and identify the segment with the highest exit rate. (SQL)

**Approach :**

We have divided the segment into five categories initially which are as below

With creditScoreSegment as (

select

case

when `CreditScore` Between '300' and '500' then 'Poor'

when `CreditScore` between '501' and '600' then 'Fair'

when `CreditScore` between '601' and '700' then 'Good'

when `CreditScore` between '701' and '800' then 'Very Good'

when `CreditScore` between '801' and '900' then ‘Excellent'

else 'Other'

End as segment,

`ï»¿CustomerId`,

`Exited`

from `bank\_churn`

),

SegmentExitRates as (

select segment ,

count(\*) as total\_Customers ,

sum(case when `Exited` = 1 then 1 else 0 end ) as exited\_Customers,

round(100 \* sum(case when `Exited` = 1 then 1 else 0 end )/count(\*) ,2)as exit\_rate

from creditScoreSegment

group by segment

)

select

segment ,

total\_Customers,

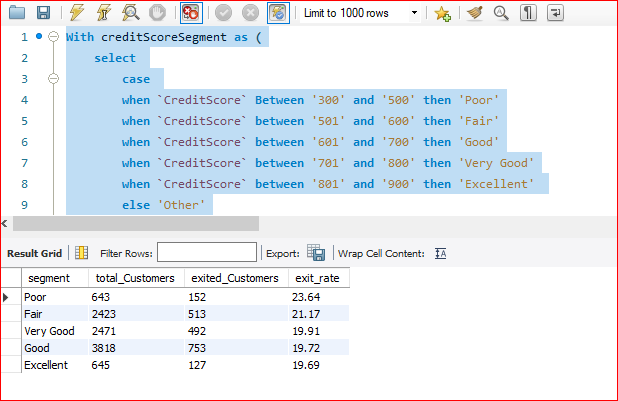
exited\_Customers,

exit\_rate

from SegmentExitRates

order by exit\_rate desc

**Conclusion:**



1. Find out which geographic region has the highest number of active customers with a tenure greater than 5 years. (SQL)

select

G.GeographyLocation,

count(\*)

from bank\_churn bc

join customerinfo CI on bc.ï»¿CustomerId = CI.ï»¿CustomerId

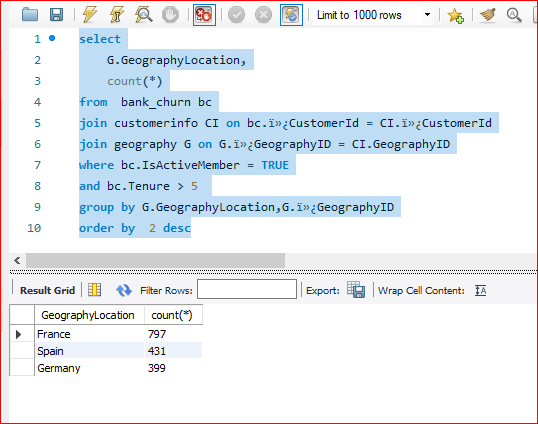
join geography G on G.ï»¿GeographyID = CI.GeographyID

where bc.IsActiveMember = TRUE

and bc.Tenure > 5

group by G.GeographyLocation,G.ï»¿GeographyID

order by 2 desc



1. What is the impact of having a credit card on customer churn, based on the available data?

In order to determine the impact of having a credit card on customer churn, we have initially found the churn rate of customers who has credit card.

select

c.Category,

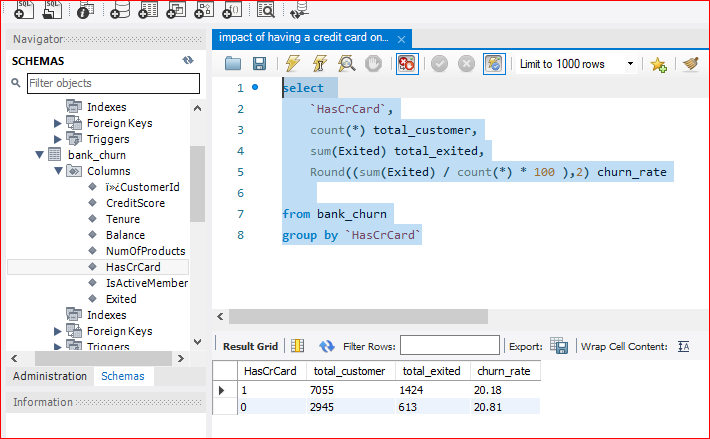
count(\*) total\_customer,

sum(Exited) total\_exited,

Round((sum(Exited) / count(\*) \* 100 ),2) churn\_rate

from bank\_churn b

join creditcard c on c.ï»¿CreditID = b.Has\_creditCard

group by 1

As per our analysis the churn rate of customers with credit card is slightly less than the customers who don’t have a credit card. This suggests that having a credit card might have a slight positive impact on customer retention.

1. For customers who have exited, what is the most common number of products they have used?

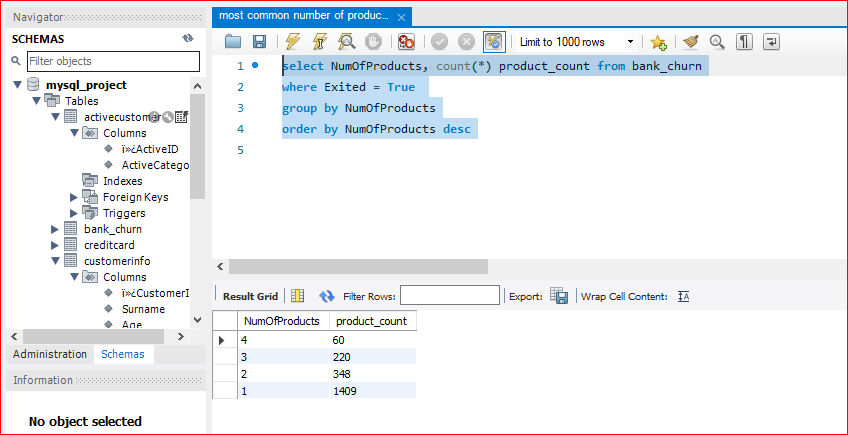
In order to answer the above question , we have grouped by NumberOfProducts and found the count by using SQL.

select NumOfProducts, count(\*) product\_count from bank\_churn

where Exited = True

group by NumOfProducts

order by NumOfProducts desc



1. Examine the trend of customers joining over time and identify any seasonal patterns (yearly or monthly). Prepare the data through SQL and then visualize it.

Now we have the DateOfJoining in date format , we can easily identify the total joining customers by month or year.

**Using Month:**

with join\_month as (

select

month(DateOfJoining) as join\_month,

count(\*) as total\_customer

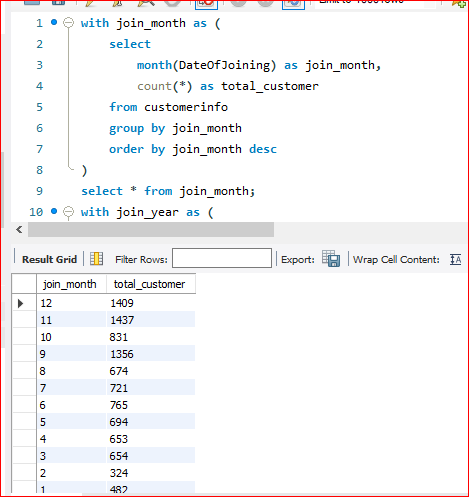
from customerinfo

group by join\_month

order by join\_month desc

)

Select \* from join\_month



**Visualization (Month wise)**

**Using Year :**

with join\_year as (

select

year(DateOfJoining) as join\_year,

count(\*) as total\_customer

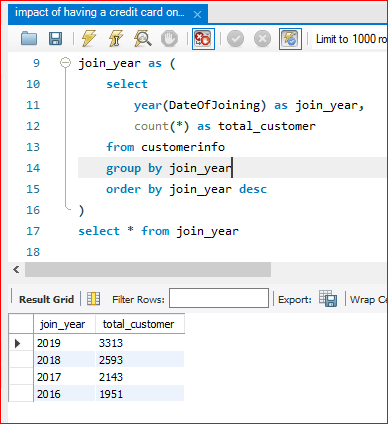
from customerinfo

group by join\_year

order by join\_year desc

)

select \* from join\_month



**Visualization (Year wise)**

1. Analyze the relationship between the number of products and the account balance for customers who have exited.

To analyze the relationship between the number of products and the account balance for customers who have exited , we will be using the NumOfProducts and EstimatedSalary from the data we are provided and group it by the NumOfProducts.

select

BC.NumOfProducts ,

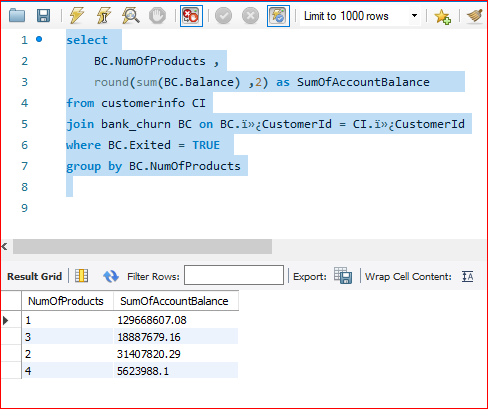
round(sum(BC.Balance) ,2) as SumOfAccountBalance

from customerinfo CI

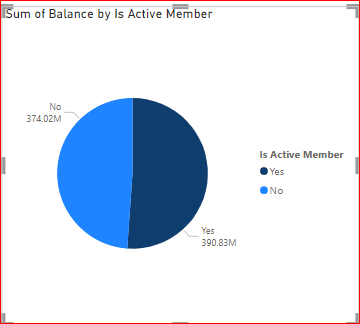
join bank\_churn BC on BC.ï»¿CustomerId = CI.ï»¿CustomerId

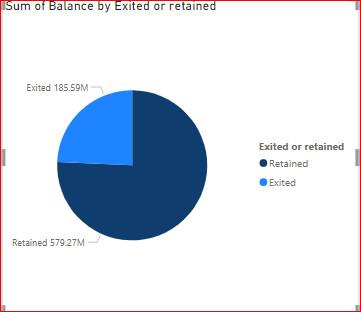
where BC.Exited = TRUE

group by BC.NumOfProducts

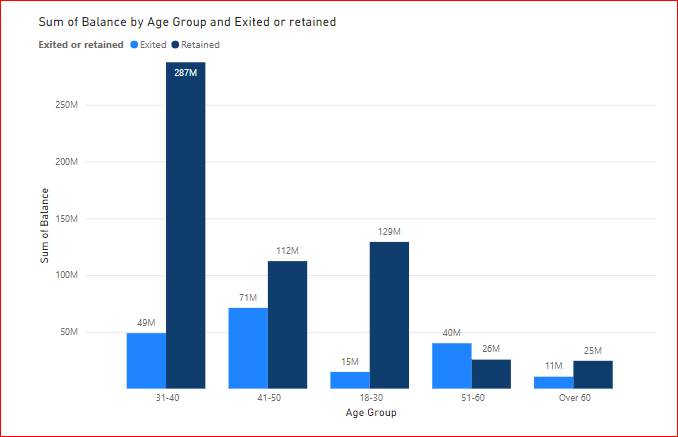


1. Identify any potential outliers in terms of balance among customers who have remained with the bank.

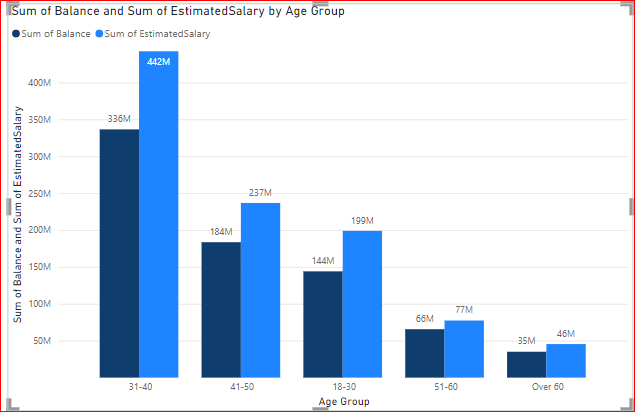


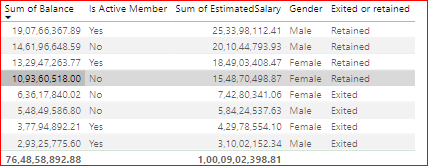


If we look at the sum of the bank balance of both the customers who remained with the bank and who have exited. We can see that those customers who remained with the bank account for 75% of the total bank balance.



