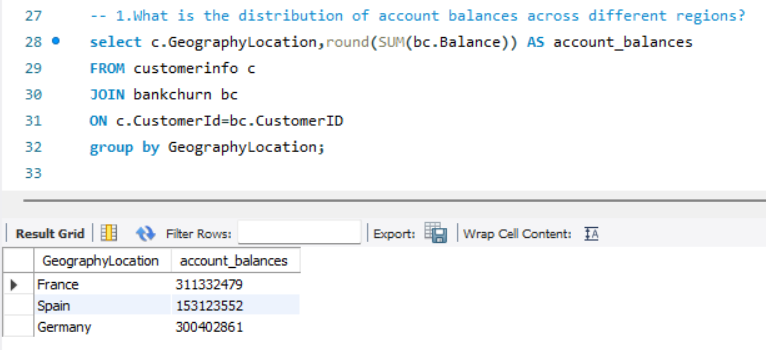
**Objective Questions:**

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

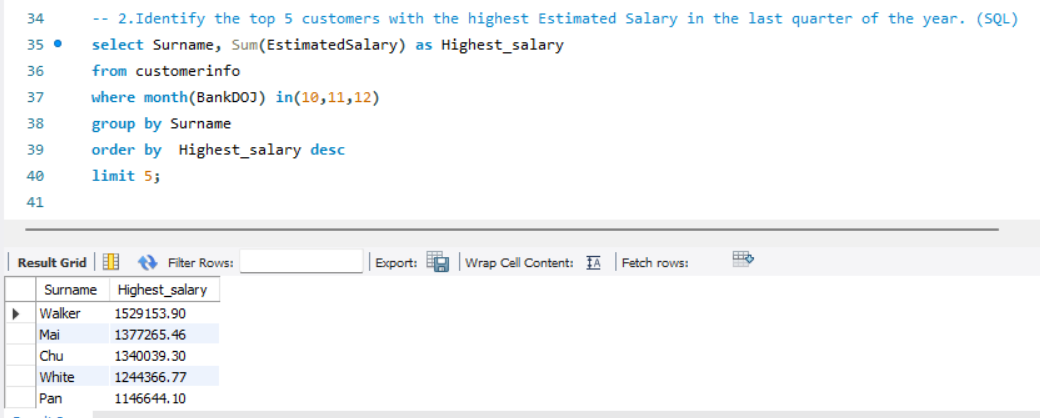
--- Here we have use SUM function to find the account balance and then we have to use group by function to get the account balances by Locations using Join function.



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

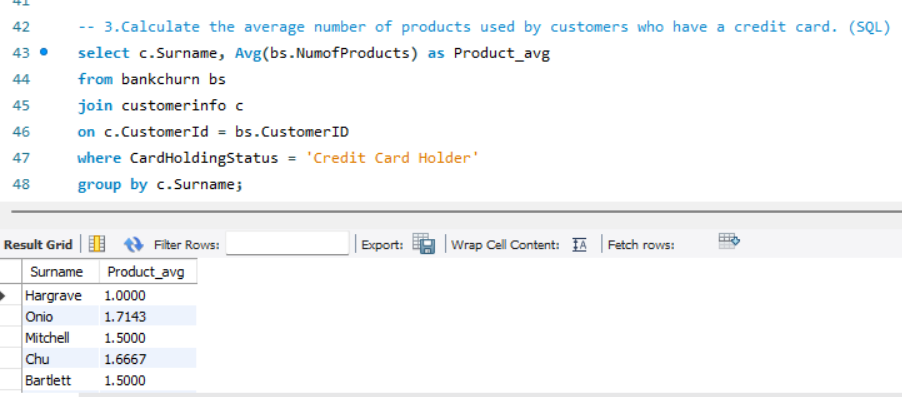
---- To get the highest Estimated Salary we have to use SUM function and then as mentioned we have to find the highest of last quarter i.e. 10,11,12 so to get that have used Month function.

To get the top customers we have to order the result of Sum function in DESC so that we can get the top customer and then limit it by 5 which will give the Top 5 customers with the highest Estimated Salary



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

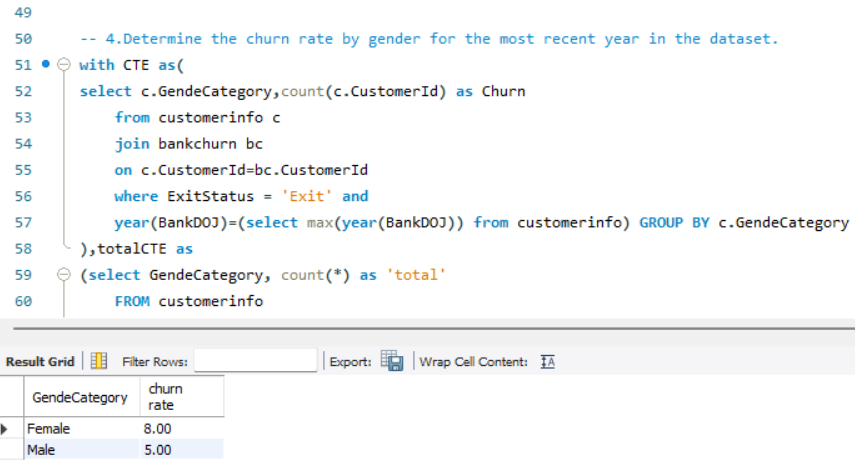
---- To get the Avg. of number of products we have used AVG function and to find out the customer who has a Credit card we have used WHERE function. And used join function to connect both the tables.



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

---- We have used CTE to calculates the number of customers and out of that to find the churned customers we have used where clause to find the status of the customer whether they have exit or retained & used MAX function to the year column to find the most recent year and group all that up by Gender Category.

