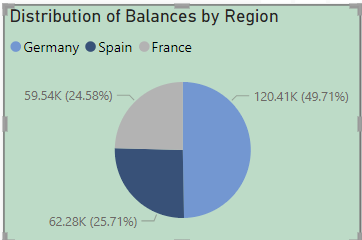
**Analytical CRM Development for a Bank**

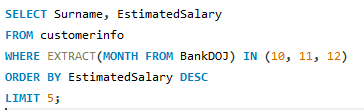
**Objective Questions:**

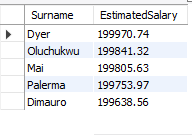
1. **What is the distribution of account balances across different regions?**



* The pie chart illustrates the proportion of account balances held by customers in various regions. It indicates that Germany, France, and Spain are the primary regions with the highest number of accounts. As a result, these regions also exhibit higher account balances, with the following distribution:
* Germany: Accounts for 25.37% of the total accounts, with an average balance of 61.82K.
* France: Represents 49.14% of the total accounts, with an average balance of 119.73K.
* Spain: Comprises 25.49% of the total accounts, with an average balance of 62.09K.
* These figures indicate that France has the highest proportion of accounts and consequently the highest average balance, followed by Germany and Spain. This suggests that there may be varying economic factors or customer behaviors contributing to these differences in account balances across regions.

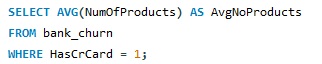
1. **Identify the top 5 customers with the highest Estimated Salary in the last quarter of the year. (SQL)**





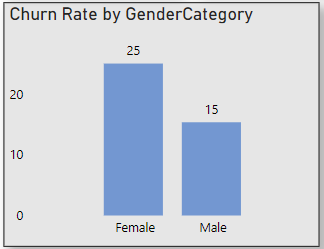
* + The query filtered records based on the date of joining the bank ('BankDOJ') to focus on the last quarter of the year.
  + Results displayed the Surname and Estimated Salary of the top 5 customers during this period.
  + Customers listed as Palerma, Oluchukwu, Mai, Dyer, and Dimauro had the highest Estimated Salaries.
  + This information aids in understanding the financial profiles of top earners and tailoring services or offers accordingly.

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



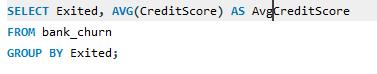
* The average number of products used by customers who have a credit card is approximately 1.53. This indicates that, on average, customers who possess a credit card utilize around 1.53 products offered by the bank.

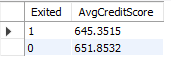
1. **Determine the churn rate by gender for the most recent year in the dataset.**



* In the most recent year, 2019, the churn rate for male customers is 25%, while for female customers, it is 15%. This suggests that male customers are more likely to churn compared to female customers in the given dataset for the year 2019.

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

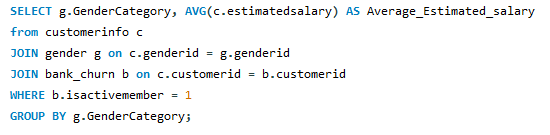


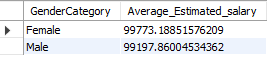


Here's the comparison of the average credit scores:

* Customers who have exited: 645.3515
* Customers who remain: 651.8532

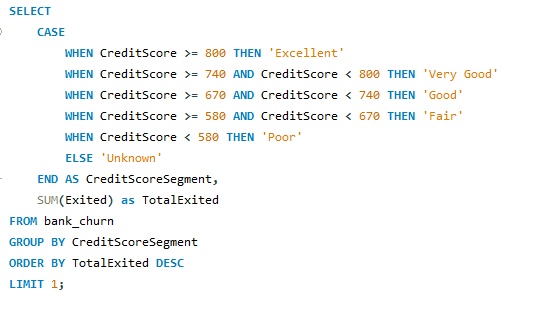
1. **Which gender has a higher average estimated salary, and how does it relate to the number of active accounts? (SQL)**





* Despite having fewer active accounts, females exhibit a higher average estimated salary.
* The average estimated salary for females is approximately $99,773.19.
* This suggests that females contribute significantly to the bank's revenue despite their smaller presence in terms of active accounts.
* Targeting and retaining female customers could be beneficial for maximizing revenue due to their higher spending potential.