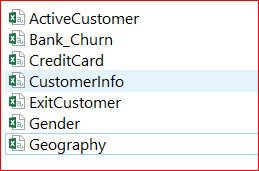
If we look at the bank balance of the customers who are retained by the company and those who exited , we can clearly see the one’s who are retained are having higher bank balance than those who exited except in the age-group of 51-60 where the Exited customer had higher bank balance.





1. How many different tables are given in the dataset, out of these tables which table only consists of categorical variables?

There are 7 different tables given to us in the dataset which are as belows



Out of these 7 tables , 5 table only consists of categorical variables which are

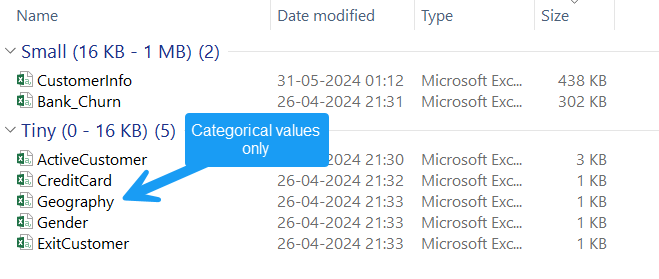
Active Customer,

CreditCard,

Exit Customer,

Gender ,

Geography



1. Using SQL, write a query to find out the gender-wise average income of males and females in each geography id. Also, rank the gender according to the average value. (SQL)

select

g.GenderCategory as Gender,

Ge.GeographyLocation,

round(avg(ci.EstimatedSalary),2) Average\_income,

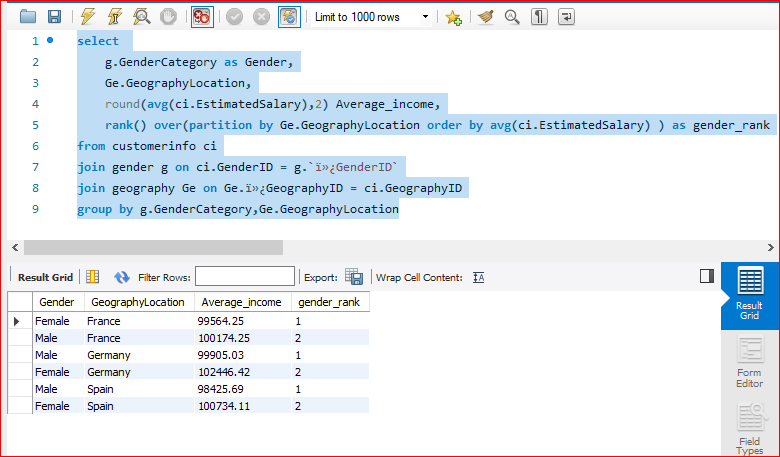
rank() over(partition by Ge.GeographyLocation order by avg(ci.EstimatedSalary) ) as gender\_rank

from customerinfo ci

join gender g on ci.GenderID = g.`ï»¿GenderID`

join geography Ge on Ge.ï»¿GeographyID = ci.GeographyID

group by g.GenderCategory,Ge.GeographyLocation



1. Using SQL, write a query to find out the average tenure of the people who have exited in each age bracket (18-30, 30-50, 50+).

select

round(avg(bc.Tenure),2) Average\_tenure,

case

when Age between 18 and 30 then 'Age Group 18-30'

when Age between 30 and 50 then 'Age Group 30-50'

when Age > 50 then 'Age group over 50'

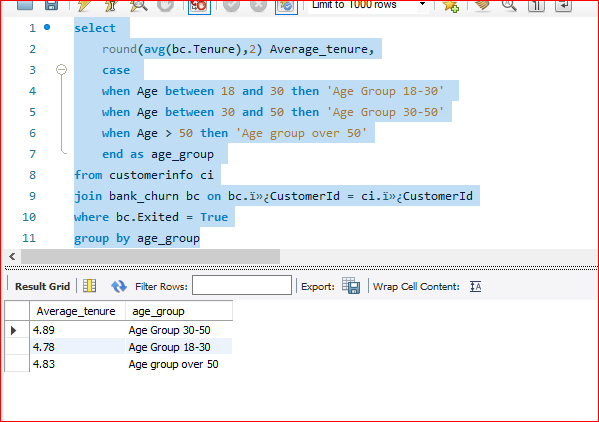
end as age\_group

from customerinfo ci

join bank\_churn bc on bc.ï»¿CustomerId = ci.ï»¿CustomerId

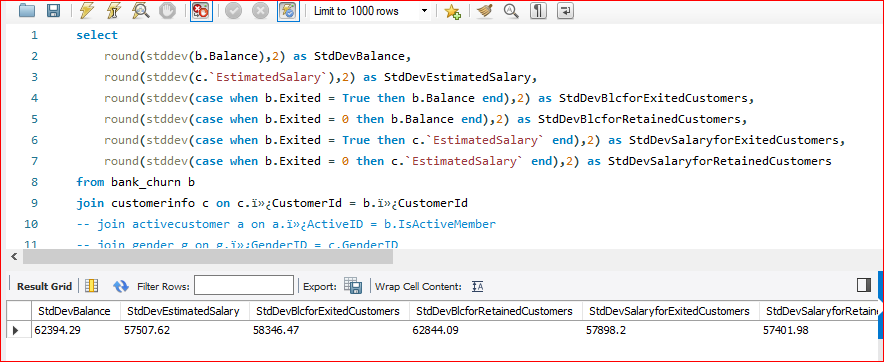
where bc.Exited = True

group by age\_group



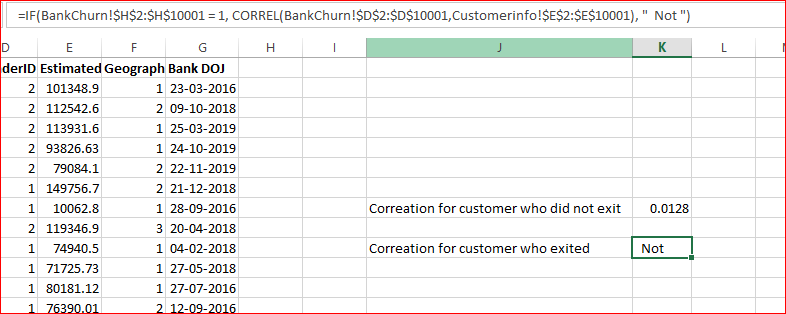
1. Is there any direct correlation between salary and the balance of the customers? And is it different for people who have exited or not?

Standard Deviation :



**Correlation between salary and the balance**

Using excel :



The correlation coefficient for the customers who remained with the bank (not exited) is **0.0128**

We have calculated it in SQL as well using below calculation

with SumStats as (

select

count(\*) as n ,

sum(c.EstimatedSalary) as SumX,

sum(b.Balance) as SumY,

Sum(b.Balance \* c.EstimatedSalary) as SumXY,

Sum(c.EstimatedSalary \* c.EstimatedSalary) as SumX2,

Sum(b.Balance \* b.Balance) as SumY2

from bank\_churn b

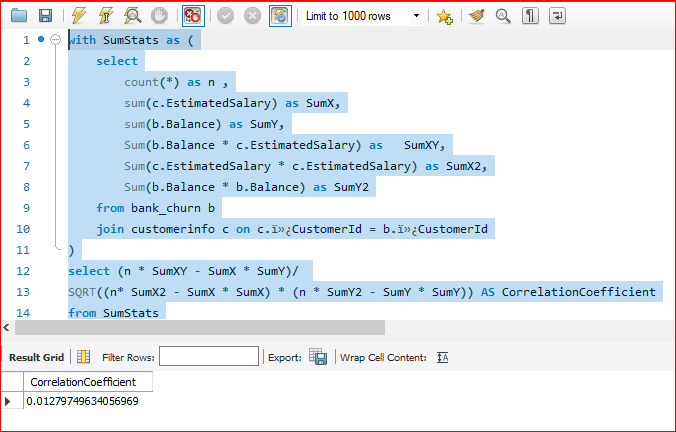
join customerinfo c on c.ï»¿CustomerId = b.ï»¿CustomerId

)

select (n \* SumXY - SumX \* SumY)/

SQRT((n\* SumX2 - SumX \* SumX) \* (n \* SumY2 - SumY \* SumY)) AS CorrelationCoefficient

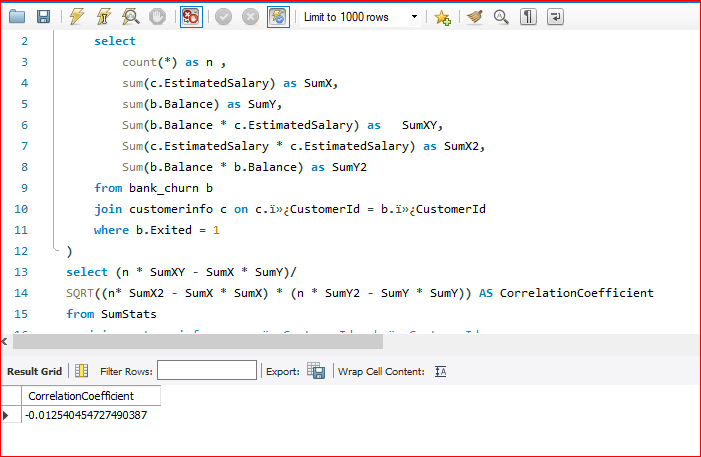
from SumStats



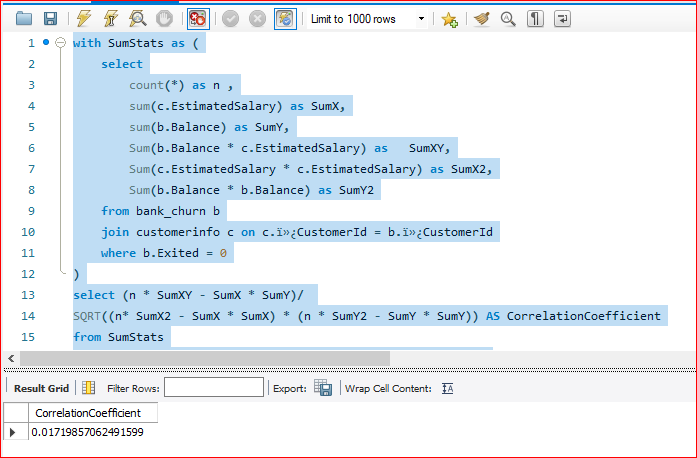
Correlation Coefficient: **0.01279749634056969**

Now we the correlation between the Salary and Balance we can fetch it for those customer who are retained or Exited by putting a WHERE condition.

For Retained Customer**: -0.012540454727490387**



For Exited Customers : **0.01719857062491599**



**Conclusion:**

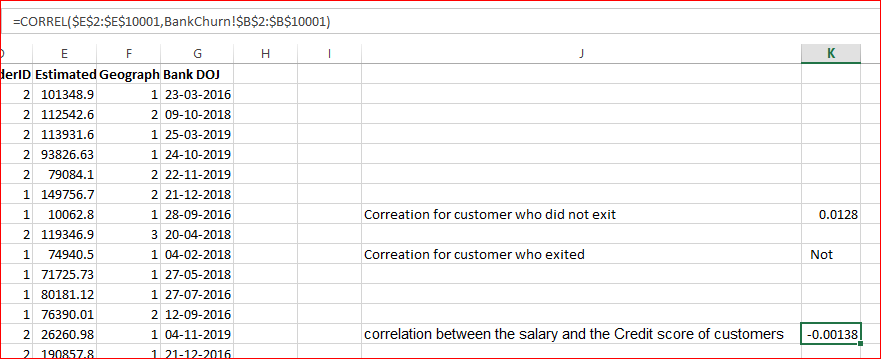
Overall Coefficient between Salary and Balance: **0.01279749634056969**

For Retained Customer: **-0.012540454727490387**

For Exited Customers: **0.01719857062491599**

:

1. Is there any correlation between the salary and the Credit score of customers?



A correlation coefficient of -0.00138 indicates a very weak negative correlation between salary and credit score. This means that there is almost no linear relationship between the two variables.

The negative sign indicates a slight tendency for credit score to decrease as salary increases, but this relationship is extremely weak.

