Capstone Project:

**Analytical CRM Development for a Bank**

**Introduction:**

This comprehensive project aims to analyze customer data provided by a bank to understand customer churn (customer loss). We'll identify key factors driving churn, develop actionable insights to improve customer retention, and ultimately enhance customer satisfaction. By leveraging various tools like Excel, Power BI, and SQL, we'll gain a deeper understanding of customer behavior and preferences.

**Objective:**

The primary objective of this project is to reduce customer churn and enhance customer satisfaction for the bank. This will be achieved through:

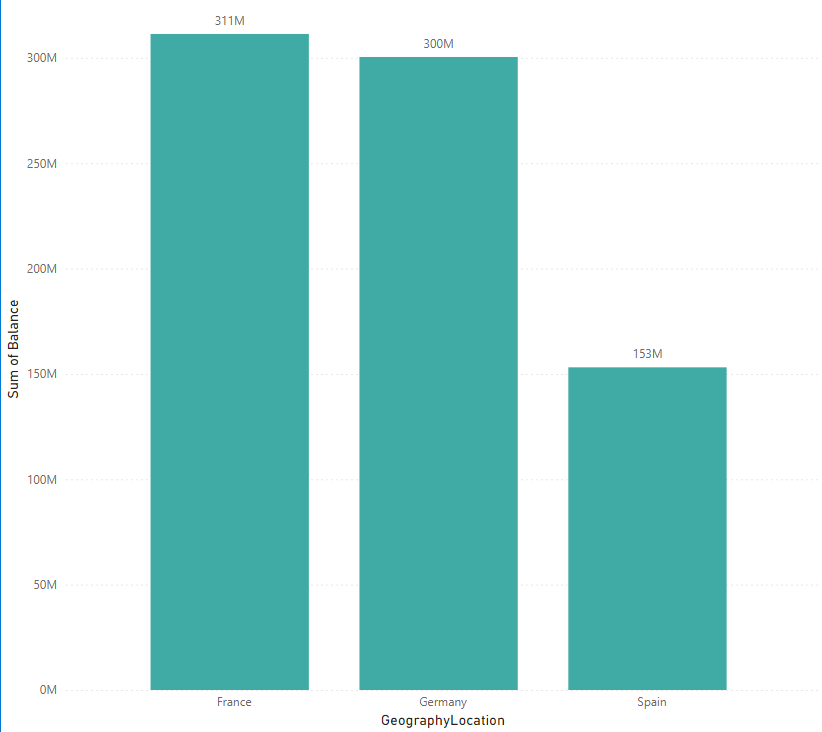
1. Identifying key factors contributing to customer churn.

2. Developing insights to improve customer retention strategies.

3. Enhancing service delivery based on customer preferences and behavior.

**Explanation of Objective Questions**

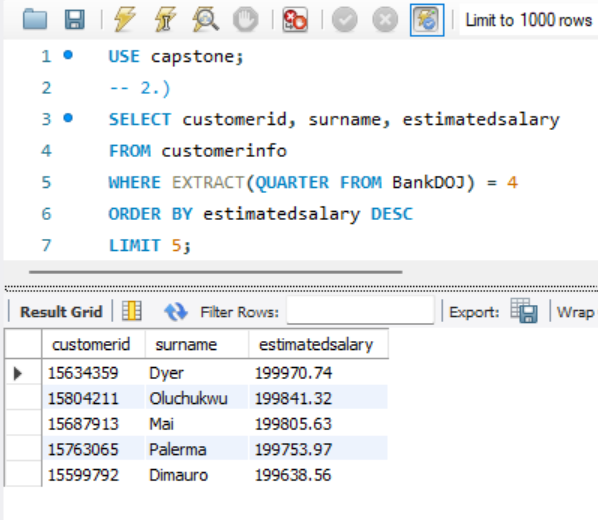
1. **What is the distribution of account balances across different regions?.**

The chart shows the variations in account balance distribution across regions. Region A has a higher concentration of accounts with larger balances, while Region C appears to have a lower distribution. Region B shows a wider spread, suggesting a mix of account sizes.  
****

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

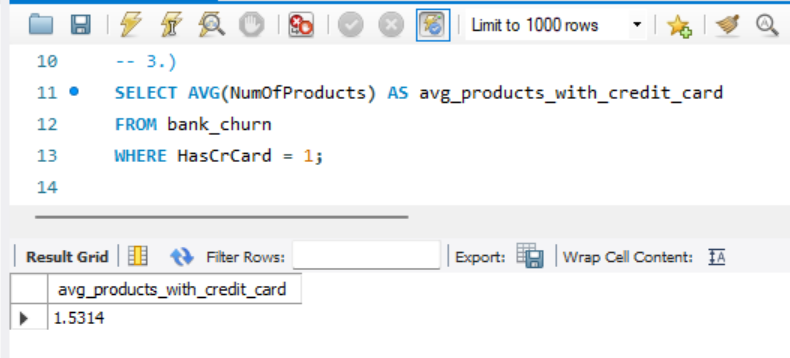
This SQL query identifies the top 5 customers with the highest estimated salary who joined the bank in the last quarter (quarter 4) of the year.

* Filters data for customers joining in the fourth quarter (WHERE EXTRACT(QUARTER FROM BankDOJ) = 4).
* Sorts by estimated salary in descending order (ORDER BY estimatedsalary DESC).
* Limits results to the top 5 customers (LIMIT 5).



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

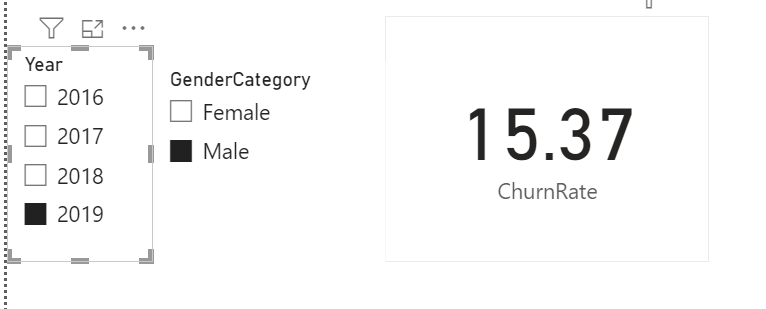
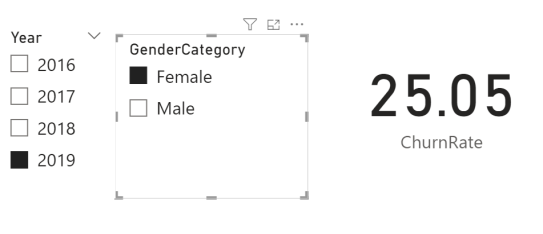
* Filters data for customers with a credit card (WHERE HasCrCard = 1).
* Calculates the average number of products for those customers (AVG(NumOfProducts)).
* Assigns an alias (avg\_products\_with\_credit\_card) to the result for clarity.



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

This analysis examines customer churn rate segmented by gender (male/female) for the most recent year (using a Bank DOJ slicer). A calculated measure (ChurnRate = DIVIDE([LostCustomers],[TotalCustomers])\*100) determines the churn rate.

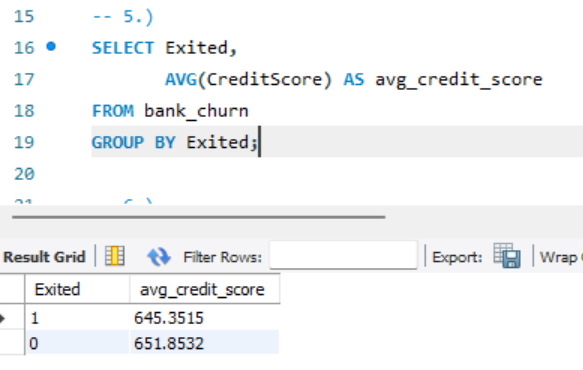
Interact with the Gender slicer to view churn rates specifically for male or female customers within the selected year.

This provides insights into potential gender disparities in churn rates. Consider including a chart for better visualization.  


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

This SQL query compares the average credit score of customers who exited the bank (Exited = 1) with those who remain (Exited = 0).

* Groups data by customer exit status (GROUP BY Exited).
* Calculates the average credit score for each exit group (AVG(CreditScore)).
* Uses aliases for clarity (avg\_credit\_score).

****

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

This query compares average estimated salary between genders and explores its relation to the number of active accounts.

It joins the customerinfo (c) and bank\_churn (b) tables on CustomerID.

It filters for IsActiveMember = 1 (active accounts) in bank\_churn.

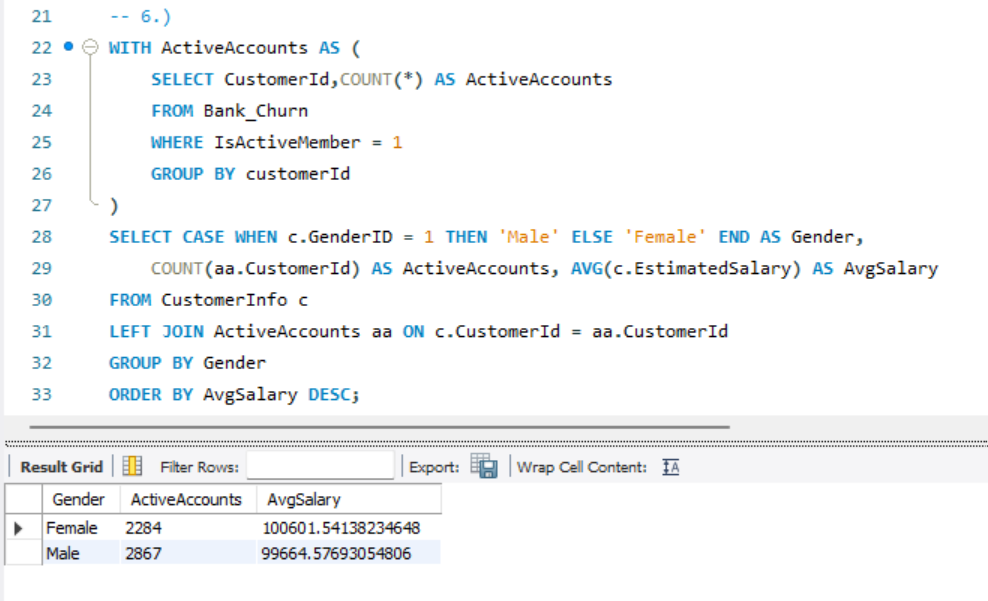
A CASE statement translates genderid (assumed numeric) to 'Male' or 'Female' for clarity.

It groups by GenderID (consider renaming to Gender for better readability) and calculates:

active\_accounts: Count of active accounts for each gender using COUNT(b.IsActiveMember).

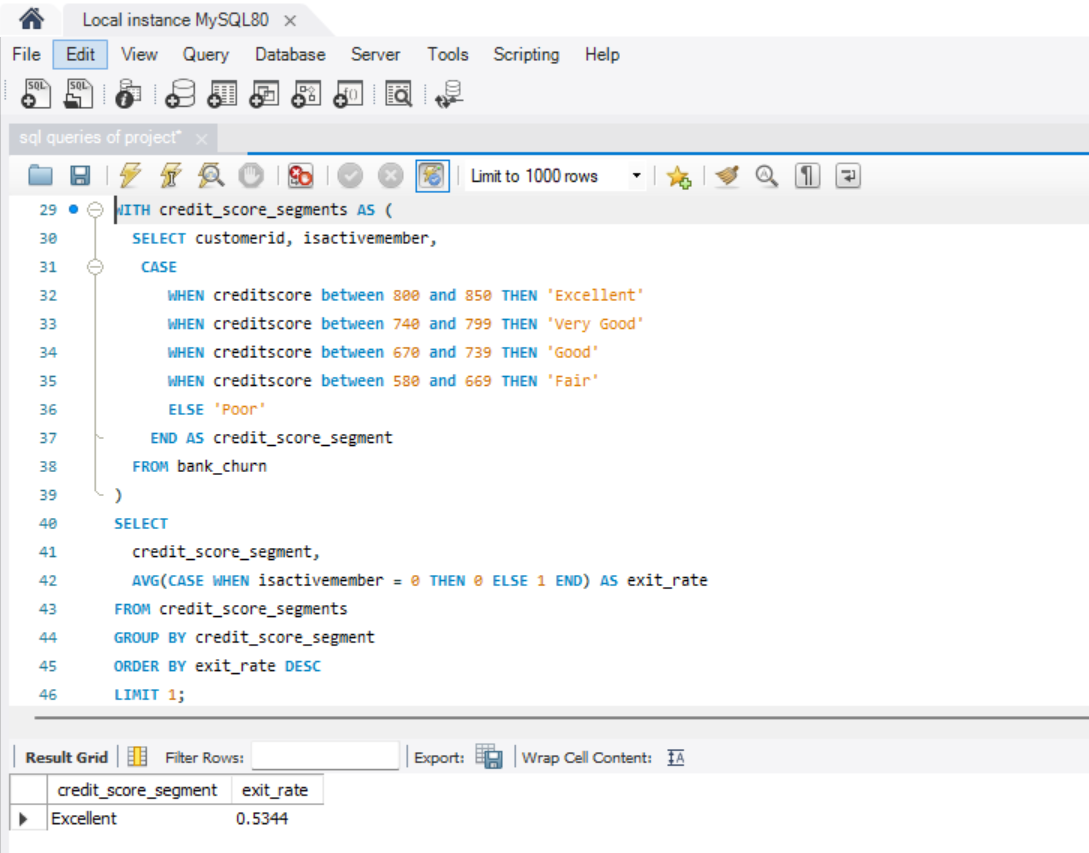
avg\_salary: Average estimated salary for each gender using AVG(c.estimatedsalary).