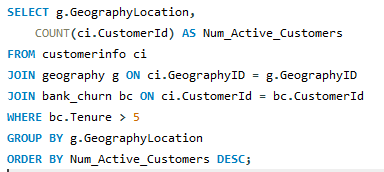
1. **Segment the customers based on their credit score and identify the segment with the highest exit rate. (SQL)**

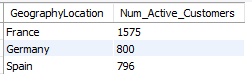




The customer's credit score has been segmented into five categories:

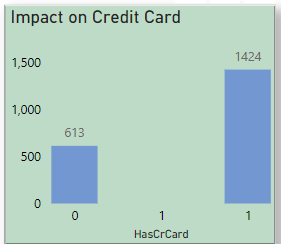
* Excellent: 800 or above
* Very Good: 740–799
* Good: 670–739
* Fair: 580–669
* Poor: 579 or less
* Among these segments, the segment with the highest exit rate is 'Fair', with an exit rate of 0.2202.
* This indicates that customers categorized in the 'Fair' credit score segment have the highest likelihood of exiting the bank compared to customers in other segments.

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



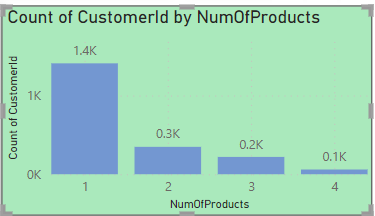
* The query identified the number of active customers with a tenure greater than 5 years in each geographic region.
* France has the highest number of active customers with a tenure greater than 5 years, totalling 1575 customers.
* Germany follows with 800 active customers, and Spain has 796 active customers meeting the specified criteria.
* This analysis helps in understanding the distribution of long-standing active customers across different regions, which could be valuable for targeted marketing or retention strategies.

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

****

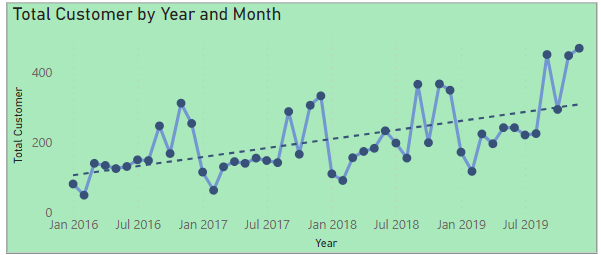
* The analysis compares the impact of having a credit card on customer churn based on available data.
* For customers without a credit card (HasCrCard = 0), the count of churned customers is 613.
* For customers with a credit card (HasCrCard = 1), the count of churned customers is 1424.
* From this data, it can be inferred that customers with a credit card (HasCrCard = 1) have a higher count of churn compared to those without a credit card (HasCrCard = 0). This suggests that the presence of a credit card may not necessarily mitigate churn and may even contribute to higher churn rates.

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

****

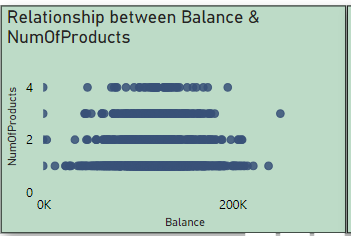
* The most common number of products used by customers who have exited the bank:
* 1 product: Majority of customers who exited had only one product with the bank.
* 2 products: Followed by a smaller proportion of customers who had two products.
* 3 products: Subsequently, there were even fewer customers with three products.
* 4 products: The least common among exiting customers, indicating a decreasing trend in product ownership.
* This trend suggests a potential correlation between the number of products owned and the likelihood of customer churn.
* Further analysis could focus on strategies to retain customers with multiple products, thereby improving overall customer retention rates.