1. Rank each bucket of credit score as per the number of customers who have churned the bank.

select

CreditScoreBucket,

Count(case when Exited = 1 then `ï»¿CustomerId` end ) as churnedCustomers,

rank() over(order by count(case when Exited = 1 then `ï»¿CustomerId` end) desc) as rank\_

from (

Select

case

when CreditScore between 300 and 500 then '(300 - 500) Poor'

when CreditScore between 501 and 600 then '(501 - 600) Fair'

when CreditScore between 601 and 700 then '(601 - 700) Good'

when CreditScore between 701 and 800 then '(701 - 800) Very good'

when CreditScore > 800 then '(Above 800 ) Excellent'

else 'Unknown'

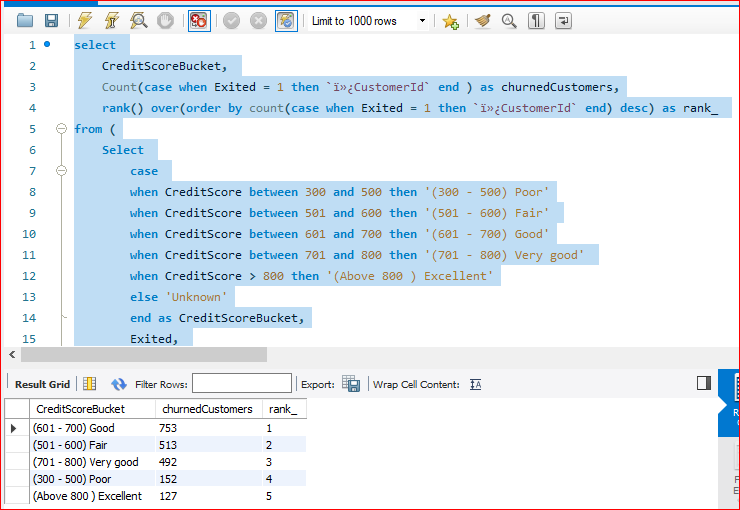
end as CreditScoreBucket,

Exited,

`ï»¿CustomerId`

from bank\_churn )as subquery

group by CreditScoreBucket



1. According to the age buckets find the number of customers who have a credit card. Also retrieve those buckets that have lesser than average number of credit cards per bucket.

Age buckets to find the number of customers who have a credit card

select

case

when Age between 18 and 30 then 'Age Group between 18-30'

when Age between 31 and 40 then 'Age Group between 31-40'

when Age between 41 and 50 then 'Age Group between 41-50'

when Age between 51 and 60 then 'Age Group between 51-60'

when Age > 60 then 'Age Group over 60'

else 'Unknown'

end as AgeBucket,

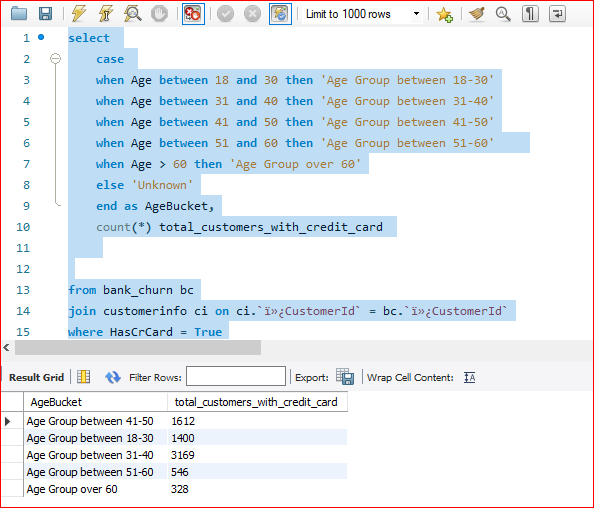
count(\*) total\_customers\_with\_credit\_card

from bank\_churn bc

join customerinfo ci on ci.`ï»¿CustomerId` = bc.`ï»¿CustomerId`

where HasCrCard = True

group by AgeBucket



To retrieve those buckets that have lesser than average number of credit cards per bucket.

WITH AgeBuckets AS (

SELECT

CASE

WHEN Age BETWEEN 18 AND 30 THEN 'Age Group between 18-30'

WHEN Age BETWEEN 31 AND 40 THEN 'Age Group between 31-40'

WHEN Age BETWEEN 41 AND 50 THEN 'Age Group between 41-50'

WHEN Age BETWEEN 51 AND 60 THEN 'Age Group between 51-60'

WHEN Age > 60 THEN 'Age Group over 60'

ELSE 'Unknown'

END AS AgeBucket,

COUNT(CASE WHEN HasCrCard = 1 THEN ci.`ï»¿CustomerId` END) AS NumCreditCardCustomers

FROM

bank\_churn bc

JOIN

customerinfo ci ON ci.`ï»¿CustomerId` = bc.`ï»¿CustomerId`

GROUP BY AgeBucket

),

AverageCreditCards AS (

SELECT

AVG(NumCreditCardCustomers) AS AvgCreditCardCustomers

FROM

AgeBuckets

)

-- Step 3: Retrieve buckets with fewer than the average number of credit cards

SELECT

AgeBuckets.AgeBucket,

AgeBuckets.NumCreditCardCustomers

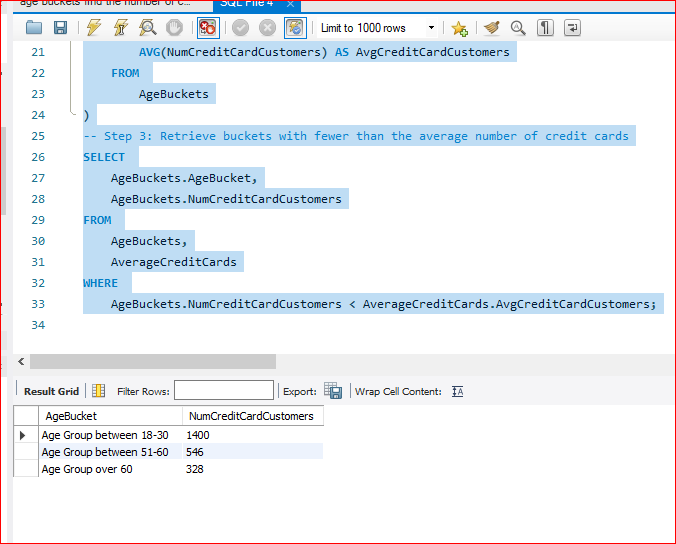
FROM

AgeBuckets,

AverageCreditCards

WHERE

AgeBuckets.NumCreditCardCustomers < AverageCreditCards.AvgCreditCardCustomers;



1. Rank the Locations as per the number of people who have churned the bank and average balance of the customers.

select

GeographyLocation,

ChurnedCustomers,

average\_Balance,

rank() over(order by ChurnedCustomers desc ) churnRank,

rank() over(order by average\_Balance desc ) BalanceRank

from (

select

g.GeographyLocation GeographyLocation,

count(case when bc.Exited = 1 then bc.`ï»¿CustomerId` end) as ChurnedCustomers ,

avg(bc.Balance) average\_Balance

from customerinfo ci

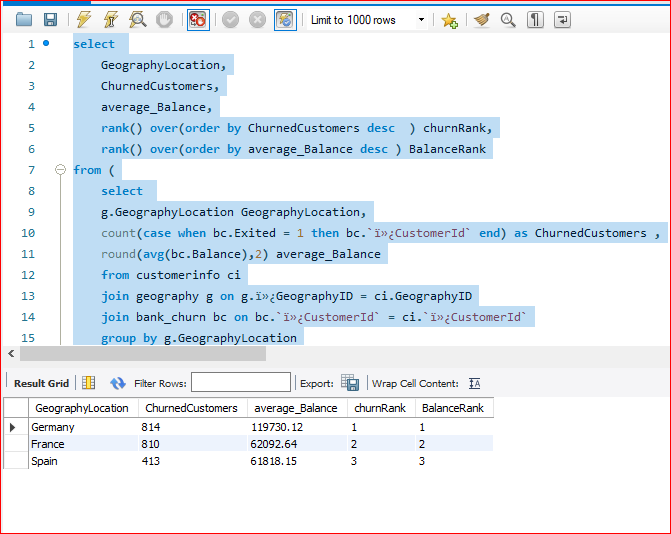
join geography g on g.ï»¿GeographyID = ci.GeographyID

join bank\_churn bc on bc.`ï»¿CustomerId` = ci.`ï»¿CustomerId`

group by g.GeographyLocation

) as location

order by churnRank, BalanceRank



1. As we can see that the “CustomerInfo” table has the CustomerID and Surname, now if we have to join it with a table where the primary key is also a combination of CustomerID and Surname, come up with a column where the format is “CustomerID\_Surname”.

**Approach :**

Initially we have created a new table

create table customers\_info (

CustomerID int ,

Surname varchar(50),

Customer\_Surname varchar(150)

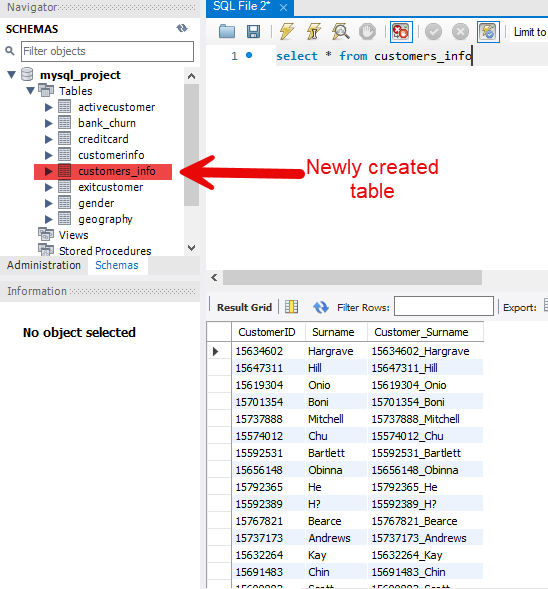
);

Then we inserted the values from the existing CustomerInfo table using the below query

insert into customers\_info( CustomerId, Surname, Customer\_Surname)

select `ï»¿CustomerId`,Surname , concat(`ï»¿CustomerId` , "\_",Surname) from customerinfo;

select \* from customers\_info;



1. Without using “Join”, can we get the “ExitCategory” from ExitCustomers table to Bank\_Churn table? If yes do this using SQL.

**Approach :**

We can use subquery to achieve it

select bc.\*,

(

select

case

when ec.`ï»¿ExitID` = 1 then 'Exited'

when ec.`ï»¿ExitID` = 0 then 'Retained'

else 'Unknown'

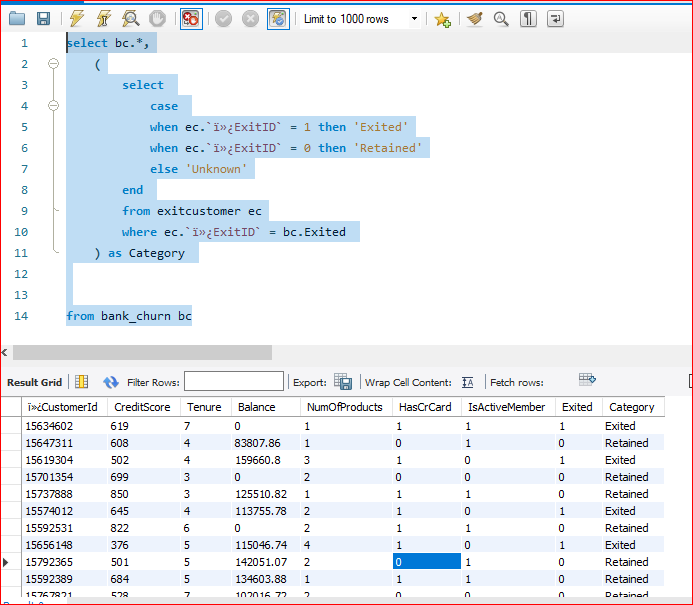
end

from exitcustomer ec

where ec.`ï»¿ExitID` = bc.Exited

) as Category

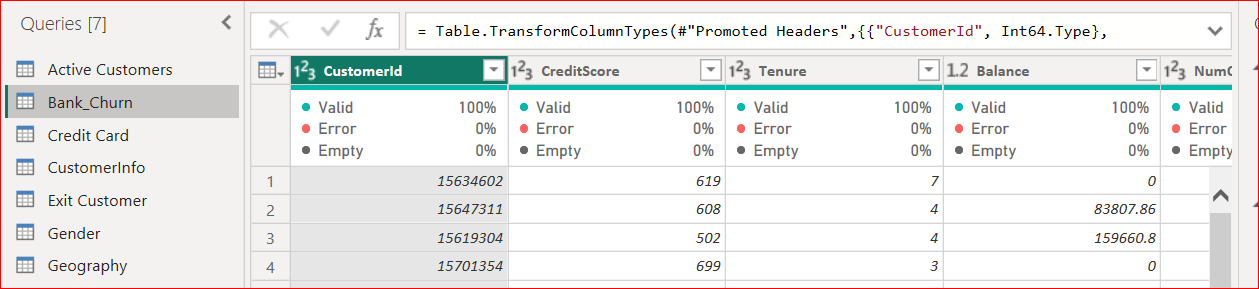
from bank\_churn bc

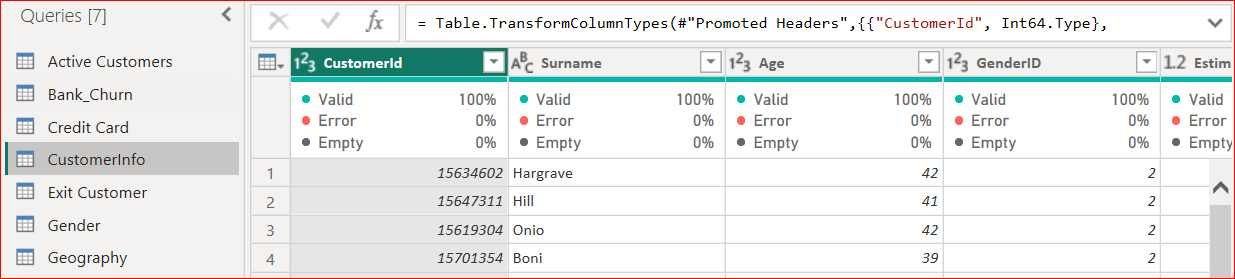


1. Were there any missing values in the data, using which tool did you replace them and what are the ways to handle them?