Again, we have used CTE to calculates the total number of customers which will help us in finding the churn rate by dividing the number of lost customers by the number of customers at start and multiplying it by 100.



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

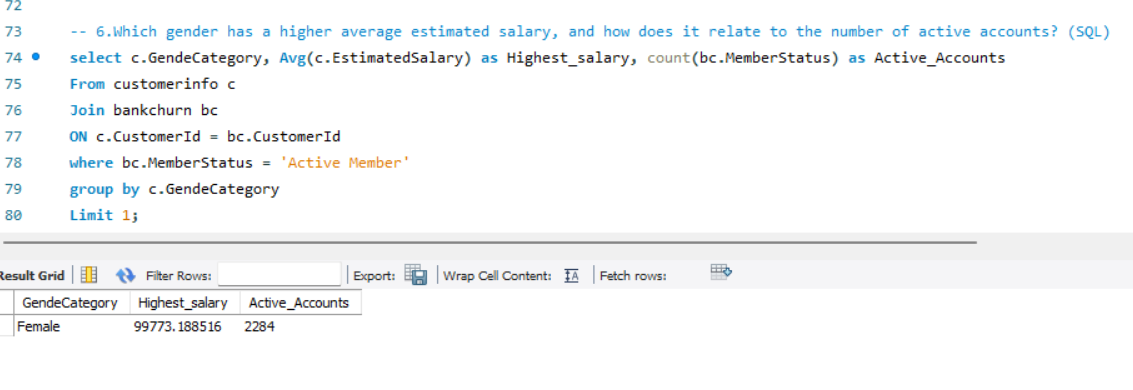
---- We have used AVG function to get the avg. of customers who has exited and remained and group them by their exit status, which will give us the Avg. of both remained and exited.



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

--- We have used AVG function to Estimated Salary and COUNT function to Member status and we have joined both the table with customer id as common ids I both table

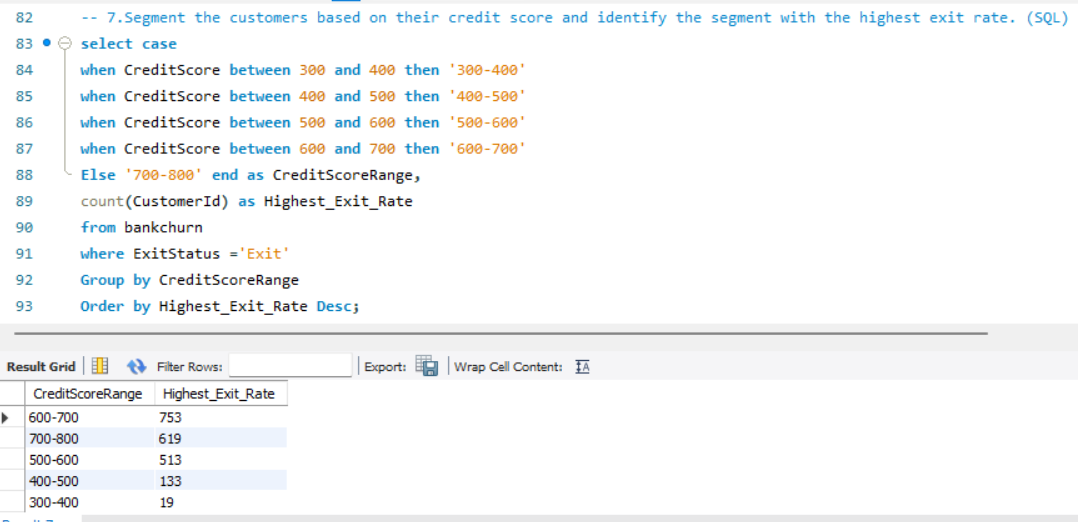
Then to find the Status of the customer whether or not they are active we have used where clause and to find which gender has the highest avg. estimated salary, we have grouped them up by gender category.



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

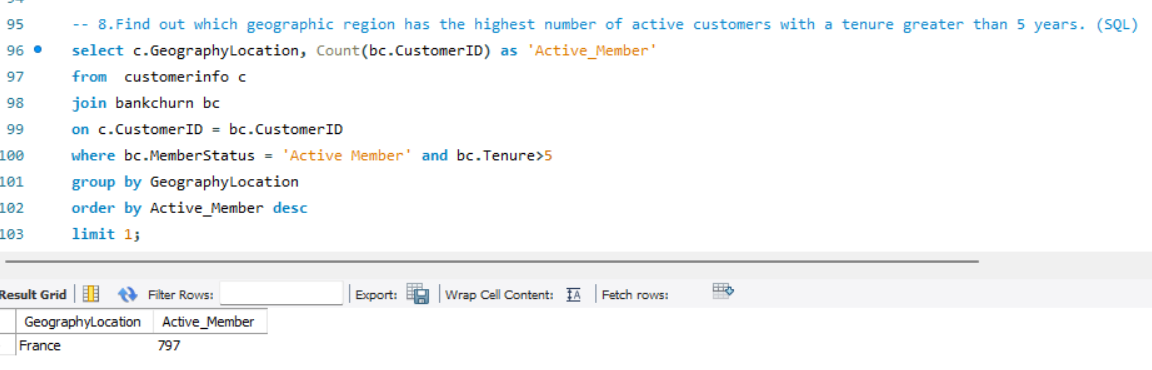
---- To segment customer based on Credit score we have used case statement in which we have made Credit score range which we will further use in to grouping it to find the which range has the highest exit rate.

But to find highest exit rate we have used CONT function to customer Id which is then filtered with where clause to find only the customers who has exited and to get the highest, we have ordered the highest exit rate by DESC.



