**--OBJECTIVE QUESTION--**

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

WITH total\_balance AS (

SELECT SUM(balance) AS total FROM bank\_churn

),

region\_balance AS (

SELECT g.GeographyLocation, SUM(b.balance) AS region\_total

FROM geography g

JOIN customerinfo ci ON ci.GeographyID = g.GeographyID

JOIN bank\_churn b ON b.CustomerId = ci.CustomerId

GROUP BY g.GeographyLocation

)

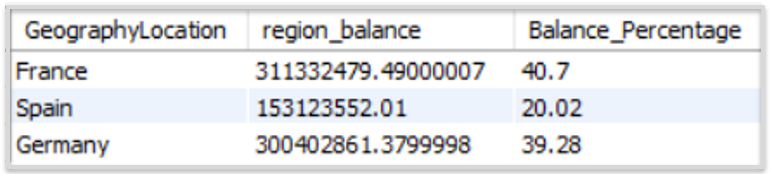
SELECT

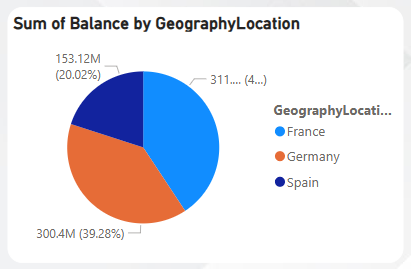
rb.GeographyLocation,

rb.region\_total,

ROUND((rb.region\_total / tb.total) \* 100, 2) AS Balance\_Percentage

FROM region\_balance rb, total\_balance tb;





**Question 2. Identify the top 5 customers with the highest Estimated Salary in the last quarter of the year. (SQL)**

With CTE as (

select CustomerId, MAX(EstimatedSalary) as Highest\_Estimated\_Salary,substr(BankDOJ,7,4) as Year from customerinfo

WHERE BankDOJ BETWEEN DATE\_FORMAT(NOW(),'01-10-%y') AND DATE\_FORMAT(NOW(),'31-12-%y')

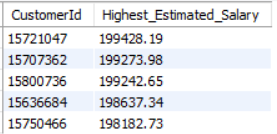
group by CustomerId,substr(BankDOJ,7,4)),

CTE2 as (select CustomerId, Highest\_Estimated\_Salary, Year,

dense\_rank() over (partition by Year order by Highest\_Estimated\_Salary desc) as Rank\_year from CTE)

select CustomerId, Highest\_Estimated\_Salary from CTE2 where Rank\_year between 1 and 5

order by year Limit 5;

****

**Question 3. Calculate the average number of products used by customers who have a credit card. (SQL)**

Select avg(NumOfProducts) as average\_number\_of\_products

from bank\_churn where HasCrCard = 1;



**Question 4. Determine the churn rate by gender for the most recent year in the dataset.**

Select g.GenderCategory,count(Case when b.Exited = 1 then 1 end) as ChurnedCount,

**-- count for those customers who left or churned --**

count(\*) as TotalCount,

Round((count(Case when b.Exited = 1 then 1 end) / count(\*)) \* 100, 2) as ChurnRatePercentage

from CustomerInfo ci

join Gender g on ci.GenderID = g.GenderID

join Bank\_Churn b on ci.CustomerId = b.CustomerId

where year(ci.BankDOJ) = (Select MAX(year(BankDOJ)) from CustomerInfo)

**-- select most recent year from dataset --**

group by g.GenderCategory;

|  |  |  |  |
| --- | --- | --- | --- |
| **GenderCategory** | **ChurnedCount** | **TotalCount** | **ChurnRatePercentage** |
| Female | 385 | 1537 | 25.0488 |
| Male | 273 | 1776 | 15.3716 |

**GenderCategory:** This column identifies the gender of the customers, classified into two groups: Male and Female.

**ChurnedCount:** This column represents the number of customers who have discontinued their service, categorized by gender. For instance, 385 female customers have churned.

**TotalCount:** This column displays the overall number of customers for each gender group. For example, there are 1,537 female customers in total.

**ChurnRatePercentage:** This column calculates the churn rate for each gender by determining the percentage of customers who churned relative to the total customers in that category. The formula used is:

Churn Rate = (ChurnedCount / TotalCount) ×100

For example:

* Among female customers, 385 out of 1,537 (approximately 25.05%) have churned.
* Among male customers, 273 out of 1,776 (around 15.37%) have churned.

This provides insights into the churn rate distribution across gender groups.

