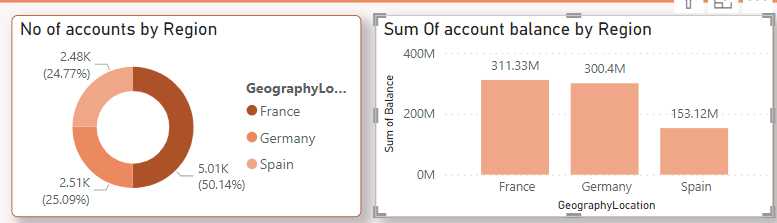
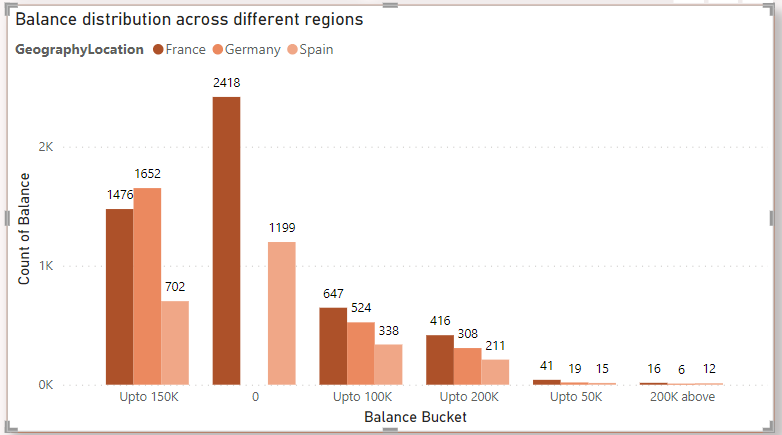
**Objective Questions:**

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

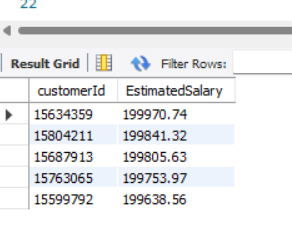




* As we can see that the France have the highest no of accounts but if we see the sum of balance for each country France and Germany have almost similar total balance despite of France having highest no of accounts.
* To further analyse this, I have created balance bucket and identifies that Germany doesn’t allow zero balance account While France and Spain do allow that and Germany have most of the customers with balance Upto150K.

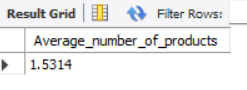
1. **Identify the top 5 customers with the highest Estimated Salary in the last quarter of the year. (SQL)**

Output-

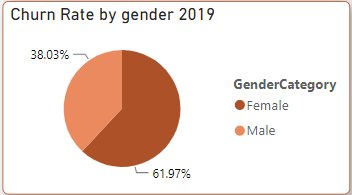


1. **Calculate the average number of products used by customers who have a credit card. (SQL)**

Output

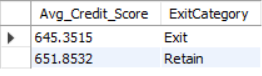


1. **Determine the churn rate by gender for the most recent year in the dataset.**



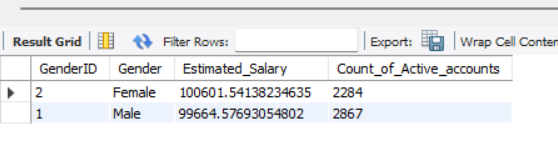
1. **Compare the average credit score of customers who have exited and those who remain. (SQL)**

Output



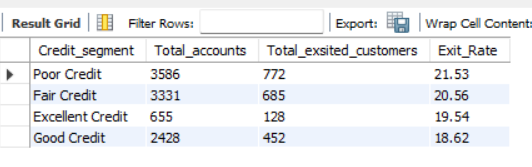
1. **Which gender has a higher average estimated salary, and how does it relate to the number of active accounts? (SQL)**

Output

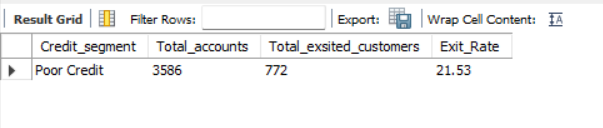


1. **Segment the customers based on their credit score and identify the segment with the highest exit rate. (SQL)**

Output-

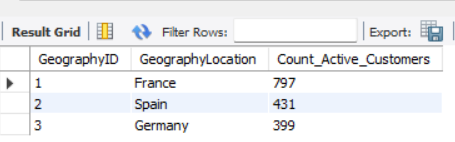


Highest Exit rate –

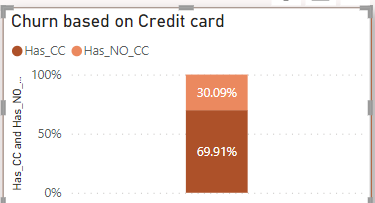


1. **Find out which geographic region has the highest number of active customers with a tenure greater than 5 years. (SQL)**

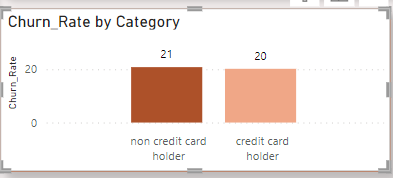
Output



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

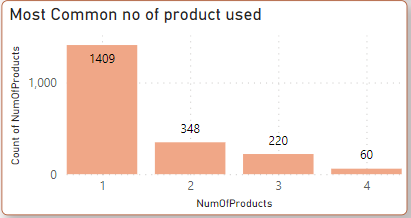


* Based on all the data 70% Customers have CC while 30% of the Customers don’t have CC



* As we can see there is very less difference between the churn rate of both but still Credit card holder churn rate is lower than the Non credit card holders.

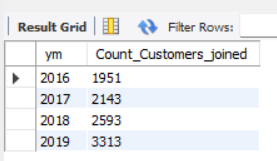
1. **For customers who have exited, what is the most common number of products they have used?**