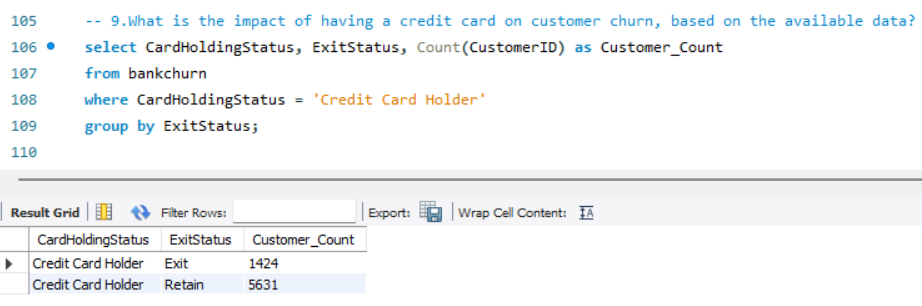
1. Find out which geographic region has the highest number of active customers with a tenure greater than 5 years. (SQL)

---- We have firstly COUNTED the customer id joined both the tables and then to find whether the customer is Active for more than 5 years or not we have used where clause. And to find which region has the highest number we have grouped up by geographic location and ordered the count of Active members in DESC which will give us the highest.



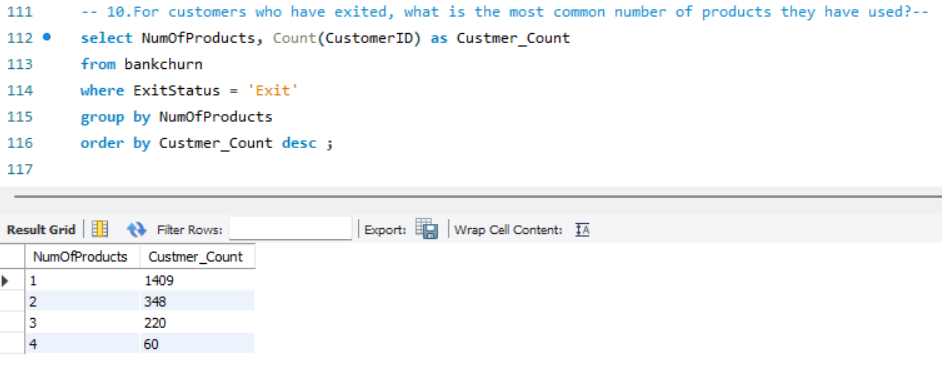
1. What is the impact of having a credit card on customer churn, based on the available data?

---- Out of total customers count there are 7055 Customers who holds Credit card out of which 5631 Customers are retain with the bank which means they are happy with the banking services. Where as 1424 customers have exited but still holds Credit Card which gives us a chance to improve our products and services which has most count of exits and churn them into retain customers.



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

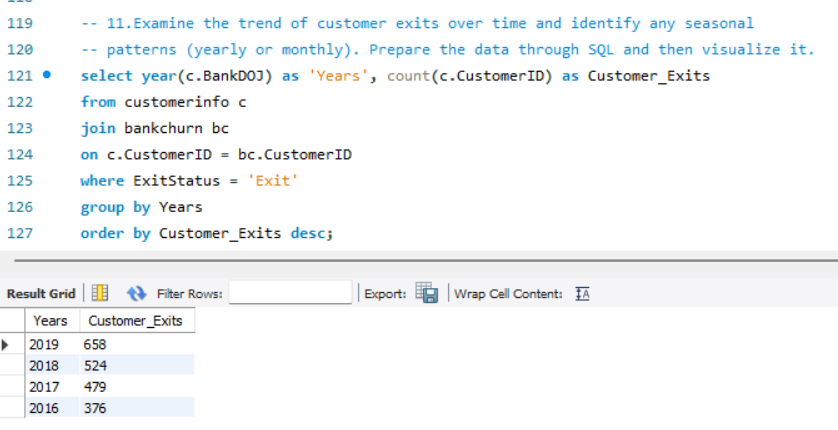
---- We have taken COUNT of all the customer and to find the availability status we have used where clause and grouped the result with number of products to get the most common product from where customers have exited which will help us in improve that product.



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.

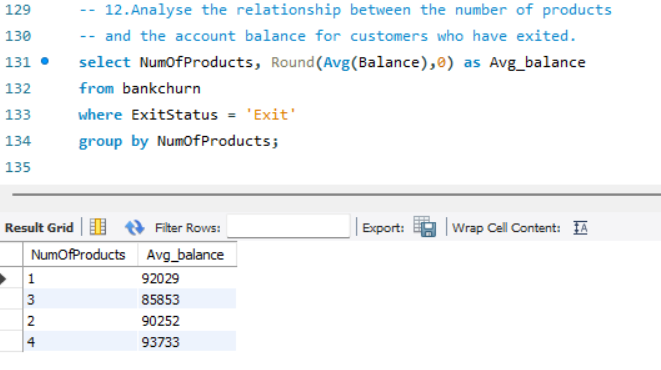
--- To find trend we have considered year and COUNTED all the customers and to find the exit availability status we have used where clause and grouped the result by yearly data which will show us the trend of customers who has joined the bank.

Monthly or quarterly seasonal pattern data will help in understanding the data more accurately and take more refined decision.



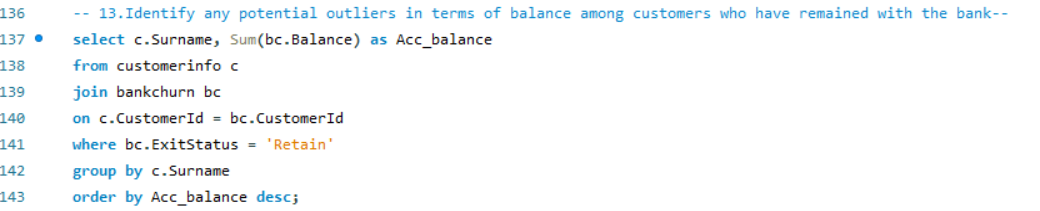
1. Analyze the relationship between the number of products and the account balance for customers who have exited.

