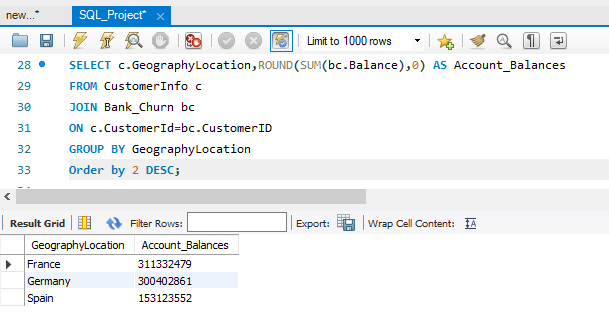
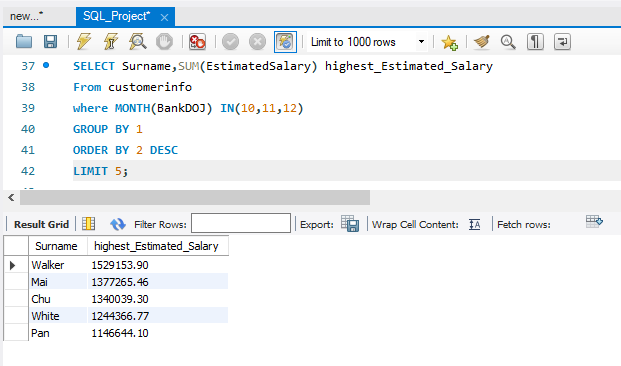
**Objective Questions:**

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

****

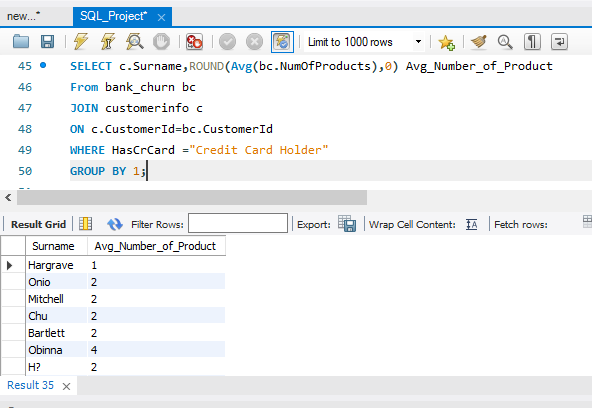
**Since there are high numbers of customers account in France, Germany and Spain. So, the account balances are also high in the following order: France, Germany, Spain.**

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

****

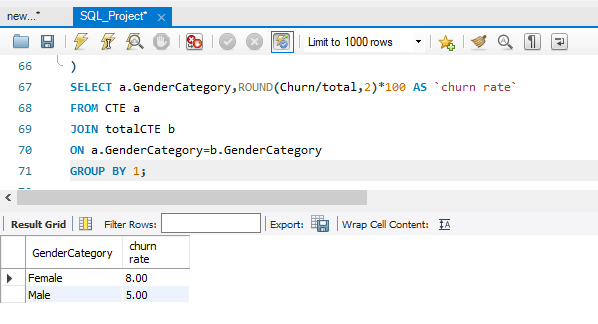
**The Top 5 customers who have highest number of transactions are Walker with 1529153.90 salary, Mai with 1377265.46 salary, Chu with 1340039.30 salary, White with 1244366.77 salary and Pan with 1146644.10.**

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

****

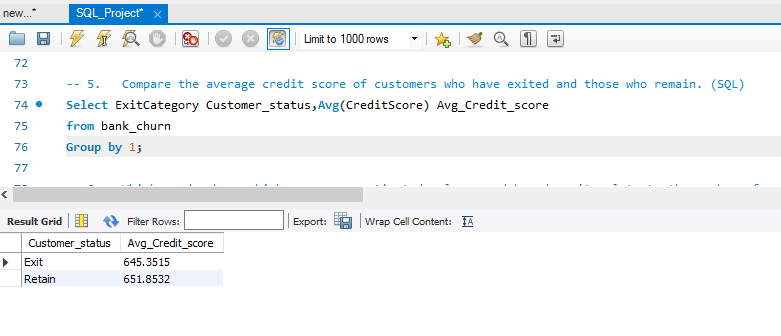
**Mostly 2 Average Number of Products have who have a Credit card.**