**Question 5. Compare the average credit score of customers who have exited and those who remain. (SQL)**

Select

Round(avg(Case when Exited = 1 then CreditScore end),2)as exitcust\_credit\_score,

Round(avg(Case when Exited = 0 then CreditScore end),2) as remainingcust\_credit\_score

from bank\_churn;

|  |  |
| --- | --- |
| **exitcust\_credit\_score** | **remainingcust\_credit\_score** |
| **645.35** | **651.85** |

**Question 6. Which gender has a higher average estimated salary, and how does it relate to the number of active accounts? (SQL)**

With GenderSalary as (Select g.GenderCategory,avg(ci.EstimatedSalary) as Avg\_Estimated\_Salary

from customerinfo ci

join gender g on ci.GenderID= g.GenderID

group by g.GenderCategory),

GenderActiveAccounts as (Select g.GenderCategory,

count(Case when b.IsActiveMember = 1 then 1 end) as Active\_Count

from CustomerInfo ci

join Gender g on ci.GenderID = g.GenderID

join bank\_churn b on ci.CustomerId = b.CustomerId

group by g.GenderCategory)

Select gs.GenderCategory,Round(gs.Avg\_Estimated\_Salary,2) as Avg\_est\_salary,gaa.Active\_Count

from GenderSalary gs

join GenderActiveAccounts gaa on gs.GenderCategory = gaa.GenderCategory

order by Avg\_est\_salary desc;



**Question 7. Segment the customers based on their credit score and identify the segment with the highest exit rate. (SQL)**

With Segments as (Select ci.CustomerId,b.CreditScore,

Case when b.CreditScore between 300 and 579 then 'Poor'

when b.CreditScore between 580 and 669 then 'Fair'

when b.CreditScore between 670 and 739 then 'Good'

when b.CreditScore between 740 and 799 then 'VeryGood'

when b.CreditScore between 800 and 850 then 'Excellent'

end as CreditScoreSegment,b.Exited

from Bank\_Churn b

join CustomerInfo ci on b.CustomerId = ci.CustomerId)

Select CreditScoreSegment,

count(Case when Exited = 1 then 1 end) as ChurnedCount,

count(\*) as TotalCount,

Round((count(Case when Exited = 1 then 1 end) \* 1.0 / count(\*)) \* 100, 2) as ExitRatePercentage

from Segments

group by CreditScoreSegment

order by ExitRatePercentage desc

Limit 1;



**Question 8. Find out which geographic region has the highest number of active customers with a tenure greater than 5 years. (SQL)**

Select g.GeographyLocation as Coutry, count(ci.CustomerId) as active\_customers

from geography g

join customerinfo ci on g.GeographyID= ci.GeographyID

join bank\_churn b on ci.CustomerId = b.CustomerID

where b.IsActiveMember = 1 and b.tenure > 5

group by g.GeographyLocation

order by active\_customers desc;

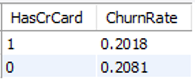
|  |  |
| --- | --- |
| **Country** | **active\_customer** |
| France | 797 |
| Spain | 431 |
| Germany | 399 |

**Question 9. What is the impact of having a credit card on customer churn, based on the available data?**

SELECT HasCrCard,AVG(Exited) AS ChurnRate

FROM bank\_churn

GROUP BY HasCrCard;



**Question 10. For customers who have exited, what is the most common number of products they had used?**

SELECT NumOfProducts,COUNT(\*) AS Frequency FROM bank\_churn

WHERE Exited = 1

GROUP BY NumOfProducts

ORDER BY Frequency DESC

LIMIT 1;



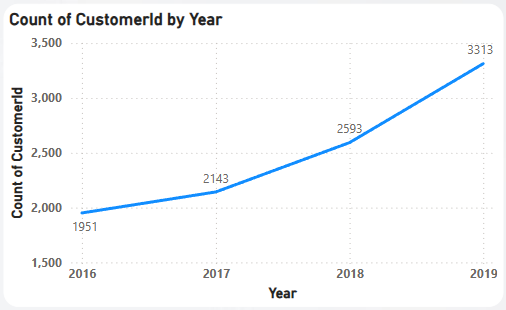
**Question 11. Examine the trend of customer joining over time and identify any seasonal patterns (yearly or monthly). Prepare the data through SQL and then visualize it.**

Select Substr(BankDOJ,1,4) as Year, count(\*) as No\_of\_Cust

from customerinfo

group by Substr(BankDOJ,1,4)

order by Year;



**Question 12. Analyze the relationship between the number of products and the account balance for customers who have exited.**

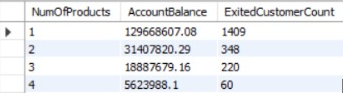
Select NumOfProducts, Round(sum(Balance),2) as AccountBalance,count(\*) as ExitedCustomerCount

from Bank\_Churn bc

where Exited = 1

group by NumOfProducts

order by NumOfProducts;