--- Here we have taken the AVG Balance of all the customers then to find the exit availability status we have used where clause and grouping the result by Number of Products which allows us to see the average balance for each product. This analysis will help understand the behaviour and characteristics of customers who have exited



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

--- To find we have the outliers we have SUM all the balance of customer & then to find the retain status we have used where clause and grouping the result by their name in this way we can get the list of customers and if we order the account balance in DESC we will get the potential outliers with abnormality.



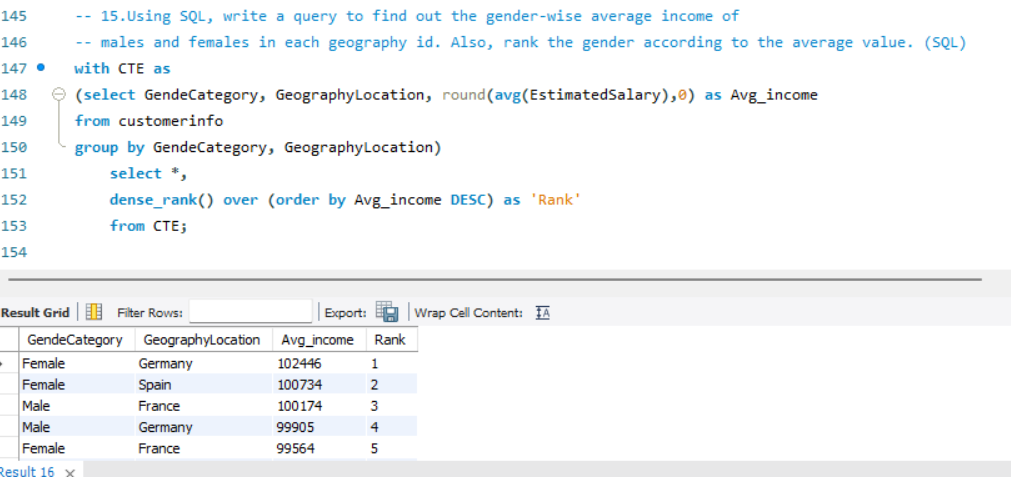
1. How many different tables are given in the dataset, out of these tables which table only consists of categorical variables?

----- There are two Tables in my dataset by name customerinfo & bankchurn but both of them contain categorical as well as numerical variables. The customerinfo table contains nominal categorical variables details such as Gender, Age.

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)

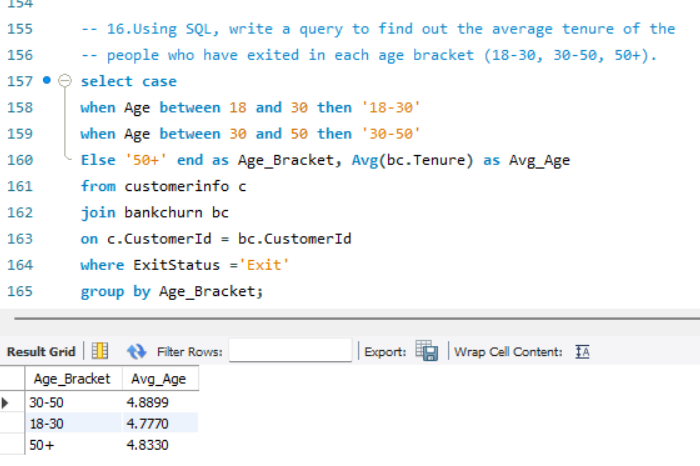
---- We have used CTE to calculates the AVG estimated salary of all the customers and grouped them by the Gender Category and location which will give us the result of gender-wise average income of males and females in each geography location

To rank the gender according to the average value we have used DENSE RANK function to rank based on Avg. Income regardless of their gender & location.



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+).

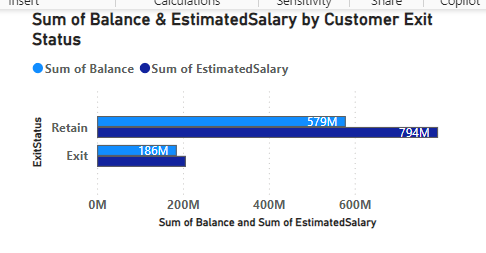
---- To get the age bracket we have used CASE statement and used AVG function to tenure & then to find the exit status we have used where clause and grouping the result by the age bracket which we have found with case statement. This will give us the average tenure of the customer in each age bracket who have exited.



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?

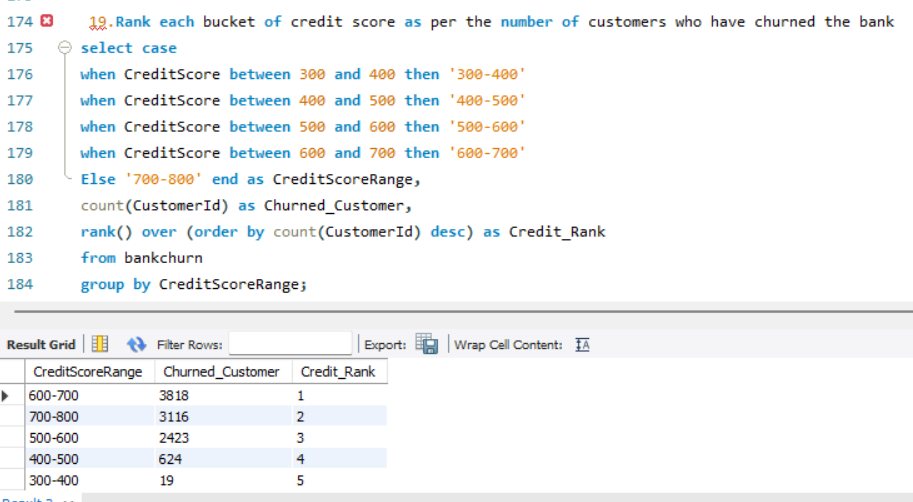
---- Yes, there is a correlation between salary & balance of the customer. There is a strong indication among customers who have stayed with the bank with higher salaries and higher balances.

