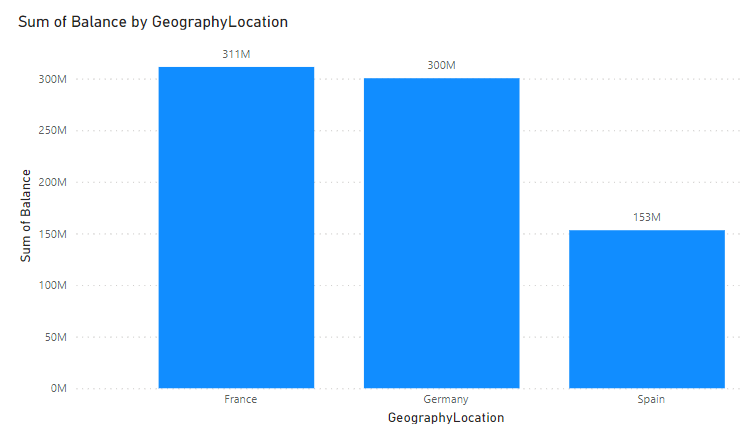
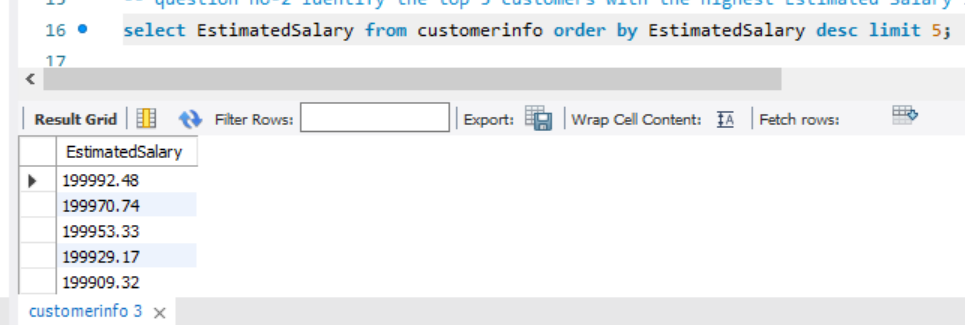
**Objective Questions**

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

**Answer:** 

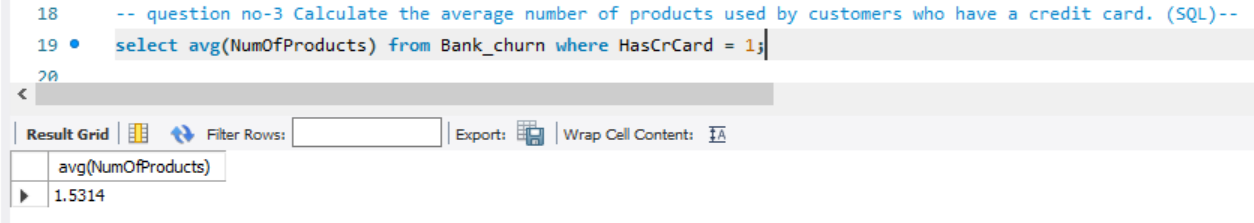
1. **Identify the top 5 customers with the highest Estimated Salary in the last quarter of the year. (SQL)**

**Answer:** The code is written in SQL file Name as (question no- 2) & the output is here



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

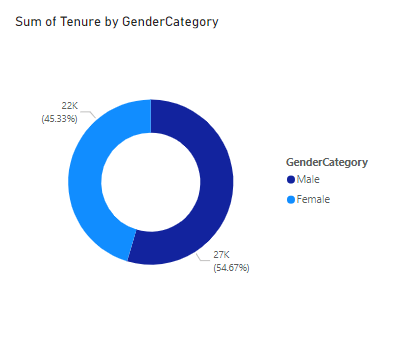
**Answer:** The code is written in SQL file Name as (question no- 3) & the output is here



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

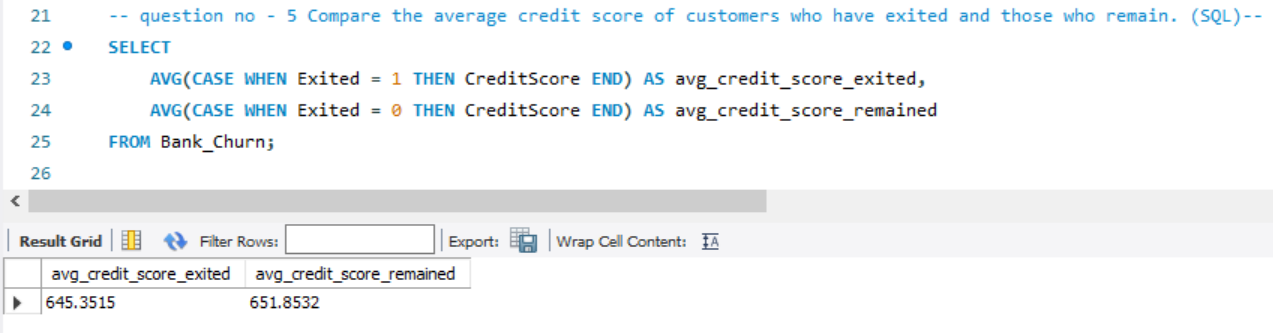
**Answer:**

I utilized a donut chart to display the distribution of gender and tenure. Additionally, I applied a filter to only include tenures greater than 7 in the visualization.