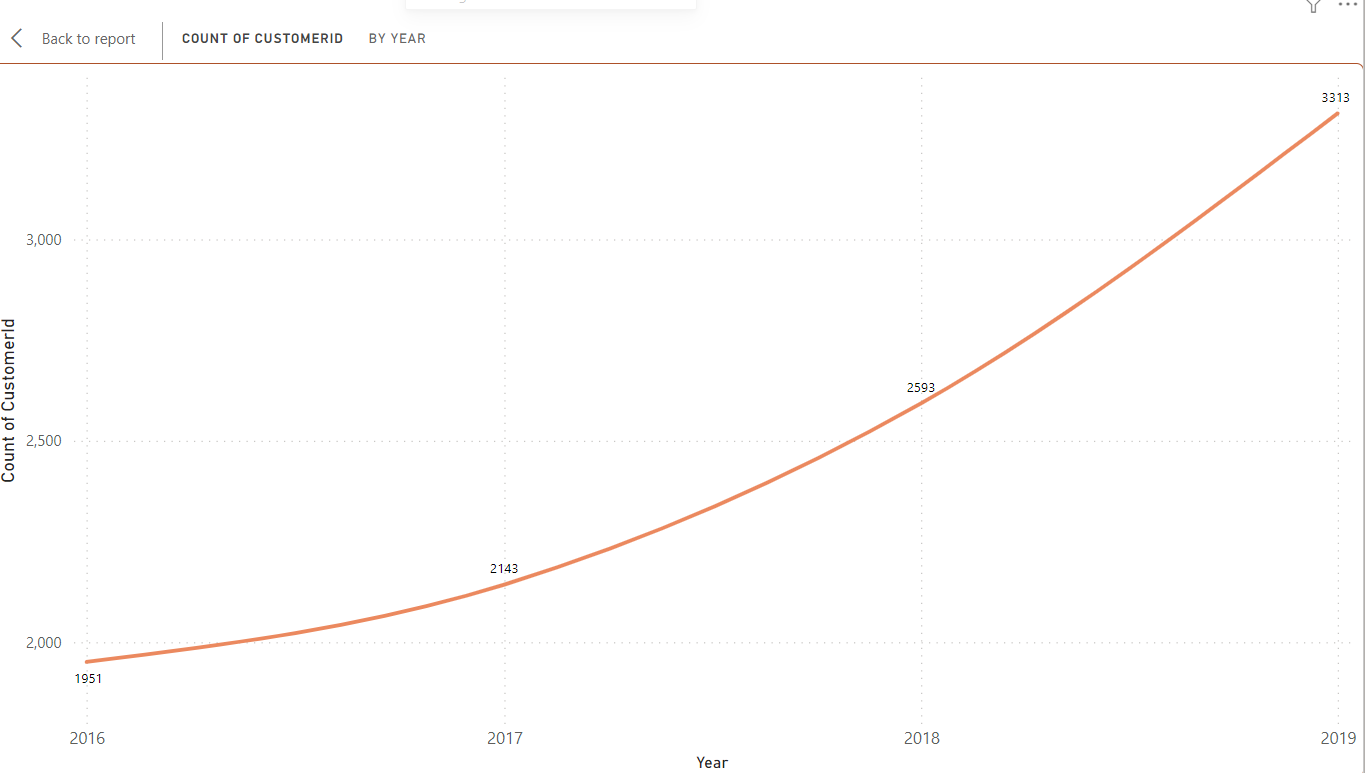
1. **Examine the trend of customers joining over time and identify any seasonal patterns (yearly or monthly). Prepare the data through SQL and then visualize it.**

* Yearly Trend of joining customers

Output

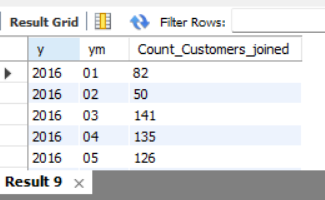


Visualization

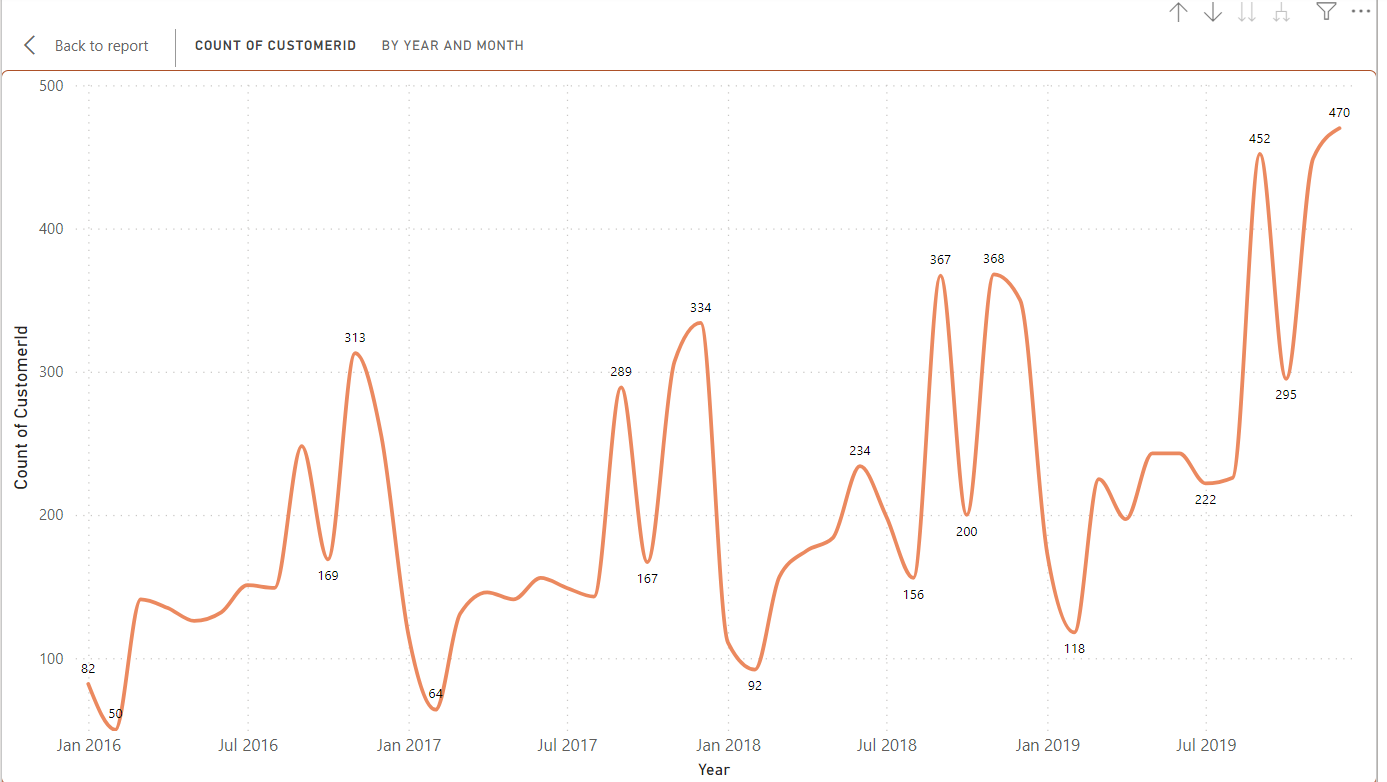


* No of customers are increasing year by year.
* Monthly Trend of joining customers for all 4 years

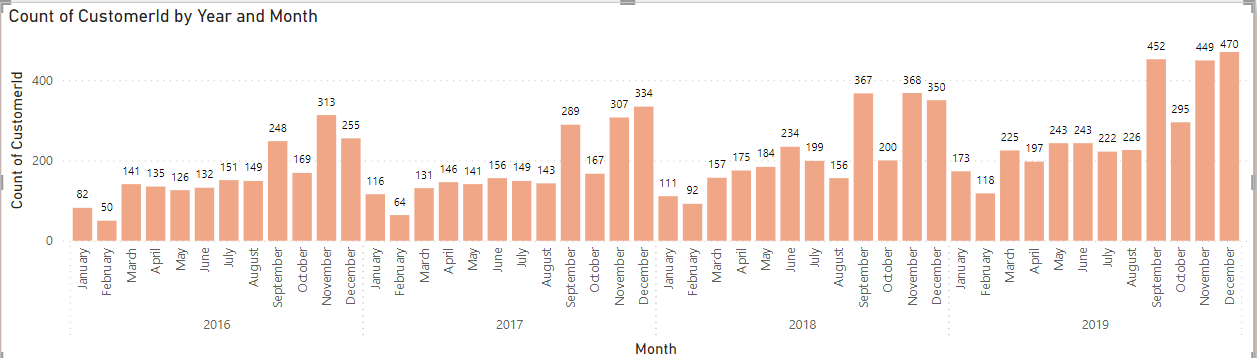
Output



Visualization



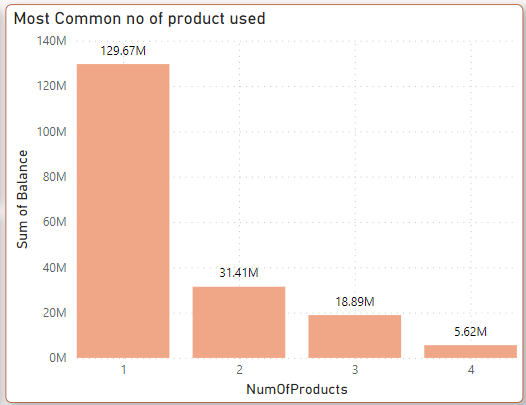
* As we can see in the graph that M shape means for every year same months the new customers are joining