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

****

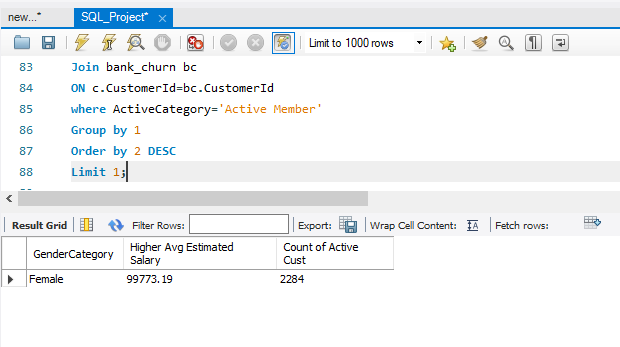
**Female has a churn rate of 8.00 and Male has a churn rate of 5.00.**

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

****

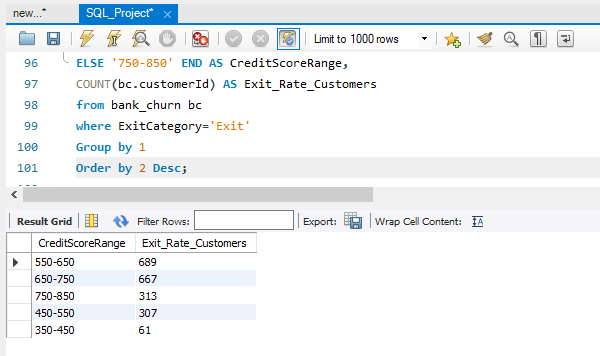
**The customers who have exited has average credit score of 645.35 and customers who have retained in the bank has average credit score of 651.85.**

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

****

**Female has the highest average estimated salary of around 99773.19 with 2284 Active Customers.**

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

****

The customer’s credit score has been Segmented from

350-450 (Very Poor)

450-550 (Poor)

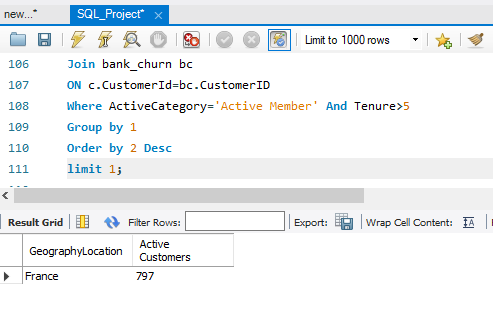
550-650(Medium)

650-750 (Good)

750-850(High)

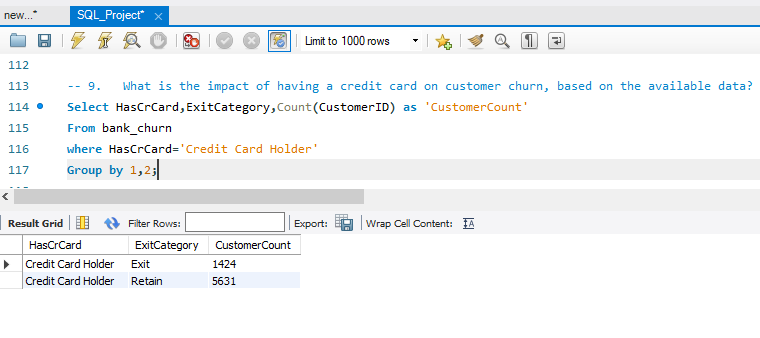
Credit Score which is having Medium Credit Range Between 550-650 has maximum 689 numbers of Exit Customers. The Very poor range of credit score which is from 350-450 has very Less Exit Rate of Customers i.e 61 Customers.

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

****

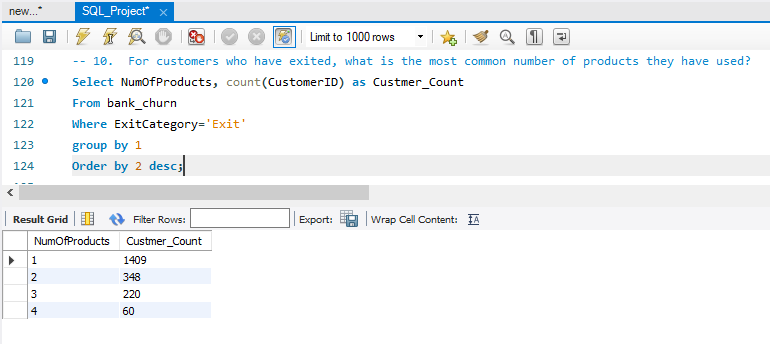
France has the highest numbers of active customers of 797 with a tenure more than 5 years.