**Approach :**

Initially we have loaded the data into Power BI and transform the data so that it can be opened in the Power Query Editor , in the view section we have added the Column quality option.





As we can see the quality is 100% in all the columns in both these tables “Bank Churn” and “CustomerInfo” . Rest all 5 tables are consists of only categorical variables .

1. Write the query to get the customer IDs, their last name, and whether they are active or not for the customers whose surname ends with “on”.

select

bc.`ï»¿CustomerId` CustomerID,

ci.Surname ,

case

when bc.IsActiveMember = 1 then 'Active'

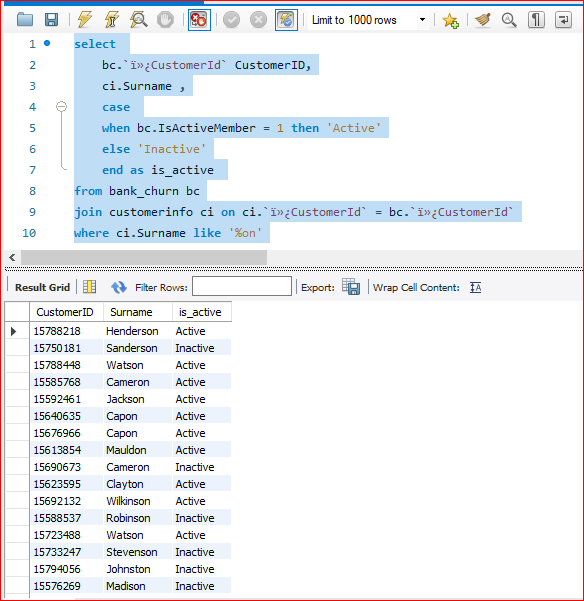
else 'Inactive'

end as is\_active

from bank\_churn bc

join customerinfo ci on ci.`ï»¿CustomerId` = bc.`ï»¿CustomerId`

where ci.Surname like '%on'



**Subjective Question:**

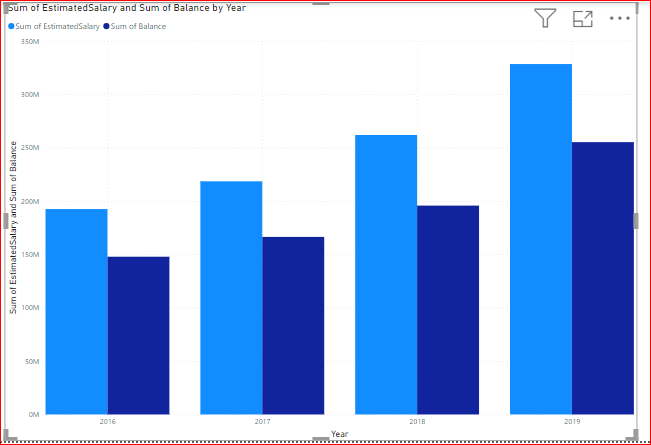
1. Customer Behavior Analysis: What patterns can be observed in the spending habits of long-term customers compared to new customers, and what might these patterns suggest about customer loyalty?

Patterns in the spending habits of long-term customers compared to the new ones can provide valuable insights about customer loyalty.

**Higher Average Spending:** Long-term customers often have higher average spending compared to the new customers as the several factors kicks in and they are more familiar with the product/services offered.

**Variety of Products :** Long-term customers often have many products compared to the new customers as he are talking in regards to the banking/financial services . Long term customer probably have more services for eg. Forex card/credit card/ debit card/ loan disbursal etc compared to the new customer also Banks also have certain policies like you can avail this particular facility once after certain time period after joining /becoming a customer with the bank .

The below clustered column chart defines the sum of estimated salary and the sum of balance year-wise



1. Product Affinity Study: Which bank products or services are most commonly used together, and how might this influence cross-selling strategies?

A product affinity study can reveal which bank products or services are frequently used together by customers.

As higher the salary, the more customer is prone to spend, having credit card is also an added advantage. Number of products customer is having is also used commonly.

**Influence on Cross-Selling Strategies**

**Targeted Marketing:**  Armed with knowledge of product affinities, banks can create targeted marketing campaigns to promote related products to customers. For example, if customers who have a savings account are frequently found to also have a retirement savings plan, the bank can target savings account holders with promotions for retirement planning services.

**Bundle Offerings**: Banks can create bundled offerings or packages that combine complementary products based on affinity patterns. Offering discounts or incentives for bundled products encourages customers to purchase multiple products from the bank, increasing their overall engagement and loyalty.

**Personalized Recommendations:** Utilizing machine learning algorithms, banks can provide personalized recommendations to customers based on their usage patterns and product affinities.

**Enhanced Customer Experience:** By proactively offering relevant products and services based on customer preferences, banks can enhance the overall customer experience and build stronger relationships with their customers.

1. Geographic Market Trends: How do economic indicators in different geographic regions correlate with the number of active accounts and customer churn rates?

There is a complex relationship between economic indicators, number of active accounts, and customer churn rates across geographic markets. As per the provided data , we don’t have enough economic indicators such as unemployment rates ,GDP growth ,inflation rates.Here's a breakdown of some potential correlations:

Positive Correlation: GDP Growth and Active Accounts

Negative Correlation: GDP Growth and Churn Rate

Other factors to consider are Industry ,Customer Base,regional differences .

select

count(case when bc.Exited = 1 then bc.`ï»¿CustomerId` end) churned\_Customer,

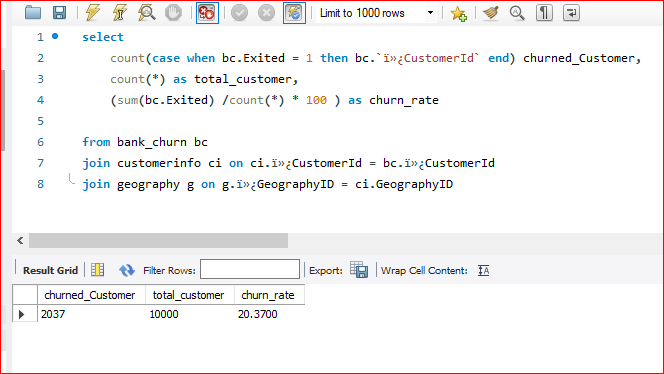
count(\*) as total\_customer,

(sum(bc.Exited) /count(\*) \* 100 ) as churn\_rate

from bank\_churn bc

join customerinfo ci on ci.ï»¿CustomerId = bc.ï»¿CustomerId

join geography g on g.ï»¿GeographyID = ci.GeographyID



1. Risk Management Assessment: Based on customer profiles, which demographic segments appear to pose the highest financial risk to the bank, and why?

select

g.GeographyLocation,

round(sum(b.Balance),2),

count(\*) as total\_customers,

round(sum(c.EstimatedSalary),2) sumofSalary,

count(case when Has\_creditCard = true then b.`ï»¿CustomerId` end ) totalCustomerWithCrCard,

count(case when b.Exited = true then b.`ï»¿CustomerId` end) churnedCustomers,

(count(case when b.Exited = true then b.`ï»¿CustomerId` end) /count(\*) \* 100) churnRate\_wrt\_total\_customer,

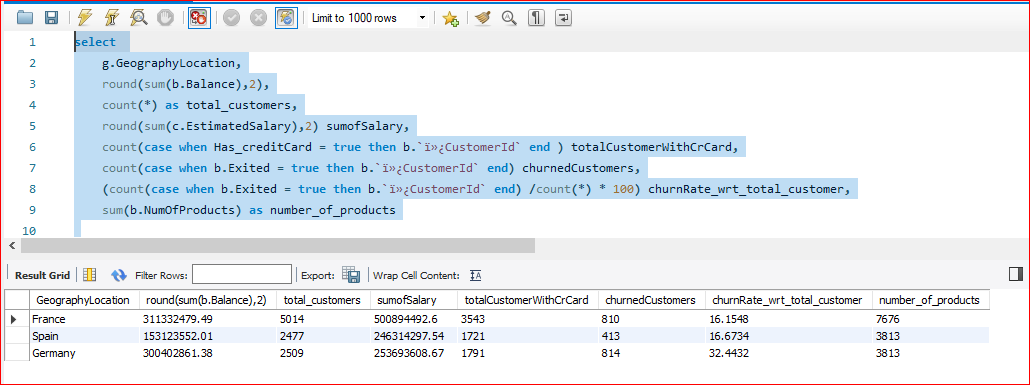
sum(b.NumOfProducts) as number\_of\_products

from customerinfo c

join geography g on g.ï»¿GeographyID = c.GeographyID

join bank\_churn b on c.`ï»¿CustomerId` = b.`ï»¿CustomerId`

group by g.GeographyLocation



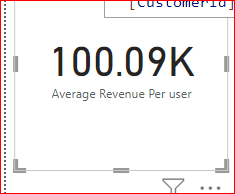
As we can see Germany is having the highest churn rate and it can probably pose the highest financial risk compared to other countries.

There are other factors as well for eg. Age Groups and retires

1. Customer Tenure Value Forecast: How would you use the available data to model and predict the lifetime (tenure) value in the bank of different customer segments?

In order to forecast the availale data to model and predict the lifetime (tenure) , we have initially calculated the

**Average Revenue Per user =** DIVIDE(SUM(CustomerInfo[EstimatedSalary]),DISTINCTCOUNT(CustomerInfo[CustomerId]))



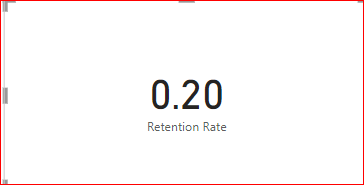
**Retention Rate =**

DIVIDE(

COUNTROWS(FILTER(Bank\_Churn, Bank\_Churn[Exited or retained ] = "Yes")),

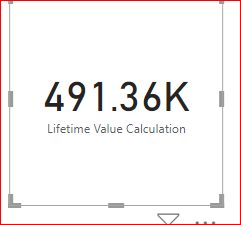
COUNTROWS(Bank\_Churn)

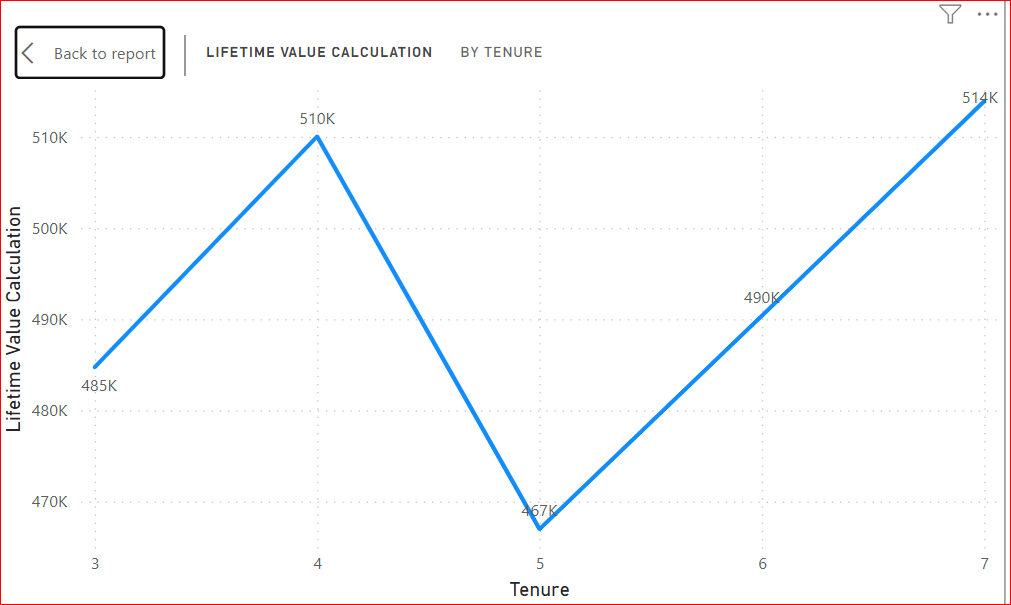
)