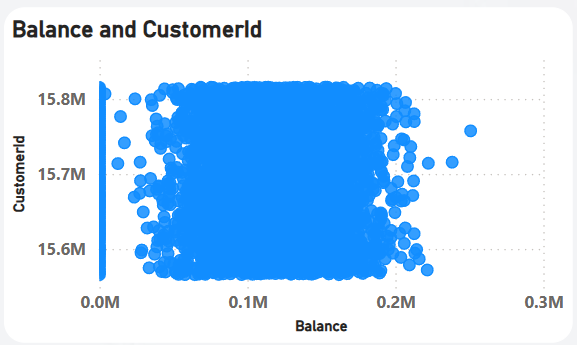


**Question 13. Identify any potential outliers in terms of balance among customers who have remained with the bank.**

We can use a scatter chart in Power BI to spot unusual balance patterns among customers who have stayed with the bank. Look for points that stand out—either much higher or lower than most others—since these could be potential outliers.

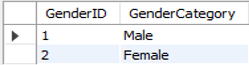
Once you identify these unusual balances, take a closer look to see if they make sense. They could be due to data entry mistakes, fraudulent activity, or just unique customer behavior. Understanding why these outliers exist can help you decide if any action is needed.

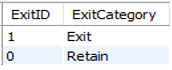
By using a scatter chart this way, you can uncover hidden patterns in customer balances and gain valuable insights into their financial behaviors.

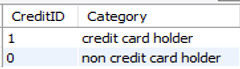


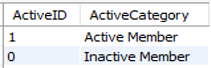
**Question 14. How many different tables are given in the dataset, out of these tables which table only consist of categorical variables?**

IN the datase have "SEVEN" different tables are given, out these seven tables which table has consist the categorical variables are "four" which is activecustomer", "creditcard", exitcustomer","gender".









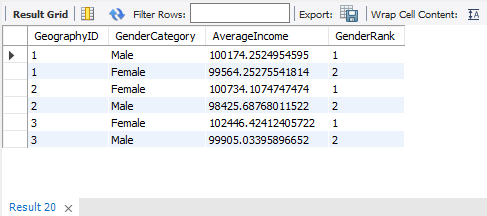
**Question 15. Using SQL, write a query to find out the gender wise average income of male and female in each geography id. Also rank the gender according to the average value. (SQL)**

Select GeographyID, GenderCategory, avg(EstimatedSalary) as AverageIncome, Rank() over (partition by GeographyID order by avg(EstimatedSalary) desc) as GenderRank

from customerinfo ci

join gender g on ci.GenderID = g.GenderID

group by GeographyID, GenderCategory;

****

**Question 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+).**

With AgeSegment as (Select Case when ci.Age between 18 and 30 then '18-30'

when ci.Age between 31 and 50 then '31-50'

when ci.Age > 50 then '50+'

end as Age\_Bracket, b.Tenure

from customerinfo ci

join bank\_churn b on ci.CustomerId = b.CustomerID

where b.Exited = 1

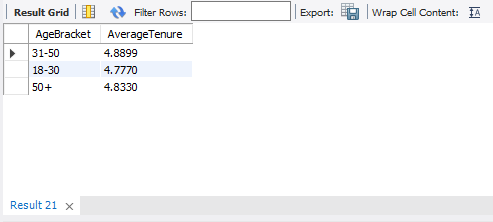
order by Age\_Bracket)

Select Age\_Bracket, Round(avg(Tenure),2) as AverageTenure

from AgeSegment

group by Age\_Bracket

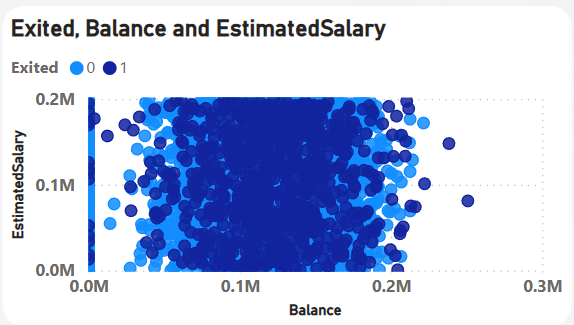
order by Age\_Bracket;