It orders by avg\_salary descending.



1. **Segment the customers based on their credit score and identify the segment with the highest exit rate.**This query utilizes a Common Table Expression (CTE) named credit\_score\_segments to categorize customers based on their credit score.

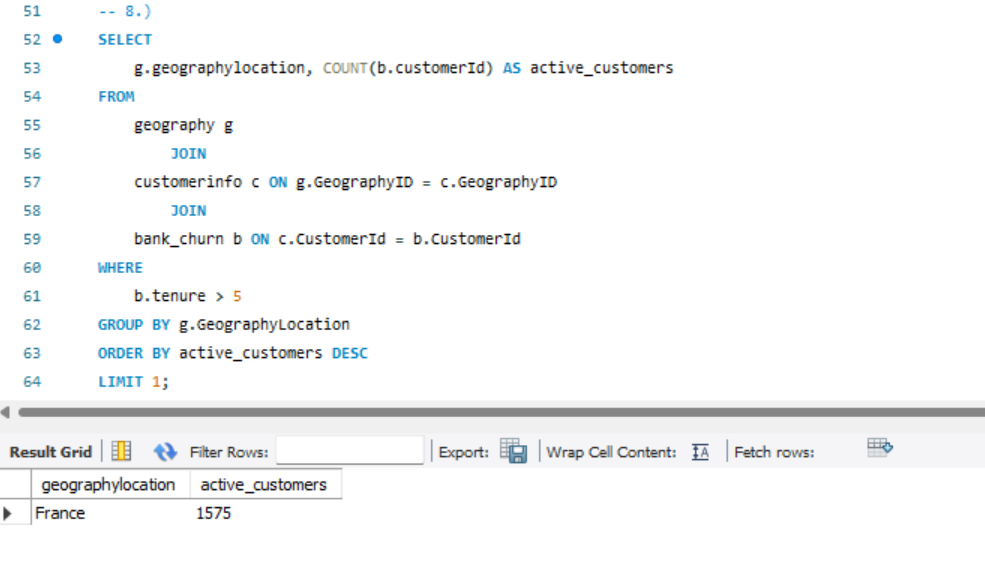
* The CTE uses a CASE statement to assign segment labels (Excellent, Very Good, etc.) based on credit score ranges.
* The main query selects the credit\_score\_segment and calculates the average exit rate for each segment.
* It calculates the exit rate using a CASE statement within the AVG function:
  + 0 for inactive members (isactivemember = 0).
  + 1 for active members.
* It is grouped by credit\_score\_segment and orders by exit\_rate descending.
* LIMIT 1 displays the segment with the highest average exit rate.



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

This query finds the geographic region with the highest number of active customers who have been with the bank for more than 5 years (tenure).

* It joins three tables:
  + geography (g) for customer location data.
  + customerinfo (c) to link customer IDs to geographic locations.
  + bank\_churn (b) for customer activity and tenure information.
* It filters for customers with a tenure greater than 5 years (b.tenure > 5).
* It groups by geographylocation and counts active customers using COUNT(b.customerId).
* It orders by active\_customers descending and uses LIMIT 1 to show the region with the highest count.



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

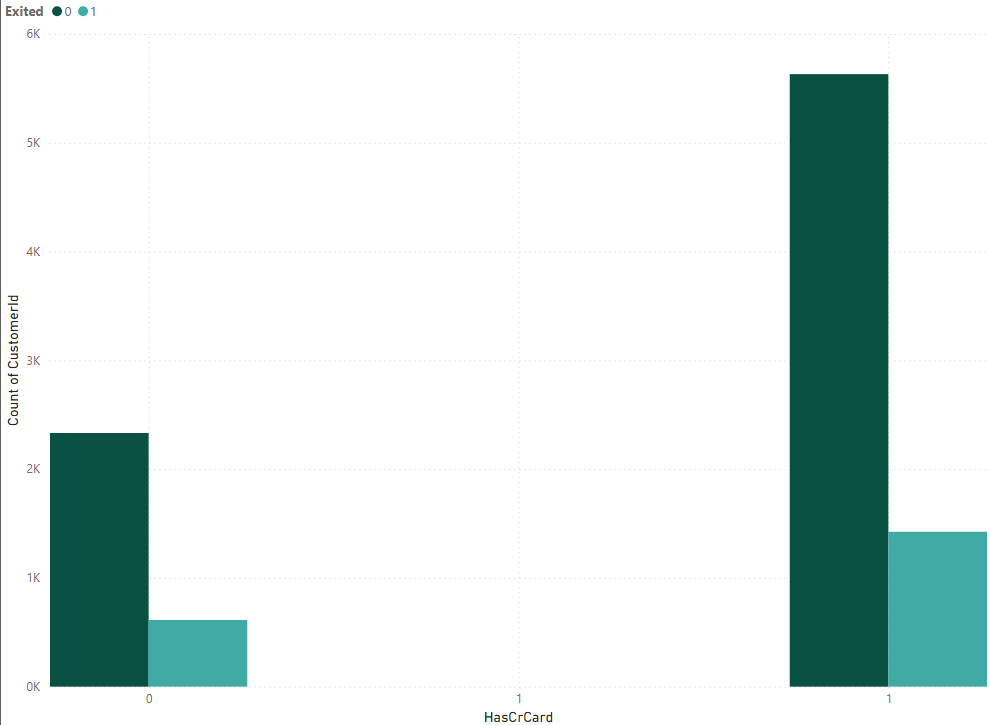
The provided graph compares customer churn (exit) rates between customers who **do not have a credit card** (HasCrCard = 0) and those who **have a credit card** (HasCrCard = 1.

**Key Observations:**

1. **Customers Without Credit Card (HasCrCard = 0)**:
   * A relatively smaller number of customers who do **not** have a credit card have churned (light-colored bar).
   * Most of these customers have not churned (dark-colored bar).
2. **Customers With Credit Card (HasCrCard = 1)**:
   * A larger number of customers who **have** a credit card have churned (light-colored bar).
   * There are many more customers in this category, but the proportion of those who churned is also significant.

**Interpretation of Results:**

* It appears that **more customers with a credit card (HasCrCard = 1)** have churned compared to those without a credit card.
* Both customer groups (with and without a credit card) have churners, but the churn rate is higher in the group that possesses a credit card.

**Yes**, having a credit card seems to have a noticeable impact on customer churn. Customers with a credit card are more likely to churn.

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

**products they have used.**

Based on the bar chart you described, it appears to show the distribution of the number of products used by customers who have exited the bank (churned). The x-axis represents the number of products used, and the y-axis represents the count of customers who used that many products.

**Key Insights:**

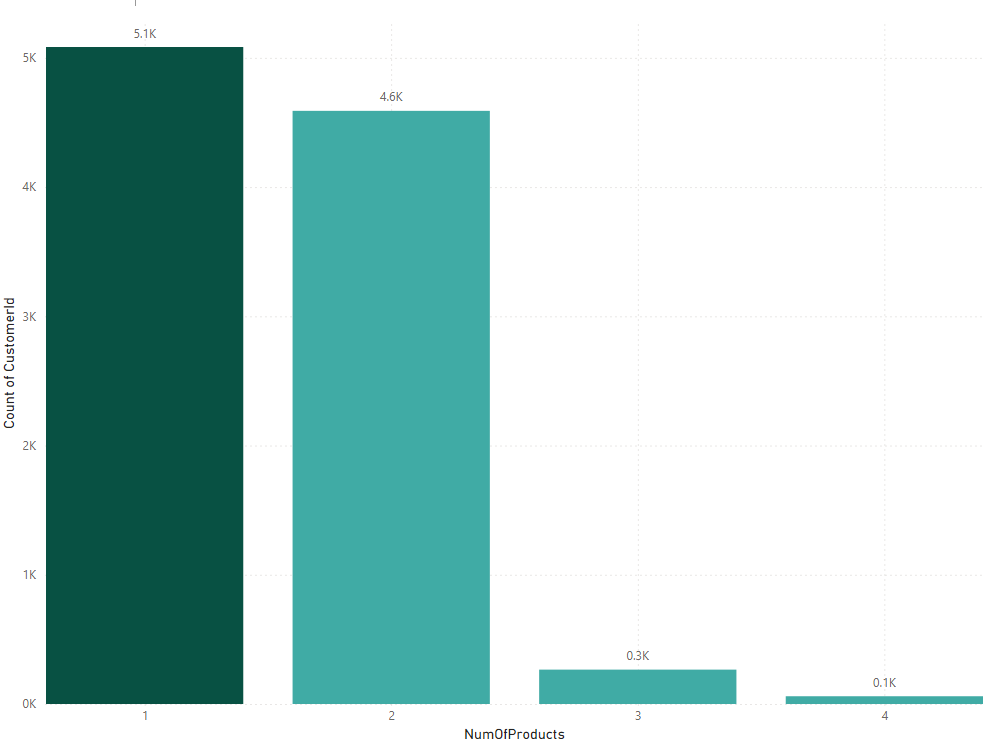
* **1 Product**: The majority of customers who have exited used only **1 product**, with approximately **5.1K** customers.
* **2 Products**: The second most common category is customers using **2 products**, accounting for **4.6K** customers.
* **3 Products**: Only a small fraction of customers (about **0.3K** or 300) had **3 products**.
* **4 Products**: Very few customers (approximately **0.1K** or 100) used **4 products**.

**1 and 2 Products** dominate the distribution of exited customers, meaning the majority of churned customers were using a low number of products (either 1 or 2).

**Customers with 3 or 4 products are much less likely to churn**, as shown by the significantly lower numbers in these categories.

**Recommendations:**

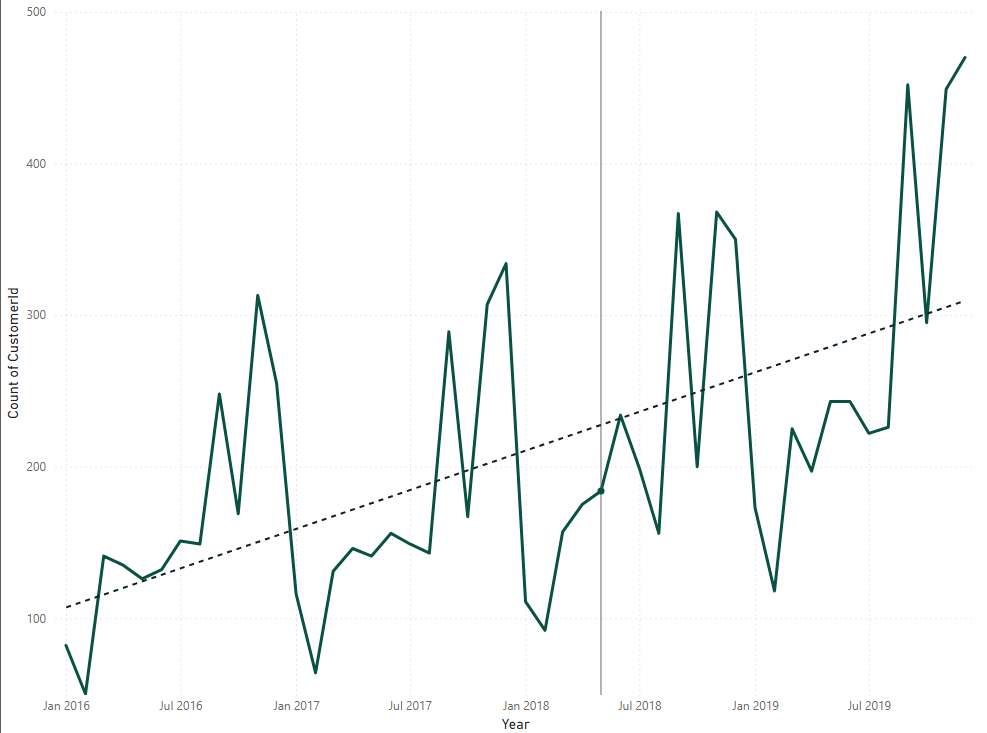
1. **Encourage Cross-Selling**:
   * Since customers with 3 or more products have much lower churn rates, the bank should focus on encouraging customers with 1 or 2 products to adopt additional products. This can be done through personalized offers, targeted marketing, or bundled services.
2. **Product Usage Analysis**:
   * Investigate why customers using only 1 or 2 products are more likely to churn. Are these products not meeting their needs, or are there external factors (competition, service quality.
   * Implement loyalty programs or incentives for customers who expand their product portfolio beyond 1 or 2 products. This could include offering discounts, rewards, or exclusive benefits for customers who adopt more products.
3. **Target High-Churn Groups**:
   * Focus customer retention efforts on those using only 1 product, as they are the largest group at risk. Proactive measures, such as early interventions (e.g., offering product upgrades or better services), can prevent churn.