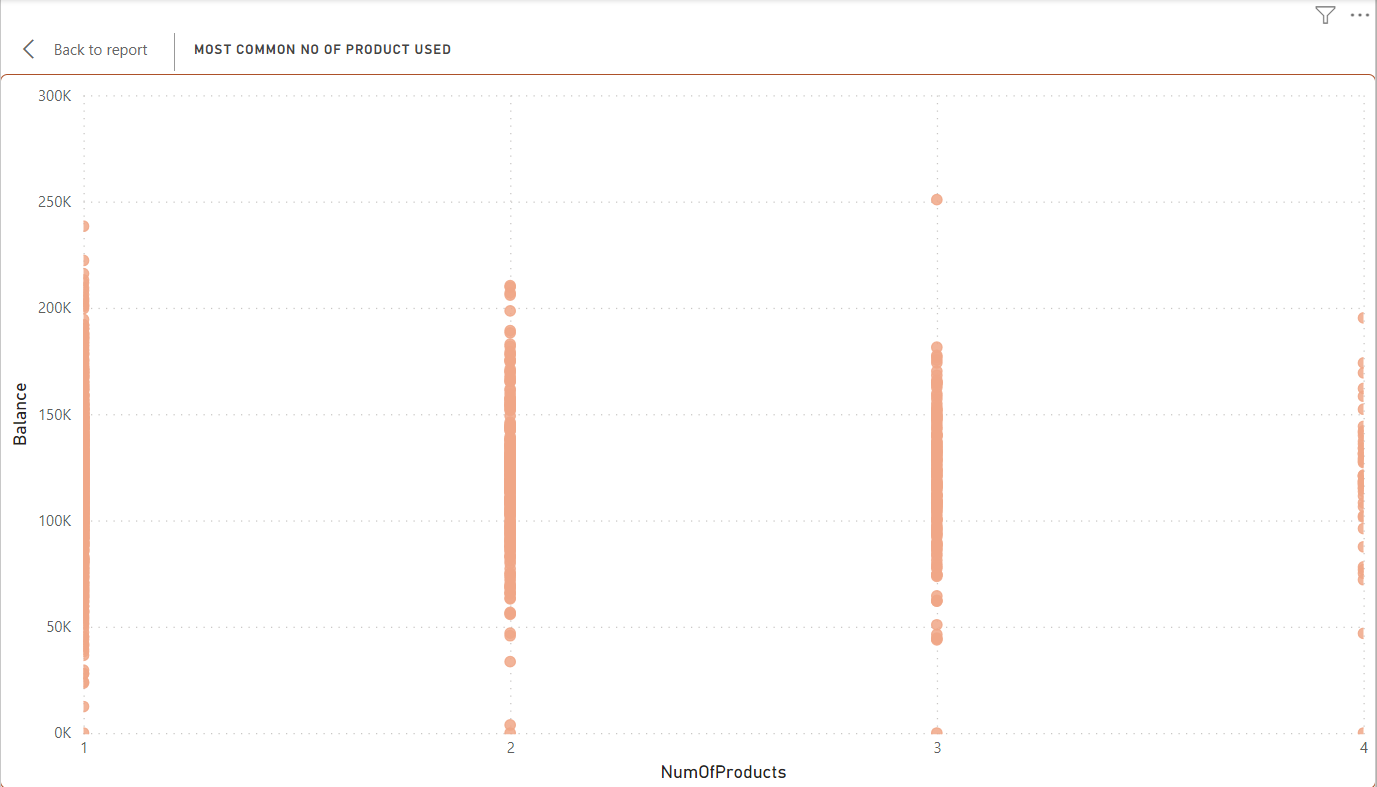


* This graph I just made for reference so that we can clearly see the months which are September, November and December are joining the bank.

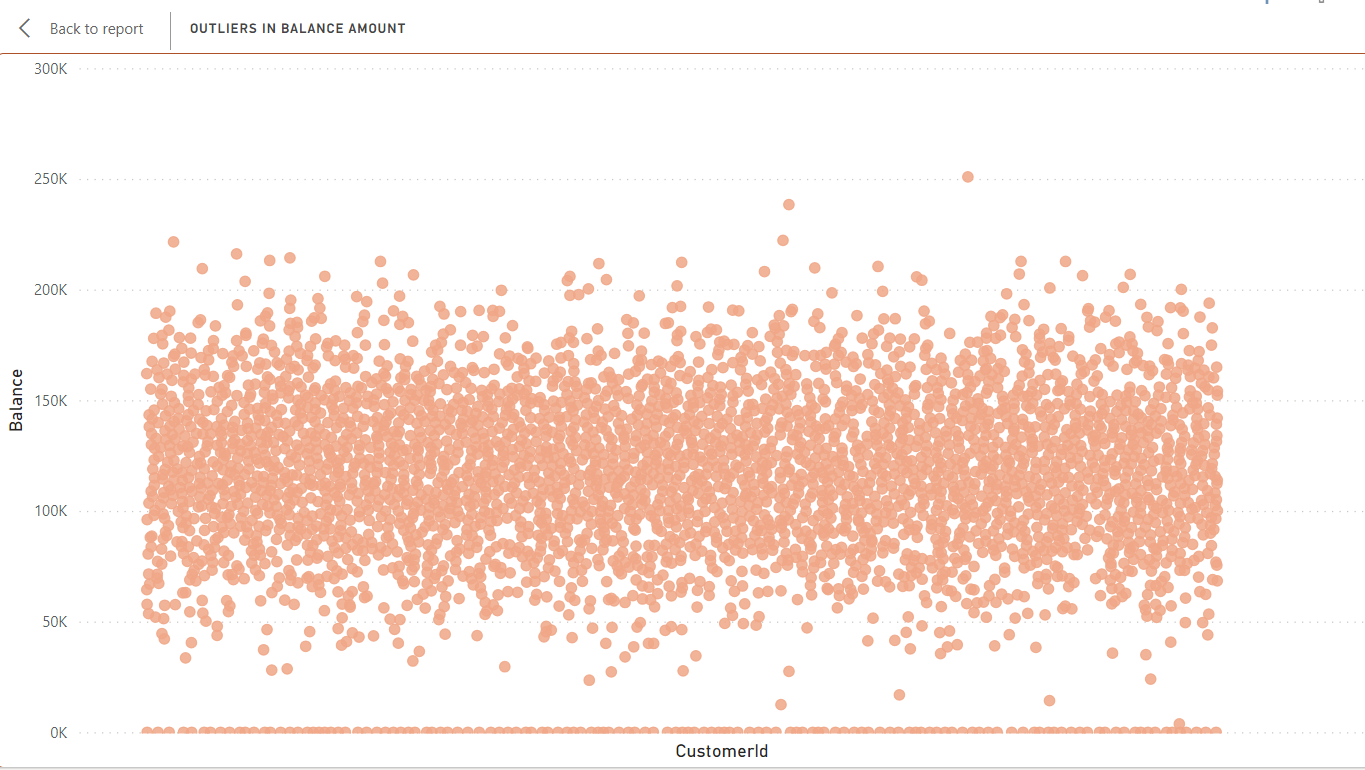
1. **Analyze the relationship between the number of products and the account balance for customers who have exited.**

* Not much difference most of the customers who have exited had used only single product and sum of those exited customers balance is high.