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.**

****

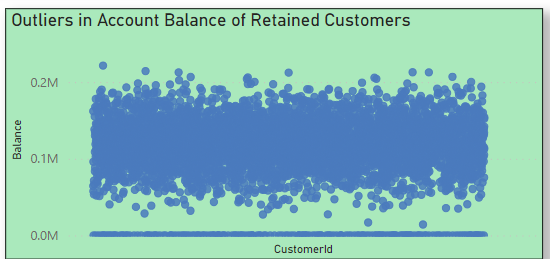
* Analysis of Customer Joining Trends Over Time:
* December 2019 witnessed a notable surge in new customer acquisitions, with 470 individuals joining the bank.
* This observation suggests a potential seasonal pattern where the end of the year experiences increased customer enrolment.
* Seasonal variations in customer joining trends can provide valuable insights for strategic planning and resource allocation within the bank.

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

****

* Analysis of Relationship between Number of Products and Account Balance for Exiting Customers:
* The analysis reveals that exiting customers exhibit a wide range of account balances, spanning from 0k to 300k.
* A trend observed indicates that customers with fewer products, particularly 1 product, tend to exit the bank more frequently.
* Conversely, customers with a higher number of products, specifically 4 products, and an account balance ranging between approximately 70k to 200k, demonstrate a lower tendency to exit.

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

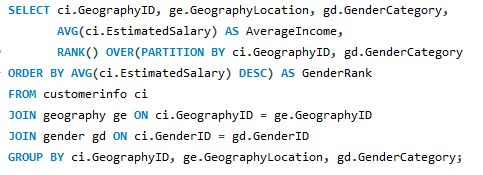
****

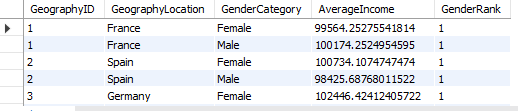
* we can identify outliers among customers who have remained with the bank by observing customer IDs with balances of *0k and less than 50k.*

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

* The dataset consists of 7 tables, each serving a specific purpose in analysing customer-related data. Among these tables, the credit card, geography, customer, and gender tables predominantly contain categorical variables. These tables facilitate the classification of various attributes related to customers, geographic locations, and gender. Although numerical identifier columns may exist within these tables, their primary function is for referencing and categorization rather than representing numerical values for analysis. Understanding the composition of these tables is crucial for conducting effective data analysis and segmentation to derive actionable insights for the bank's CRM development.

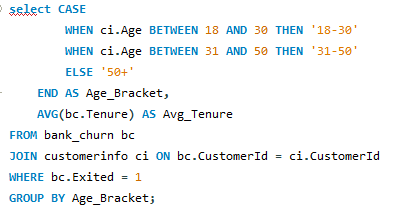
1. **Using SQL, write a query to find out the gender-wise average income of males and females in each geography id. Also, rank the gender according to the average value. (SQL).**

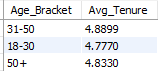




* The output illustrates the average income of males and females within each geographic location, accompanied by their respective rankings based on income. Notably, females generally exhibit a higher average income than males across all regions. This disparity in income distribution suggests potential gender-based discrepancies favouring females within the dataset.

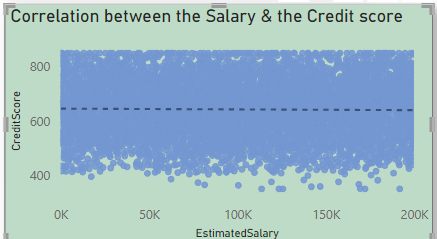
1. **Using SQL, write a query to find out the average tenure of the people who have exited in each age bracket (18-30, 30-50, 50+).**





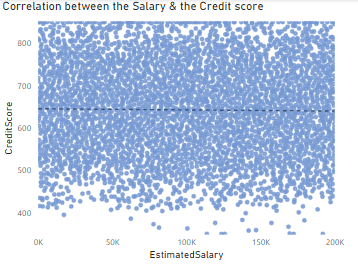
* The output provides the average tenure of customers who have exited the bank within different age brackets. Here's the breakdown:
* Customers aged 18-30 have an average tenure of approximately 4.7770.
* Customers aged 31-50 have an average tenure of approximately 4.8899.
* Customers aged 50 and above have an average tenure of approximately 4.8330.

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?**

****

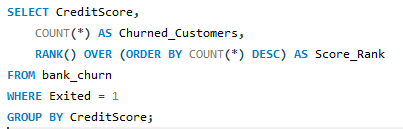
* The provided data indicates a positive correlation between salary and balance for both retained and exited customers. This means that as salary increases, the balance tends to increase as well, regardless of whether the customer has exited or not. This relationship is evident in the scatter plot, which visually demonstrates the positive correlation between salary and balance.