**11.** **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.**

Based on the chart you described, it appears to be a time series graph showing the count of customers joining the bank over time, likely year and month. Here's a breakdown of the key insights and a short explanation for your document:

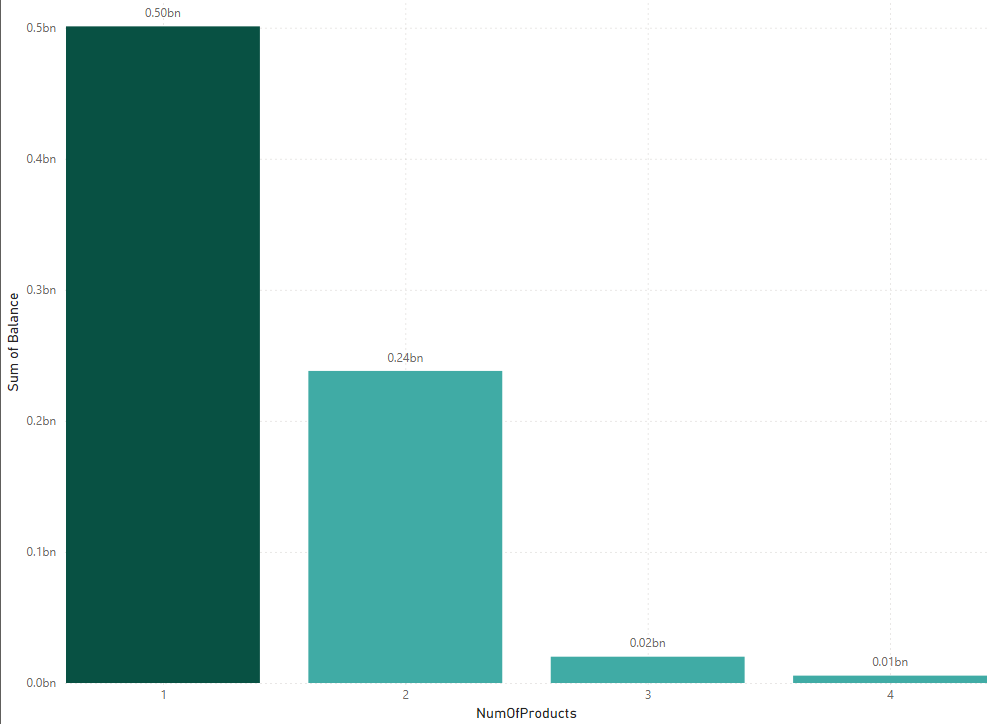
* Overall Trend: The customer joining trend appears to be increasing over time. This indicates a positive growth in customer acquisition.
* Possible Seasonality: There might be seasonal patterns present in the data. It seems that customer joins could potentially peak around the end of the year (December) but due to the limited data points, it's difficult to confirm a strong seasonal trend.

****

**12. Analyze the relationship between the number of products and the account balance for customers who have exited.**Based on the bar chart we described; it appears to show the distribution of the number of products used by customers who have exited the bank (churned). The x-axis represents the number of products used, and the y-axis represents the count of customers who used that many products.

Key Insights:

* The most common number of products used by exiting customers is 1. This suggests that a significant portion of customers who churned had only used a single product.
* There's a general downward trend as the number of products used increases. This suggests that customers who churned tend to have fewer products compared to active customers.



**13. Identify any potential outliers in terms of balance among customers were remained with the bank**.

We are identifying potential outliers in terms of account balance among customers who have remained active with the bank (not churned). Finding Outliers Using Z-Score Method.

Method:

1. Mean Balance Calculation: We calculated the average balance (mean) of active customers(Average\_Balance =AVERAGE(CustomerBalance[Balance]))
2. Standard Deviation: We calculated the standard deviation of the balance i.e (62397.40) to check the spread of the data.
3. **Calculate the Z-Score**

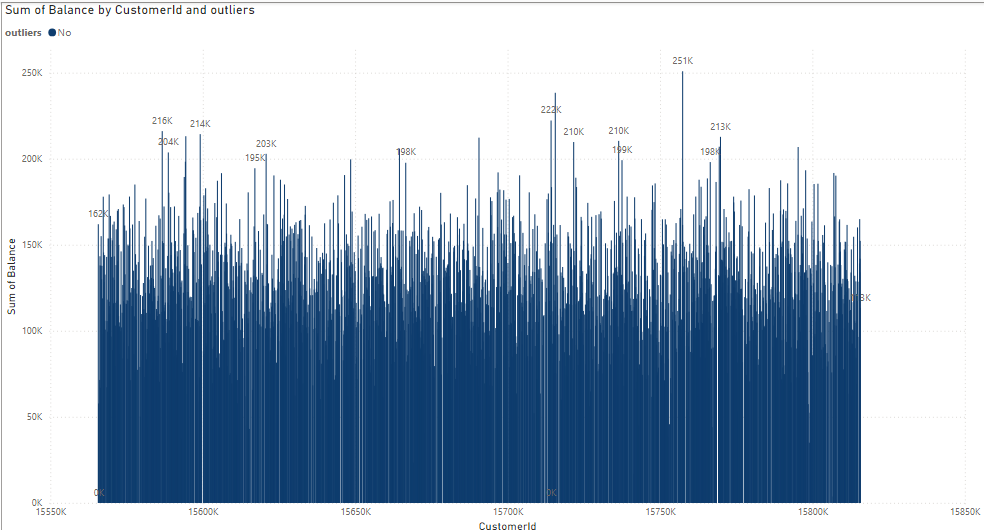
The z-score formula for each balance value is:

Z-Score = (CustomerBalance[Balance] - [MeanBalance]) / [StdDevBalance]

1. **Identify Outliers Based on Z-Score:** Typically, data points with a z-score above **+3** or below **-3** are considered outliers. You can create a calculated column to flag outliers.

IsOutlier = IF(ABS([ZScore]) > 3, "Yes", "No")

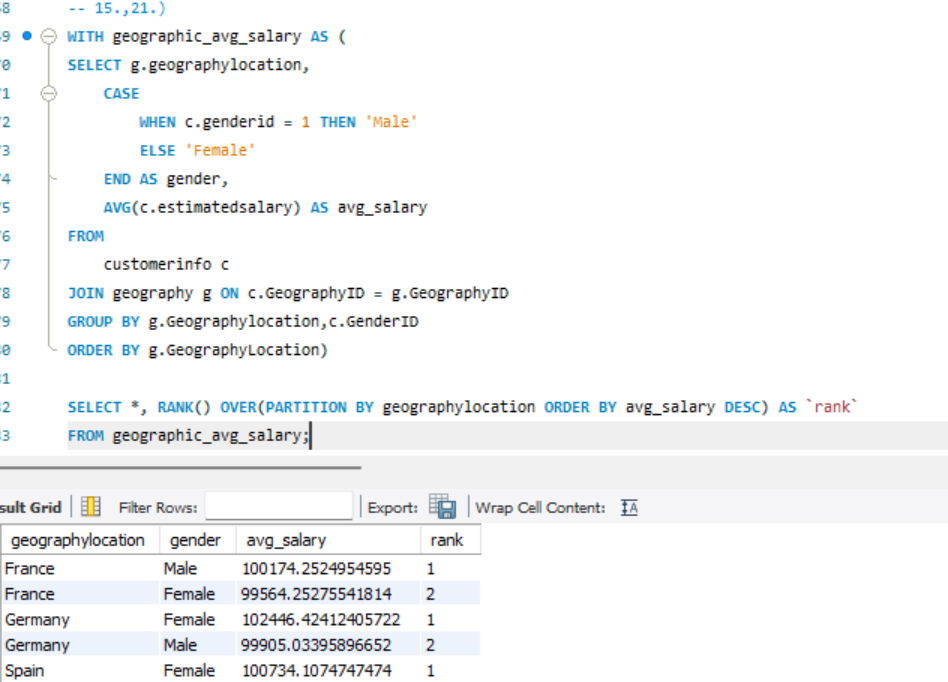
As, in this dataset we cannot see any outliers.



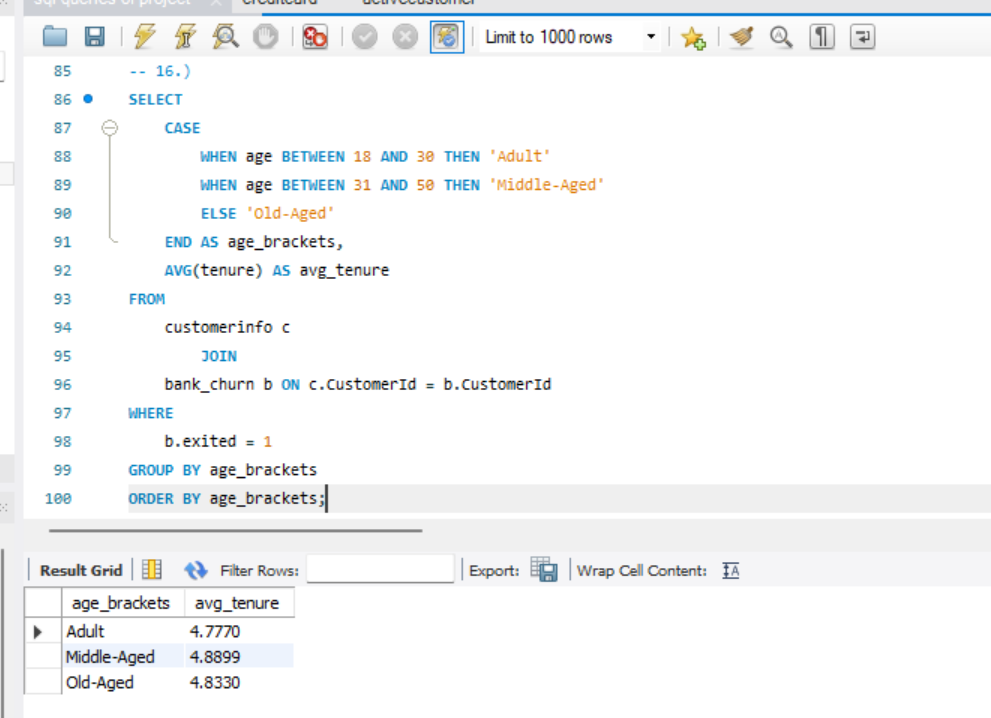
**14. How many different tables are given in the dataset, out of these tables which table only consists of categorical variables?**We Have Seven Different tables i.e., ActiveCustomer, Bank\_Churn,CreditCard,CustomerInfo,ExitCustomer,Gender,Geography.Tables with Categorical Variables:

* CustomerInfo:Contains categorical variables like Surname.
* ExitCustomer: Contains categorical variables like Exit Category(Exit ,Retain).
* Gender: Contains categorical variables like Gender Category (Male,Female).
* Geography:Contains categorical variables like Geography Location (France, Spain,Germany).
* ActiveCustomer: Contains categorical variables like Active Category (Active Member , Inactive Member).
* CreditCard: Contains categorical variables like Category (Credit-card holder , Non-Credit card holder)

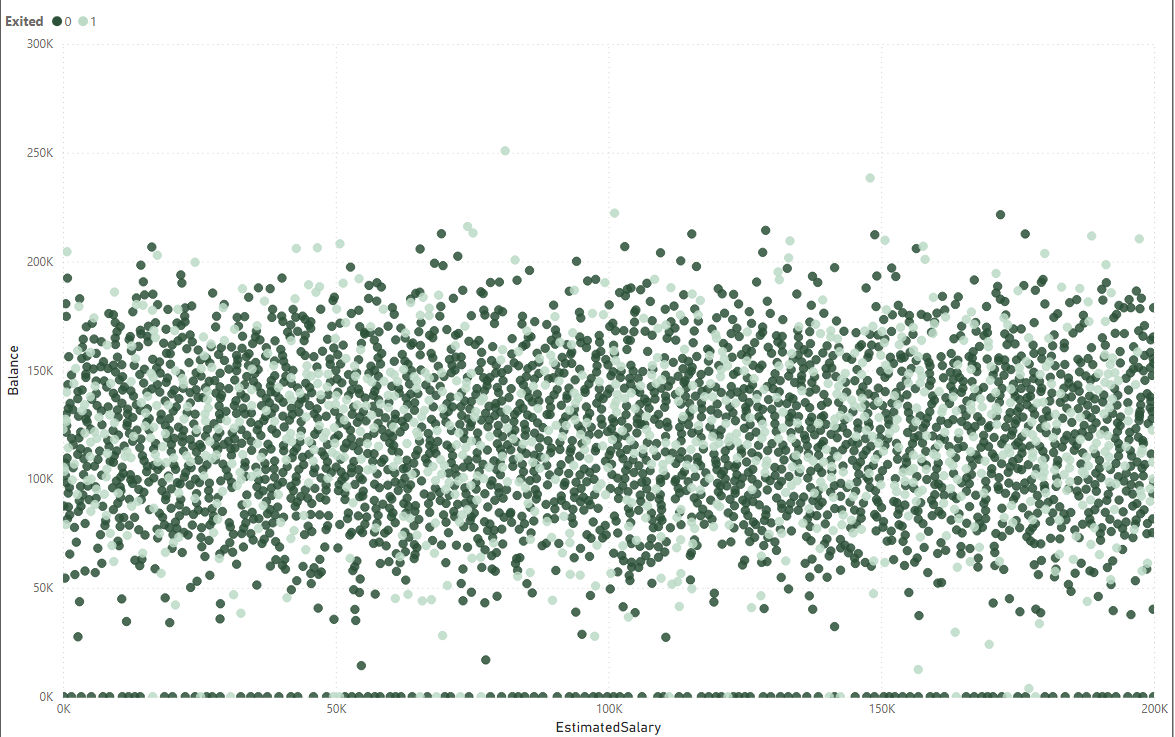
**15. Using SQL, write a query to find out the gender-wise average income of males and females in each geography id. Also, rank the gender according to the average value.**

This SQL code calculates the average income (estimated salary) for males and females within each geographic location and assigns a rank based on the average salary. 