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

****

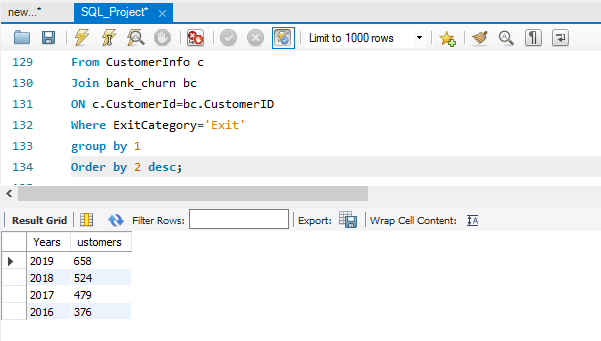
Totally 1424 customers have exited from the bank who have credit card. But customers with credit card who remained in the bank is higher than the exited people. So, there is no impact of having a credit card on customer churn.

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

****

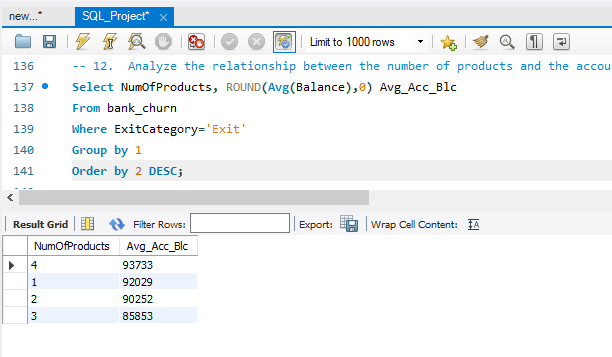
The exited customers have used 1 as their most common number of products.

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

****

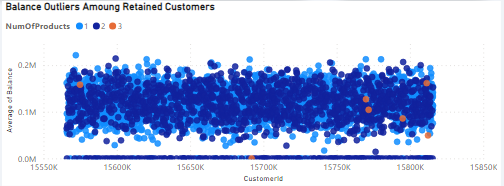
According to the given dataset, the most recent year 2019 have the highest exit rate of customers. And also, the outlier is from the year 2016 to 2019 we are losing customers higher when compared to the previous year.

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

****

The customers who have 4 as their number of products has the highest average of account balance (93733) compared to others followed by one product (92029), two products (90252.) and three products (85853).

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

****

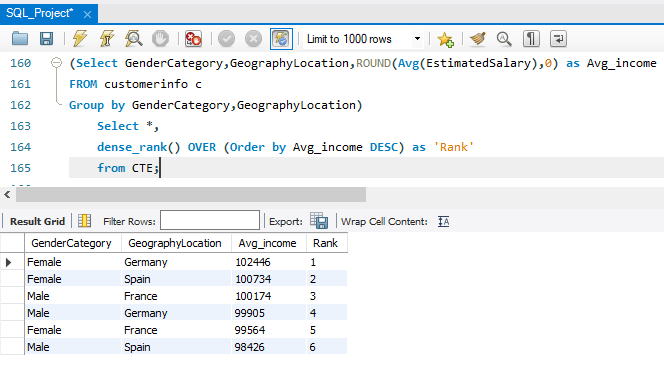
The customers who are not having 4 number of products are mostly retained in the bank.

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

There are two Tables in the dataset but in Both tables contain a mix of categorical and numerical variables. The only categorical variables are table is not there.