****

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

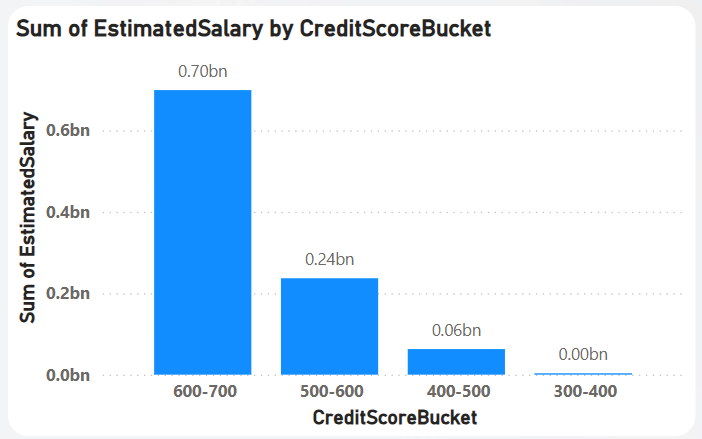
If we want to understand the relationship between salary and balance among customers and whether this differs between those who have left and those who have stayed, we can use a scatter chart in Power BI. To make the comparison clearer, we can add a category that differentiates between customers who have exited and those who are still active. Using different colors or shapes will help us visually separate the two groups. By analyzing the chart, we can look for patterns—do customers with higher salaries generally have higher balances? Is there a noticeable difference between those who have stayed and those who have left? Observing where the data points cluster can provide valuable insights.

By studying these trends, we can better understand how salary and balance interact and whether they influence customer retention. Factors like income levels, spending habits, and financial behavior might be contributing to these differences. This analysis can help us uncover key insights into customer behavior, leading to more informed retention strategies and financial planning decisions.



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

Let's analyze how credit scores vary across different salary ranges. We should look for patterns or connections between them, but keep in mind that this chart won’t directly show a correlation between salary and credit score. To visualize this, we can use a stacked column chart. First, we need to group salaries into estimated ranges to create clear categories. Then, within each salary range, we’ll calculate how credit scores are distributed and stack them in the chart. This will help us see how credit scores are spread across different income levels and whether any noticeable trends exist.



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

-- Step 1: Segment customers into credit score buckets

With Segments as (Select CreditScore,

Case when CreditScore between 300 and 579 then 'Poor'

when CreditScore between 580 and 669 then 'Fair'

when CreditScore between 670 and 739 then 'Good'

when CreditScore between 740 and 799 then 'VeryGood'

when CreditScore between 800 and 850 then 'Excellent'

end as CreditScoreSegment,Exited

from Bank\_Churn),

-- Step 2: Count churned customers in each credit score bucket

ChurnedCustomers as (Select CreditScoreSegment,COUNT(\*) AS ChurnedCount

from Segments

where Exited = 1 -- Only include churned customers

group by CreditScoreSegment

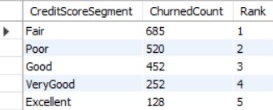
)

-- Step 3: Rank buckets by number of churned customers

Select CreditScoreSegment,ChurnedCount,

Rank() over (order by ChurnedCount desc) as 'Rank'

from ChurnedCustomers;

****

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

-- Step 1: Segment customers into age buckets

With AgeBuckets as (Select Case when ci.Age between 18 and 30 then '18-30'

when ci.Age between 31 and 50 then'31-50'

when ci.Age > 50 then '50+'

end as AgeBucket,

count(\*) as CreditCardCount

from customerinfo ci

join bank\_churn bc on ci.CustomerId = bc.CustomerId

where bc.HasCrCard = 1 -- Only include customers who have a credit card

group by AgeBucket),

-- Step 2: Calculate the average number of credit cards across all buckets

AverageCreditCardCount as (Select Round(avg(CreditCardCount),2) as AvgCreditCardCount

from AgeBuckets)

-- Step 3: Retrieve buckets with fewer than average credit cards

Select ab.AgeBucket,ab.CreditCardCount,ac.AvgCreditCardCount

from AgeBuckets ab

cross join AverageCreditCardCount ac

where ab.CreditCardCount < ac.AvgCreditCardCount;



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

With Locations as (Select g.GeographyLocation,count(Case when bc.Exited = 1 then 1 end) as ChurnedCount,

avg(bc.Balance) as AvgBalance

from bank\_churn bc

join customerinfo ci on bc.CustomerId = ci.CustomerId

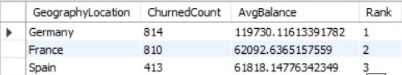
join geography g on ci.GeographyID = g.GeographyID

group by g.GeographyLocation)

Select GeographyLocation,ChurnedCount,AvgBalance,

Rank() over (order by ChurnedCount desc, AvgBalance desc) as 'Rank'

from Locations;

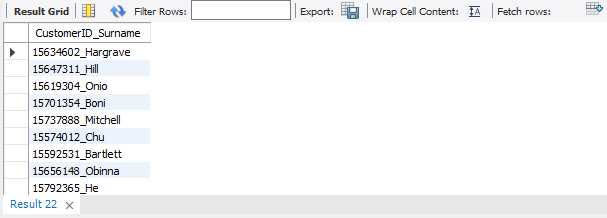


**Question 22. As we can see that the “CustomerInfo” table has the CustomerID and Surname, now if we have to join it with a table where the primary key is also a combination of CustomerID and Surname, come up with a column where the format is “CustomerID\_Surname”.**