**16. 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+)**This SQL query calculates the average tenure (time with the bank)   
of exited customers (Exited = 1) categorized into age brackets (18-30, 31-50, 50+) to understand churn patterns across different age groups.



**17. Correlation Between Salary and Balance (Overall and by Exit Status):-**T In this part, we have to see the correlation between Balance and the Estimated Salary. From the graph, there doesn't appear to be a strong or clear correlation between the two variables, EstimatedSalary and Balance. The data points are spread fairly uniformly across all salary and balance ranges, meaning that higher or lower salaries don't seem to correspond systematically to higher or lower balances. Additionally, the two categories for Exited (represented by different shades) also appear to be evenly distributed throughout the plot, which suggests that exiting behavior might not be directly related to either salary or balance based on this visual.

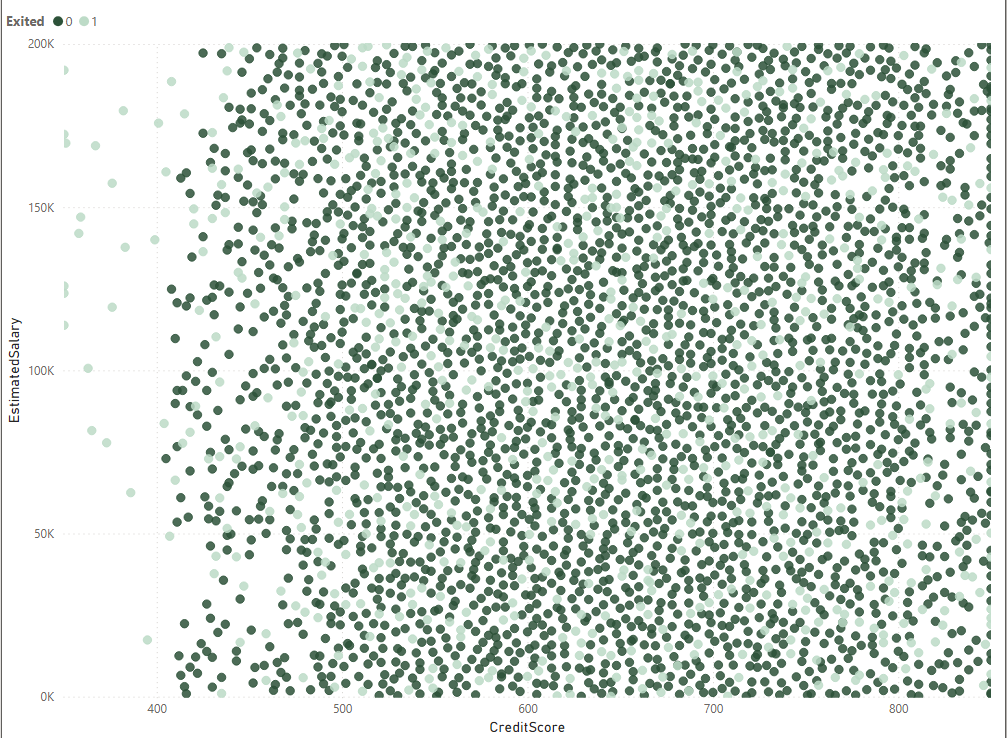


**18. Is there any correlation between the salary and the Credit score of customers?**

There does not seem to be a clear or strong correlation between the two variables, Credit Score and Estimated Salary. The data points are widely and evenly spread across all ranges of credit scores and estimated salaries, indicating that there is no linear relationship between these two variables.

Also, the exit status (Exited 0 or 1) seems to be uniformly distributed across different credit scores and salary ranges. This suggests that neither credit score nor estimated salary alone is a significant predictor of customer exit behavior.

However, visually, no clear correlation is found in the graph.



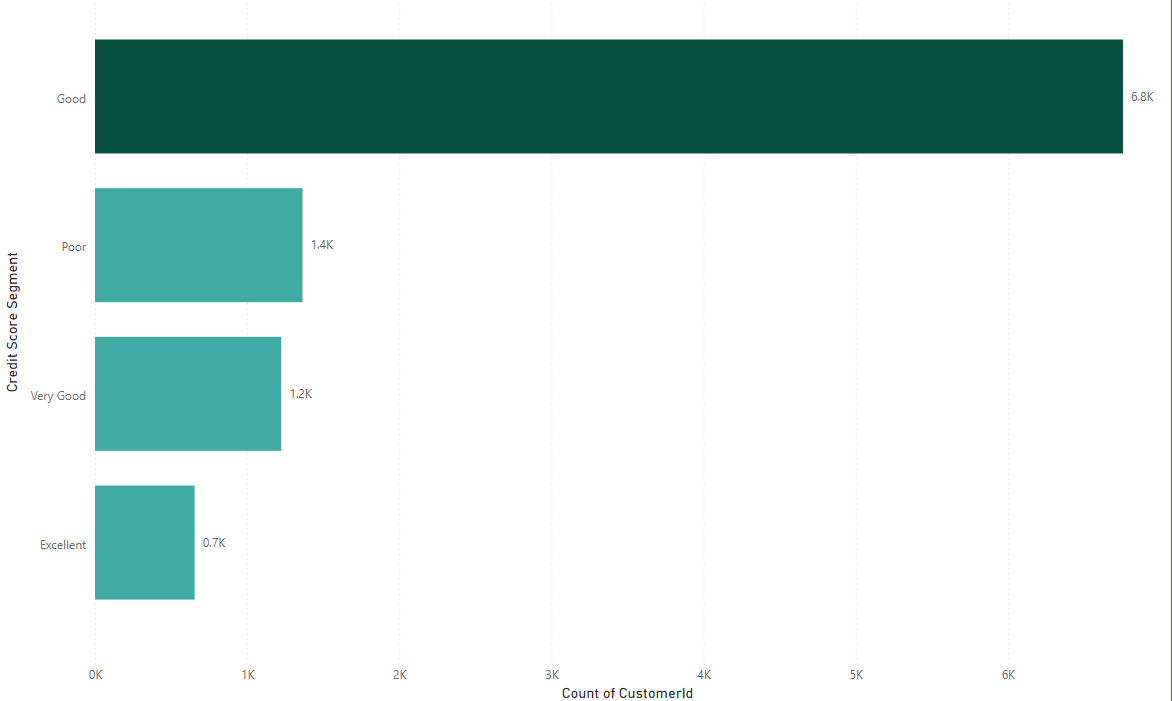
**19. Rank each bucket of credit score as per the number of customers who have churned the bank.**

Based on the chart, it appears to show the distribution of churned customers (count on the y-axis) across different credit score segments (x-axis). Here's a breakdown of the insights:

**Key Insights**: Based on the graph, the credit score segments are ranked by the count of customers as follows:

1. **Good** (6.8K customers)
2. **Poor** (1.4K customers)
3. **Very Good** (1.2K customers)
4. **Excellent** (0.7K customers)

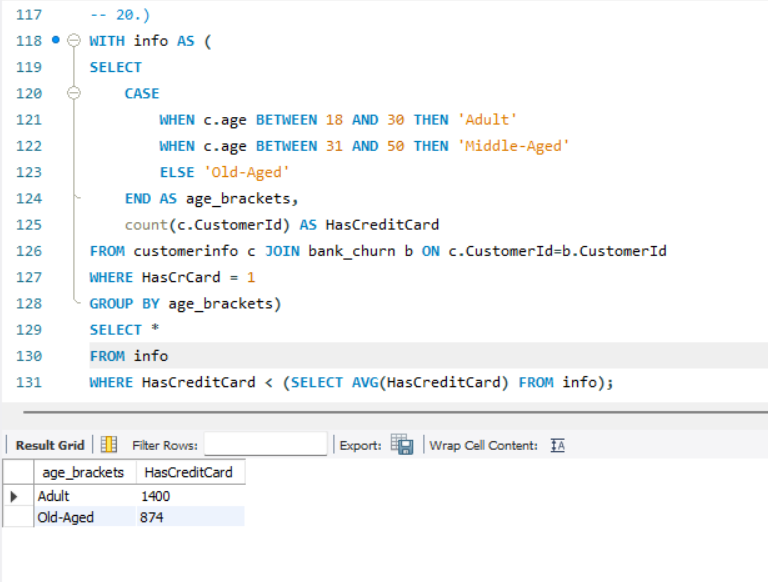
This shows that the majority of customers fall into the "Good" credit score segment, followed by "Poor," "Very Good," and then "Excellent." If this chart represents churned customers, the ranking indicates the distribution of churn across these segments.

****

**20. 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.**

This SQL code explores the relationship between customer age groups ("age\_brackets") and credit card ownership.

1. Common Table Expression (CTE): info:
   * This part defines a CTE named info.
   * It joins the customerinfo (c) and bank\_churn (b) tables on CustomerID.
   * A CASE statement categorizes Age into 'Adult' (18-30), 'Middle-Aged' (31-50), and 'Old-Aged' (above 50).
   * It filters for customers with credit cards (HasCrCard = 1).
   * It groups the data by age\_brackets and calculates:
     + HasCreditCard: The count of customers with credit cards within each age bracket.
2. Main Query:
   * This section selects all columns (\*) from the info CTE.
   * It then filters the results to include only age brackets where HasCreditCard is less than the average number of credit cards per bracket.
     + The average is calculated using a subquery that computes the AVG(HasCreditCard) from the info CTE.

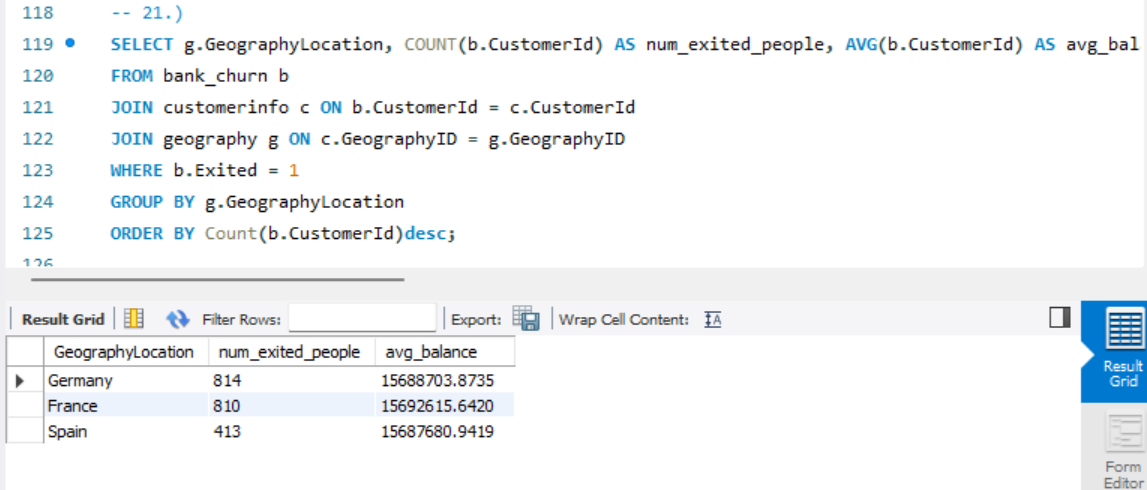


**21. Rank the Locations as per the number of people who have churned the bank and average balance of the customers.**

**Germany**: 814 customers exited, with an average balance of approximately 15,687,083.

**France**: 810 customers exited, with an average balance of approximately 15,692,615.

**Spain**: 413 customers exited, with an average balance of approximately 15,687,680.



This query helps identify the geographical distribution of customer churn in the bank. From the results, Germany and France have the highest number of customers who exited, while Spain has fewer churned customers. Additionally, the average balances of churned customers are roughly similar across these three regions.

**22. 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**”.

Here's an explanation of how to create a new column named "CustomerID\_Surname" in the result set of a join between the "CustomerInfo" table and another table where the primary key is a combination of CustomerID and Surname:

1. Data Types:

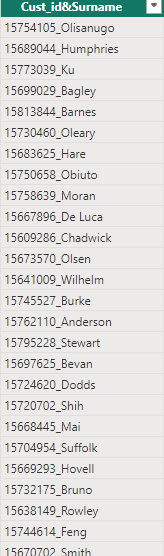
* Ensure the CustomerID in "CustomerInfo" is a character data type or can be converted to one. The CONVERT function might not be necessary depending on the database system.