Whereas there is a lower relationship between salary and balance for customers who have exited still follows the same positive trend but at a lower scale.



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

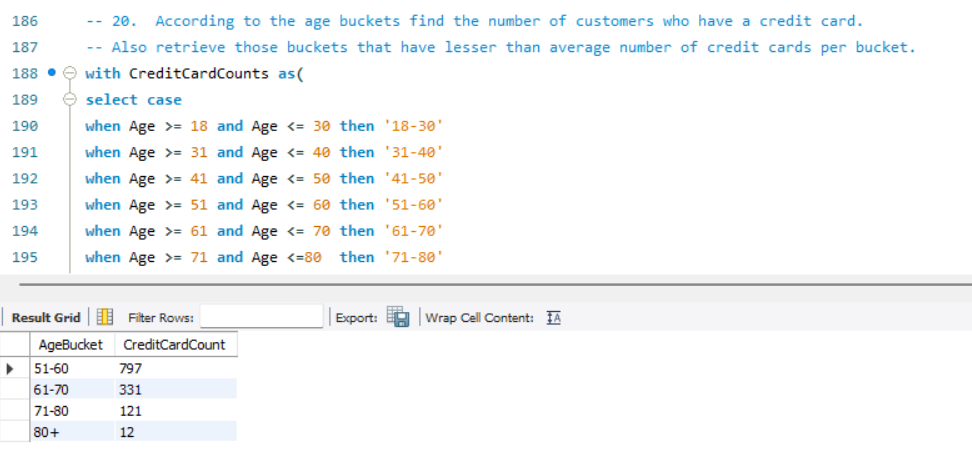
---- To get the credit score bracket we have used CASE statement and used COUNT function to all the customers & then ranked them based of count of customer in DESC to the highest rank. And the result is then grouped by the credit score range which we have founded using case statement. In this way we have got the count of customers in each credit score range.



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.

---- To get the age bracket we have used CASE statement and used COUNT function with CASE statement to find only the COUNT of customers who hold Credit Card & named it as CreditCardCount and grouping the result by age bracket which we had founded using case statement.

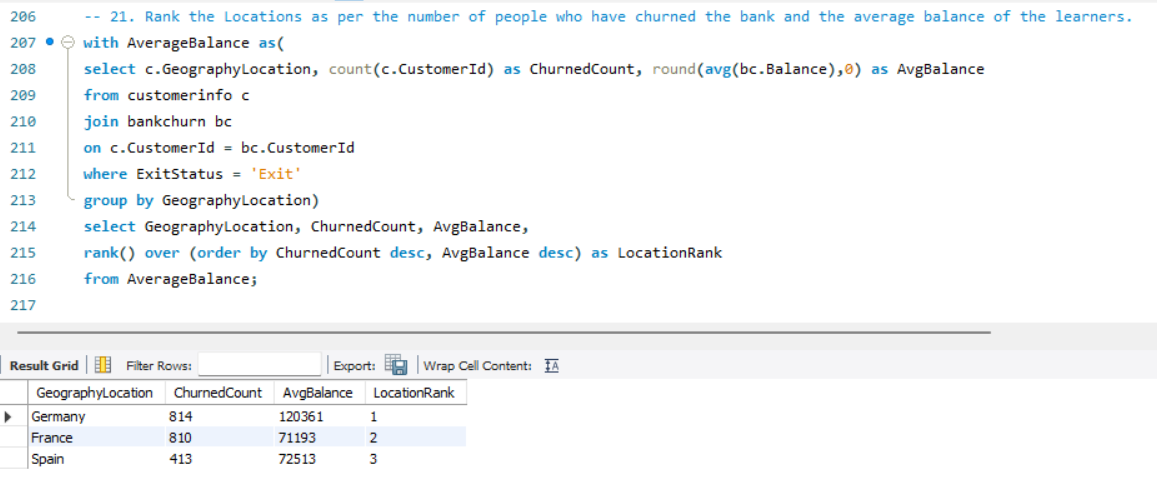
And in the subquery, we have used Credit Card holders Count to compare with the AVG count of credit card to display it.



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

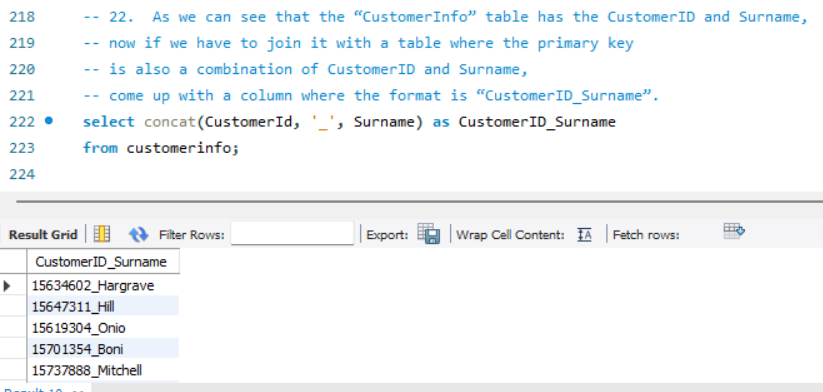
--- To find the average balance we have used AVG function to balance column & to find number of customers who have churned the bank we have used COUNT function to get the count of all the customers and the result of all is used to find churned customer we have used WHERE clause.