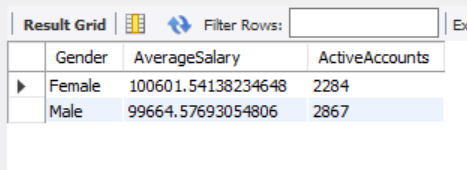
1. **Compare the average credit score of customers who have exited and those who remain. (SQL)**

**Answer :** Code is written in SQL file Name as (question no- 5) & the output is here



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

**Solution** : Code is written in SQL file Name as (question no- 6) & the output is here



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

**Answer:** you can find output in SQl.file

SELECT

CreditScore,

AVG(CAST(IsActiveMember AS DECIMAL)) AS ExitRate

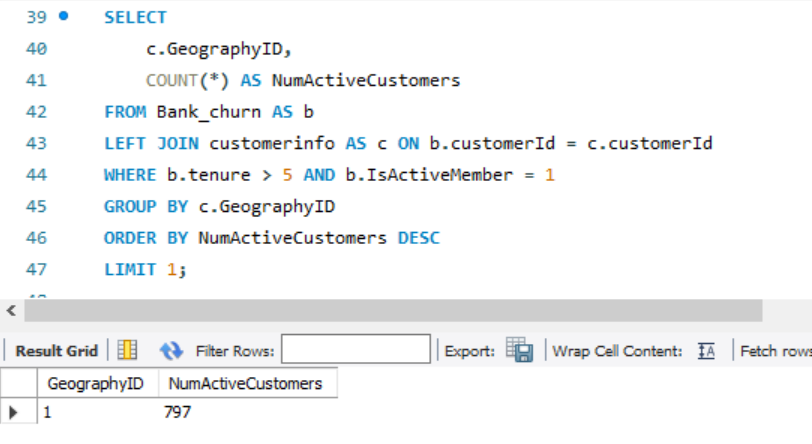
FROM Bank\_Churn

GROUP BY CreditScore

ORDER BY ExitRate DESC;

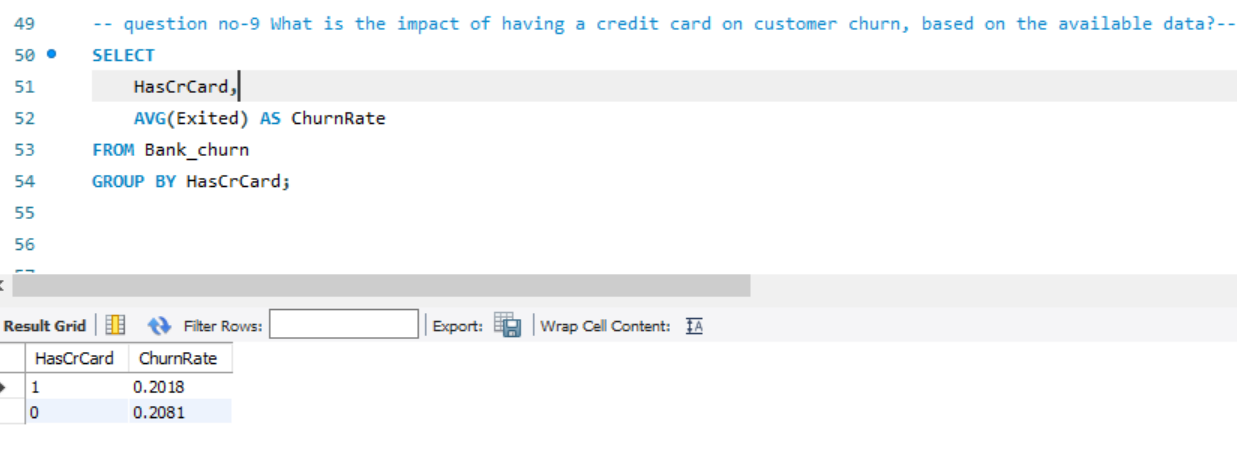
1. **Find out which geographic region has the highest number of active customers with a tenure greater than 5 years. (SQL)**

**Answer :** The code is written in SQL file Name as (question no- 8) & the output is here



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

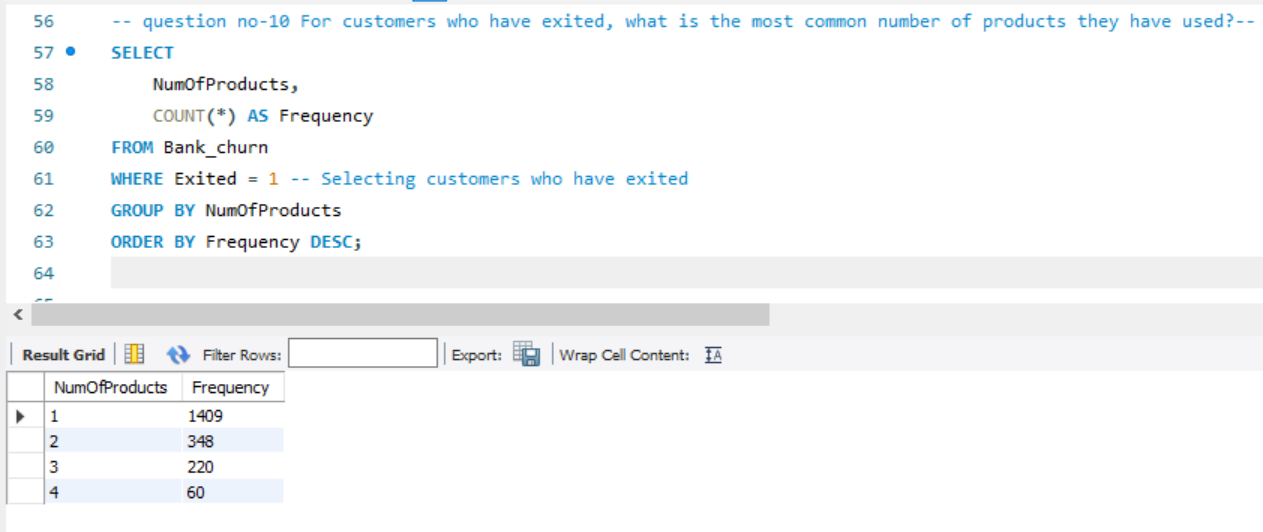
**Answer :** The code is written in SQL file Name as (question no- 9) & the output is here