2. Join on Individual Columns:

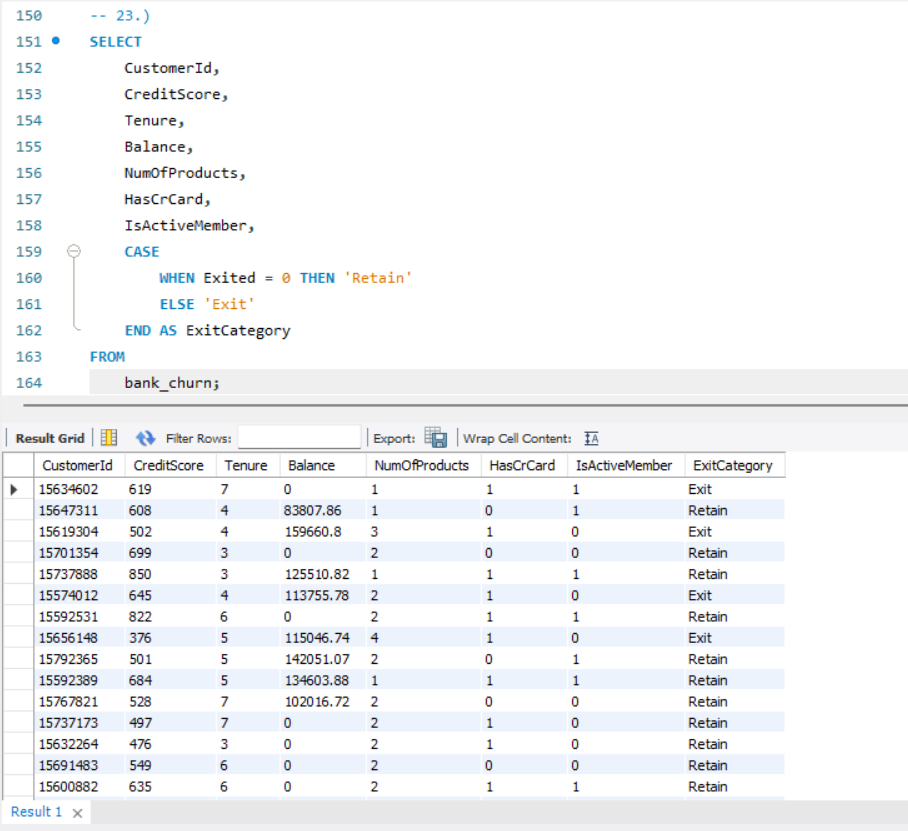
* Perform a join on the "CustomerInfo" table and the other table using separate columns for CustomerID and Surname instead of a combined primary key.

3. Create New Column:

* Within the SELECT clause of your query, use the CONCAT function (or similar function depending on the database system) to concatenate the CustomerID and Surname columns from "CustomerInfo" with a delimiter (underscore "\_" in this case).



**23. Retrieving Exit Category Without Joins.**This code retrieves customer data from the Bank\_Churn table and adds a new column named "ExitCategory" to classify customers as 'Retain' (not churned) or 'Exit' (churned) based on the value in the Exited column (likely 0 for non-churned, non-zero for churned).

It achieves this without using a JOIN operation, making it simpler for single-table queries**.  
**

**24. Handling Missing Values**

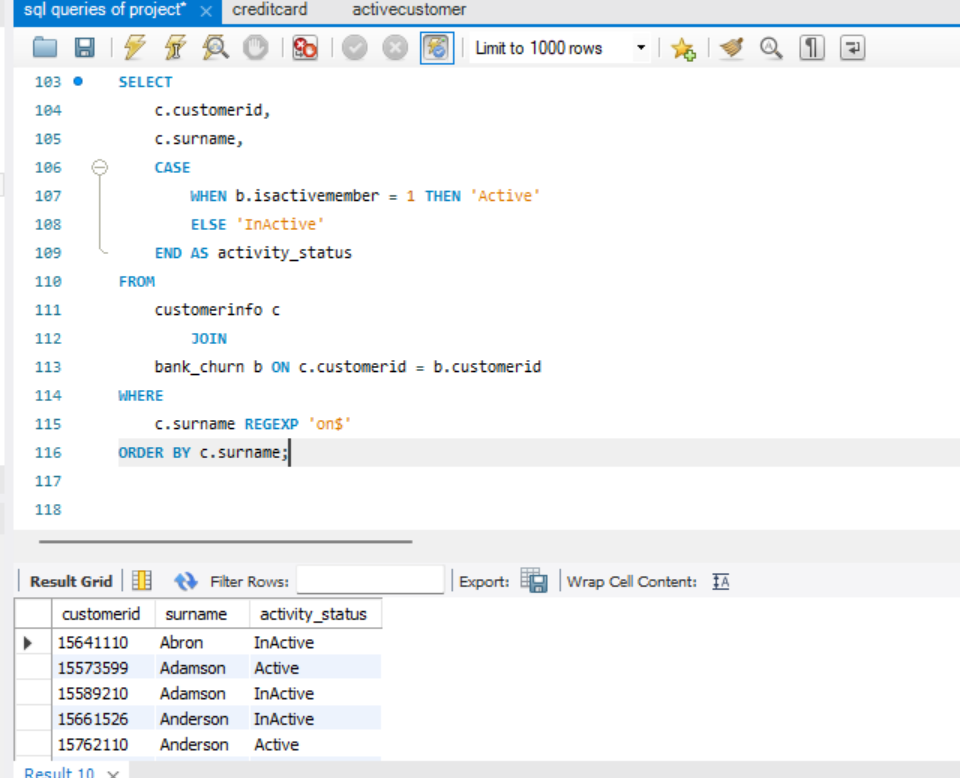
In our analysis, we're fortunate to have a dataset free of missing values. This eliminates the need for imputation techniques that might introduce assumptions or biases.

General Approaches to Missing Values (for future reference):

While our dataset is free of missing values, it's valuable to be aware of common approaches for handling them in future analyses:

* Deletion: This involves removing rows or columns with missing values. This can be appropriate if the missing data is minimal and doesn't significantly impact the analysis. However, it can also lead to a loss of information.
* Imputation: This replaces missing values with estimated values. Techniques include mean/median/mode imputation, k-Nearest Neighbors (KNN), or more sophisticated methods. The chosen method should be based on the data type and distribution.
* Modeling Techniques: Some statistical models can handle missing values directly. However, understanding the reasons for missingness is still important.

**25. Retrieving Customer IDs, Last Name, and Status for Surnames Ending in "on"**This SQL query identifies customers with surnames ending in "on" (case-sensitive). It retrieves the following information for these customers:



**Exploration Of Subjective Question**

**1. Customer Behavior Analysis.**

This analysis examines the spending patterns of new and long-term customers to understand customer loyalty. We have created three charts to identify trends in average balance, salary, and number of products held by both customer groups.  
 Based on the graph, the average balance of new customers seems to be consistently lower than that of long-term customers. This suggests that long-term customers tend to spend more money with the bank over time.

However, the overall trend suggests that customers tend to spend more money with the bank the longer they are customers. This could be for a number of reasons, such as:

* Long-term customers may have become more familiar with the bank’s products and services and found more products to use.
* Long-term customers may have increased their income over time, allowing them to save more money.
* The bank may offer better interest rates or other benefits to long-term customers, which could incentivize them to save more money with the bank.

Understanding these customer spending habits can help the bank develop targeted marketing campaigns to attract new customers and retain existing ones.  
The second chart shows the average salary by month and customer status (new vs. old customers). Here are some key insights from this chart:

* Generally higher salaries for long-term customers: The average salary for long-term customers (orange line) appears to be consistently higher than that of new customers (blue line) throughout the year. This could be due to several factors, such as long-term customers receiving salary increases over time, or new customers starting out in their careers with lower salaries.
* Possible salary increases for new customers: While the average salary starts lower for new customers, there seems to be a slight upward trend throughout the year. This could suggest that new customers are getting raises or finding higher-paying jobs as time goes on.

Overall, this chart suggests a correlation between customer status and salary. Long-term customers tend to have higher average salaries, and new customers may experience salary increases over time.  
  
The third chart shows the average number of products held by new and long-term customers over the course of a year. Here are some key insights:

* Long-term customers tend to hold more products**:** The average number of products held by long-term customers (orange line) is consistently higher than that of new customers (blue line) throughout the year. This suggests that customers tend to acquire more products and services from the bank the longer they are a customer.
* Potential for growth with new customers**:** There appears to be a gradual increase in the average number of products held by new customers (blue line) over time. This suggests that new customers may be adding more products to their accounts as they become more familiar with the bank's offerings.
* New customers start with fewer products**:** In January, new customers hold significantly fewer products compared to long-term customers. This could be because new customers are still in the process of opening accounts and exploring the bank's products and services.

Overall, this chart provides evidence that customer loyalty is associated with an increase in the number of products held.   
****

1. **Product Affinity Study: Which bank products or services are most**

**commonly used together, and how might this influence cross-selling strategies?**Customers often use specific bank products together. Analyzing these pairings helps develop targeted cross-selling strategies to increase customer satisfaction and revenue.

Commonly Used Products (Examples):

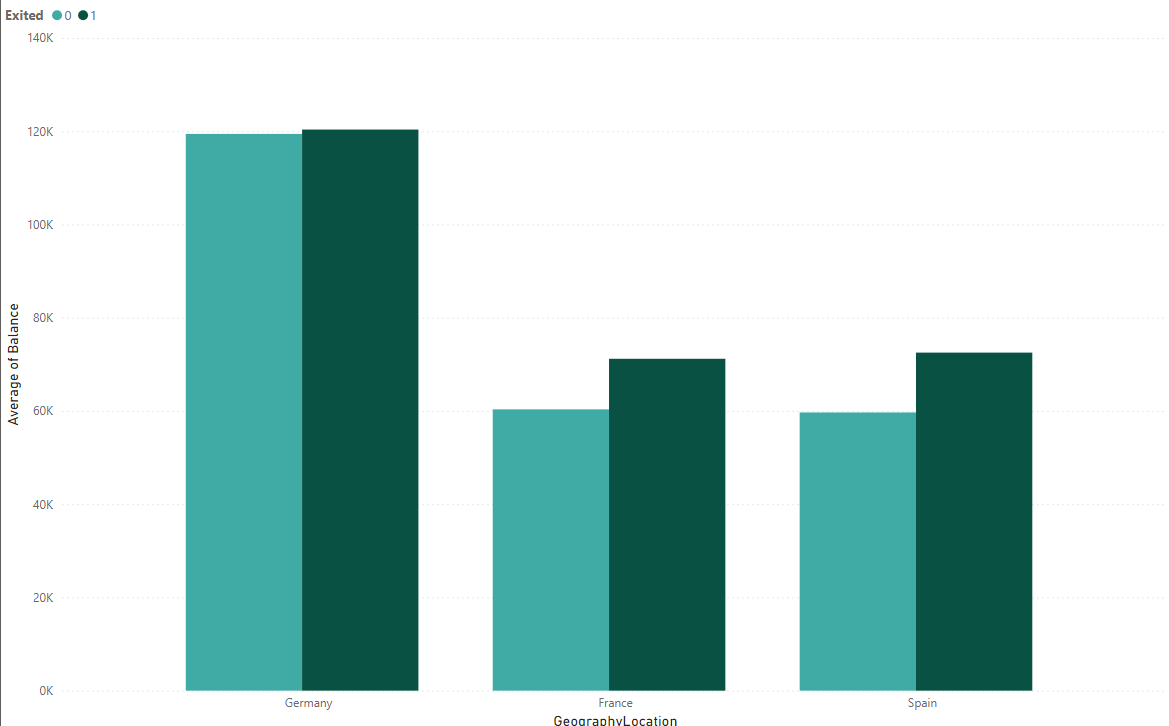
* + - * Checking Accounts: Core for everyday transactions.
* Debit Cards: Linked to checking, providing convenient access to funds.
* Savings Accounts: Grow savings and often earn interest.
* Credit Cards: Offer a line of credit for purchases, requiring repayment with interest.
* Loans: Tailored financial solutions like mortgages or auto loans.