In the subquery we have used the RANK function to rank the highest Churned Customer Count and the Avg. balance of that customers based on the geographic location.



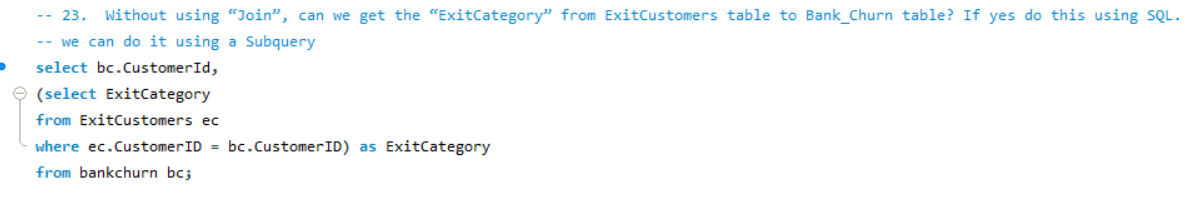
1. As we can see that the “CustomerInfo” table has the CustomerID and Surname, now if we have to join it with a table where the primary key is also a combination of CustomerID and Surname, come up with a column where the format is “CustomerID\_Surname”.

--- As both the columns are in a single table, we have used CONCAT function to join the CustomerID & Surnme and named as CustomerID\_Surname



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

--- We can do it with the help of Subquery as both the tables have CustomerID as primary keys (Common) and joining them with a WHERE clause.

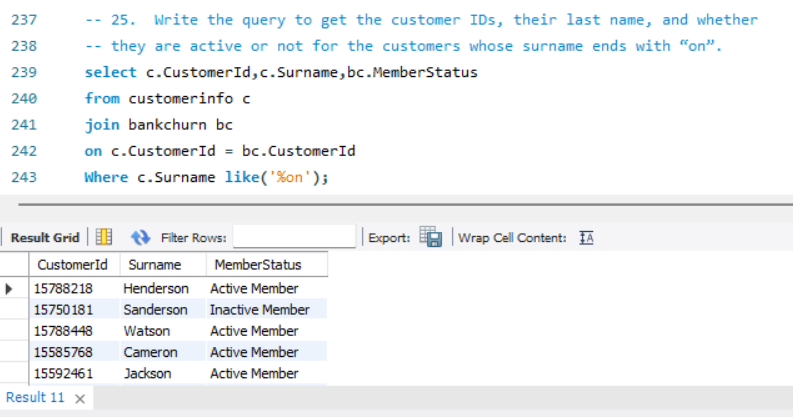


1. Were there any missing values in the data, using which tool did you replace them and what are the ways to handle them?

--- If there are missing values in the SQL data, i.e. NULL values. To handle these, I used SQL queries to identify and replace them. The COALESCE function is commonly used to replace NULL values with a specified value. Other methods to handle missing values include using ISNULL & IFNULL functions. Additionally, missing values can be handled by deleting rows with NULL values or by substituting them with mean, median, Avg. or mode of the column, which can be done using SQL functions

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”.

--- To find the Surnames of all the customers who have ‘on’ at the end we have to use wild card character with LIKE operator.



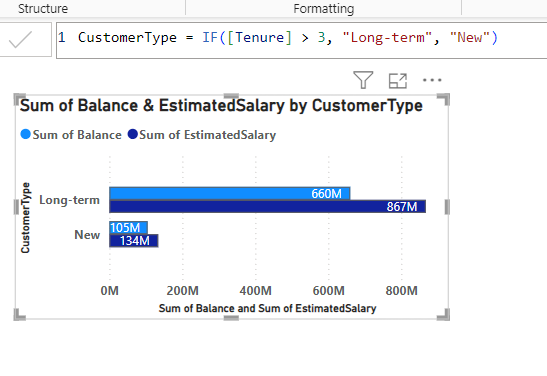
1. Can you observe any data discrepancy in the Customer’s data? As a hint it’s present in the IsActiveMember and Exited columns. One more point to consider is that the data in the Exited Column is absolutely correct and accurate.

--- I didn’t find any data discrepancy in the Customer’s data as I had used VLOOKUP Function in excel to fetch the data correctly and to avoid confusion, I had also renamed the options in IsActiveMember and Exit columns

**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?

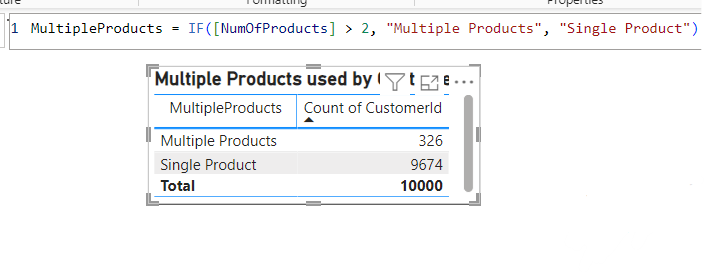
--- Here we have used DAX function in which we have considered Tenure > 3 years as Long-Term Customers and for analysing spending habits we have considered Account balance & Estimated salaries of customers and to visualize it we have used Clustered Bar Chart. From the visualization it is seen that long term customers remained with the bank for a longer period, which shows a strong indication that higher salaries are associated with higher balances



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

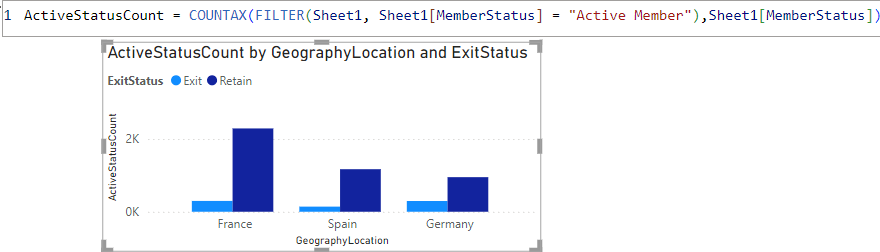
--- As there are no specific bank products provided in the datasets, we can’t conclude the cross-selling strategies.