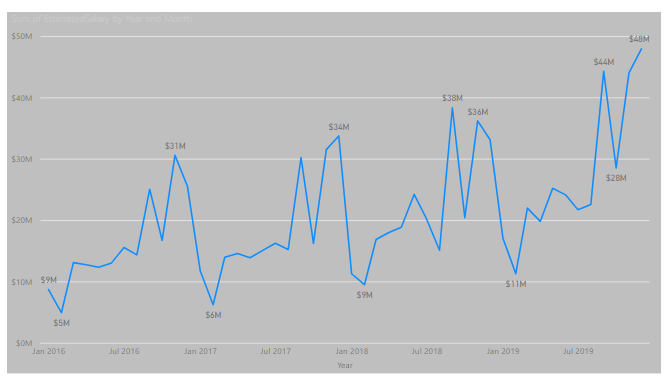


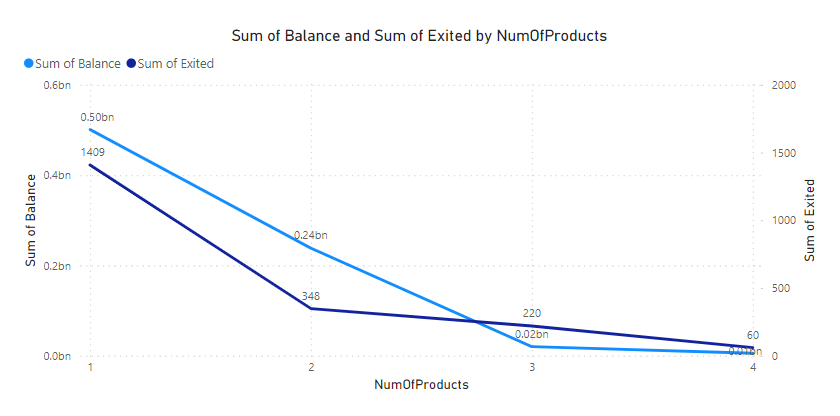
1. **For customers who have exited, what is the most common number of products they have used?**

**Answer:** The code is written in SQL file Name as (question no- 10) & the output is here

The most common number of products = 1



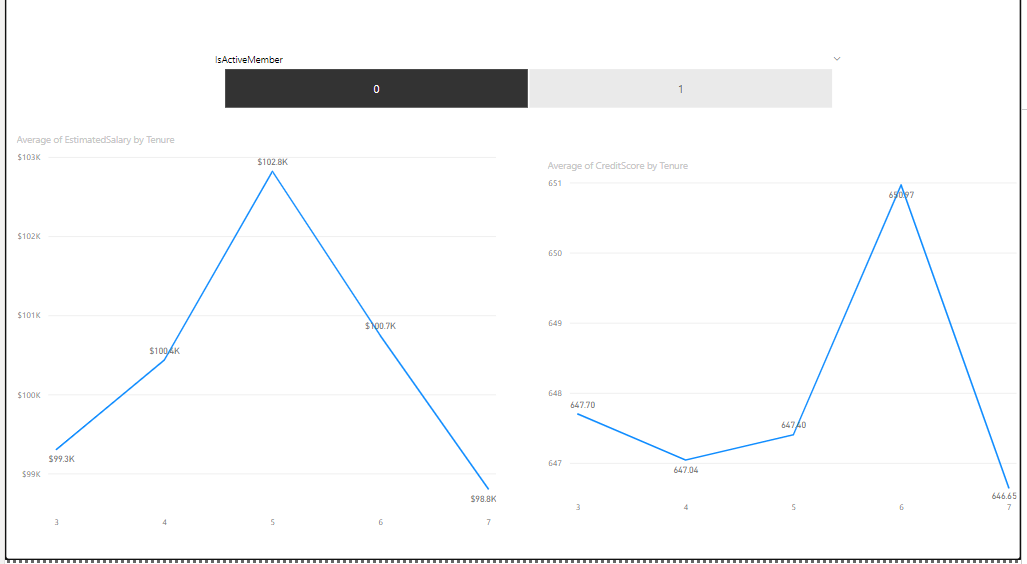
1. **Examine the trend of customers joining over time and identify any seasonal patterns (yearly or monthly). Prepare the data through SQL and then visualize it.**
2. **Analyze the relationship between the number of products and the account balance for customers who have exited.**



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

**Solution :**

1. Filter Customers: Isolate long-term customers who have maintained active membership status.
2. Credit Score Analysis: Assess credit score trends among these long-term, active members to observe if scores generally improve or remain stable over time.
3. Spot Discrepancies: Identify instances where long-term customers exhibit low credit scores, signaling potential financial challenges or shifts in behavior that require further investigation.

****

****

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

**Answer:** I have 7 Tables  
 **Categorial data** means, which only contains a text value

Example:

Gender (categories: "Male," "Female," "Other")

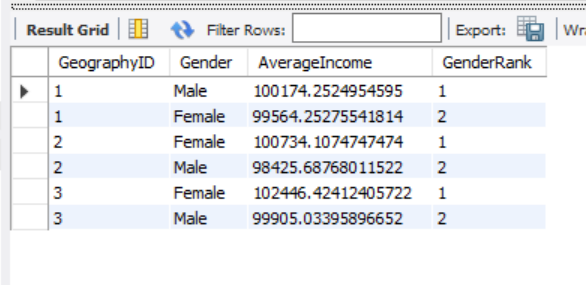
Marital status (categories: "Married," "Single," "Divorced," "Widowed")

And, **we have six columns that contain Categorial data**

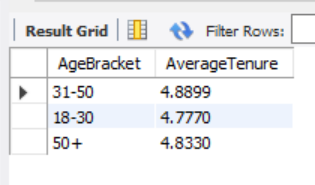
1. **ActiveCustomer (ActiveCategory)**
2. **CreditCategory (Category)**
3. **CustomerInfo (Surname)**
4. **Exit\_table (ExitCategory)**
5. **Gender (GenderCategory)**
6. **Geography (GeographyLocation)**