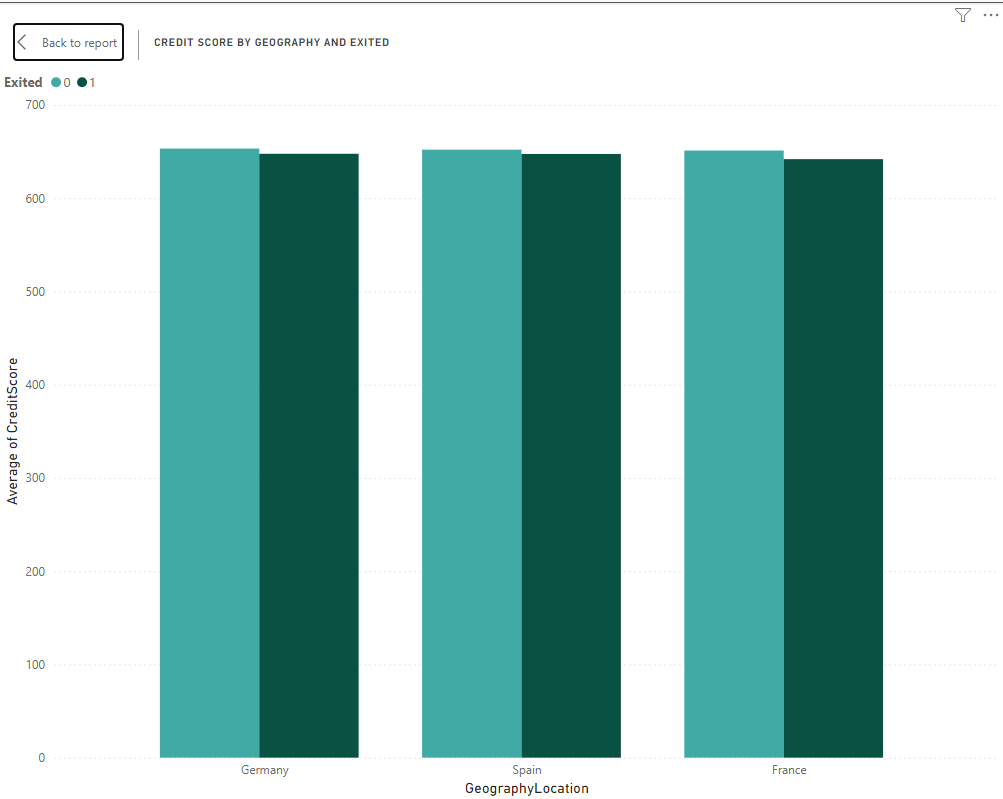
Cross-Selling Strategies:

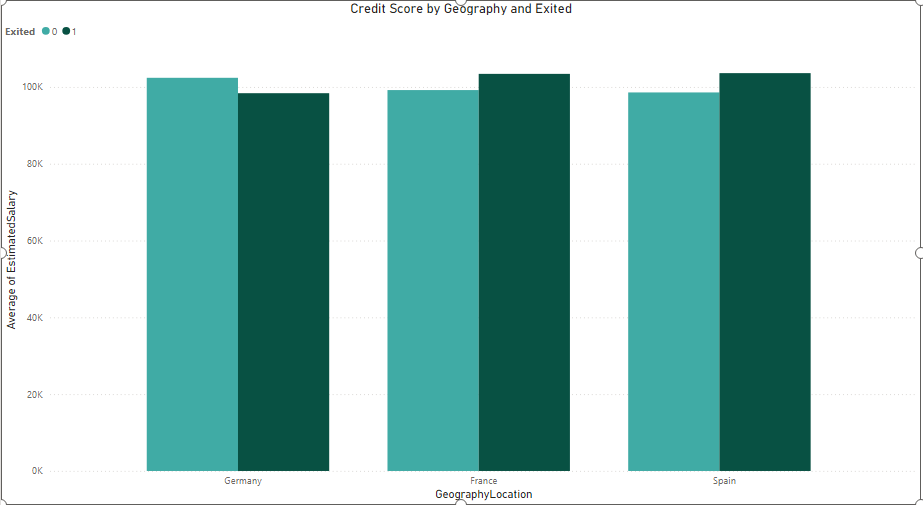
* Recommend complementary products: Checking account users might benefit from a debit card and online banking for easy management. Savings account holders could be recommended automatic transfers to boost saving habits or higher-interest options like CDs for larger balances.
* Personalize based on usage: Credit card users with travel habits could benefit from travel rewards cards. Loan seekers might be interested in bundled insurance options.
* Leverage digital platforms: Promote paperless statements and bill autopay through online/mobile banking. Offer investment options or financial tools accessible through these platforms.

By understanding product usage patterns and tailoring recommendations, banks can create a win-win situation: increased revenue and customers who feel their financial needs are being met.

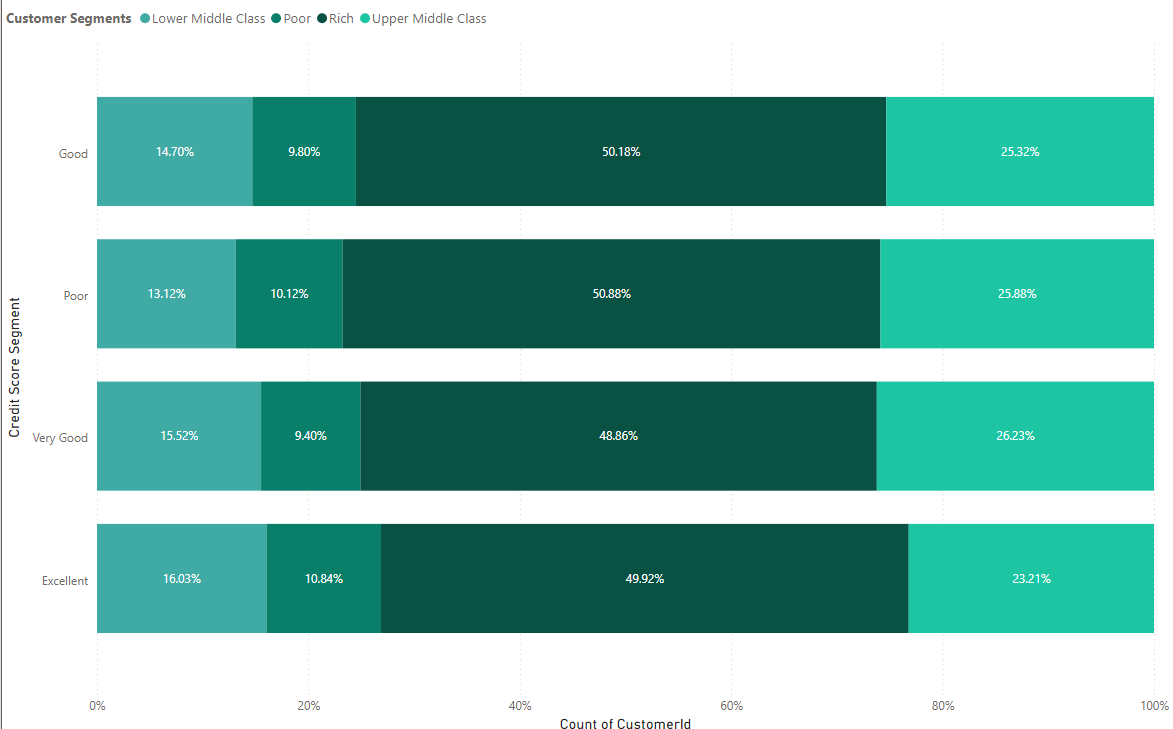
1. **Geographic Market Trends: How do economic indicators in different geographic regions correlate with the number of active accounts and customer churn rates?**This section explores how economic indicators vary across geographic regions and their potential correlation with customer churn rates. We'll analyze three key economic indicators: average balance, average salary, and credit score.

The first chart examines the average balance by geographic location.  
The chart shows average balance by location, with Germany having the highest. While it doesn't directly address churn, lower balance regions might have higher churn if customers there are more cost-sensitive.  
  
  
The second chart you sent shows the average credit score by geographical location.  
The credit score chart shows variations by location, with Germany having the highest average. This, along with churn rate data, could help identify a link between higher credit scores and lower churn (customers are more creditworthy and less likely to switch).

  
  
The third chart shows the average salary by geographical location.   
Similar to the credit score chart, average salary varies by location (Germany highest). This, along with churn data.



**4. Risk Management Assessment.**Based on the chart, the demographic segments with the most customers are those with lower credit scores ("Lower Middle Class" and "Poor"). These same segments also represent the largest portion of the bank's customers. Therefore, based on the data presented in the chart, we can say that the demographic segments with the most customers are also the ones that pose the highest potential financial risk to the bank.

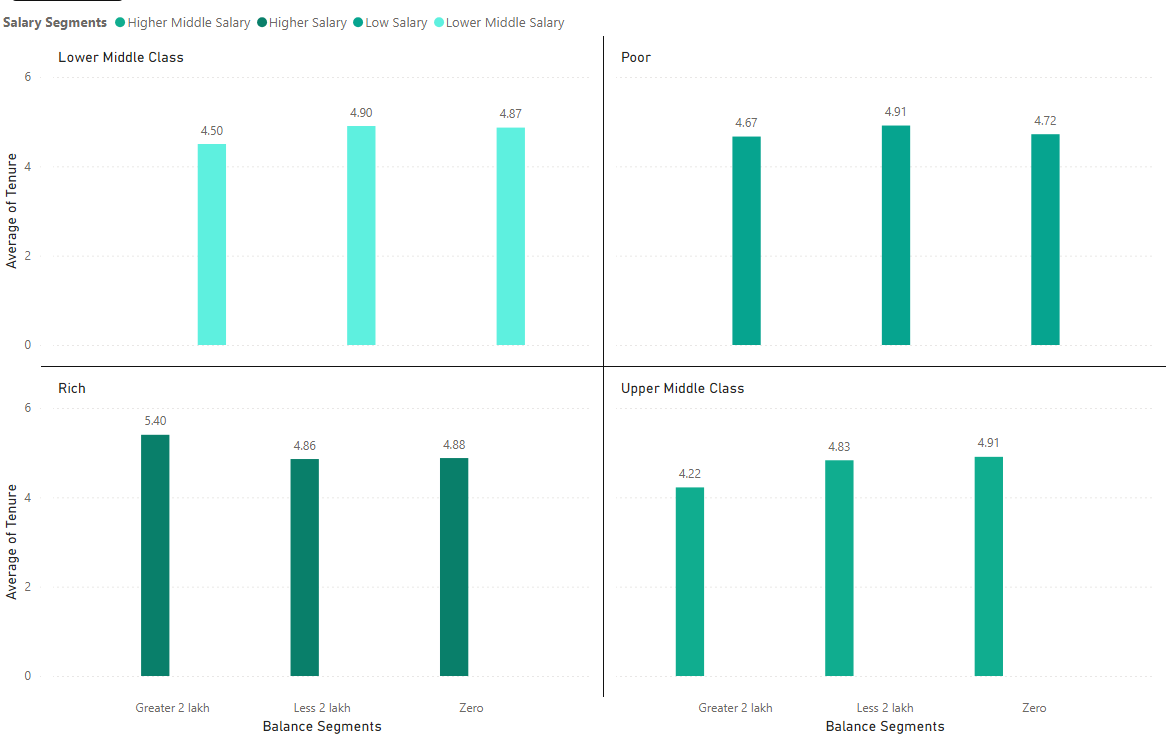
There are other factors to consider when assessing customer risk besides creditworthiness, such as income and employment history. However, creditworthiness is a strong indicator of a customer's ability to repay loans. Customers with lower credit scores are statistically more likely to default on loans than customers with high credit scores. This means that the bank is more likely to lose money on loans made to customers in the "Lower Middle Class" and "Poor" segments than on loans made to customers in other segments.  
****

**5. Customer Tenure Value Forecast.**The chart you provided shows the average tenure by balance segments, salary segments, and customer segments. Here are some key insights and predictions you can make based on the data:

* Balance Segments: Customers with higher average balance segments tend to have a longer tenure with the bank. This could be because they are more invested in the bank's products and services, or because the bank offers them better benefits to retain them.
  + High Salary & Greater than 2Lac: This segment has the highest average tenure (around 5.4 years).
  + Zero Balance: Customers with zero balance tend to have the lowest tenure across all salary segments (around 4.2 years on average).
* Salary Segments: There is a weak trend between salary segments and tenure. On average, customers in higher salary segments seem to have slightly longer tenures. However, the differences are not substantial across most segments.

Here's a prediction of the average tenure for different salary segments based on the data:

* High Salary: Customers in this segment are likely to have a tenure around 4.8 - 5.4 years.
* Low Salary & Lower Middle Class: Customers in this segment are likely to have a tenure around 4.5 - 4.9 years.
* Upper Middle Class & Rich: Customers in this segment are likely to have a tenure around 4.8 - 4.9 years.

Overall, the data suggests that balance segments are a more important factor in predicting tenure than salary segments. Banks can leverage this information to target retention efforts towards customer segments with higher balance segments and develop strategies to improve customer satisfaction and product adoption across all segments.  


**6. Marketing Campaign Effectiveness.**To assess the impact of marketing campaigns on customer retention and acquisition within a dataset, you would typically use a combination of data analysis and statistical techniques. Here's a general approach you could take:

1. Define Metrics: Define key metrics for customer retention and acquisition. For retention, you might use metrics like customer churn rate or retention rate. For acquisition, you might use metrics like new customer acquisition rate or customer acquisition cost (CAC).

2. Segment Data: Segment the data based on different marketing campaigns. This will allow you to analyze the impact of each campaign separately.

3. Calculate Metrics: Calculate the defined metrics for each segment and for each time period (e.g., monthly, quarterly, annually). This will help you understand how each campaign is affecting customer retention and acquisition over time.

4. Compare Results: Compare the metrics across different campaigns to identify which campaigns are most effective at retaining and acquiring customers.