Top of Form

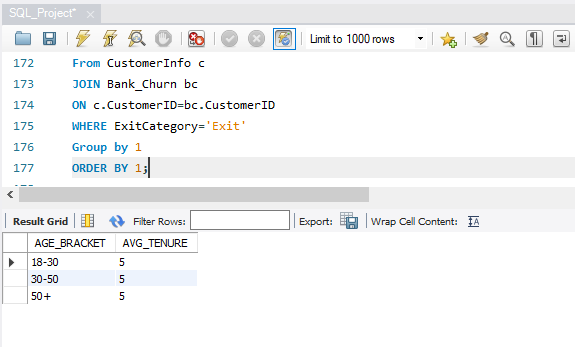
1. **Using SQL, write a query to find out the gender-wise average income of males and females in each geography id. Also, rank the gender according to the average value. (SQL)**

****

By using Dense Rank Function, I have ranked the gender-wise average income of males and females in each geography id.

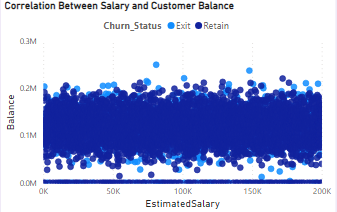
* First Rank - Female in Germany has average income value 102446.
* Second Rank - Female in Spain has average income value 100734.
* Third Rank - Male in France has average income value 100174.
* Fourth Rank – Male in Germany has average income value 99905.
* Fifth Rank – Female in France has average income value 99664.
* Sixth Rank – Male in Spain has average income value 98426

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+).**

****

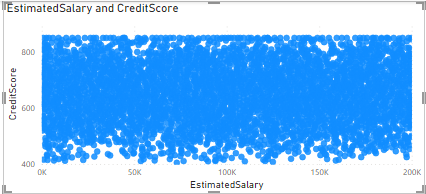
The Average tenure of the customers with all age bracket of are Same i,e 5 .

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



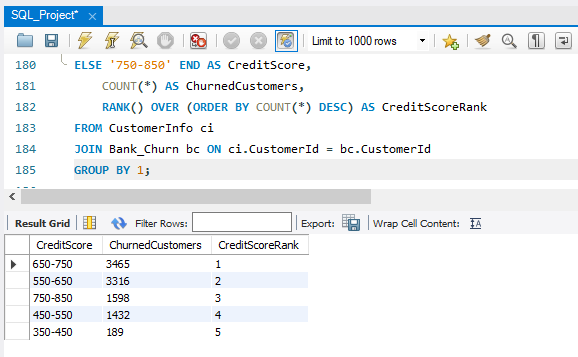
From the above picture, we can get the conclusion that there is No Correlation between the Estimated Salary and Customer Balance with the people who have exited or not.

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

****

From the column chart mentioned above, we can understand that there is no correlation between salary and credit score.

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

****

The customer’s credit score has been Segmented from

350-450 (Very Poor)

450-550 (Poor)

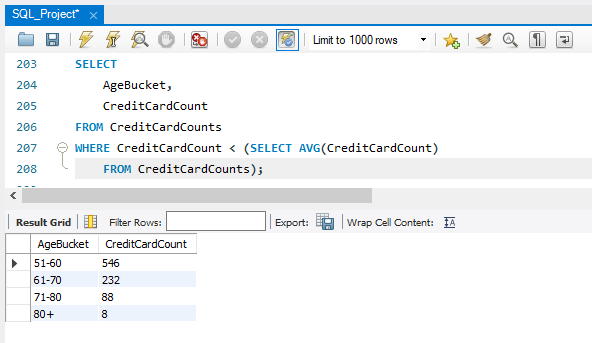
550-650(Medium)

650-750 (Good)

750-850(High)

I have ranked each bucket of credit score as per the number of customers who have churned the bank in SQL. From that, I have Good range of Credit card is worthiness (i.e.).,650-750 got the first rank with 3465, customers followed by Medium Range of Credit card worthiness (i.e.).,550-650 got the second rank with 3316 customers, High range of Credit card worthiness (i.e.).,750-850 got the third rank with 1598 customers, Poor range of Credit card worthiness (i.e.).,450-550 got the fourth rank with 1432 customers and very poor range of Credit card worthiness (i.e.).,350-450 got the fifth rank with 189 customers.