**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. (SQL)**

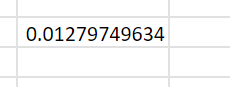
**Solution :** Code is written in SQL file Name as (question no- 15) & the output is here



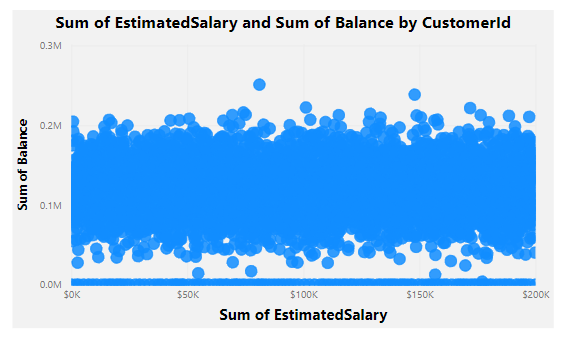
**16. Using SQL, write a query to find out the average tenure of the people who have exited in each age bracket (18-30, 30-50, 50+).**

**Solution**: Code is written in SQL file Name as (question no- 16) & the output is here   


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

**Solution: the correlation between balance and salary is **

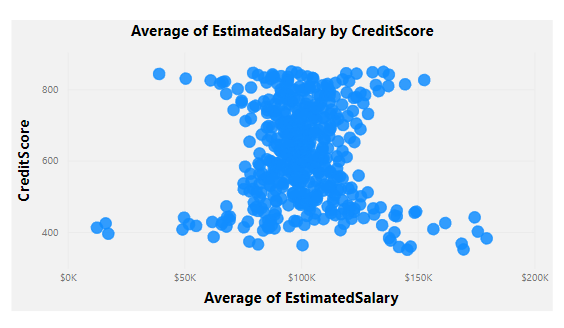
**=CORREL(Bank\_Churn!$D$2:$D$10001,CustomerInfo!$E$2:$E$10001)**

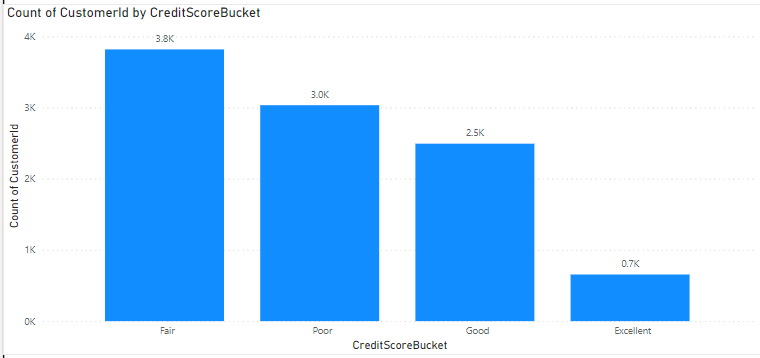


**18. Is there any correlation between the salary and the Credit score of customers?**

**Solution: the correlation between Credit Score and Salary is **

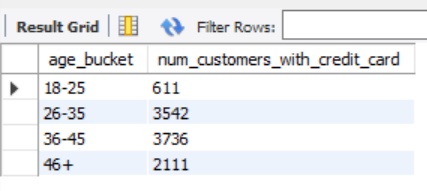
**=CORREL(CustomerInfo!$E$2:$E$10001,Bank\_Churn!$B$2:$B$10001)**

****

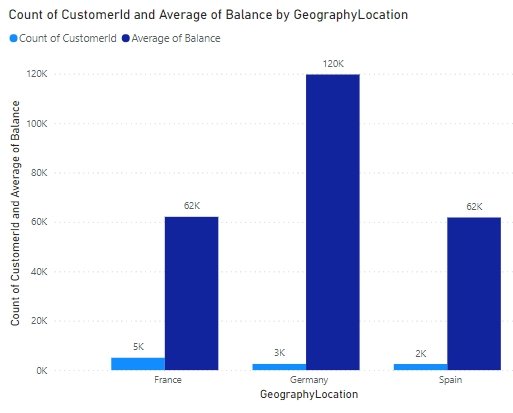
**19. Rank each bucket of credit score as per the number of customers who have churned the bank.**

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

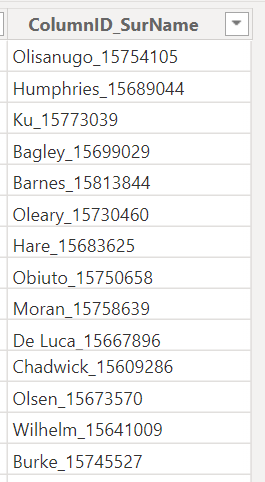
**Solution :** Code is written in SQL file Name as (question no- 20) & the output is here

****

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

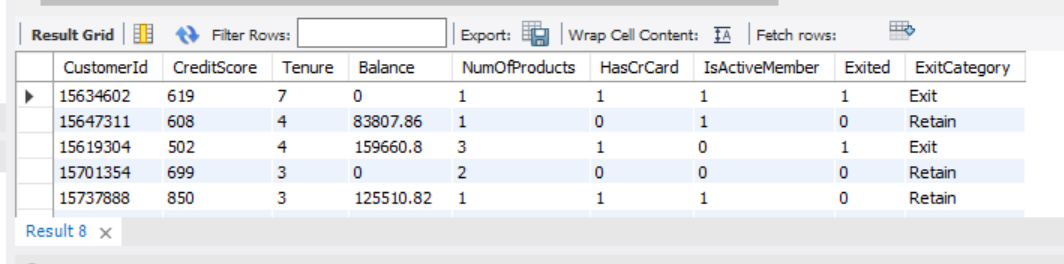


**22. As we can see that the “CustomerInfo” table has the CustomerID and Surname, now if we have to join it with a table where the primary key is also a combination of CustomerID and Surname, come up with a column where the format is “CustomerID\_Surname”.**



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

**Solution**: Code is written in SQL file Name as (question no- 23) & the output is here



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

**Answer:**

the best tool depends on our specific needs and requirements.