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



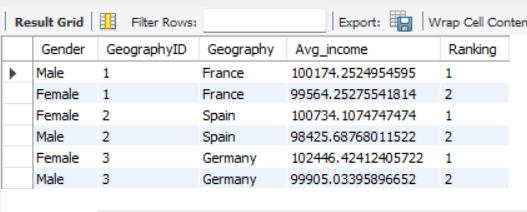
* Customer id who has 0K balance and greater than 200k balance are the potential outliers.

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

* There is total 7 table in the dataset and they are connected to each other to form meaning full relationship between all the column values.
* In these tables we have 5 categorical tables have name as below
* ActiveCustomer
* CreditCard
* ExitCustomer
* Gender
* Geography

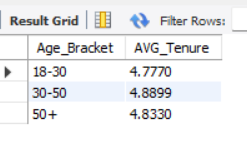
1. **Using SQL, write a query to find out the gender-wise average income of males and females in each geography id. Also, rank the gender according to the average value. (SQL)**

Output

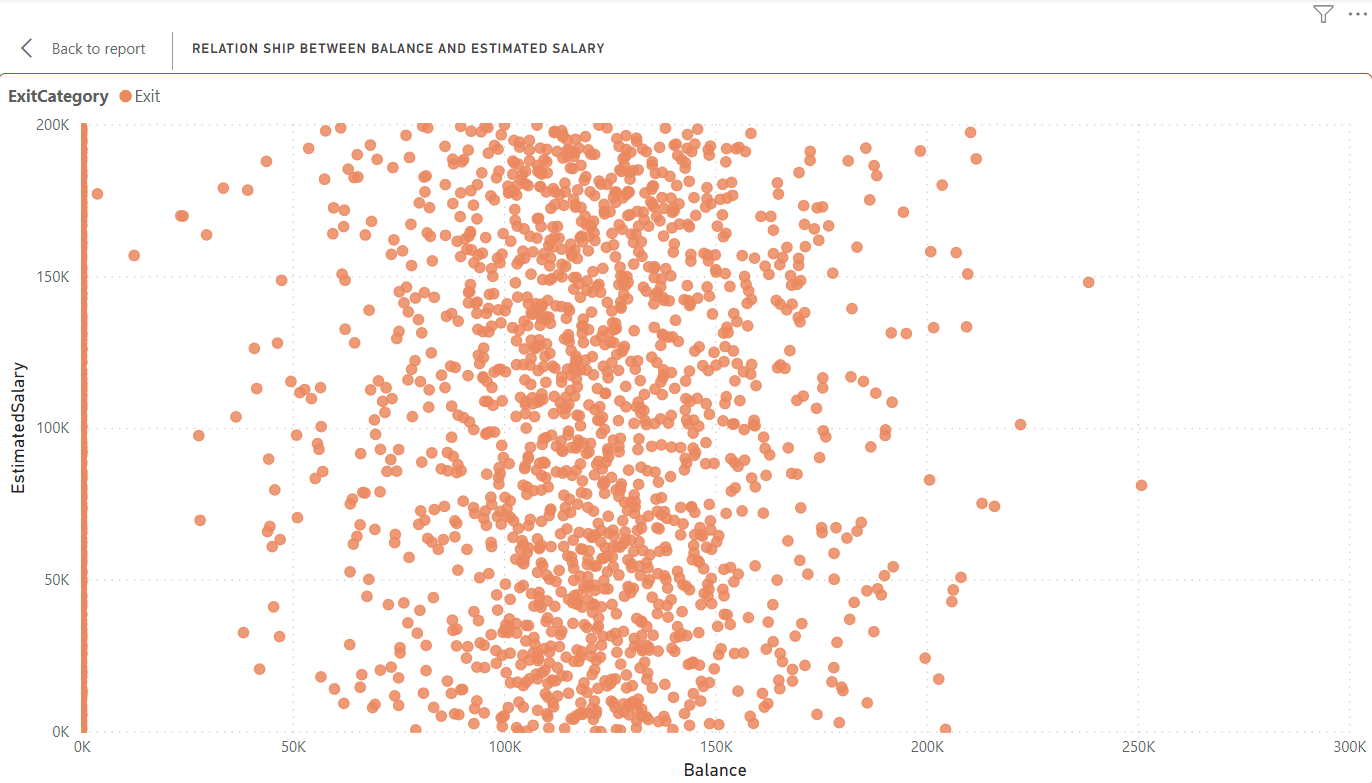
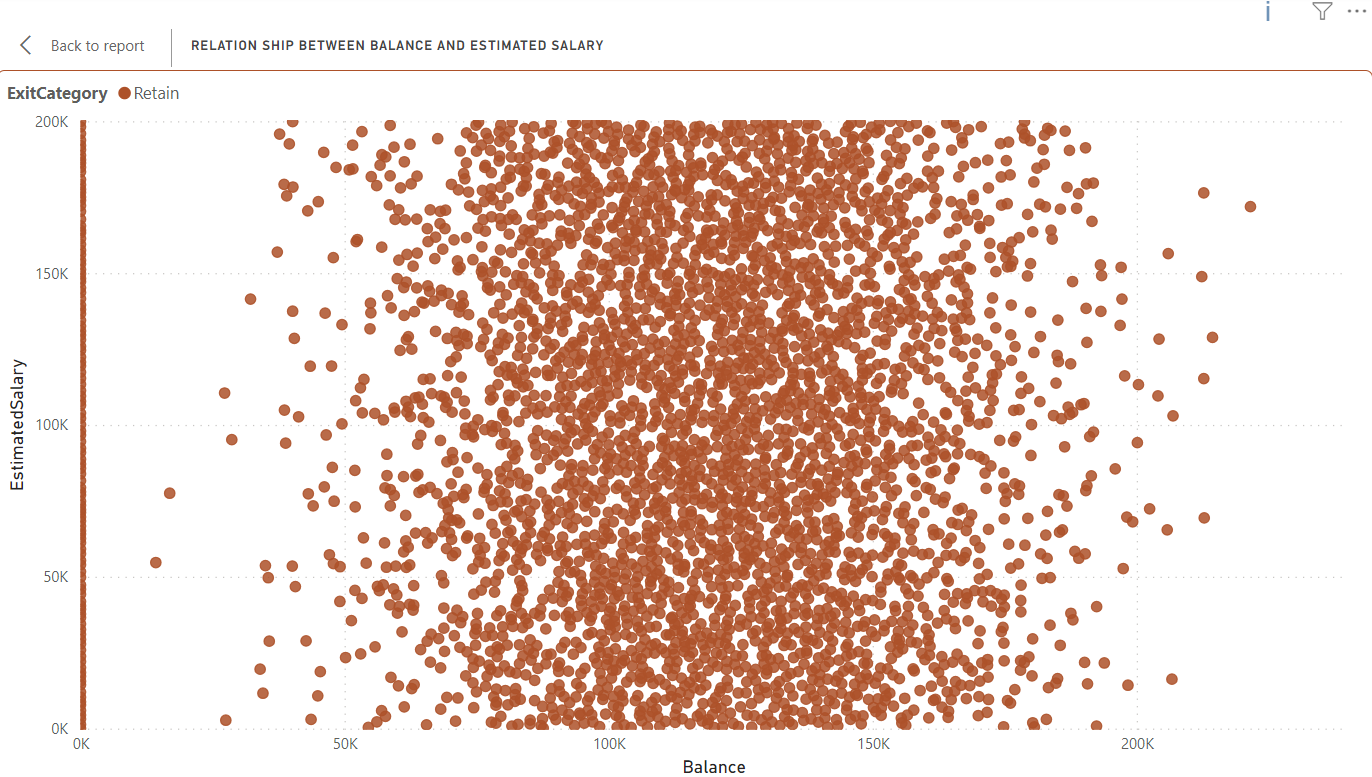