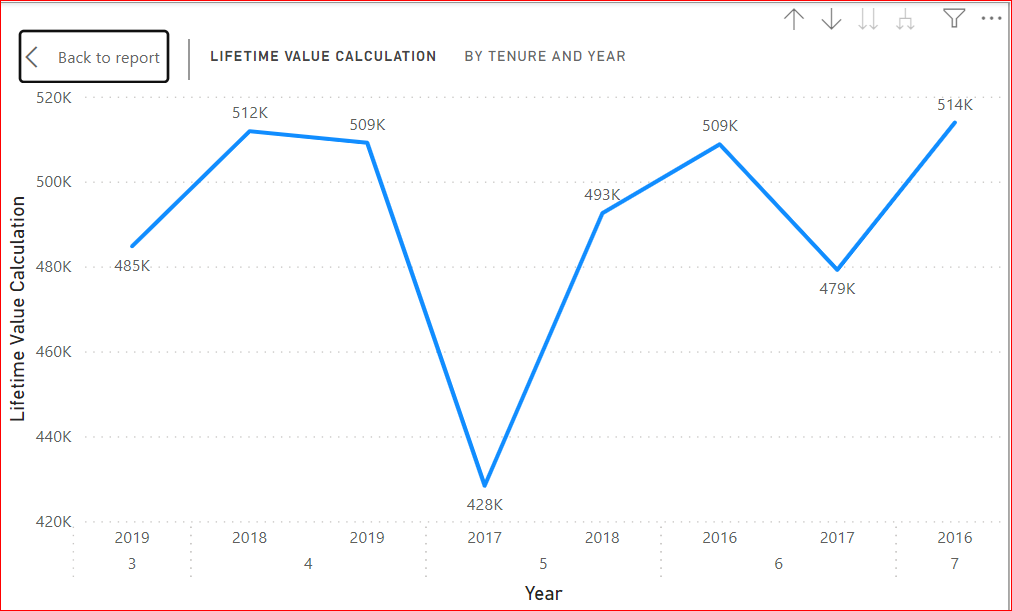


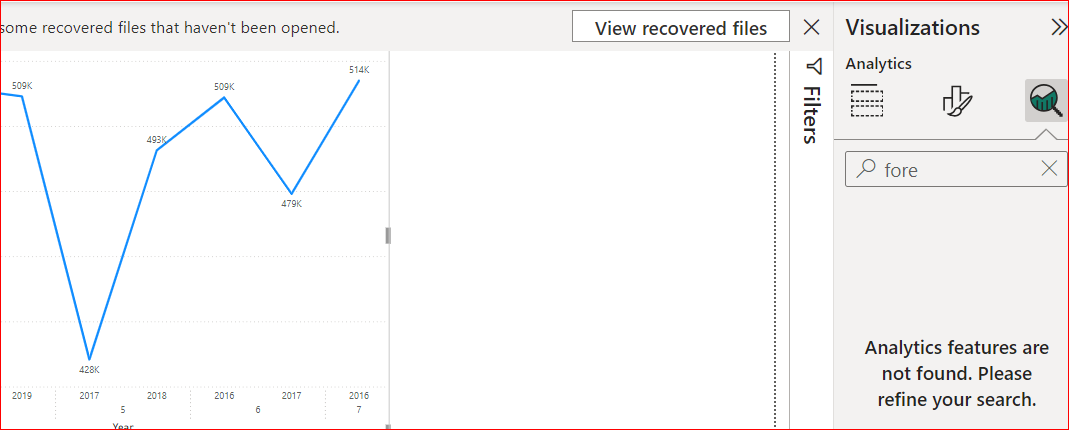
**Lifetime Value (LTV) Calculation**: Create a measure to calculate the average revenue per customer over a specific period

**Lifetime Value Calculation** = [Average Revenue Per user]/Bank\_Churn[Retention Rate]







Since the option to forecast it isn’t available on the analytics tab, we haven’t forecasted it further. 

1. Marketing Campaign Effectiveness: How could you assess the impact of marketing campaigns on customer retention and acquisition within the dataset? What extra information would you need to solve this?

**Approach :**

To assess the impact of marketing campaigns on customer retention and acquisition, you'll need to analyze how customers' behaviors and statuses (churned, retained, or acquired) are influenced by these campaigns. We have to define those key metrics .

key metrics :

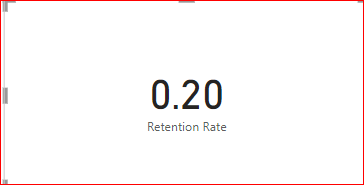
**Retention Rate**:

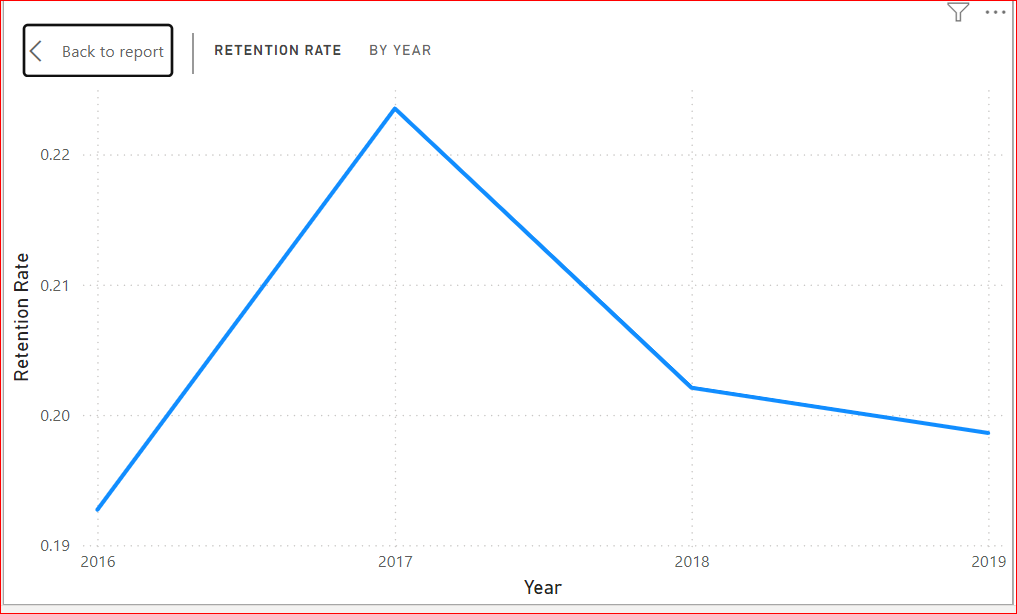
DIVIDE(

COUNTROWS(FILTER(Bank\_Churn, Bank\_Churn[Exited or retained ] = "Yes")),

COUNTROWS(Bank\_Churn)

)

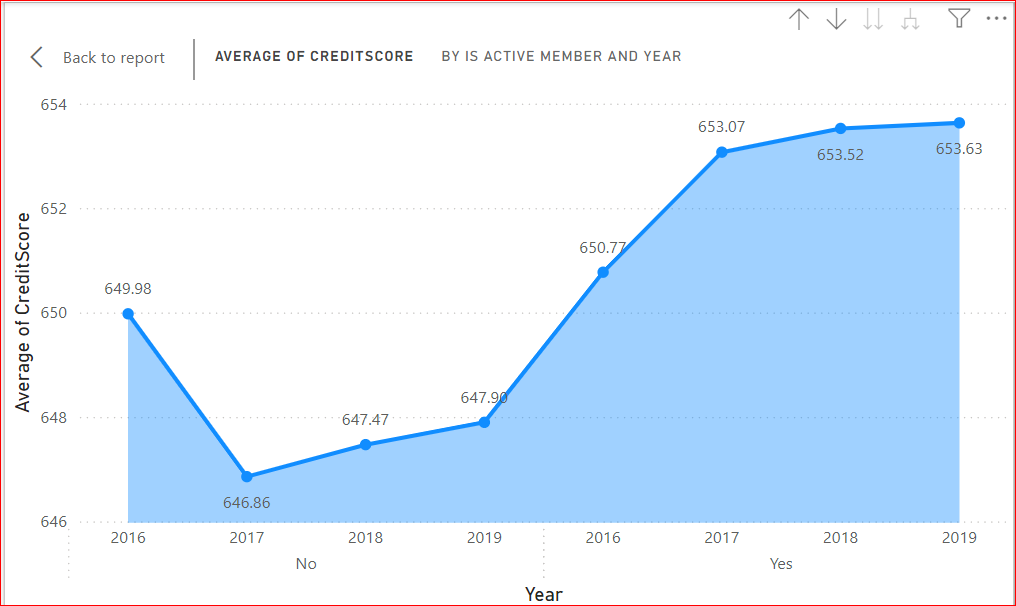


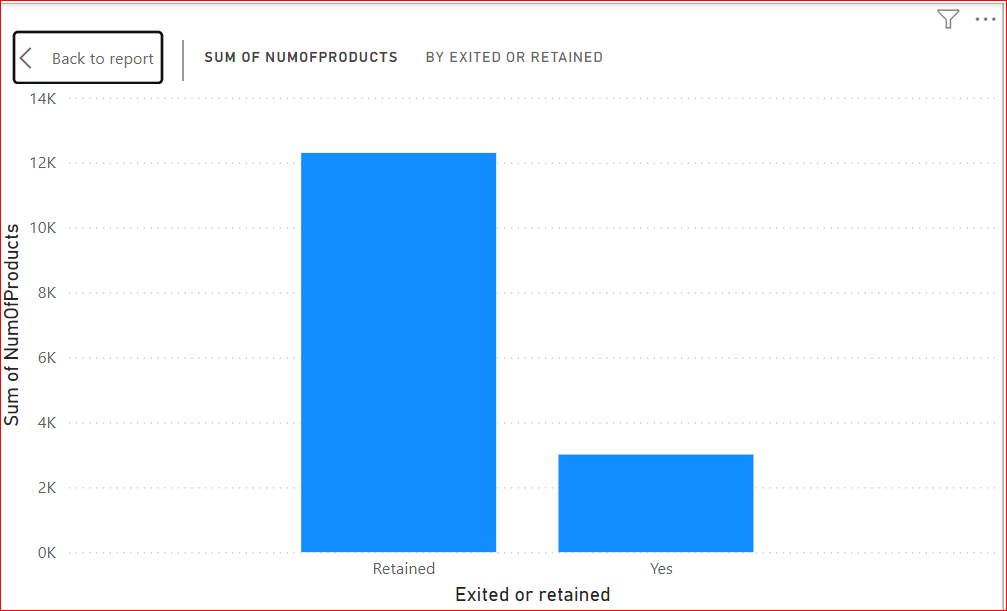


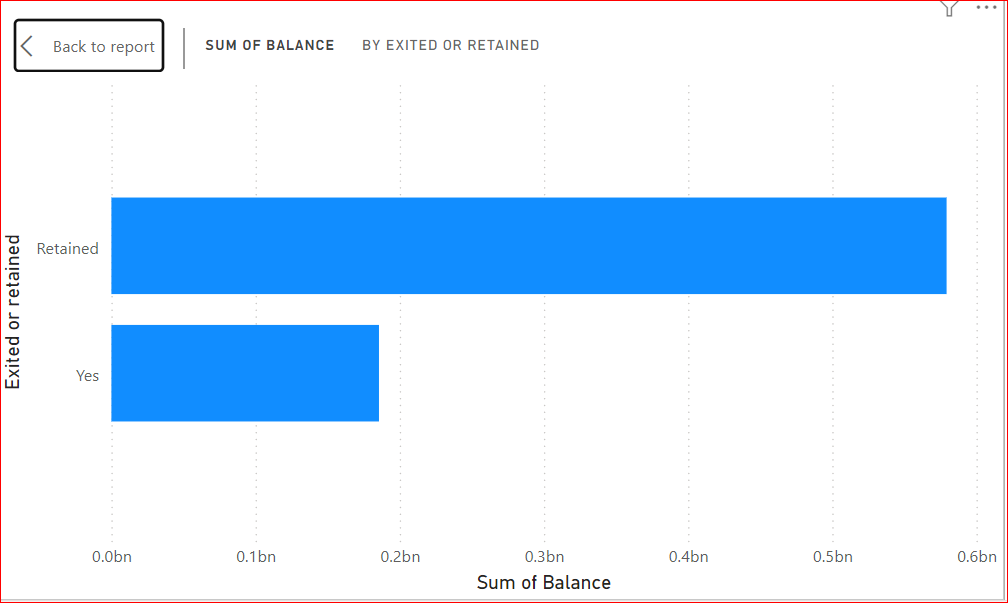
The extra information we are going to need is of Marketing campaigns in the data. If we have those information we can derive some comparative insights about before and after those marketing compaigns. Acquisition information , Churned information .