**Exce**l is suitable for basic data analysis and Basic visualization tasks,

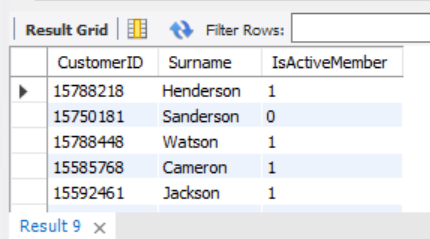
**Power BI** offers advanced visualization and connectivity features for larger datasets and more complex analyses, and

**SQL** provides powerful capabilities for data manipulation and querying in relational databases. Consider factors such as ease of use, scalability, and performance.

**For Handling Missing Value** ,Excel is always my first choice for simplicity and But Again it’s depends on the size of the dataset

**25. 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”.**

**Solution:** Code is written in SQL file Name as (question no- 25) & the output is here



**Subjective Questions(Start Here)**

1. **Customer Behavior 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?**

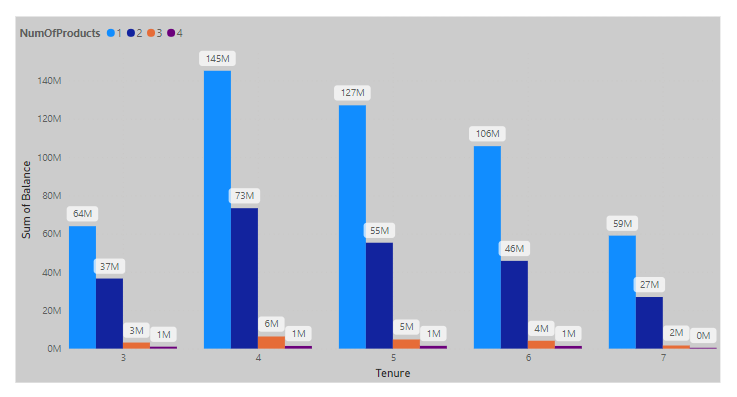
**Solution :** I take 3 columns for Customer Behaviour Analysis

1. **Balance**
2. **Number Of Product**
3. **Tenure**

While spending more than three years with a bank can indicate a higher likelihood of customer loyalty, it's essential to consider that loyalty is influenced by various factors beyond just the duration of the relationship. Here are a few points

* **Behavioral Patterns**: It's important to analyze the actual spending habits.
* Customers with **longer tenure** tend to have **higher balances**, indicating a potential correlation between loyalty and account activity

**Conclusion:** A customer who spends actively within their first year of tenure is likely to become a long-term customer.

****

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

**Solution:**

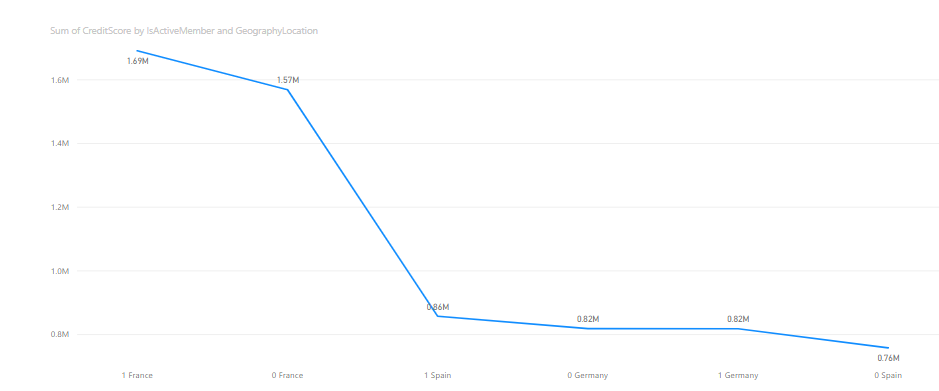
we need more information about that to calculate this question.

To perform a product affinity study using the provided data, we would need **transactional data or information about the banking products/services** used by customers.

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

**Solution:**

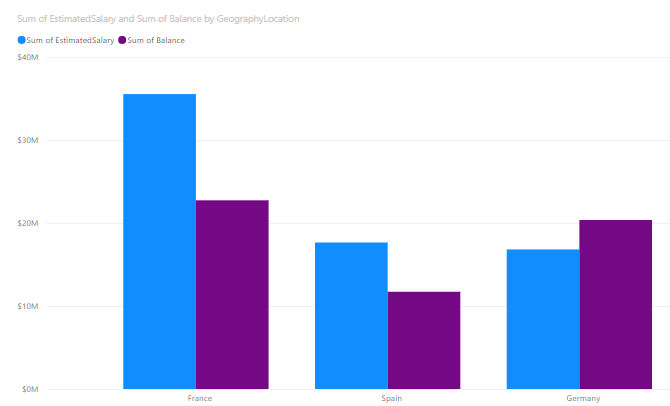
1. The visual analysis indicates that active bank customers generally have higher credit scores compared to non-active customers.
2. This trend is particularly noticeable in Spain, where active customers exhibit significantly higher credit scores than non-active customers.
3. The observed disparity underscores the importance of customer engagement and retention strategies, especially in regions like Spain where non-active customers show notably lower credit scores.
4. Understanding these patterns can guide targeted efforts to improve customer engagement, mitigate risk, and enhance overall financial health in specific geographic markets.