5. Statistical Analysis: Use statistical tests (e.g., t-tests, ANOVA) to determine if the differences in metrics between campaigns are statistically significant.

6. Additional Information: To perform a comprehensive analysis, you may need additional information such as:  
- Customer demographics: To understand if certain demographics respond better to certain campaigns.

- Campaign details: To understand the specifics of each campaign (e.g., duration, channels used, messaging).

- Competitor data: To understand the competitive landscape and how it might be impacting your results.

- External factors: Such as economic conditions, seasonality, or industry trends that might affect customer behavior.

By following this approach and gathering the necessary information, we can assess the impact of marketing campaigns on customer retention and acquisition within your dataset.

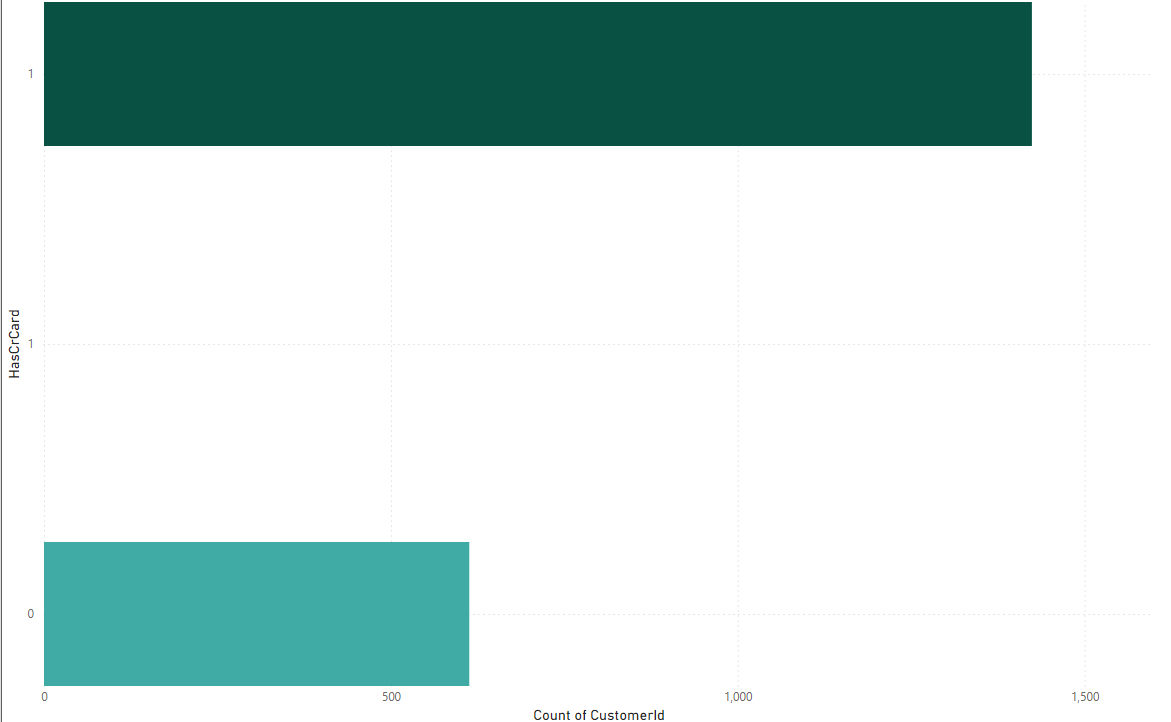
**7. Customer Exit Reasons Exploration.**

Sure, the question is asking us to identify common characteristics or trends among customers who have exited the bank (churned) in order to understand why they left. You identified two possible characteristics:

1. Credit card ownership: The idea here is that customers who have credit cards are more likely to churn than those who don't.
2. Number of products purchased: Customers who buy fewer products from the bank are more likely to churn than those who buy more products.

The bar chart titled "COUNT OF CUSTOMERID BY HASCRCARD AND EXITED". The chart seems to partially support your first characteristic about credit card ownership.

Analysis of Credit Card Ownership:

* Does the chart support the idea that customers with credit cards are more likely to churn? The chart shows the number of customers who exited (1) broken down by whether they have a credit card (HasCrCard) or not. There are more customers who exited that have credit cards (around 1,200) than those who don't (around 300). This suggests that there could be a correlation between having a credit card and exiting the bank.
  + However, it is also important to consider the total number of customers with and without credit cards. If there are many more total customers with credit cards than without, then the higher number of churned customers with credit cards could simply reflect the larger population.  
    

Sure, the chart is a bar chart titled "COUNT OF CUSTOMERID BY NUMOFPRODUCTS AND EXITED". The chart shows the number of customers who exited the bank (1) broken down by the number of products they purchased (NumOfProducts).

Analysis of Number of Products Purchased:

* Does the chart support the idea that customers who buy fewer products are more likely to churn? Yes, the chart supports this idea. The number of customers who exited the bank is highest for customers who purchased zero products (around 800) and then steadily declines as the number of products purchased increases. There are very few exited customers (around 60) who purchased four or more products. This suggests that customers who use a wider range of the bank's products are less likely to churn.

**8. Predicting Customer Churn.**

This analysis aims to assess whether factors like tenure (time with the bank), number of products held, active membership status, and estimated salary can predict customer churn (leaving the bank). We have created four charts to visualize trends in these factors for both exiting and non-exiting customers.

The chart shows the count of customer IDs (number of customers) on the y-axis and the number of products held by the customer (NumOfProducts) on the x-axis. It appears to be a stacked bar chart where the blue bars represent exiting customers (Exited = 1), and the orange bars represent customers who remained (Exited = 0).

Key Insights:

* There is no clear consistent pattern between the number of products a customer holds and their likelihood of leaving the bank.
* In some product ranges (e.g., 1 product, 3 products), there seem to be more exiting customers than those who stayed.
* Conversely, in other ranges (e.g., 2 products, 4 or more products), there appear to be more customers who stayed than those who exited.

Importance of NumOfProducts for Prediction:

Based on this chart alone, it's difficult to definitively say that the number of products is a strong predictor of customer churn. There seems to be no clear trend, and the number of exiting customers fluctuates across the different product ranges.

However, it's possible that NumOfProducts could still be a relevant factor in conjunction with other customer attributes. Here's why:

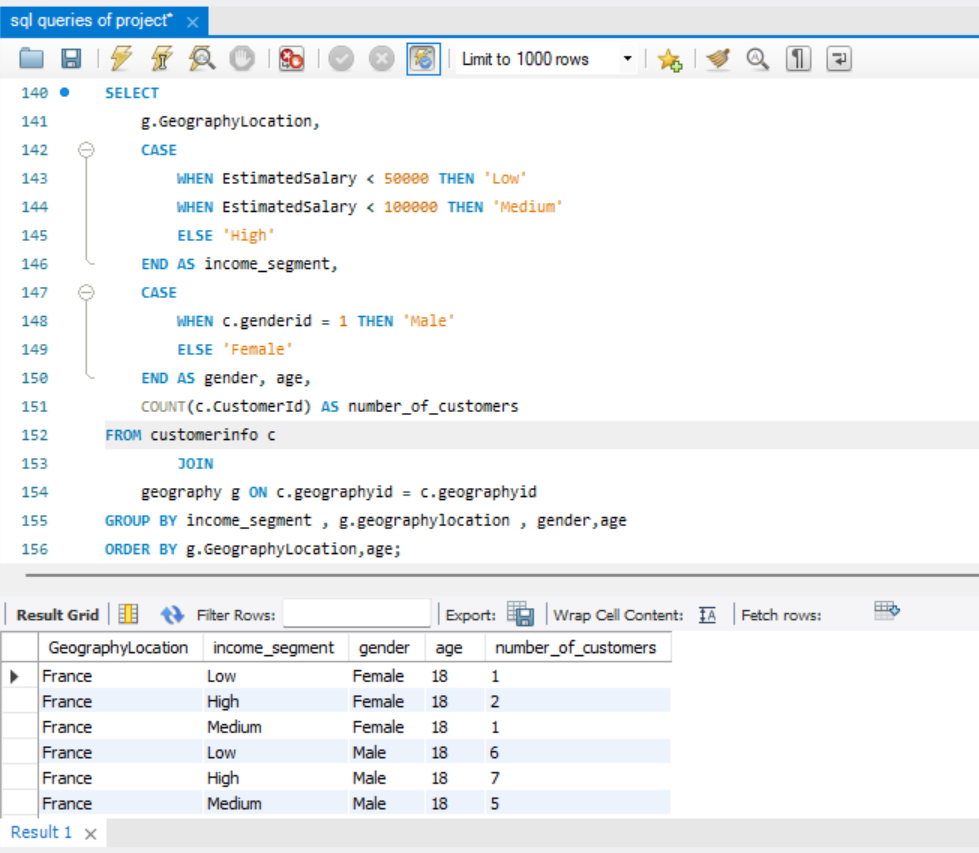
* Customer Needs: Customers with more products might have their banking needs well-met, potentially increasing their satisfaction and reducing churn.
* Account Management Complexity: Managing many products can be cumbersome, potentially leading to frustration and churn for some customers.

**9. Customer Segmentation with SQL.**  
This code segments customers based on:

* Income: categorized as Low, Medium, or High based on estimated salary.
* Location: uses the GeographyLocation field (country/region).
* Gender: assigned as Male or Female based on gender id values.
* Age: included for further analysis within segments.

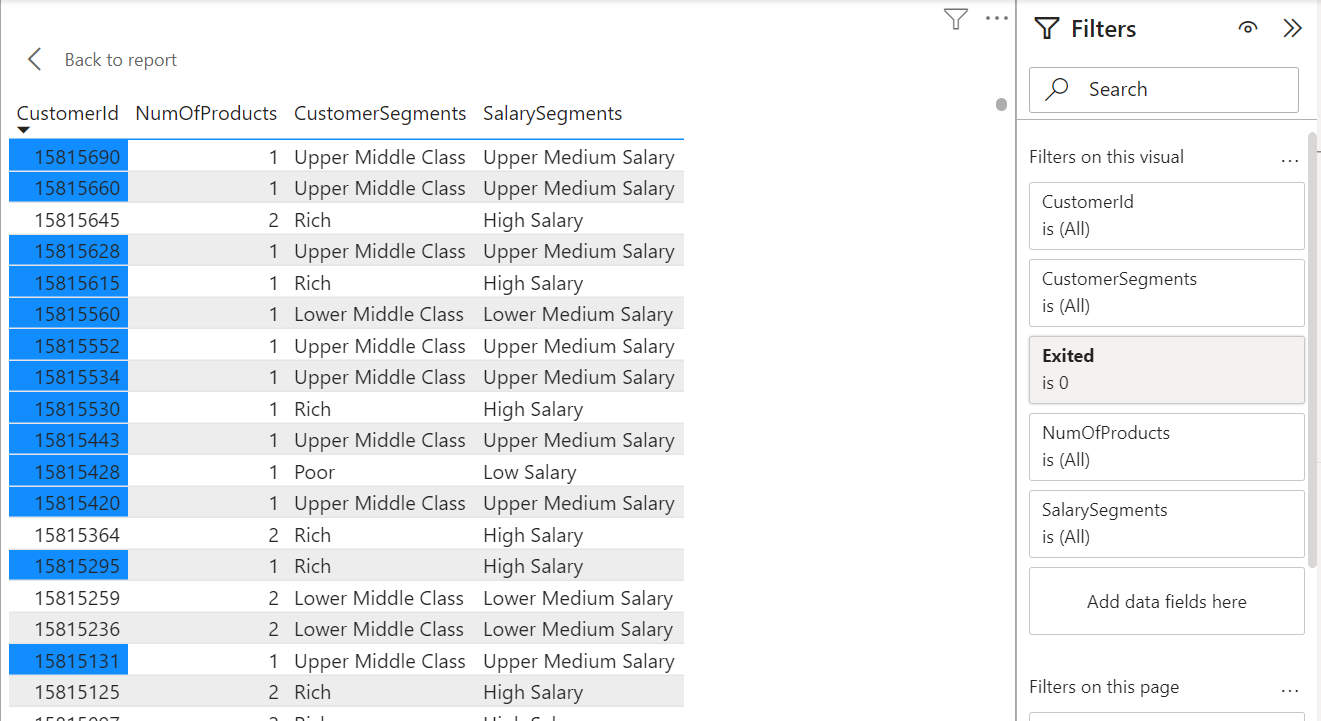
It achieves this by:

* Joining customer data with geographic information (if available).
* Grouping customers by income segment, location, gender, and age.
* Counting the number of customers within each specific segment.
* Ordering the results by location and then age for easier exploration.

This allows you to analyze customer characteristics across various demographics and potentially geographic regions.  