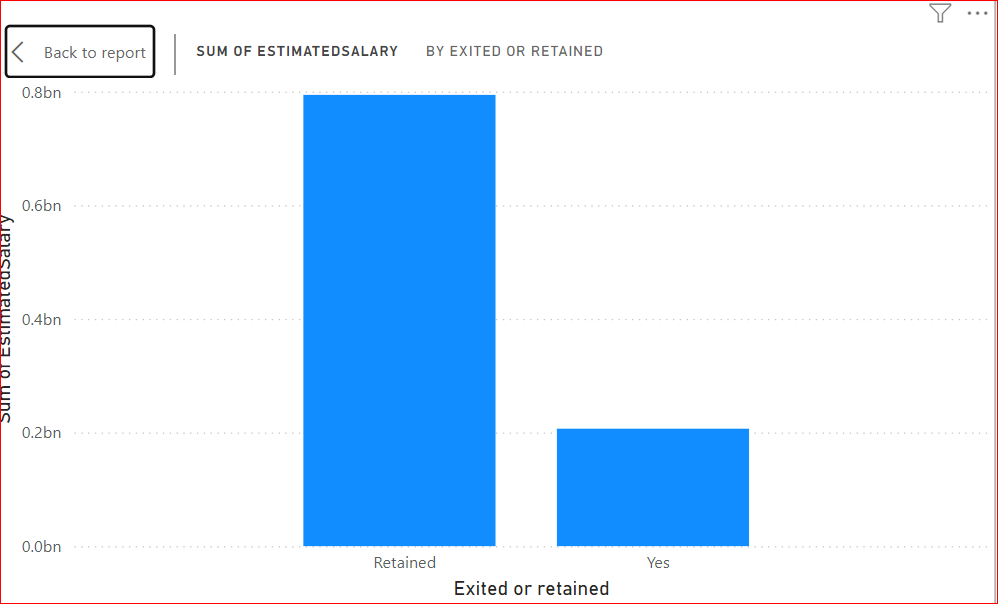
1. Customer Exit Reasons Exploration: Can you identify common characteristics or trends among customers who have exited that could explain their reasons for leaving?

To identify common characteristics or trends among customers who have exited, you can analyze various attributes of these customers and compare them with those who have stayed. This analysis will help uncover patterns that could explain why customers are leaving. Here’s how you can approach this:





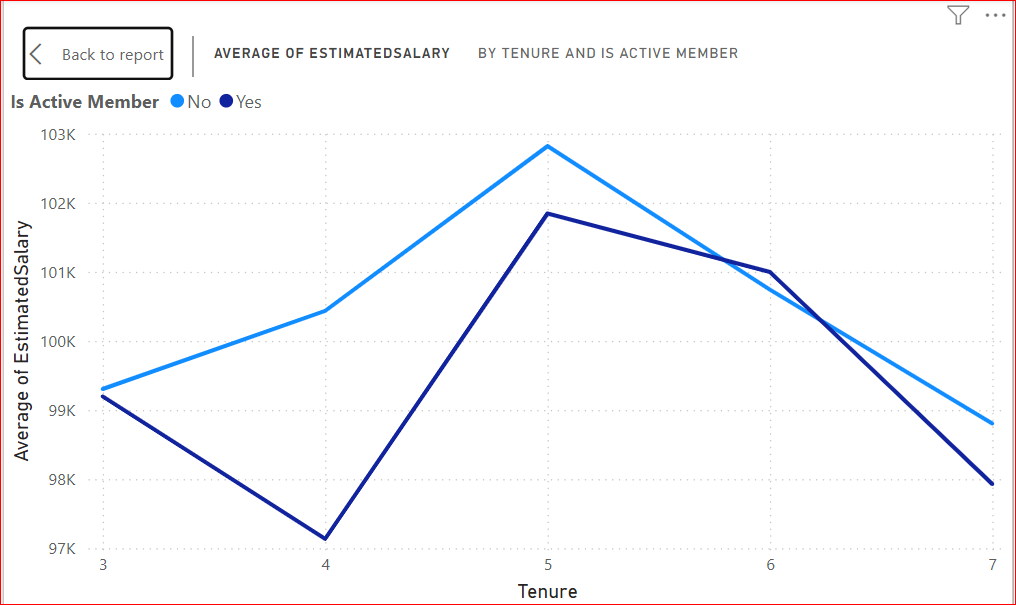


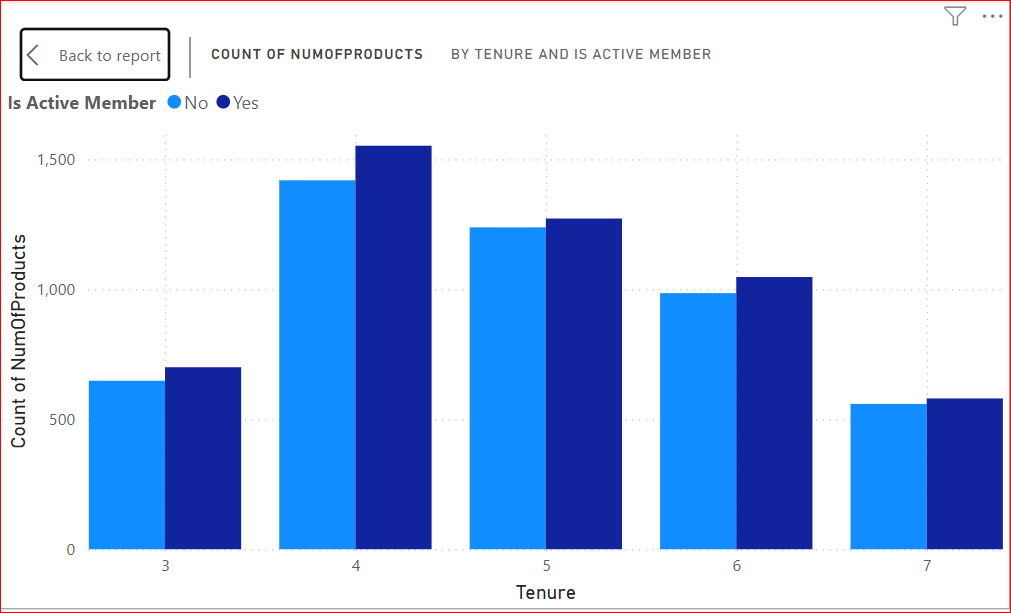


As we seen from the above reports/visualizations that the customers have exited are usually having less number of resources comparing to the one who are retained .

They have less number of products , their balance , estimated salary , credit score all are less comparing to the one who are retained.   
It may be tough for them to to be the bank customer with less number of resources and there is some charges are levied on by the bank as well. It can be one of the reason for them to exit .

1. Are 'Tenure', 'NumOfProducts', 'IsActiveMember', and 'EstimatedSalary' important for predicting if a customer will leave the bank?





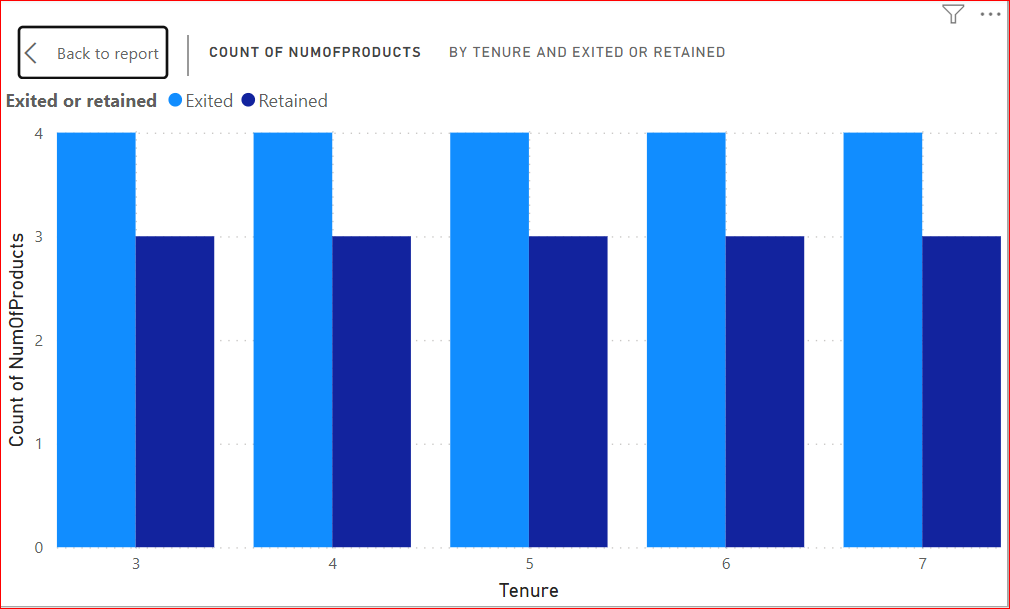
Yes, all of the factors mentioned (Tenure, NumOfProducts, IsActiveMember, and EstimatedSalary) are likely important for predicting if a customer will leave the bank (customer churn).

**Tenure:** Customers with a longer tenure at the bank have a greater chance of being familiar with the products and services offered, and may have developed a sense of loyalty. This makes them less likely to churn.

**NumOfProducts:** Customers who use a variety of the bank's products (accounts, credit cards, etc.) are more invested in the bank's ecosystem. This increased investment can make them less likely to switch to another bank.

**IsActiveMember**: Active members who regularly use the bank's services (online banking, transactions, etc.) are more likely to find value in the bank's offerings. This can make them less susceptible to churn compared to inactive members.

**EstimatedSalary:** Customers with a higher estimated salary might have a greater banking need and potentially a higher balance. This can make them more valuable to the bank and potentially influence efforts to retain them. However, it's important to note that a high salary might also indicate the ability to switch to a competitor offering premium services.



1. Utilize SQL queries to segment customers based on demographics and account details.

**Segmentation by Age and Geography**

select

b.ï»¿CustomerId as CustomerID,

c.Surname,

c.Age,

g.GenderCategory,

ge.GeographyLocation,

b.Balance,

b.NumOfProducts,

b.Tenure,

case when c.Age < 25 then 'Youth'

when c.Age Between 25 and 40 then 'Adult'

when c.Age Between 41 and 60 then 'Middle Aged'

else 'Senior'

end as AgeSegment ,

case when b.IsActiveMember = True then 'Yes'

when b.IsActiveMember = false then 'No'

else 'NA'

end as ActiveMember,

case when b.Exited = true then 'Yes'

when b.Exited = false then 'No'

else 'NA'

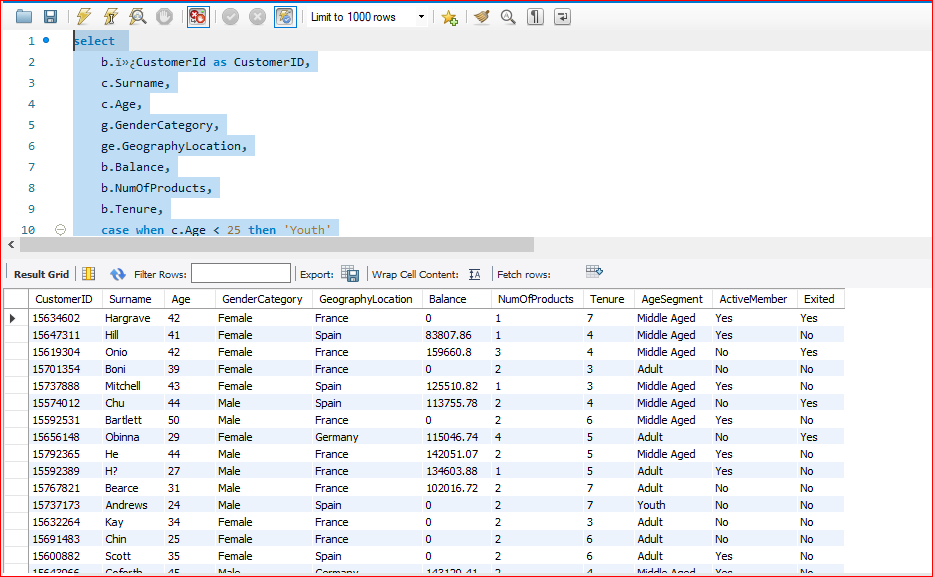
end as Exited

from bank\_churn b

join customerinfo c on c.ï»¿CustomerId = b.ï»¿CustomerId

join gender g on c.`GenderID` = g.`ï»¿GenderID`

join geography ge on ge.`ï»¿GeographyID` = c.`GeographyID`



**Combining multiple Segmentation Balance Segment ,Product Segment, Age Segment ,Active Member , Exit Segment**

select

b.ï»¿CustomerId as CustomerID,

c.Surname,

c.Age,

g.GenderCategory,

ge.GeographyLocation,

case when b.Balance < 50000 then 'Less than 50000'

when b.Balance between 50000 and 100000 then 'Between 50-100k'

when b.Balance between 100000 and 150000 then 'Between 100-150k'

when b.Balance between 150000 and 200000 then 'Between 150-200k'

else 'Over 200k'

end as BalanceSegment ,

case when b.NumOfProducts = 1 then 'Single Product'

when b.NumOfProducts = 2 then 'Two Products'

when b.NumOfProducts = 3 then 'Three Products'

when b.NumOfProducts = 4 then 'More than 3 Products'

else 'NA'

end as ProductSegment,

b.Tenure,

case when c.Age < 25 then 'Youth'

when c.Age Between 25 and 40 then 'Adult'

when c.Age Between 41 and 60 then 'Middle Aged'

else 'Senior'

end as AgeSegment ,

case when b.IsActiveMember = True then 'Yes'

when b.IsActiveMember = false then 'No'

else 'NA'

end as ActiveMember,

case when b.Exited = true then 'Yes'

when b.Exited = false then 'No'

else 'NA'