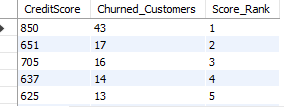
1. **Is there any correlation between the salary and the Credit score of customers?**



* We use scatter plot chart to show mentioned above, we can understand that there is correlation between salary and credit score.

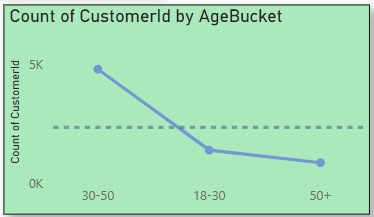
1. **Rank each bucket of credit score as per the number of customers who have churned the bank.**





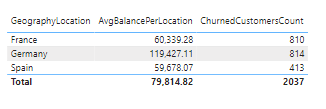
* The output provides a breakdown of credit scores alongside the corresponding count of customers who have churned from the bank. Each credit score bucket is ranked based on the number of churned customers, with the highest count receiving a rank of 1. This ranking enables the identification of credit score ranges with the highest churn rates, facilitating further analysis and strategic decision-making.

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.**

****

* In this outcome, we can see the age buckets that have the number of customers with the number of credit cards. The 50+ age bucket has a lesser than average number of credit cards.
* The age bucket "18-30" consists of 1,968 customers, out of which 1,400 have a credit card.
* Age bucket "50+" consists of 1,261 customers, out of which 874 have a credit card.
* These age buckets have fewer customers with credit cards compared to the average number of credit cards per bucket.

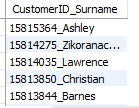
1. **Rank the Locations as per the number of people who have churned the bank and average balance of the customers.**



* France is ranked first due to having the highest number of people who exited the bank (810) and an average balance of $60,339.28.
* Germany is ranked second with 814 people who exited the bank and an average balance of $119,427.11.
* Spain is ranked third with 413 people who exited the bank and an average balance of $59,678.07.
* This analysis ranks the locations based on the combination of the number of people who exited the bank and the average balance of the customers.

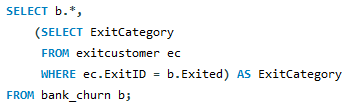
1. **As we can see that the “CustomerInfo” table has the CustomerID and Surname, now if we have to join it with a table where the primary key is also a combination of CustomerID and Surname, come up with a column where the format is “CustomerID\_Surname”.**

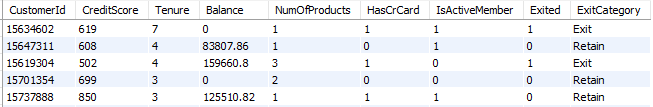




* This format allows for a convenient way to uniquely identify each customer within the dataset, combining their identification number with their surname. It can be useful for various data operations, including joining with other tables based on the CustomerID\_Surname column.

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



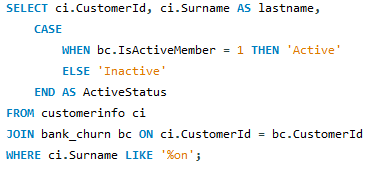


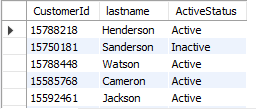
* Yes, we can retrieve the "ExitCategory" from the "ExitCustomers" table without using the JOIN keyword by using a subquery**.**

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

* Identified missing values in the dataset.
* Utilized Power Query Editor in Power BI to address missing values.
* Used Power Query Editor to locate empty cells or areas with missing data.
* Replaced missing values either with specific predefined values or by copying values from adjacent non-blank cells.
* This process ensured that the analysis was conducted on complete data, improving the accuracy of insights derived from the dataset.