Select CONCAT(ci.CustomerID, '\_', ci.Surname) AS CustomerID\_Surname

from CustomerInfo ci

join bank\_churn bc ON ci.CustomerID = bc.CustomerID;

****

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

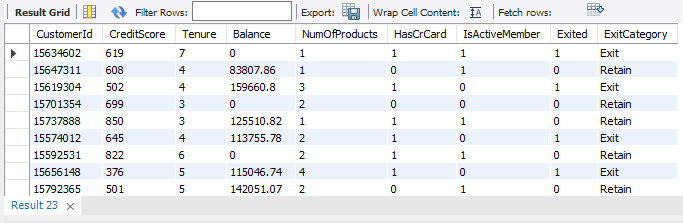
Select b.\*,

(Select ExitCategory

from ExitCustomer ec

where ec.ExitID= b.Exited) as ExitCategory

from bank\_churn b;

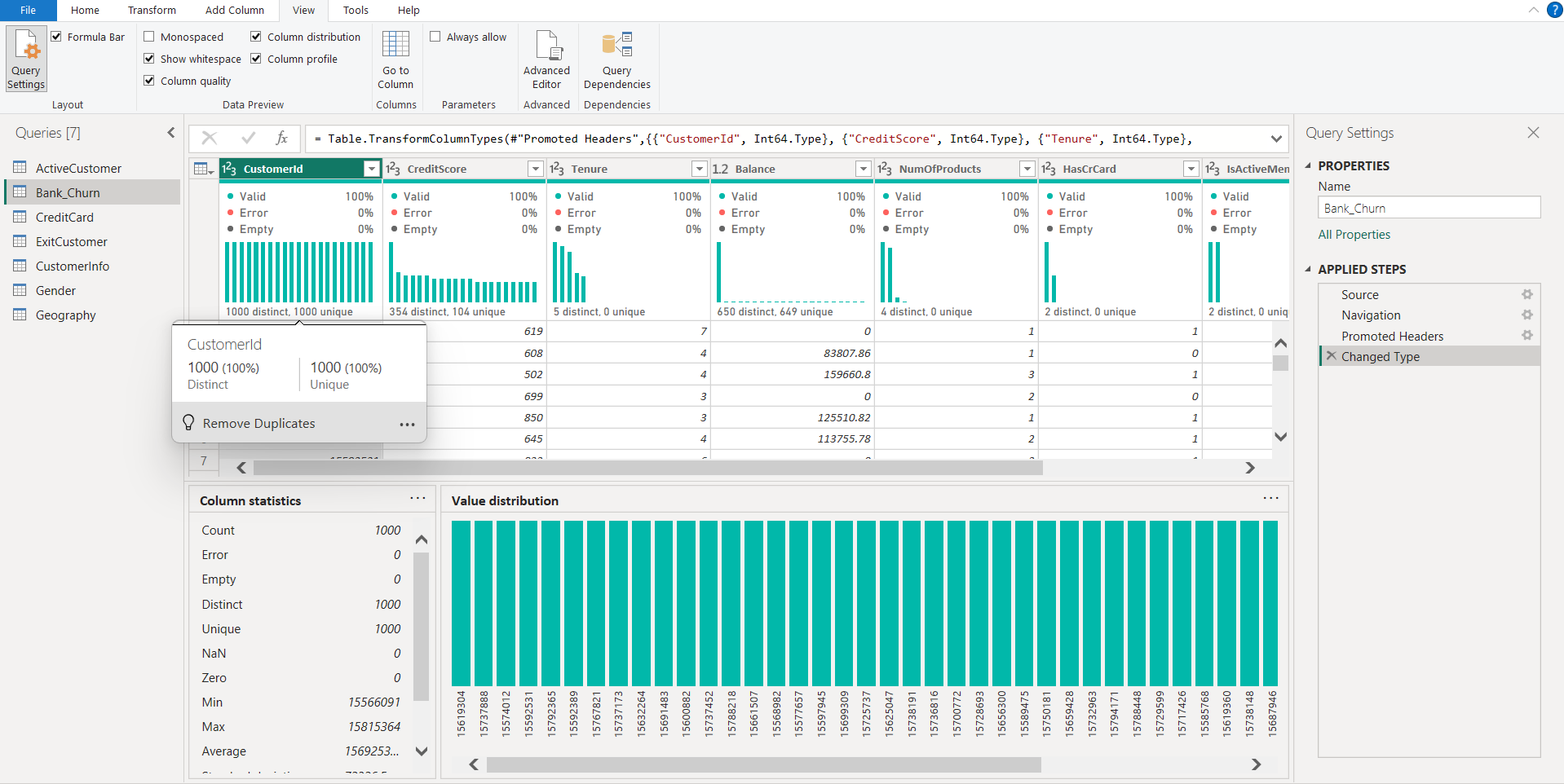