1. **Using SQL, write a query to find out the average tenure of the people who have exited in each age bracket (18-30, 30-50, 50+).**

Output

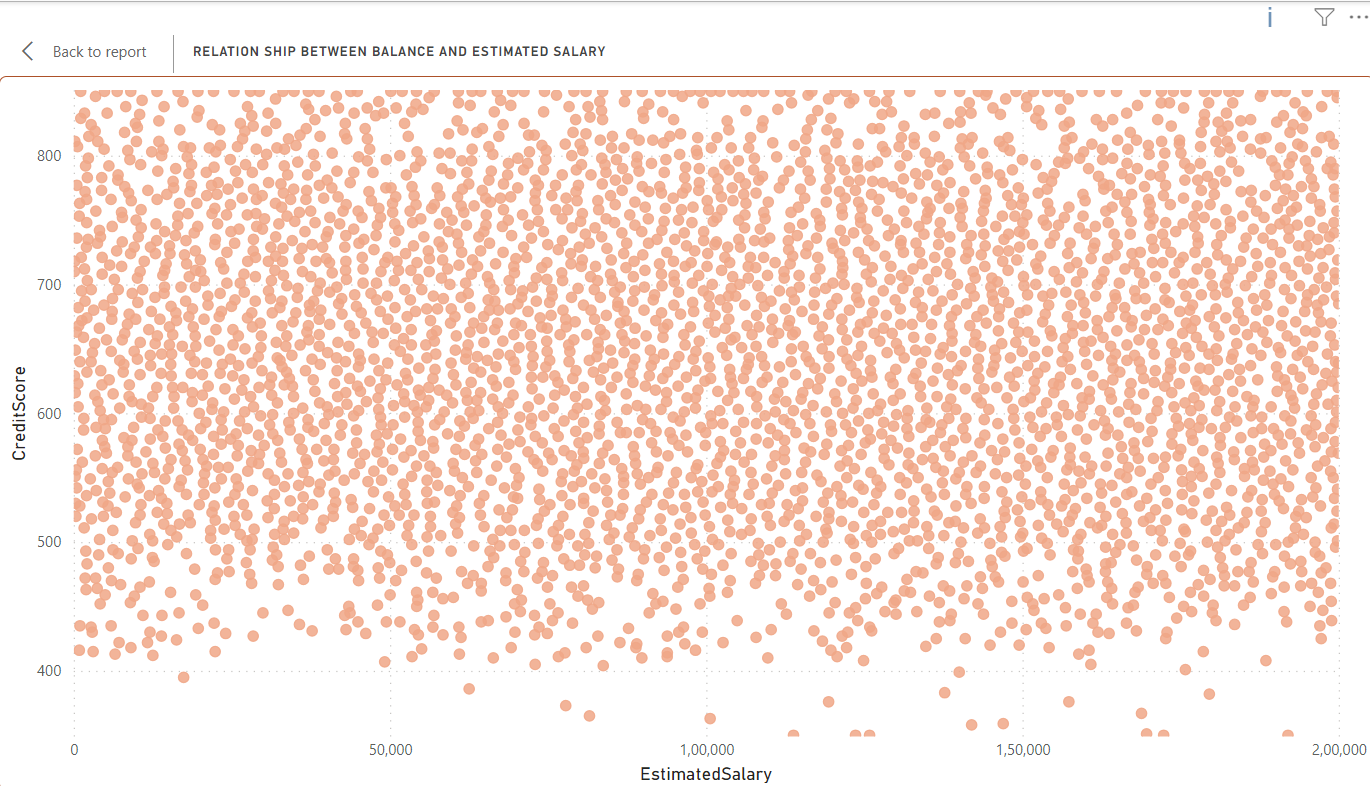


1. **Is there any direct correlation between salary and the balance of the customers? And is it different for people who have exited or not?**

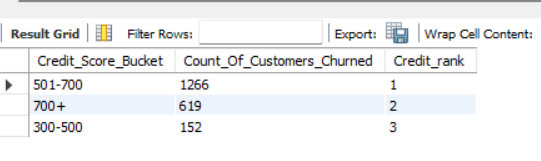
* For both the Exit and Retain Customers there was not much relation between the balance and salary.
* But there is one that based on salary people used to spend because at the end they need to pay the bills.

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

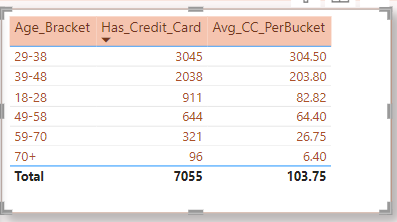


* Credit Card Score depends on various factors like spending habits , loan on your card , how much active you are etc but salary value is helpful to get the credit card the higher the salary is the higher credit card limit you can get.

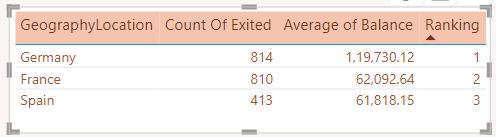
1. **Rank each bucket of credit score as per the number of customers who have churned the bank.**



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.**



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

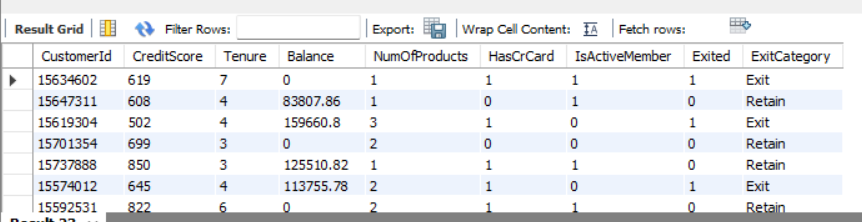


1. **As we can see that the “CustomerInfo” table has the CustomerID and Surname, now if we have to join it with a table where the primary key is also a combination of CustomerID and Surname, come up with a column where the format is “CustomerID\_Surname”.**



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

Output-



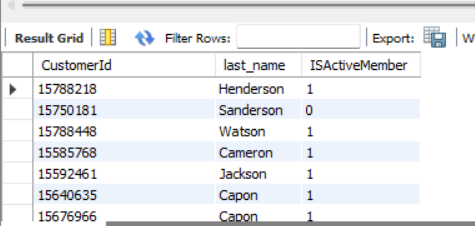
1. **Were there any missing values in the data, using which tool did you replace them and what are the ways to handle them?**

No, there was no missing value in the data, by using Power query tool we had confirmed that there were no missing values

If there were a missing value then in power query editor there is option of replacing a missing value with anything you want which can make more sense to the data or you can replace those missing value with imputed value using dax.

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”.**

Output



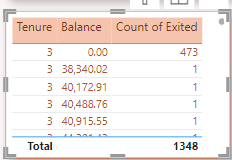
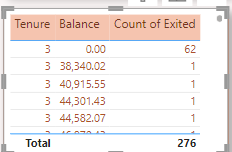
**Subjective Question:**

1. **Customer Behaviour 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?**

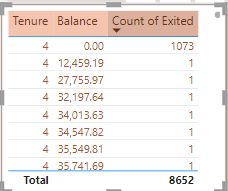
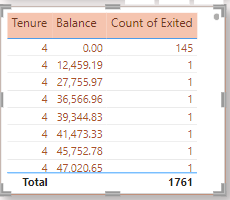
* Long term and new customers can be identified by tenure column
* As the average of tenure column is 4. I am taking customer with less than 4 as new customer and greater than 4 as a long-term customer.
* Spending habit of new customer is lesser than the long-term customer.
* Long term customers spend more.