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

****

The Age Bracket are Categories as below

Age '18-30'

Age '31-40'

Age '41-50'

Age '51-60'

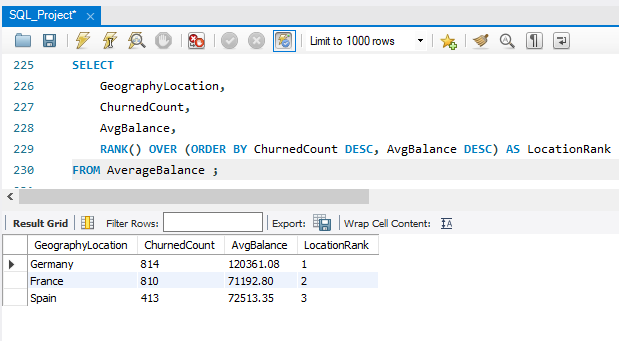
Age '61-70'

Age '71-80'

Age '80+'

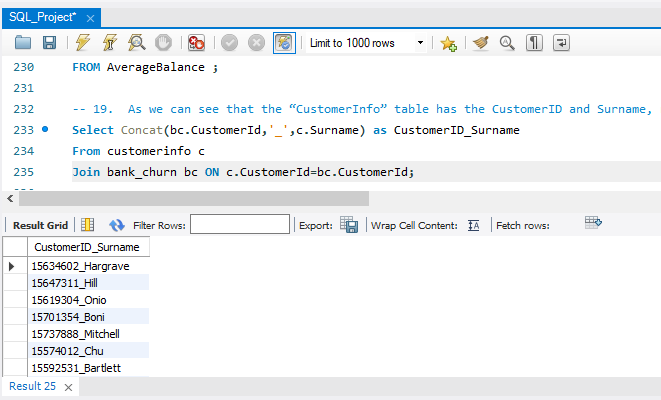
The count of customers with the age bracket 51-60 is 546, the count of customers with the age bracket 61-70 is 232, customers with the age bracket 71-80 is 88 and the customers with the Age 80+ is 8. These are all the buckets having lesser than the average number of credit cards per bucket.

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

****

* Germany got the first rank because of having 814 number of people and average balance of 120361.08 who exited the bank.
* France got the second rank because of having 810 number of people and average balance of 71192.80 who exited the bank.
* Spain got the third rank because of having 413 number of people and average balance of 72513.35 who exited the bank.

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”.**



By Using CONCAT Function we can Joined the two String i.e CustomerID and Surname as Shown in the above Figure.

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

Yes, you can retrieve the "ExitCategory" from the "ExitCustomers" table and incorporate it into the "Bank\_Churn" table without using explicit JOINs by using a correlated subquery or a subquery in the SELECT statement. Here's how you can do it using a subquery:

**SELECT**

**bc.\*,**

**(SELECT ExitCategory**

**FROM ExitCustomers ec**

**WHERE ec.CustomerID = bc.CustomerID) AS ExitCategory**

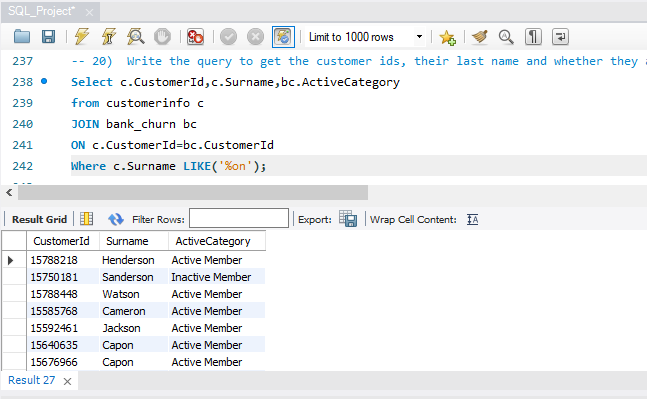
**FROM**

**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?**

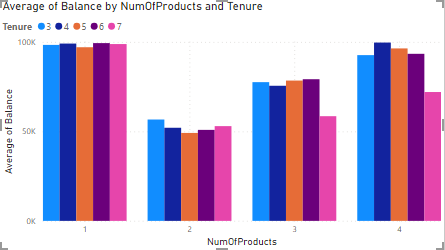
No there was no missing Values in the data. We can handle missing values with the mean, median, or mode of the column.