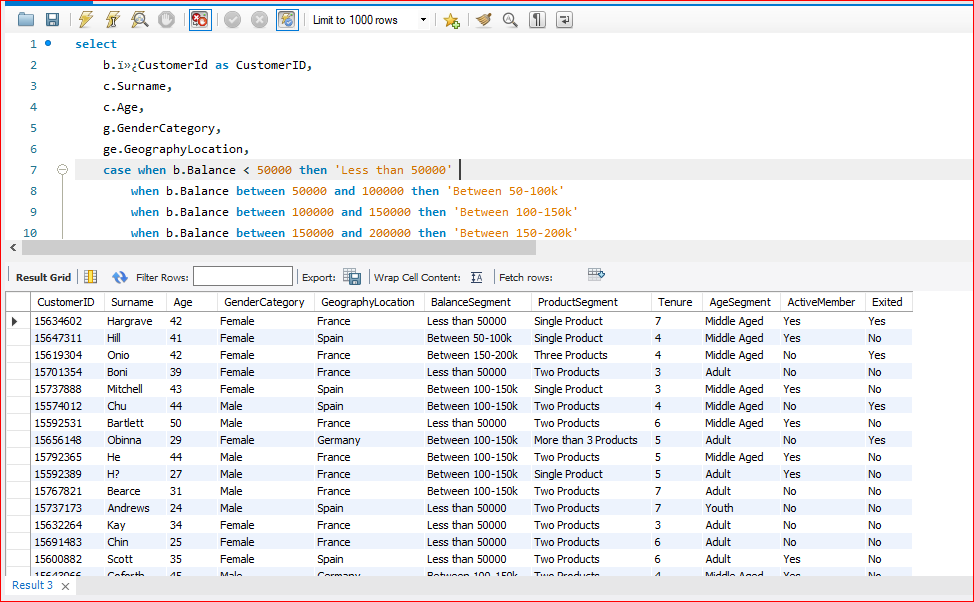
end as Exited

from bank\_churn b

join customerinfo c on c.ï»¿CustomerId = b.ï»¿CustomerId

join gender g on c.`GenderID` = g.`ï»¿GenderID`

join geography ge on ge.`ï»¿GeographyID` = c.`GeographyID`



1. How can we create a conditional formatting setup to visually highlight customers at risk of churn and to evaluate the impact of credit card rewards on customer retention?

**Approach:**

Firstly we have to calculate the churn risk , we can deduce from the data that if any customer is having a tenure of less than two years and he is having only one product and also is not an active member , he’s likely to be leaving the bank Therefore it’s safe to say that the churn risk here is high.

If he is having a tenure of less than two years and he is having only one product but an active member ,the churn risk is medium and in other cases it’s low.

This is what we have done from the below SQL query.

select

\* ,

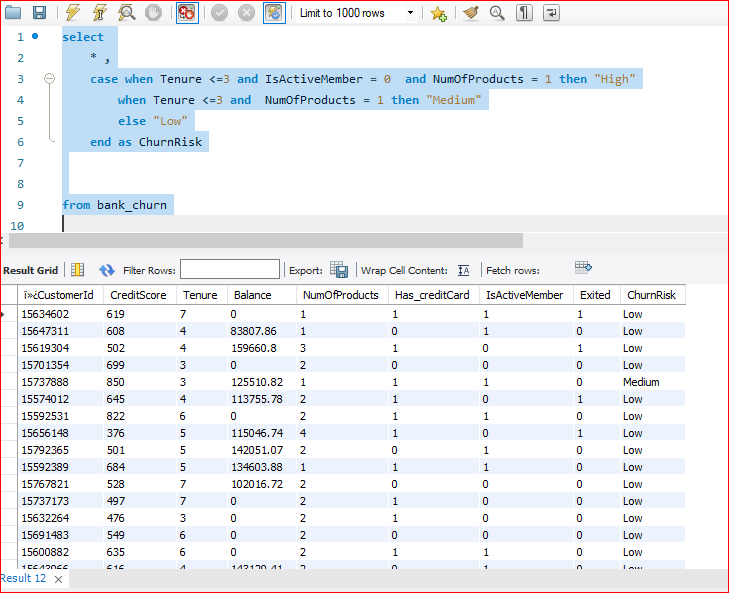
case when Tenure <=3 and IsActiveMember = 0 and NumOfProducts = 1 then "High"

when Tenure < =3 and NumOfProducts = 1 then "Medium"

else "Low"

end as ChurnRisk

from bank\_churn



We got the table now we have exported it and will load the same on the power BI .

Yes we can do it in powerBI as well but there seems to be some issue so I’ve exported it from SQL .

Then we have calculated the Retention Rate with or without credit cards using DAX function, but in order to calculate the rate , we need to calculate the retention with or without rewards and the total with or without rewards.

Then only we can calculate the rate ., we need to calculate the retention with or without rewards and the total with or without rewards.

Then only we can calculate the rate .

**Total With Rewards** = CALCULATE(COUNTROWS('Bank\_Churn'),Bank\_Churn[Has Credit Card] = "Yes")

**Total without Rewards =** CALCULATE(COUNTROWS('Bank\_Churn'),Bank\_Churn[Has Credit Card] = "No")

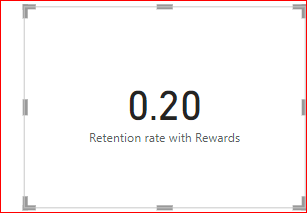
**Retention without rewards** = CALCULATE(COUNTROWS(Bank\_Churn),Bank\_Churn[Has Credit Card] = "No" ,Bank\_Churn[Exited or retained] = "Retained")

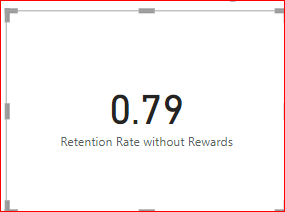
**Retention with rewards** =

CALCULATE( COUNTROWS('Bank\_Churn'),Bank\_Churn[Has Credit Card] = "Yes",Bank\_Churn[Exited or retained] = "Exited")

**Retention Rate without Rewards** =

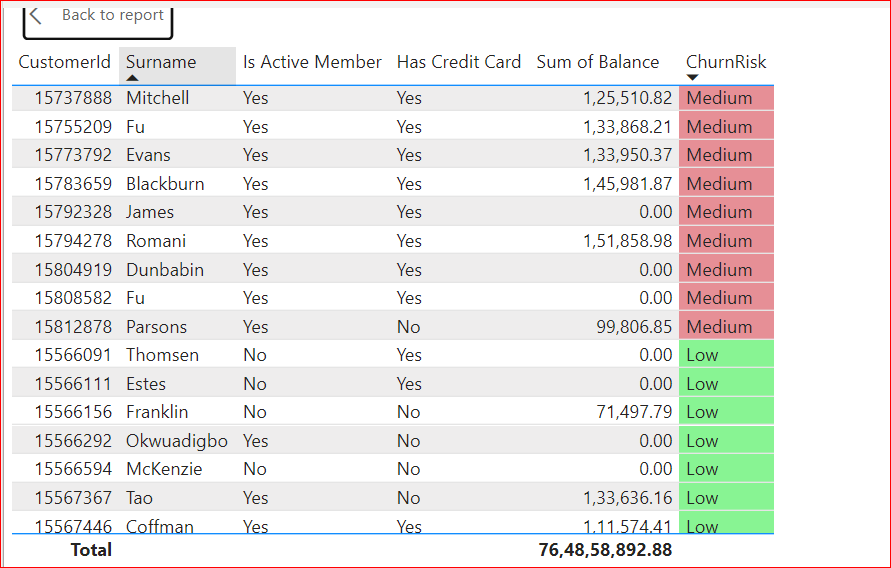
DIVIDE([Retention without rewards],[Total without Rewards])





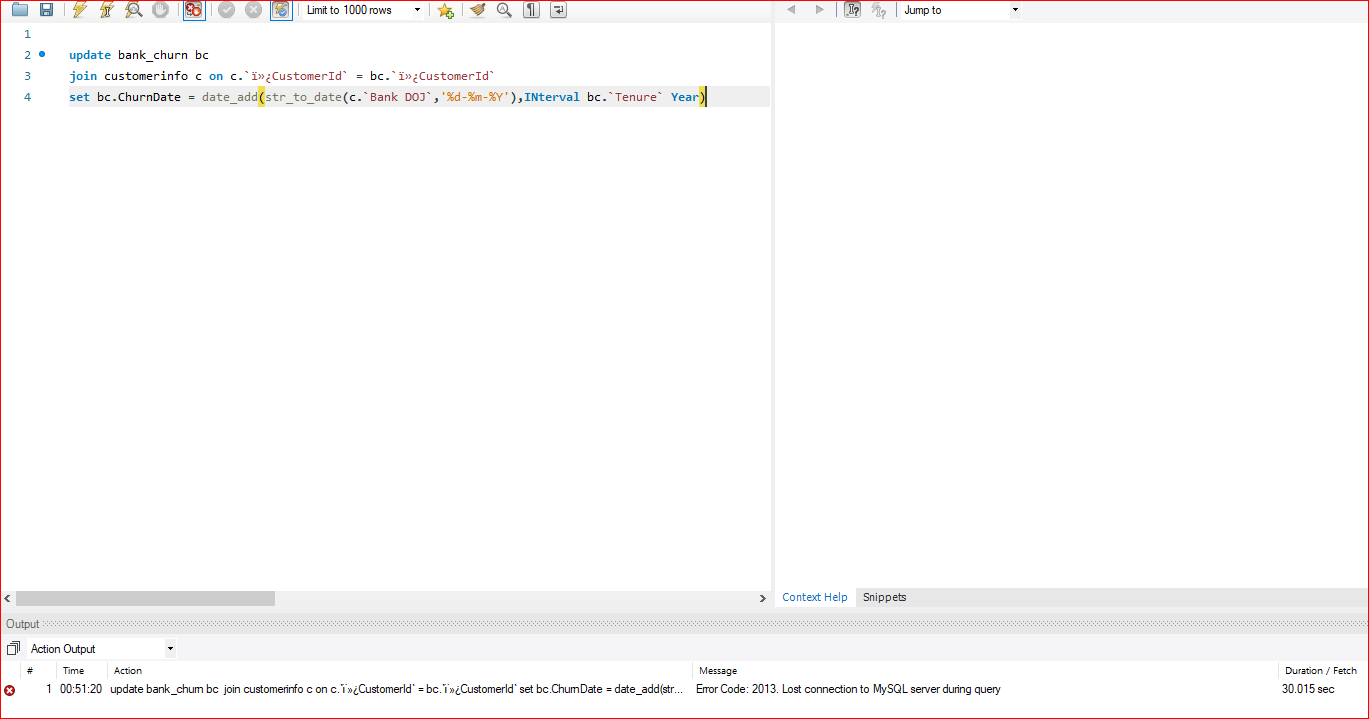
We have created a table visual and selected the required fields in the table visual and in the visual setting of the table reported ,under section Cell Elements ,selected the series as Churn risk and defined the rules like

High risk will be in red color , low will be in somewhat green .



1. What is the current churn rate per year and overall as well in the bank? Can you suggest some insights to the bank about which kind of customers are more likely to churn and what different strategies can be used to decrease the churn rate?