**Churn Rate**

* The churn rate is higher in long term customers then new customers.
* New customer- there is total 1348 new customers and within that 276 has exited and in those 276 customers 62 has 0 balance accounts.

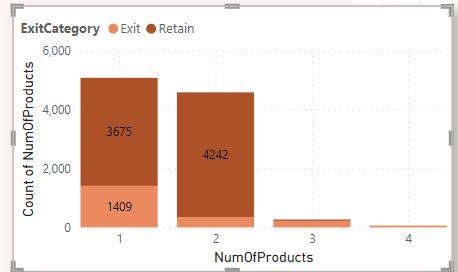
* Long term customers- there is total 8652 new customers and within that 1761 has exited and in those 1761 customers 62 has 0 balance accounts and 8652 has 1073 customers with balance account.

* No of product use- there is no difference in the useability of product between new and long-term customers.

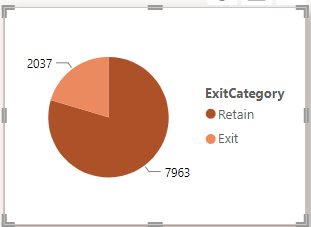
1. Product Affinity Study: Which bank products or services are most commonly used together, and how might this influence cross-selling strategies?

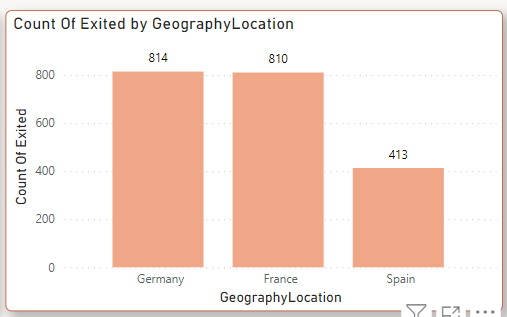
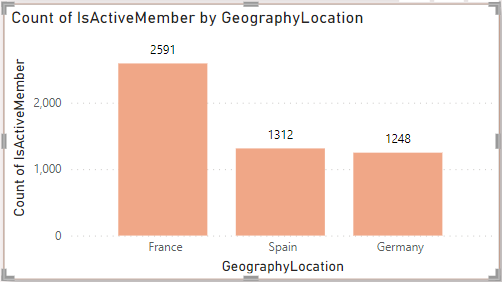
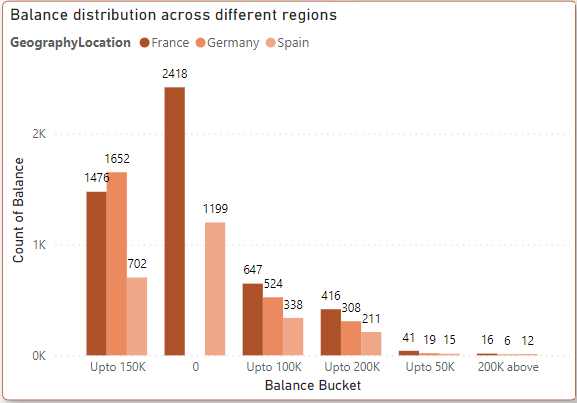
(Refer Repot 4 for the visualizations)



* 50% of the customers used only single product.
* 4590 customers used 2 products.
* Most of the customers how have exited the bank had used only 1 product.
* Customers who are with bank are using 2 products.
* We can target the customers who are using single product and let them know about the different products that are available with the bank.
* For example, if customers frequently use a checking account and have a mortgage, offering them a home equity line of credit might be a suitable cross-selling strategy.
* By conducting a Product Affinity Study and leveraging the insights gained, banks can enhance their cross-selling efforts, improve customer satisfaction, and drive revenue growth.

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



As we can see in the above graph that

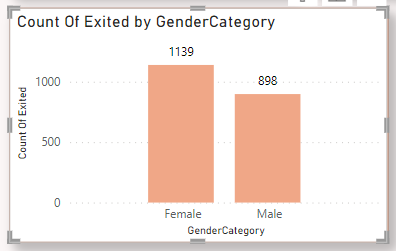
* Germany has the highest no of exit counts and less no of Active account this can be due to that Germany doesn’t have 0 balance account. We can also say that the lower no of active accounts contributes to higher exit count.
* France has the 2nd highest exit count but it has the greater no of Active accounts this can be due to because it allows to have 0 balance account. For France the higher the no. of active accounts the lower is the exit count.
* Spain have less no of exit account and active account greater than Germany which indicated that their customers are satisfied and most of the account for this are od 0 balance accounts.

**Relationship between Active account and exit count/chrun**

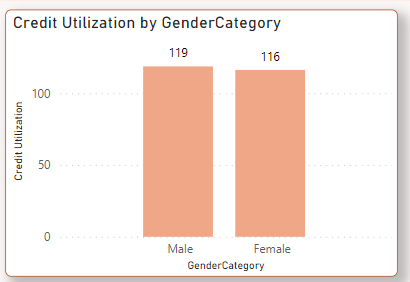
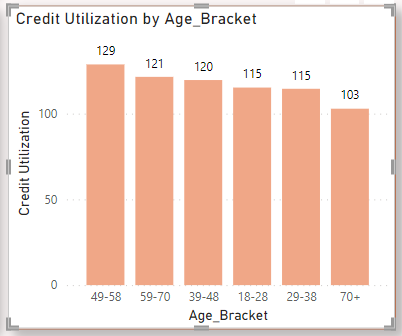
* As the count of Active account increases there is decrease in the Exit count or churn.

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

* **Gender based insight**

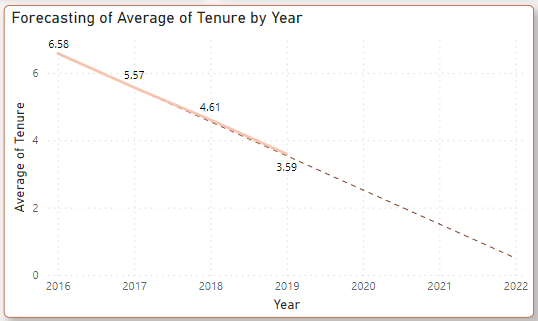


* As we can see that the females have the highest no of churn than male. So, we can focus on understanding what would be reasons for their exiting here we can understand that females are more prone to leave.
* **Credit utilization by gender and age**
* To find out credit utilization we need to have credit limit column but, in our dataset, we don’t have credit limit column we have credit score column that can we use assuming that higher the credit scores the higher is the credit limit and vice versa.

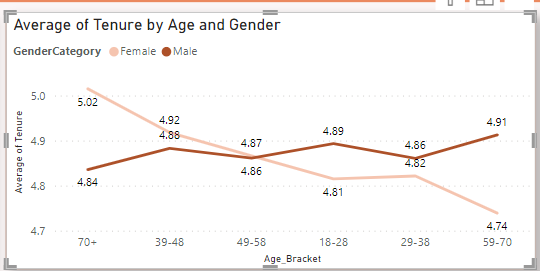
* As we can see that female and male have same credit utilization but the age 49-58 have higher credit utilization, they might become financial risk for the bank as their utilization is more.

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?