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”.**



**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?**

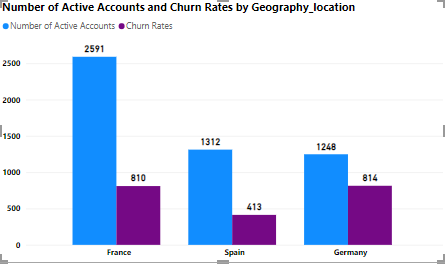


From the above visualization, we can notice that the Customers who have Purchased 4 Products are the loyal customers.

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

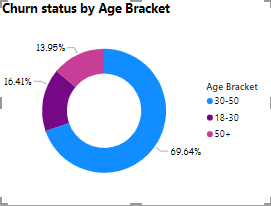
Since there are no bank products provided in the datasets, we can’t come to conclusion that which factor influence the cross-selling strategies.

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

****

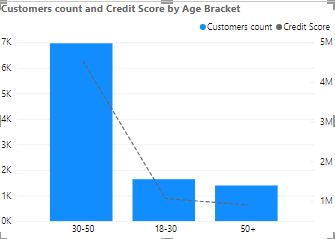
From the above visualisations, we can analyse that the France has the highest number of active accounts and second largest churn rates. On the other hand, Germany has lowest number of active accounts but also the highest churn rates. Spain has the lowest churn rates and second smallest number of active accounts.

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



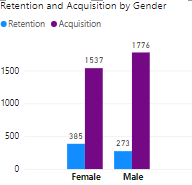
The reason why I choose Age bracket as demographic segment is because age factor is common for both male and female which has high impact on bank’s financial risk. From the above visualizations, middle age bracket people have the highest rate of churn. At the age bracket between 30 and 50, all the people will try to put the investments, but at the time customers are exiting from the bank. It could cause high financial risk to the bank.

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



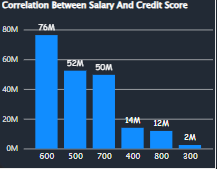
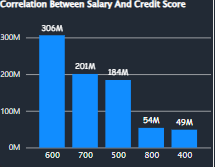
Since the question is asked for the life time value of customer, I have chosen to present the customer count and their credit score for their life time which is age bracket. By this visual, we can understand that credit score for the customer joined at the age of 18-30 will have low credit score since they have just joined. Moving to the age bracket of 30-50, customers are having high credit score which seems to be profitable for the bank. And at the age of 50+, Customers are not ready to spend on banks so that the credit score decreases.

1. **Marketing Campaign Effectiveness: How could you assess the impact of marketing campaigns on customer retention and acquisition within the dataset?**



We can say that the acquisition of customer for the last year is higher than the rate of retention for both the male and female. To decrease the retention rate, we may offer some offers on the old customers they could have gain some credit score additionally. So that the new customers also remain for the long period of time.

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

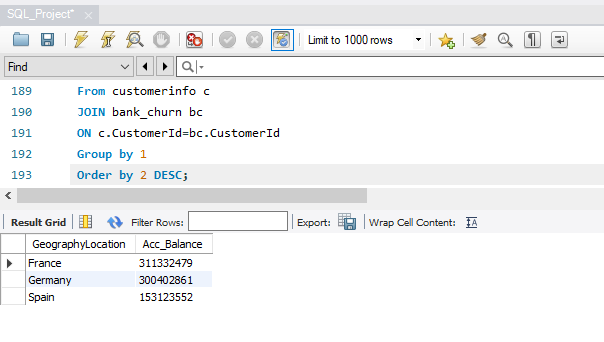
**Exited Retained**

From the above visuals, we can say that the salary and so credit score are the main characteristics or trends among the customers who have exited. Because as the salary is low, they are not able to maintain the credit score. So that they may planned to exit the bank.

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

Yes, Tenure, NumOfProducts, IsActiveMember and EstimatedSalary are important for predicting if a customer will leave the bank. Because these columns will help us to find the customers can buy product with his estimated salary and can settle down the payment within the tenure time and after paying things will he be available as active member in the bank.