But, with the available dataset and information we have used a DAX function in which the Number of products used by customers are greater > 2 then we have used them as Multiple Product customers.



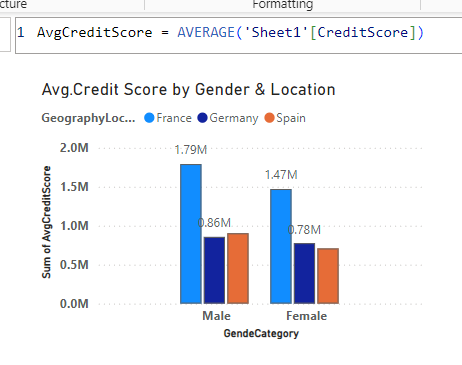
1. Geographic Market Trends: How do economic indicators in different geographic regions correlate with the number of active accounts and customer churn rates?

--- Here we have used DAX function in which we have considered COUNT function to the Active customers which are then visualize using stacked column chart based on geographic location & and Availability status.



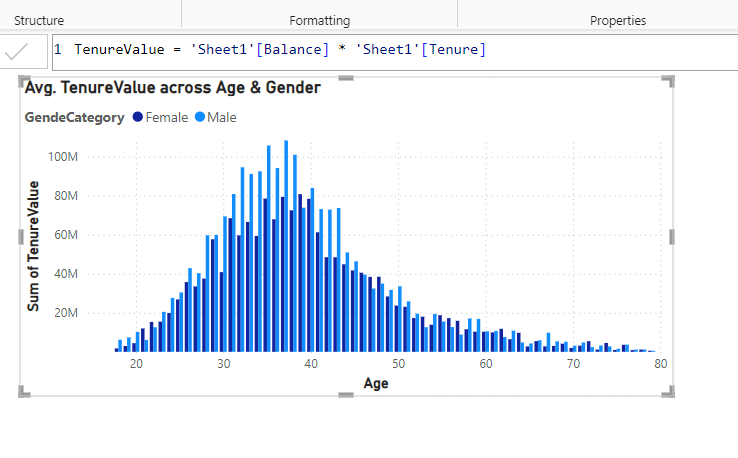
1. Risk Management Assessment: Based on customer profiles, which demographic segments appear to pose the highest financial risk to the bank, and why?

--- As per the dataset highest financial risk to the bank would be Credit score & Exit status. So, we have taken AVG credit score of all the customers and visualized them based on Demographic information such as Gender & Geography. Here, we have used Clustered column chart to shown the trend.



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 build a lifetime value in the bank we have used customer characteristics like age & gender. Here we have founded TenureValue which is calculated based on balance & tenure of a customer and we have used age and gender to allocate the trend. This visualisation conveys the customers relationship with bank, segmented by both age and gender. This helps to understand how different age groups relate to the length of time they remain with bank

.

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?

--- As per the available data we have only founded Retain customers and compared them with the total customers over time which will show us the Marketing Campaign Effectiveness over time by using retention count.



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

--- There may be several reasons for customers exit like Age, Estimated Salary, Services at different location, etc. So, we have considered Age between 18-35 as the reasons to exit the bank and founded the Exit customer rate. We have used Donut chart to visualize the rate of exit customer by age and location.

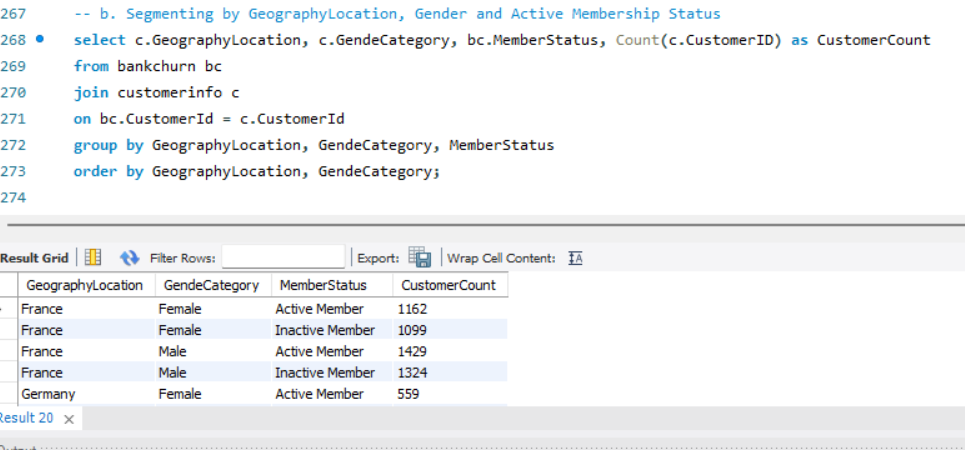


1. Are 'Tenure', 'NumOfProducts', 'IsActiveMember', and 'EstimatedSalary' important for predicting if a customer will leave the bank?