**10. Conditional Formatting for Churn Risk.**Sure, based on the chart, which appears to be filter window, we can 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 by following these steps:

1. Identify the churn criteria: Define the criteria to identify customers at risk of churn. This could be based on a combination of factors, such as:
   * Customers with a low number of products purchased (NumOfProducts)
   * Customers with low balance segments
   * Customers who have recently exited (Exited) in the past (e.g., in the last 6 months)
2. Conditional formatting based on churn criteria: Apply a conditional formatting rule to highlight cells that meet the churn criteria. You can format the cells with a different background color or font to make them visually distinct.
3. Filter by Credit Card ownership: Create a filter for the "HasCrCard" field. This will allow you to segment customers by whether they have a credit card or not.
4. Evaluate churn rate by Credit Card ownership: Analyze the churn rate (percentage of customers who exited) for customers with and without credit cards. You can calculate this by comparing the number of exited customers (where Exited = 1) to the total number of customers in each segment (HasCrCard = Yes or No).



**11. Churn Rate and Insights.**

This analyzes customer churn rates and identifies segments most susceptible to churn. It also proposes strategies to decrease churn and improve customer retention.

Churn Rate:

* The overall churn rate for the bank is 20.37%.
* Year-on-year churn rates show some fluctuations:
  + 2016: 19.27%
  + 2017: 22.35% (highest)
  + 2018: 20.21%
  + 2019: 19.86% (lowest)

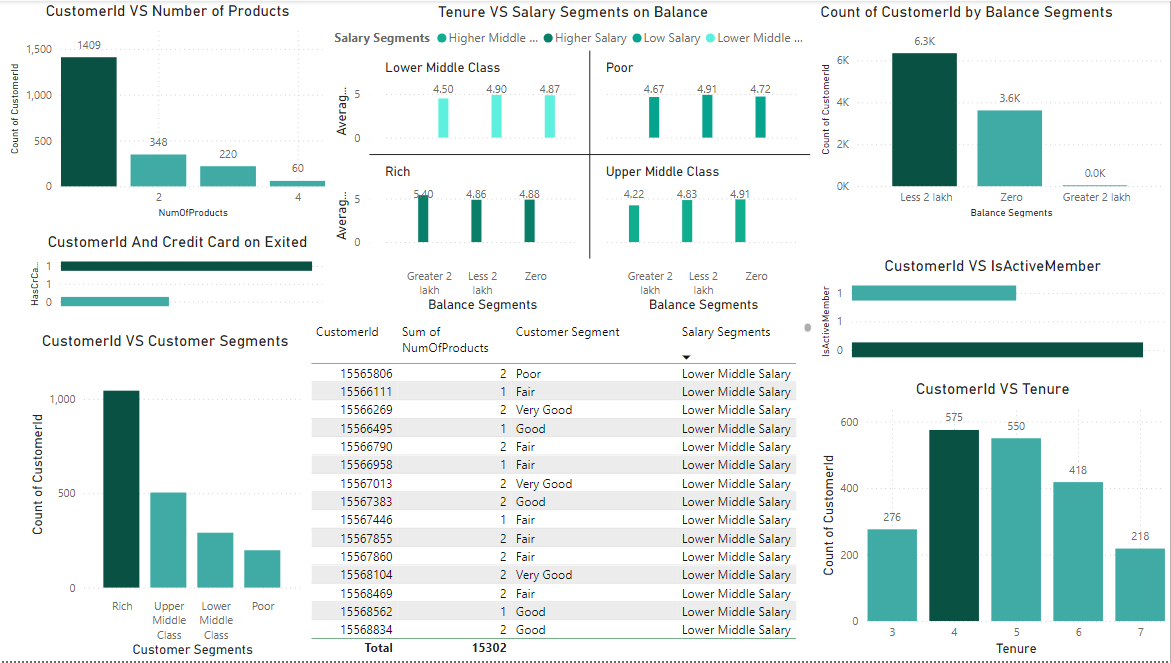
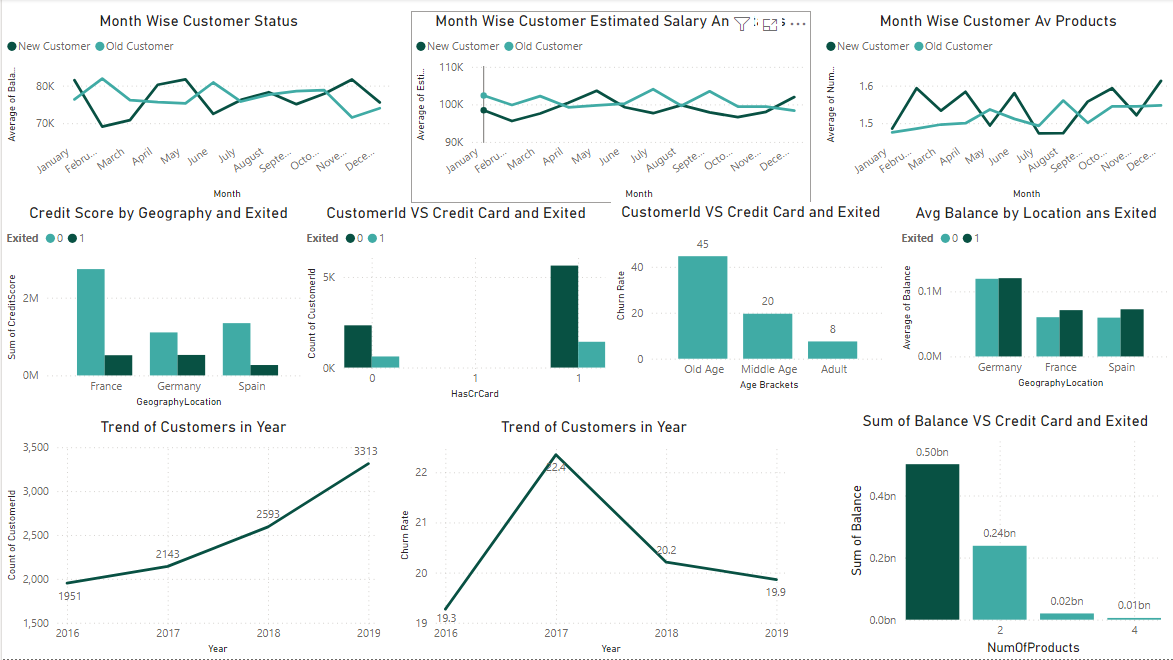
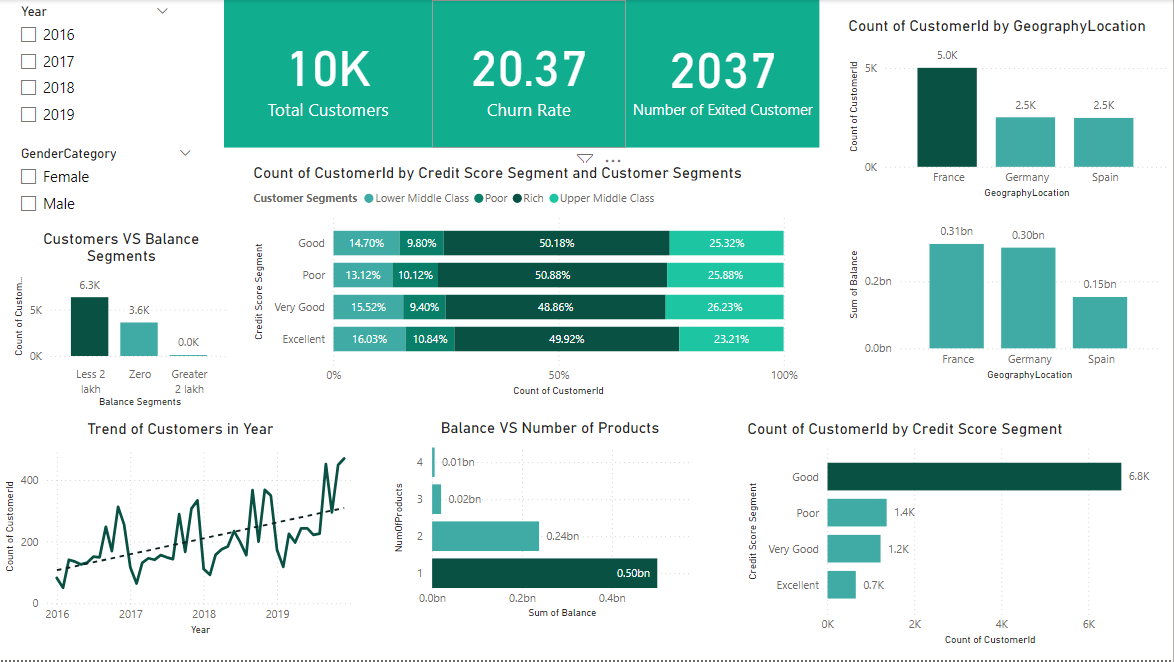
Customer Segments Prone to Churn:

Data analysis suggests a customer segment with a higher likelihood of churn:

* Purchases 1 product: Customers who only use one product by bank might not find enough value compared to competitors offering wider ranges or integrated services.
* Has credit card:Potential reasons for churn among credit card holders could be:
  + Limited credit limits not meeting their needs.
  + Lack of rewards programs that incentivize them to keep the card.
  + High credit card fees.
* Tenure of 4-5 years: Customers with this tenure might be nearing the end of introductory offers or discounts, making them susceptible to competitor offers with better rates or features.
* High salary: High earners might have more options and be more likely to switch for a slightly better interest rate or benefit elsewhere.

Recommendations to Reduce Churn:

* Targeted Product Bundles: Create product bundles that cater to specific customer segments and needs. Offer these bundles to customers who only use one product, highlighting the additional benefits and potential cost savings.
* Enhanced Credit Card Rewards: Improve credit card rewards programs for existing customers. This could involve:
  + Increasing credit limits based on customer history and creditworthiness.
  + Offering rewards programs aligned with spending habits (e.g., travel rewards, cash back for specific categories).
  + Reducing or eliminating annual fees, especially for high-value customers.
* Retention Offers for Existing Customers: Proactively reach out to customers nearing the end of introductory offers with personalized retention deals. This could include extending introductory rates or offering discounts on other products or services.
* Customer Satisfaction Surveys: Regularly conduct customer satisfaction surveys to understand why customers churn. This can help identify areas for improvement and tailor retention strategies accordingly.
* Relationship Management for High-Value Customers: Develop dedicated relationship managers for high-value customers to provide personalized service, address their specific needs, and offer exclusive benefits.

**12. Dashboard Creation (Power BI)**Here is Dashboard created in Power BI using a given dataset by Bank.  
****

**13. Problem-Solving Approach Without Objectives.**Absolutely! Even without explicit objective and subjective questions, We can effectively approach a problem by following a process that involves generating hypotheses, asking your own questions, and deriving insights. Here's a breakdown of this approach:

1. Hypothesis Generation:

* Start by making assumptions about the data or problem at hand. These hypotheses can be based on your understanding of the domain, industry best practices, or even initial observations of the data.

2. Question Formulation:

* Based on your hypotheses, formulate questions that can be answered using the available data. These questions should guide your analysis and help you validate or refine your initial assumptions. Here are some examples of questions you could ask in the absence of predefined questions:
  + Customer Churn Analysis:
    - Are there any demographic patterns (age, income) associated with customer churn?
    - Does account balance or number of products held influence churn rates?
    - How does customer activity (transactions, logins) correlate with churn?
  + Marketing Campaign Analysis:
    - Which marketing channels (email, social media) are most effective at reaching target audiences?
    - Is there a correlation between ad spend and campaign performance?
    - How does campaign messaging impact customer engagement and conversion rates?

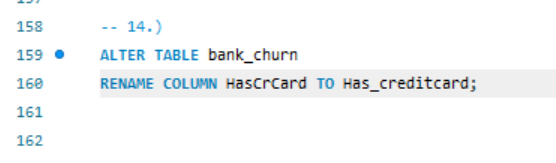
3. Data Exploration and Analysis:

* Use techniques like data visualization and statistical analysis to answer the formulated questions.

4. Insights and Recommendations:

* Based on the findings from your analysis, draw insights that can inform decision-making. These insights could relate to customer behavior, marketing strategies, product development, or other areas relevant to the problem.

**14. Renaming a Column in SQL.**

This code changes the name of the "HasCrCard" column to   
"Has\_creditcard" in the "Bank\_Churn" table. It improves clarity by using a more descriptive name.  


**Conclusion**This project aimed to analyze various customer-related datasets provided by a bank to gain insights into customer churn, improve service delivery, and enhance customer satisfaction. By examining factors like demographics, transaction details, customer exit information, and active customer profiles, we were able to uncover valuable information.

Our analysis revealed key factors contributing to customer churn, including credit score, account balance, and product usage. We identified profitable customer segments and explored potential reasons for customer exits. Additionally, we investigated the relationship between various customer attributes and churn rates.  
The findings from this project can be used by the bank to develop targeted strategies to reduce churn, retain valuable customers, and optimize product offerings. Here are some specific recommendations:

* Develop targeted marketing campaigns for customer segments identified as high churn risk.
* Offer incentives like credit cards or loyalty programs to encourage product usage and increase customer engagement.
* Investigate reasons behind customer exits and address areas where the bank can improve its services.
* Continuously monitor customer behavior and churn rates to refine strategies over time.

This project provided a valuable starting point for understanding customer behavior and churn. Further analysis can be conducted to delve deeper into specific areas, such as product affinity and customer lifetime value. By using the insights gained from data analysis, the bank can develop data-driven strategies to improve customer relationships and achieve its business goals.