1. **Utilize SQL queries to segment customers based on demographics, account details, and transaction behaviours.**

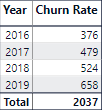


1. **How can we create a conditional formatting setup to visually highlight customers at risk of churn and to evaluate the impact of credit card rewards on customer retention?**

|  |  |  |
| --- | --- | --- |
| **Credit Card Status** | **Churn Status** | **Count of Customer Id** |
| **Credit Card Holder** |  | **7055** |
|  | Exit | 1424 |
|  | Retain | 5631 |
| **Non-Credit Card Holder** |  | **2945** |
|  | Exit | 613 |
|  | Retain | 2332 |
| **Grand Total** |  | **10000** |

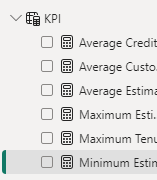
In this, we can select the count of Customer Id column an select the conditional formatting which is present in the home tab. In that we can format things according to that data by the MS Excel.

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?**



Current churn rate we can say that 658 for the year 2019 which is current churn rate per year. And the total churn rate is 2037. The age bracket of 30-50 has the highest percent of churn rate. For those people, we can give some extended tenure period so that he could able tackle his issue by the extended time.

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

****

As mentioned above, I have created a dashboard and created some KPI’s such as Average Credit Score, Average Customer Balance, Average Estimated Salary, Maximum and Minimum of all those categories.

1)Average Credit Score = AVERAGE(CustomerDetails[CreditScore])

2)Average Customer Balance = AVERAGE(CustomerDetails[Balance])

3)Average Estimated Salary = AVERAGE(CustomerDetails[EstimatedSalary])

4)Maximum Estimated Salary = MAX(CustomerDetails[EstimatedSalary])

5)Maximum Tenure Period = MAX(CustomerDetails[Tenure])

6)Minimum Estimated Salary = MIN(CustomerDetails[EstimatedSalary])

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

If the objective and subjective questions weren't given, and the task is to analyze the dataset without specific guidance, here's a general approach you could take:

**Data Exploration**: Begin by exploring the datasets ("CustomerInfo" and "Bank\_Churn"). Understand the structure of the data, including the columns, data types, and any missing values. Use descriptive statistics and visualizations to gain insights into the distribution of variables and identify any patterns or anomalies.

**Data Cleaning**: Preprocess the data to handle missing values, outliers, and inconsistencies. This may involve imputing missing values, removing duplicates, and standardizing formats.

**Feature Engineering**: Create new features or transform existing ones to extract meaningful information from the data. For example, you could derive additional variables such as age groups, tenure categories, or financial ratios.

**Exploratory Data Analysis (EDA)**: Conduct exploratory data analysis to uncover relationships and trends within the data. Visualize relationships between variables using plots such as histograms, scatter plots, and correlation matrices. Explore how different variables are distributed and whether there are any correlations or associations.

**Customer Segmentation**: Segment customers based on demographic, geographic, or behavioral characteristics. Use clustering techniques such as K-means clustering or hierarchical clustering to group customers with similar attributes together.

**Churn Analysis**: Analyze churn patterns to understand factors contributing to customer attrition. Identify characteristics or behaviors associated with churned customers and explore potential predictors of churn.

**Predictive Modeling**: Build predictive models to forecast customer churn or predict customer behavior. Train machine learning models such as logistic regression, decision trees, or random forests using historical data. Evaluate model performance using metrics such as accuracy, precision, recall, and ROC curve analysis.

**Interpretation and Insights**: Interpret the results of your analysis and derive actionable insights for the business. Communicate your findings effectively through reports, presentations, or dashboards. Highlight key insights, recommendations, and areas for further investigation.

**Iterative Process**: Data analysis is often an iterative process. Continuously refine your approach, experiment with different techniques, and validate your findings to ensure robust and reliable results.

**Documentation and Reporting**: Document your analysis methodology, assumptions, and findings in a clear and organized manner. Provide detailed documentation to facilitate collaboration and knowledge sharing with stakeholders.

By following this approach, you can systematically analyze the dataset, uncover valuable insights, and provide actionable recommendations to support business decision-making.

Documentation and Reporting

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;

**Conclusion**

**Benefits of CRM implementation in Banking:**

* Customer relationship Management plays a vital role for competition in the banking industries.
* Customer satisfaction, Customer focus and retention Improve customer service.
* Understanding customer value better relationship marketing Benefits for employees.
* Conclusion bank CRM is an scientific approach for building and sustaining long term business with customers to satisfying them. It is focused on improvement and enhancement of business processes

**Solution**

* By increasing the tenure period for the customers who is having low salary, Customers could able to stay in the bank.