****

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

### Identifying Missing Values in Power BI

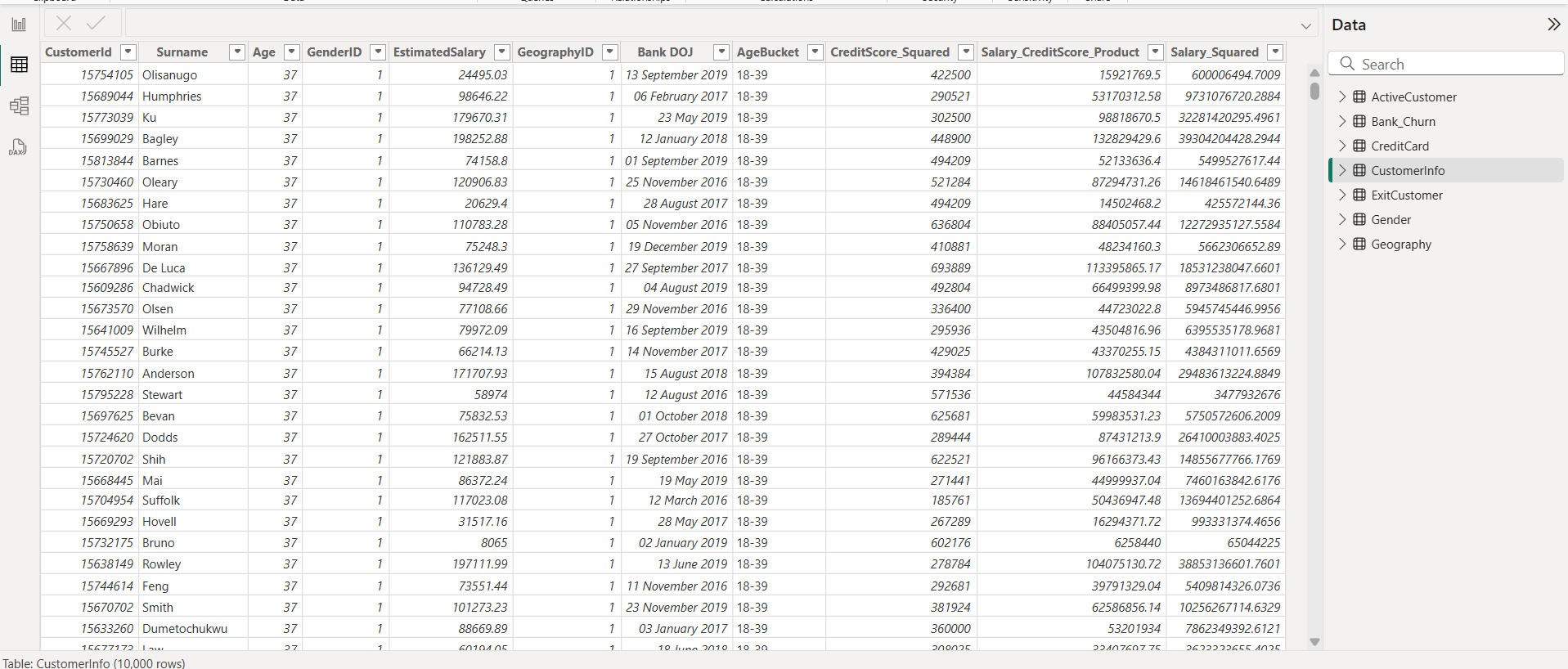
There are several ways to find missing values in Power BI:

* **Data Profiling:** Power BI has built-in data profiling tools that help detect missing values in a dataset. To use this, go to **"Transform Data"**, then select **"Column Quality"** to see summary statistics. This will highlight any missing values in your data.