****

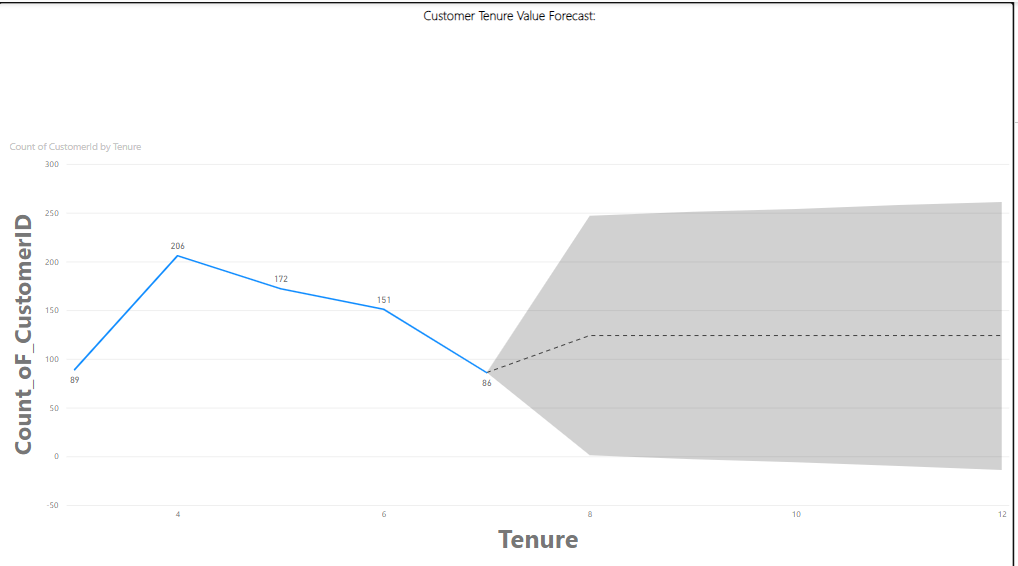
1. **Risk Management Assessment: Based on customer profiles, which demographic segments appear to pose the highest financial risk to the bank, and why?**

**Solution:**

1. Visual analysis indicates that customers from Spain demonstrate lower income levels and maintain smaller account balances compared to other regions.
2. These characteristics suggest a higher susceptibility to financial instability and may increase the likelihood of defaulting on loans or missing payments.
3. Consequently, customers from Spain pose a heightened financial risk for the bank, potentially leading to increased loan delinquencies and credit losses.
4. To mitigate this risk, the bank should implement targeted risk management strategies tailored to address the unique challenges faced by customers from Spain, such as adjusting lending criteria and offering financial education programs.

****

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?**

****

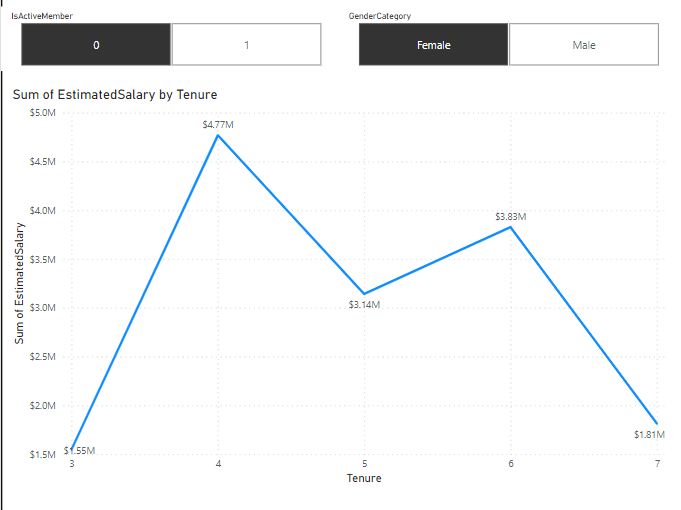
1. **Marketing Campaign Effectiveness: How could you assess the impact of marketing campaigns on customer retention and acquisition within the dataset? What extra information would you need to solve this?**

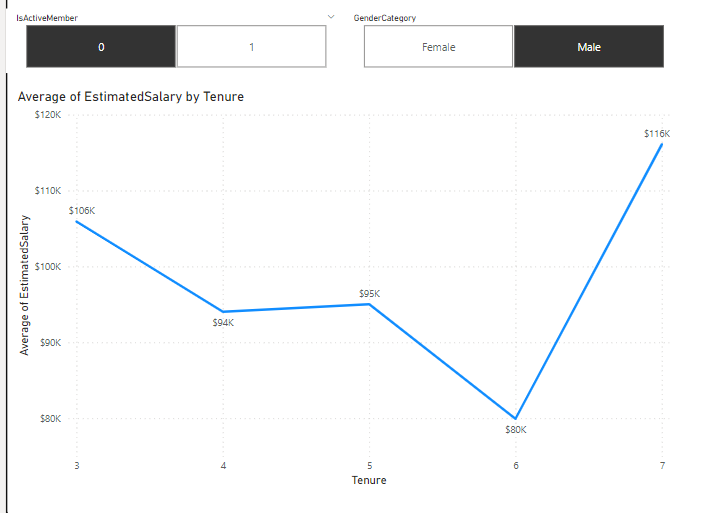
**Solution :**To assess the impact of marketing campaigns on customer retention and acquisition within the dataset, we would need:

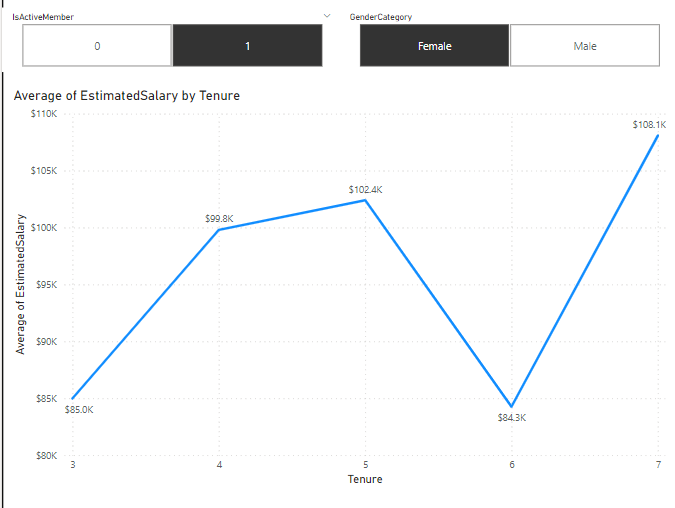
1. **Campaign Data:** Details about the campaigns conducted, including type, duration, channels, messaging, and offers.
2. **Customer Interaction Data:** Information on customer responses to the campaigns, such as clicks, purchases, or account openings.
3. **Customer Behavior Data:** Historical data on customer behavior, including transactions, product usage, and churn rates.
4. **Customer Segmentation Data:** Demographic and behavioral information to analyze campaign effectiveness across different customer segments.