* To Forecasting the data, we need date time column so here I have used bank date of joining column and predicted what will be the Average tenure with the upcoming year.
* As you can see the results are bad as it is decreasing year by year the forecasting model also gave decreases in the Average tenure.

Based on Different customer Segments: -



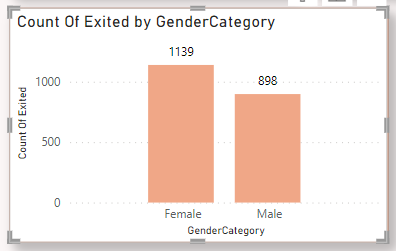
* I have chosen two segment Age and Gender as we can see that the Average tenure is high for 70+ Females followed by 39-48 females and then the 59-70 males.

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?

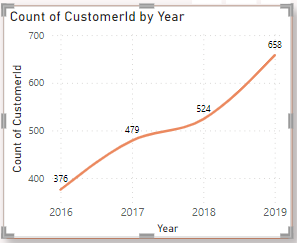
Here is an approach-

1. First, we need to have a data about the marketing campaigns like start time , end time type of campaigns.
2. By getting the campaign that we can check we CRM data that which customers saw which campaign.
3. Now we can see that what is the retention rate and acquisition rate.
4. Here we can compare the data of retention rate with acquisition rate to get the idea about the any effects.
5. Then we can compare the data with customers who didn’t saw the campaign and who saw the campaign.
6. We can even take the feedback of the campaign from the customers who saw it.
7. Customer Exit Reasons Exploration: Can you identify common characteristics or trends among customers who have exited that could explain their reasons for leaving?

* Gender – Female tend to Exit more than man



* Age – Age group of 40-60 tend to exit.
* Product Used – Most of the customers who have Exited used single product.
* Time based churn- the churn has increased from year 2016.



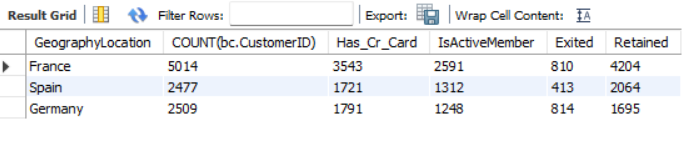
* Tenure based – Long term customer basically tenure >4 tend to exit more than the new customers.
* Active Account – People with Inactive account tend to Exit more.

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

* Yes, it is above column are import to predict the value weather the customers will leave or not because we found out how long-term customers tend to Exit and customers with 1 product used also left the bank whereas inactive members also tend to leave more.

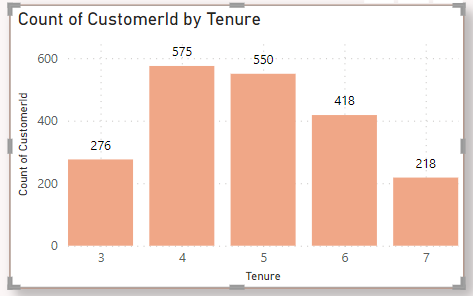
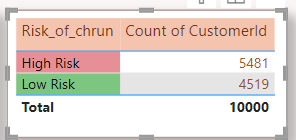
1. Utilize SQL queries to segment customers based on demographics and account details.

Output

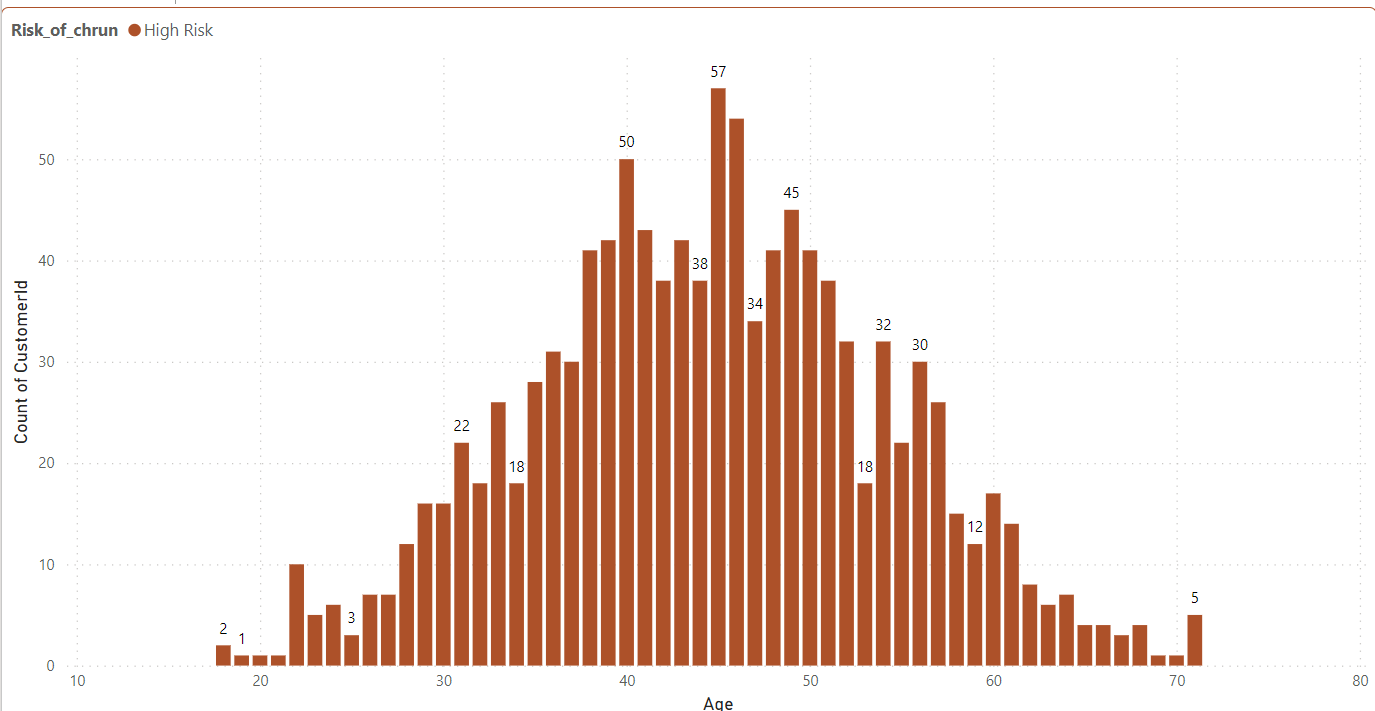
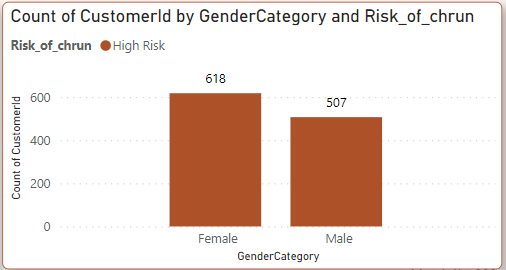


1. How can we create a conditional formatting setup to visually highlight customers at risk of churn and to evaluate the impact of credit card rewards on customer retention?

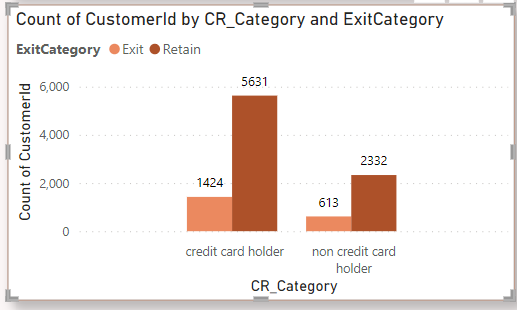
* To differentiate the customer in High Rish churn and Low risk churn I have use tenure column as we know from the above analysis that long term customers are very likely to Exit the bank.  
    