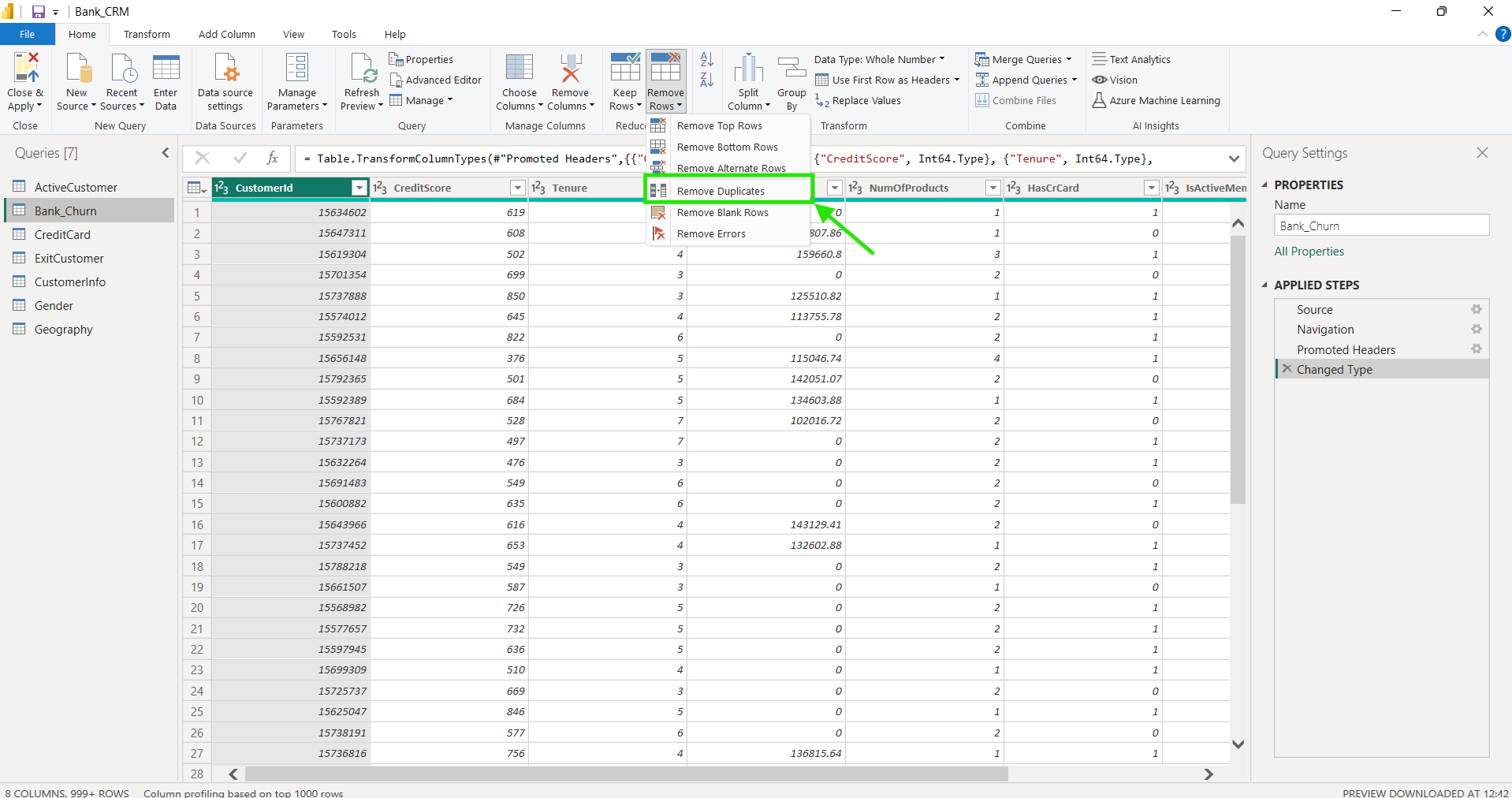
### **Data View:**

In Power BI Desktop, the **Data View** allows us to visually check for missing values. By scanning through the dataset, we can spot blank cells or unusual patterns that might indicate missing data. This helps in identifying gaps before performing further analysis.



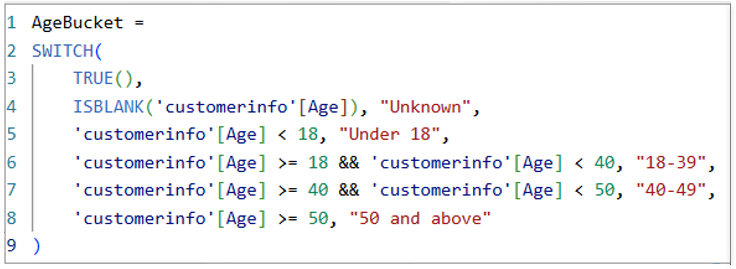
### **Data Cleaning Tools:**

Power BI provides several data cleaning options within the **"Transform Data"** window. Features like **"Replace Values"** allow us to fill in missing data with specific values, while **"Fill Down"** helps carry forward values from the previous row. These tools make it easier to handle missing values and ensure data consistency.



### **DAX Expressions:**

We can use **DAX expressions** to manage missing values programmatically in Power BI. Functions like **ISBLANK()** or **IF()** help us detect and replace missing values with default or calculated values. This approach provides more control over how missing data is handled in reports and calculations.