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”.**

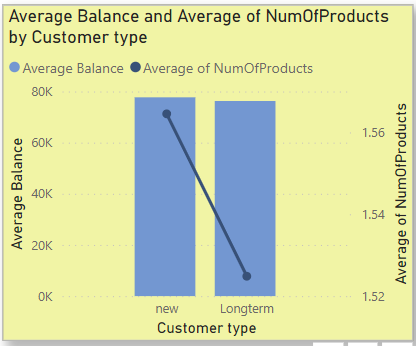




* This query will provide the customer IDs, and their last name with active or inactive customers whose surname ends with “on”.

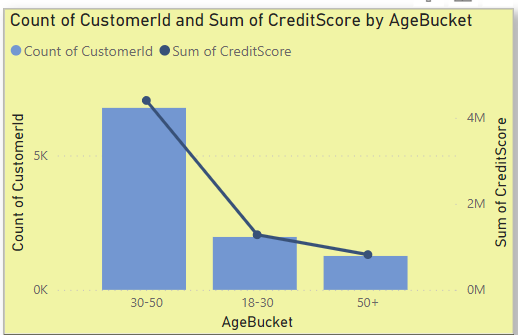
**Subjective Question:**

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?**

****

* New customers have a higher average number of bank products (1.58 vs. 1.46) and maintain slightly higher average balances (92k vs. 91k) compared to long-term customers.
* This indicates that new customers exhibit more active spending habits than long-term customers.
* The minimal difference in average balances suggests that new customers' financial standing isn't significantly higher than long-term customers.
* New customers may represent a segment with greater growth potential and deeper engagement over time.
* Focus on nurturing relationships with new customers, as they show a willingness to engage with a wider range of products.
* Provide tailored services and personalized experiences to encourage long-term loyalty and enhance customer satisfaction.

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

****

* 1.4k customers are assumed to use "Savings Account" alone.
* 0.3k customers are assumed to use both "Savings Account" and "Credit Card."
* 0.2k customers are assumed to use "Savings Account," "Credit Card," and "Personal Loan."
* 0.1k customers are assumed to use all four products: "Savings Account," "Credit Card," "Personal Loan," and "Insurance."
* While these product combinations were assumed for the analysis, they offer insights into potential cross-selling opportunities.
* Understanding potential affinities between products can guide future data collection and analysis to validate these assumptions and refine cross-selling strategies accordingly.
* Once validated, develop targeted cross-selling strategies based on the confirmed product affinities to maximize the effectiveness of marketing campaigns and promotions.
* Continuously monitor customer behavior and feedback to iterate and optimize cross-selling efforts over time.

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

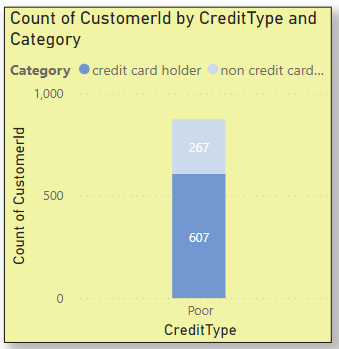


* Across Germany, France, and Spain, there are notable differences in economic indicators and banking metrics. Germany has the highest average balance and churn rate, France leads in credit card usage, and Spain has the highest average estimated salary.

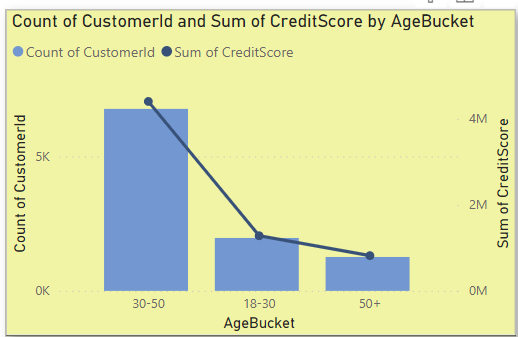
1. **Risk Management Assessment: Based on customer profiles, which demographic segments appear to pose the highest financial risk to the bank, and why?**

In the low credit score range (300-579):

* 1,648 customers hold credit cards, while 714 do not.
* Among these, 607 credit card holders and 267 non-credit card holders have a 0 balance in their account.
* Identify demographic segments, such as young adults with limited credit history or individuals with low income and high debt levels, that may pose a heightened financial risk.
* These segments may exhibit behaviors like irregular payment patterns, high credit utilization, or insufficient savings, indicating potential difficulty in meeting financial obligations.