--- It potentially plays a little role in predicting customer exits. But the actual impact is quite less because each variable neds to be evaluates under different data-driven techniques, which are discussed below:-

1. Tenure: Generally, longer-tenured customers are less likely to leave, as they have established a relationship with the bank. However, new customers might leave if they find a better offers, and long-tenured customers might leave due to dissatisfaction over a product.
2. NumOfProducts: Customers with more products are usually more engaged and may find it more difficult to leave due to the complexity of moving multiple accounts. Conversely, a low number of products might lower the engagement and likelihood of exit.
3. IsActiveMember :- Active members are typically more satisfied and engaged with the bank's services, reducing their chances of leaving. Whereas, Inactive members might be less loyal or satisfied.
4. EstimatedSalary: Income levels can influence financial behaviour & product usage. For example, higher-income customers might have different expectations or might lock for premium products.
5. Utilize SQL queries to segment customers based on demographics and account details.

---- Different Customer demographic variables are include in the dataset out of which Location & gender we have used & for the account details we have select Customer Status with the bank



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?

---- Conditional formatting in Power BI can be used to visually highlight key insights, such as identifying customers at risk of churn and evaluating the impact of credit card rewards on customer retention. Here's how you can set up conditional formatting for these purposes

1. Identifying Customers at Risk of Churn: To highlight customers at risk of churn, you can create a measure based on criteria that indicate higher churn risk these includes factors like low tenure, few products, inactive membership, or low salary.

2. Evaluating the Impact of Credit Card Rewards on Customer Retention: To evaluate the impact of rewards, we can use the CrCardStatus that indicates their card holding & their rewards status. This aims to compare churn rates between credit card holders who use rewards and those who don't.

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?

--- I) To calculate the churn rate per year and overall, we need more information in our dataset like:

1. Total number of customers at the beginning of each year.
2. Number of customers who have exited during each period.

II) – Insights for Customer which are likely to Churn

There are different customer aspects and their correlation with churn.

1. Demographic Factors

* Age: Younger customers might be more likely to switch banks
* Gender: Genders can reveal target groups for retention strategies.

2. Account Details

* Tenure: Customers with shorter tenures may be more likely to churn, possibly due to lack of loyalty.
* Number of Products: Customers with fewer products are at risk of leaving, as they may not have a engagement with the bank.
* Activity Level: Inactive members may not see the value in staying, leading to higher churn rates.

3. Financial Status

* Estimated Salary: Income levels can influence customer behavior and satisfaction.

III) Strategies to Decrease Churn Rate

1. Enhance Customer Engagement:

* Use personalized messages and offers to keep customers informed and engaged.
* Implement programs that reward customers for continued use of bank services.

1. Improve Product Offerings:

* Offer products that cater to the specific needs of customer in different segments (e.g., young professionals, families, retirees).
* Inspire customers to use multiple products by offering bundled services at a discounted rate.

1. Customer Support and Experience:

* Improve Customer Service to address issues promptly.
* Implement regular feedback to understand customer concerns and improve services accordingly.

1. Proactive Retention Strategies:

* Identify customers likely to churn and proactively reach out to them.
* Provide special offers or incentives to customers at-risk.

1. Financial Education and Advisory Services:

* Offer financial planning and advisory services to help customers manage their finances better.
* Provide educational content to help customers understand the value of different financial products.

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

----- Key Performance Indicators (KPIs) and Metrics

1. Churn Rate: It Measures the rate at which customers are leaving the bank over the years, which is crucial for understanding customer satisfaction and retention.

2. Retention Rate: Shows the percentage of retained customers, providing insights into customer loyalty.

3. Customer Lifetime Value (CLV): Predicts the estimated salary of a customer which will generate over their lifetime, helping prioritize high-value customers.

4. Customer Segmentation Metrics: Helps in understanding different customer age groups, enabling us to serve different product offerings.

5. Risk Score: Identifing high-risk customers, based on TenureScore, ProductScore, ActiveScore, SalaryScore, which enable us to make proactive retention strategies.

6. Average Account Balance: Indicates financial engagement and potential for upselling based on different age groups

7. Geographic Trends: Analyses how different regions are performing, which can guide regional marketing strategies.

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

---- When no specific guidance or objectives are provided, here's how you can approach to create a Power BI dashboard:

1. Understand the Dataset: Review the dataset & understand its structure, data type, and patterns. Look at the columns and few rows.
2. Business Objectives: Consider common objectives such as sales performance, customer behaviour, operational efficiency, or financial health, etc…
3. Data Cleaning: Identify and clean any missing values, or outliers. Ensure that data is properly formatted.
4. Basic Analysis: Perform basic statistics and visualizations to understand the dataset and what can be extracted from it. Correlate relationships with different columns.
5. Idealize Possible KPIs: Based on your basic analysis, visualize potential key performance indicators (KPIs) that might be relevant to various stakeholders.
6. Create assumptions: Based on the data, guess about what might be valuable to explore (e.g., trends over time, different categories).
7. Basic Visualizations: Create basic visualizations like charts, tables, and graphs to display key metrics and trends. Add slicers and filters to allow users to interact with the data and explore different aspects of the dataset.
8. Engage with Stakeholders & working on feedbacks: Share the initial dashboard with stakeholders or team members to get feedback and discuss, any additional information they might need. Refine the dashboard based on feedback, adding or adjusting visualizations and metrics as necessary.
9. Findings & Insights: Provide explanations for the visualizations and metrics included in the dashboard. Prepare a brief summary or presentation of the insights derived from the dashboard to help stakeholders understand the findings.
10. In the “Bank\_Churn” table how can you modify the name of the “HasCrCard” column to “Has\_creditcard”?

---- Please find the below way to modify the column name in SQL & Power BI

1. MySQL: ALTER TABLE Bank\_Churn RENAME COLUMN HasCrCard TO Has\_creditcard;

2. Power BI: Open the file in Power BI Desktop, select the HasCrCard column in the DATA icon in the right sidebar. Right-click on the column and select Rename. Type “Has\_creditcard” and press Enter.