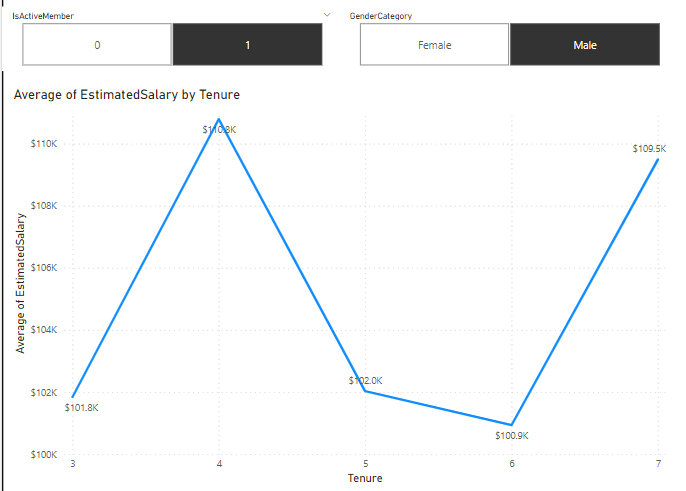
Analyzing these datasets can help evaluate changes in retention and acquisition rates before and after campaigns, compare results between exposed and unexposed groups, and identify which customer segments respond best to the campaigns.

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

****

****

****

****

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

**Solution:**

Yes, 'Tenure', 'NumOfProducts', 'IsActiveMember', and 'EstimatedSalary' are all important factors for predicting if a customer will leave the bank.

* **Tenure:** This data provides information about the number of years a customer has been associated with the bank. Longer tenure may indicate a stronger relationship with the bank and potentially lower likelihood of churn.
* **NumOfProducts**: The number of products a customer holds with the bank gives insights into their level of engagement and loyalty. Customers with multiple products are likely more integrated into the bank's ecosystem and may be less inclined to leave.
* **IsActiveMember:** The status of whether a customer is an active member or not is crucial. Active members are more likely to have ongoing interactions with the bank and may demonstrate higher loyalty, reducing the likelihood of churn.
* **EstimatedSalary:** Estimated salary provides an indication of the customer's financial capacity and ability to afford banking products and services. Customers with higher estimated salaries may be more financially stable and less likely to switch banks.

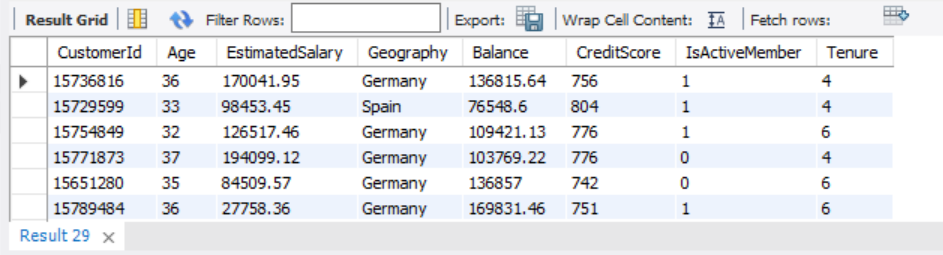
In Conculsion , each of these factors contributes valuable information for predicting customer churn. They offer insights into the customer's relationship with the bank, level of engagement, financial capacity, and overall likelihood of leaving. Therefore, considering all these factors collectively improves the accuracy of churn prediction models and enables banks to proactively address customer retention strategies.

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

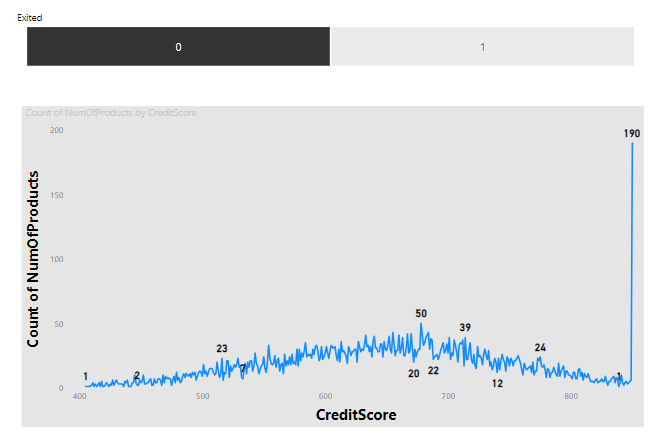
**Solution:**

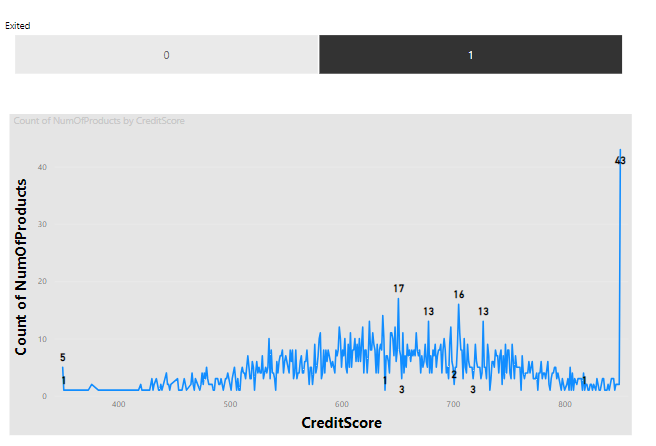
1. **Demographics:**
   1. Age: Customers aged between 30 and 40 years old (CI.Age >= 30 AND CI.Age <= 40).
2. **Account Details:**
   1. Account Balance: Customers with a balance greater than $10,000 (BC.Balance > 10000).
   2. Credit Score: Customers with a credit score higher than 650 (BC.CreditScore > 650).