****

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?**

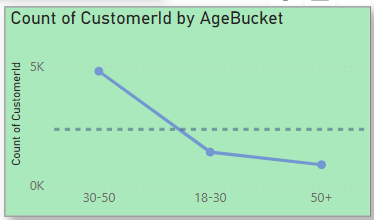
****

* 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? What extra information would you need to solve this?**

* Segmentation of the customer base based on demographics, behavior, and product usage will allow us to analyze the impact of marketing campaigns on different customer groups. This segmentation will help identify which segments are most responsive to the campaigns and where adjustments may be needed.
* To conduct a comprehensive analysis of marketing campaign effectiveness, additional information may be needed, such as:
* Detailed data on marketing campaign spend, channels used, and messaging.
* Customer response data, including click-through rates, conversion rates, and campaign attribution.
* Customer feedback and sentiment analysis from social media, reviews, and customer support interactions.
* External factors that may influence customer behavior, such as economic conditions, competitive landscape, and industry trends.

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

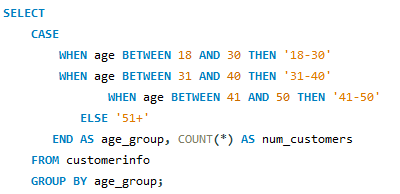
****

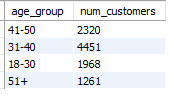
* Visualizing agebucket can help identify any disproportionate representation of certain groups among existing customers. Analyzing transactional data salary may reveal differences in behavior between existing and retained customers.

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

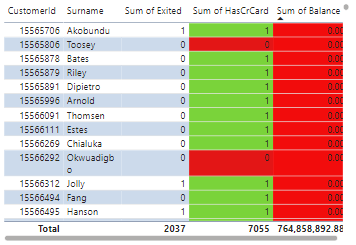
* Yes, 'Tenure', 'NumOfProducts', 'IsActiveMember', and 'EstimatedSalary' are important predictors for determining if a customer will leave the bank. These columns provide valuable insights into a customer's ability to purchase products within their estimated salary, settle payments within their tenure, and maintain active membership status in the bank.

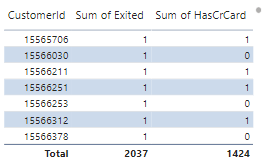
1. **Utilize SQL queries to segment customers based on demographics and account details.**





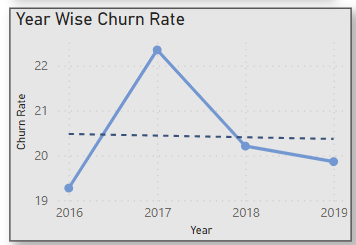
1. **How can we create a conditional formatting setup to visually highlight customers at risk of churn and to evaluate the impact of credit card rewards on customer retention?**