We tried to add the exit year by adding a new column in the bank churn table but somehow it didn’t work



So we have done it in PowerBI using DAX functions

ChurnDate =

VAR JoinDate =

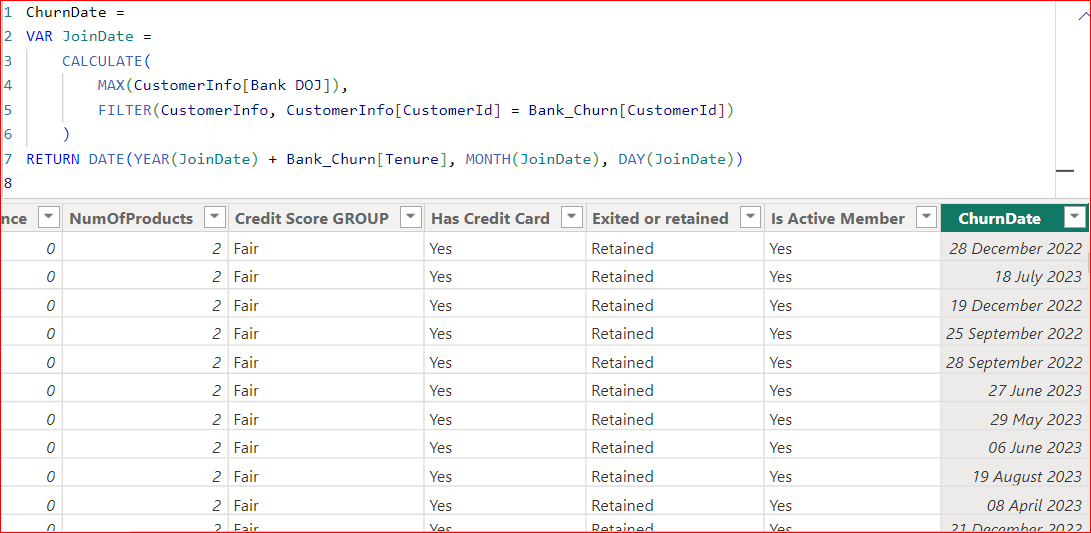
CALCULATE(

MAX(CustomerInfo[Bank DOJ]),

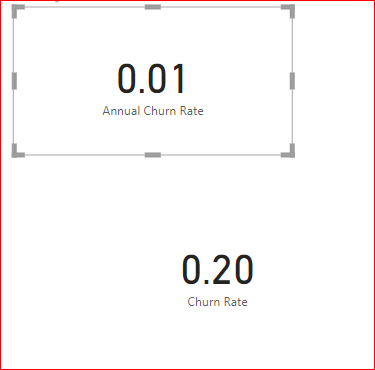
FILTER(CustomerInfo, CustomerInfo[CustomerId] = Bank\_Churn[CustomerId])

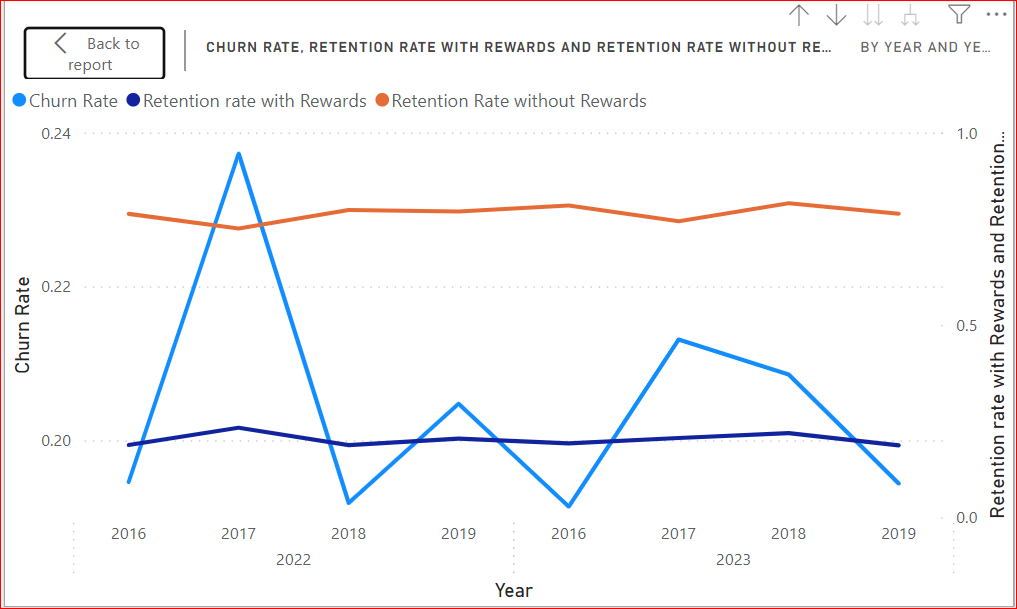
)

RETURN DATE(YEAR(JoinDate) + Bank\_Churn[Tenure], MONTH(JoinDate), DAY(JoinDate))

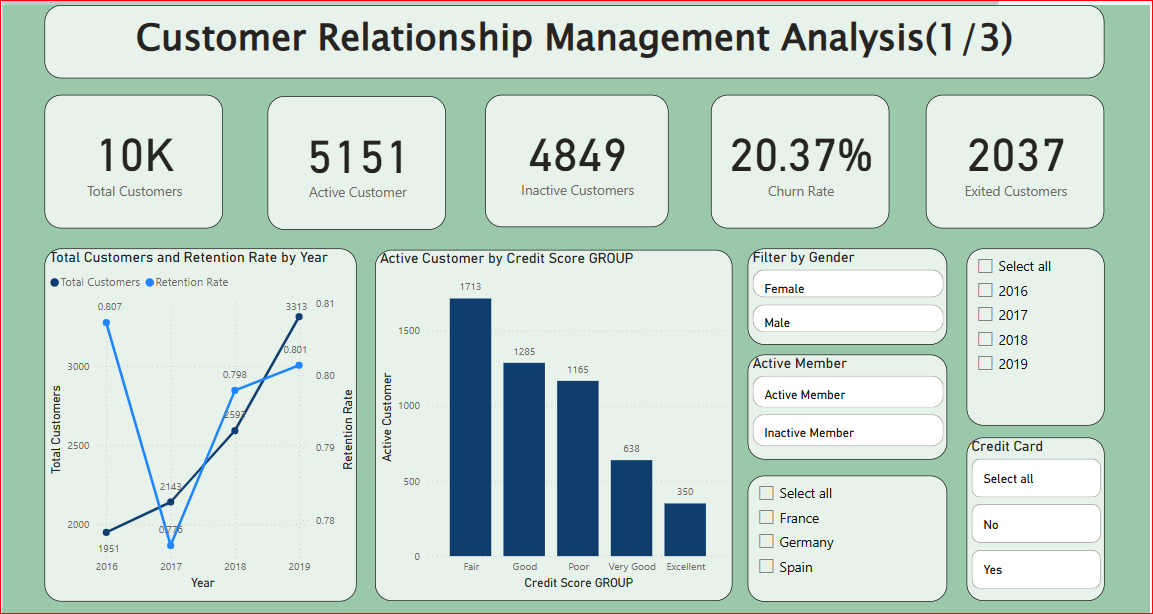


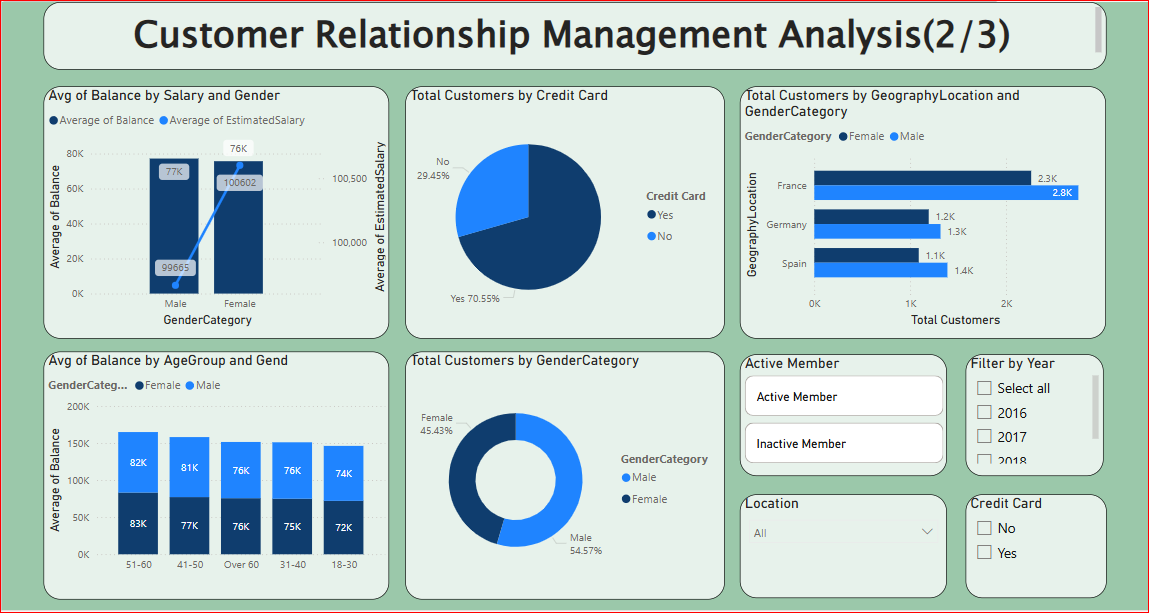
**Annual Churn Rate =** DIVIDE(COUNTROWS(Bank\_Churn),CALCULATE(COUNTROWS(Bank\_Churn),ALL(Bank\_Churn[ChurnDate]))\* 100)

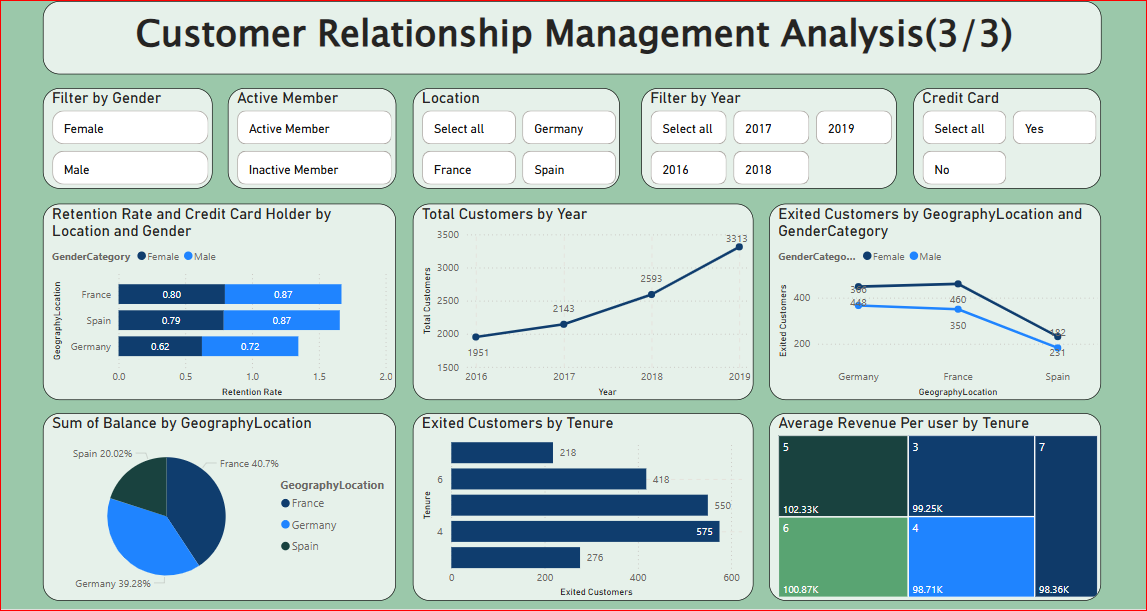




1. Create a dashboard incorporating all the KPIs and visualization-related metrics. Use a slicer in order to assist in selection in the dashboard.







1. How would you approach this problem, if the objective and subjective questions weren't given?

Approaching a problem without predefined objective and subjective questions can be tricky and complex as it involves a structured and strategic method to identify key areas of focus.

Certainly, We won’t be able to pull this much information if these pre-defined objective and subjective questions were not given as it have a wide range of question from Analytical approach to hypothesis . But even if the question weren’t given ,we can certainly extract information such as

* *Average balance across different geographical regions and how it’s expanding over the time,*
* *Distribution of wealth among male and female*
* *Customers with highest estimated salary/balance on gender basic and ranking them in descending order*
* *How many active customers are there*
* *Churn rate and new customers over the year and so on.*

*And so on…*

1. In the “Bank\_Churn” table how can you modify the name of the “HasCrCard” column to “Has\_creditcard”?

alter table bank\_churn

rename column HasCrCard to Has\_creditCard;

select \* from bank\_churn;