Code is written in SQL file Name as (question no- 9(Subjective\_Questions ) & the output is here

****

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?**

****

****

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?**

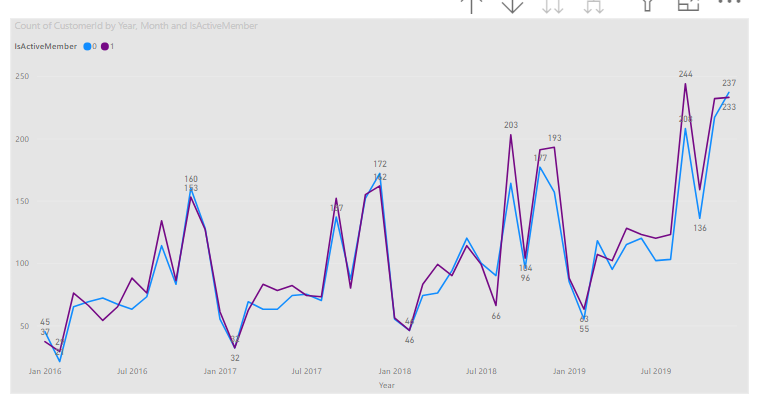
**Solution:**

**Insights:**

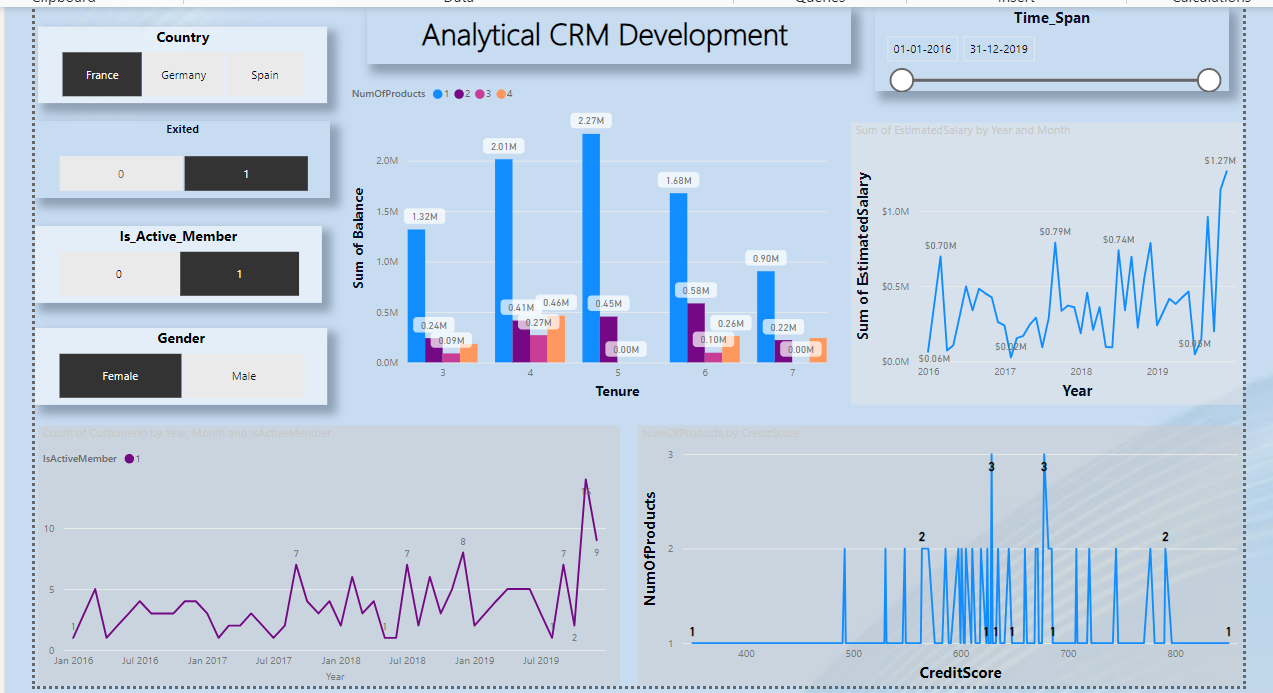
1. **Variable Churn Rates:** Churn rates fluctuated over the years, peaking at 8% in 2018 and reaching a low of 2% in 2019.
2. **Overall Stability:** Despite fluctuations, the overall churn rate remained relatively stable at an average of 5% from 2017 to 2021.

**Suggestions to Decrease Churn Rate:**

1. **Identify Root Causes:** Conduct thorough investigations into the factors driving high churn rates in specific years, possibly stemming from policy changes, economic shifts, or customer dissatisfaction.
2. **Segmentation for Targeting:** Segment customers based on demographics and behavior to tailor retention strategies, addressing individual preferences and pain points effectively.
3. **Enhance Customer Experience:** Prioritize improvements in the overall customer journey, emphasizing user-friendly digital platforms, streamlined processes, and responsive customer support.
4. **Personalization and Proactivity:** Utilize data analytics to offer personalized products and services, accompanied by proactive financial guidance to meet diverse customer needs and foster long-term relationships.
5. **Continuous Monitoring and Adaptation:** Establish a robust monitoring system to track churn metrics and customer feedback, enabling agile adjustments to retention initiatives and ensuring ongoing improvement in customer retention efforts.

****

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?**

**Solution :**

1. **First step -** Understanding of data & their columns & ready for the data cleaning process.
2. **Second step** - find the connection between the two columns, and how they will help me to find the insights
3. **Third step** - find which visual is better for better visualization.
4. **In the “Bank\_Churn” table how can you modify the name of the “HasCrCard” column to “Has\_creditcard”?**

**Solution :**

1. **Right-click on the "HasCrCard" column header and select "Rename".**
2. **Enter the new name "Has\_creditcard" and press Enter.**

****