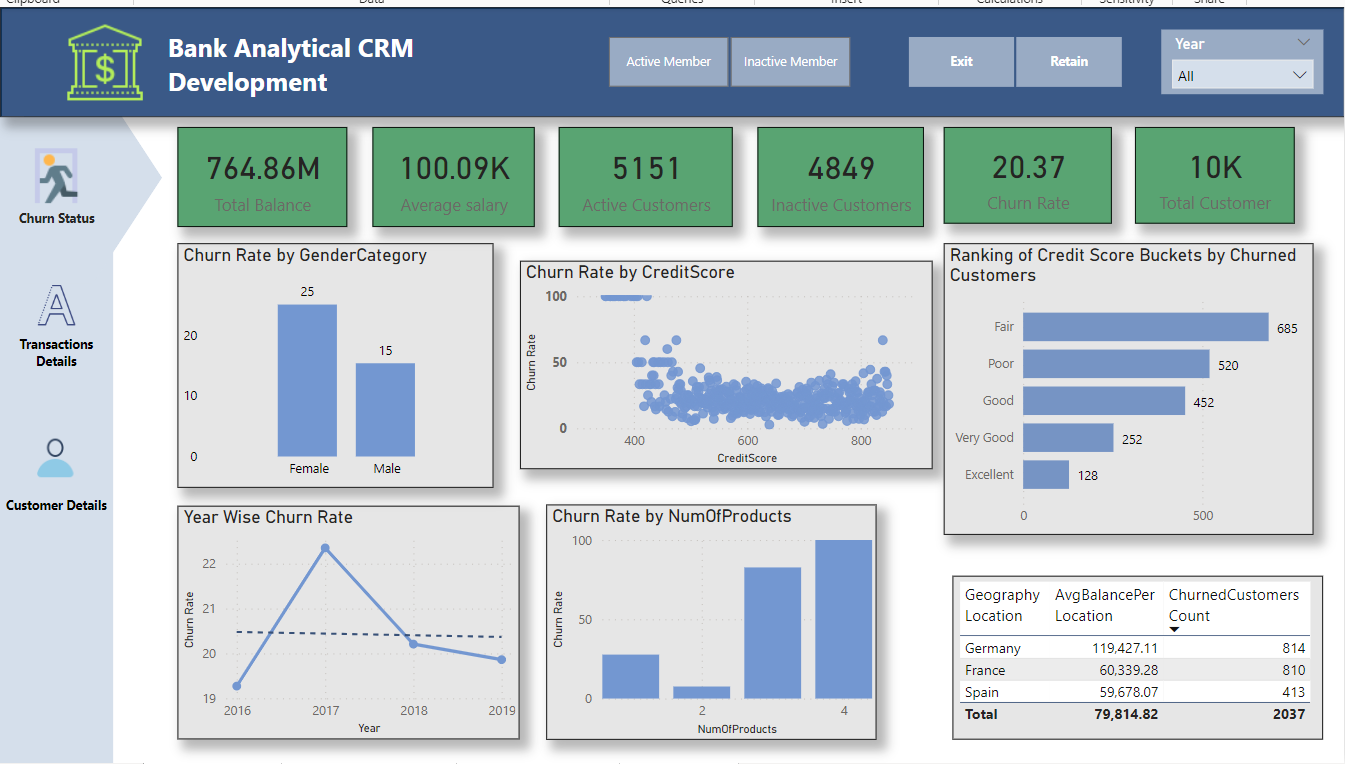
* By visually highlighting customers with 0 credit cards and 0 balance, the bank can quickly identify individuals who may be at risk of churn. This allows for targeted retention efforts to encourage account activity and mitigate churn. It indicating by red color.
* Comparing churn rates between customers with and without credit cards helps assess the effectiveness of credit card rewards in retaining customers. A lower churn rate among customers with credit cards indicates a positive impact on retention, informing strategies to enhance credit card rewards programs and foster customer loyalty.

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?**

--Overall Churn Rate

* Based on the current data analysis, the churn rate per year and overall, in the bank can be seen from the above visual. Insights to the bank about the types of customers more likely to churn can be gleaned from patterns observed in the data. For instance, customers with lower tenures, fewer products, inactive memberships, or lower estimated salaries may be more prone to churn. Strategies to decrease the churn rate could include targeted marketing campaigns to re-engage inactive customers, personalized offers to incentivize loyalty, improved customer service experiences, and proactive retention efforts such as loyalty programs or rewards programs. Additionally, analysing customer feedback and conducting satisfaction surveys can provide valuable insights into areas for improvement and help tailor retention strategies accordingly.

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

****

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

* If the objective and subjective questions were not provided, and I needed to approach the problem of analyzing customer churn and related factors in a banking dataset, I would follow a structured approach to gain insights and create a meaningful analysis. Here's how I would approach it:
* Familiarize with the dataset and understand its structure, variables, and meanings.
* Exploring data distributions, patterns, and relationships between variables using visualizations and statistical summaries.
* Creating new features or transforming existing ones to enhance analysis.
* Analyze churn rates, trends, and factors influencing churn using statistical tests or machine learning models.
* Develop interactive dashboards to visualize key insights and trends effectively.
* Summarize findings and provide actionable recommendations to reduce churn and improve customer satisfaction.
* Iterate on the analysis based on feedback and new insights, updating the dashboard and recommendations accordingly.

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



* This SQL statement will rename the column "HasCrCard" to "Has\_creditcard" in the "Bank\_Churn" table. It effectively changes the column name without affecting the data stored in the column.