  So to differentiate I have created a new column and as we can see in the below visual of Tenure graph that customers with 4-5 tenure have Exited the bank. So that condition I have used in creating the table.

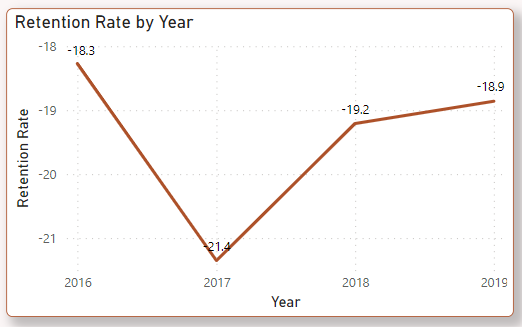
* As we can see in above graph that Tenure of 4-5 customers is Exiting the bank so this can be a check point of risk, we can use this insight while evaluating for the next set of data. So, we found out that 5481 customers have the high risk of churn.

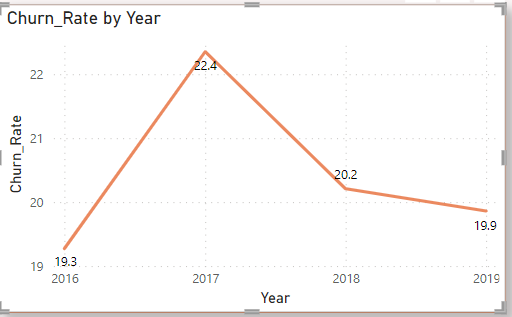
* In those 5481 customers the Age group 40-47 have the highest Exit accounts and the 47-52 have the highest churn of risk and in that also female are more likely to Exit the bank.
* Impact of Credit Card on retention- As we can see in the below graph that the total credit card holder is 7055 and non-credit card holders are 2945.
* In Credit Card holders and Non-credit card holders the Exit Rate is 20% (1424/7055\*100) and (613/2945\*100)
* In Credit Card holders and Non-credit card holders the Retain Rate is 80% (5631/7045\*100) and (2332/2945\*100)
* So, by this We can say that Having Credit card doesn’t not effect on Exit Rate and Retain Rate



* Retention Rate Over the Years –



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?

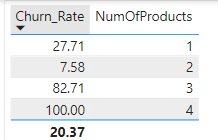
sss

* 2017 has the highest Churn rate or we can say that most of the year have closely similar churn rates.

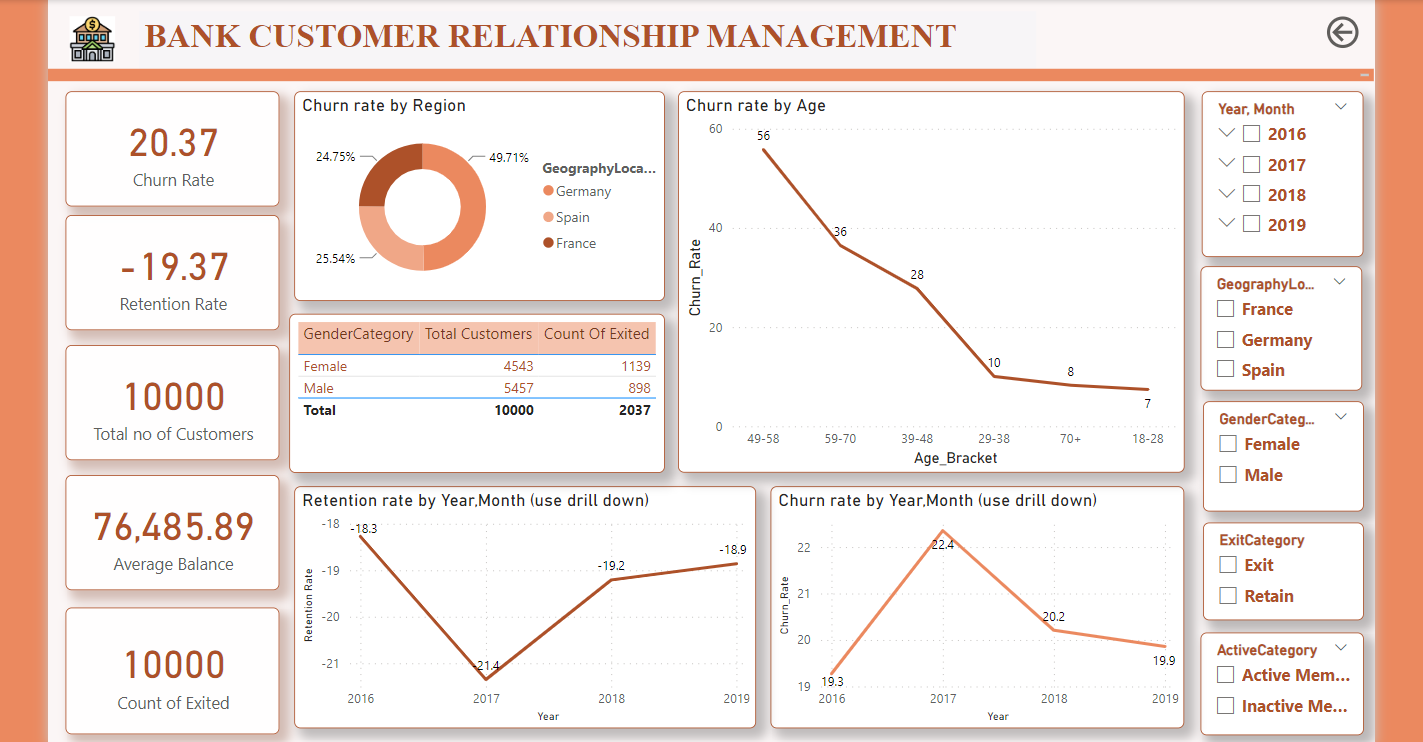
**Insights –**

* Location based Churn – Germany has the highest churn rate.
* Gender based churn – Females have the highest churn rate
* Age based Churn- age between 40-52 have the highest churn.
* Tenure based Churn – Tenure of 4-5 customers has most of the churn.
* Is Active Accounts churn - lower no of active accounts contributes to higher exit count.
* See at the below visualization.



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?

Steps: -

* Understanding the data – It is very measure part as this will give whole a lot information about the columns and how they are related.
* After understanding the data crate relationship between the data then using Model view between the tables present in the sheet to understand that how all the column and tables are related to each other this will make easy to understand and access the data.
* Then will move to the transformation of the data to check for missing value, errors, column have corrected data type or not, format of the values in the column, replacing null values, making data look more sense adding columns if requires we can do that all in Power query editor etc.
* Now, will start with making visual to understanding the relationship between the column like how one column is affecting the other or how or not affecting at all we can even create measures and new column according to our requirement using DAX expressions which come very handy to show KPI and even important data dynamically.
* By doing this research and understanding we would have known that churn is depending on many columns in the dataset and would have approach that why is this churring happening, what is the reason behind that churning.

In technical terms we can say that, Description analysis, Correlation analysis, Data Exploration, Data transformation, Segment analysis, Dynamic Calculations.

s

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

As we already know Power Bi is ETL tool so by using power bi Transform data feature we can do this easily.

1. Load the data
2. Transform data
3. Select the column cell/header which you want to rename.
4. You can just double click on it a change it or either you can right click and choose the option to rename.
5. Then click on Apply and close know the name has been changed.