**Question 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”.**

Select ci.CustomerID,ci.Surname,

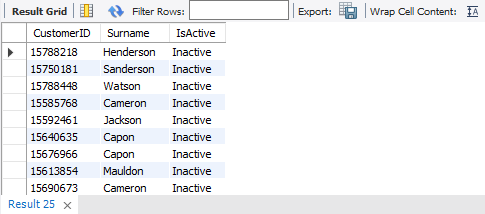
Case when ac.ActiveID = 1 then 'Active' else 'Inactive'

end as "IsActive"

from CustomerInfo ci

left join ActiveCustomer ac on ci.CustomerID = ac.ActiveID

where ci.Surname LIKE '%on';



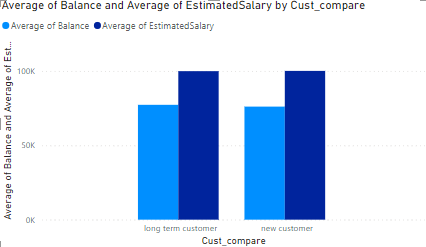
**-- SUBJECTIVE QUESTION--**

**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?**

**Solution:-** We can see that long-term customers tend to maintain a higher average balance despite having a lower average salary. This suggests that they are more loyal compared to new customers. If a customer continues to keep a high balance even with a lower income, it indicates they spend less compared to newer customers.

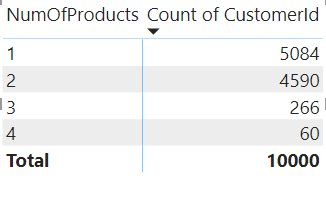
A clear pattern emerges when comparing spending habits while new customers generally have a higher average salary, their average balance is lower. On the other hand, long-term customers have a lower average salary but a higher average balance.

This pattern highlights customer loyalty, showing that long-term customers are more financially stable and committed to maintaining their relationship with the company.



**Question 2. Product Affinity Study: Which bank products or services are most commonly used together, and how might this influence cross-selling strategies?**

**Solution**:- The majority of customers tend to use **Product 1** more than any other product. This could be influenced by cross-selling strategies, as Product 1 appears to be the most preferred choice among customers compared to other products.



**Question 3. Geographic Market Trends: How do economic indicators in different geographic regions correlate with the number of active accounts and customer churn rates?**

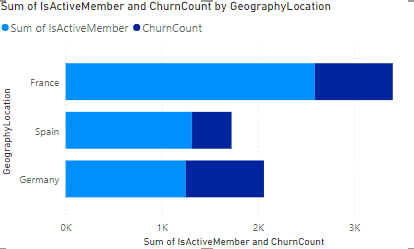
**Solution:-** To understand market trends in different regions, we can compare average balance, active customers, and churn rate.

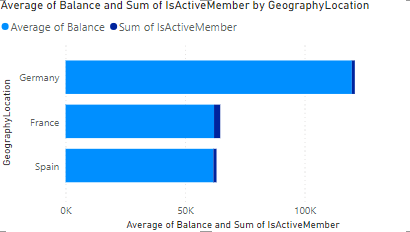
* The first bar chart shows how many customers are still active and what the churn rate looks like in each region.
* The second bar chart provides economic insights to help us see the bigger picture.

We can look at two key relationships:

1. How the number of active customers relates to the churn rate in different locations.
2. How the average balance compares to the number of active customers across regions.

By analyzing these trends, we can get a clearer picture of customer behaviour and how different factors influence market performance in various locations.





**Question 4. Risk Management Assessment: Based on customer profiles, which demographic segments appear to pose the highest financial risk to the bank, and why?**

**Solution:-** The **"Fair"** credit score group seems to carry the highest financial risk for the bank. This is because it has the largest number of customers, **3,818**, compared to other credit score categories.

### **Why This Group Poses a Higher Risk**

Certain factors make this segment riskier:

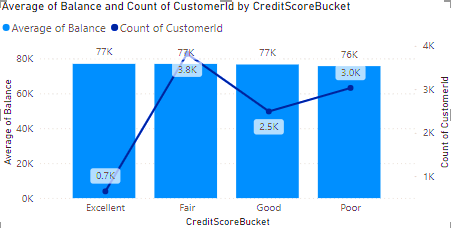
* **Higher debt compared to income**, which increases financial strain.
* **Unstable employment**, making loan repayment uncertain.
* **Economic challenges in some regions**, affecting financial stability.

### **Key Signs of High Risk**

To identify customers at higher financial risk, we can look at:

* **Low credit scores**, which indicate past financial struggles.
* **High account balances**, which might show over-reliance on borrowed money.
* **Short banking history**, making it harder to predict future financial behavior.
* **Using multiple bank products**, which could mean heavy dependence on credit.

By recognizing these warning signs, the bank can take proactive steps to manage risks and support customers effectively.

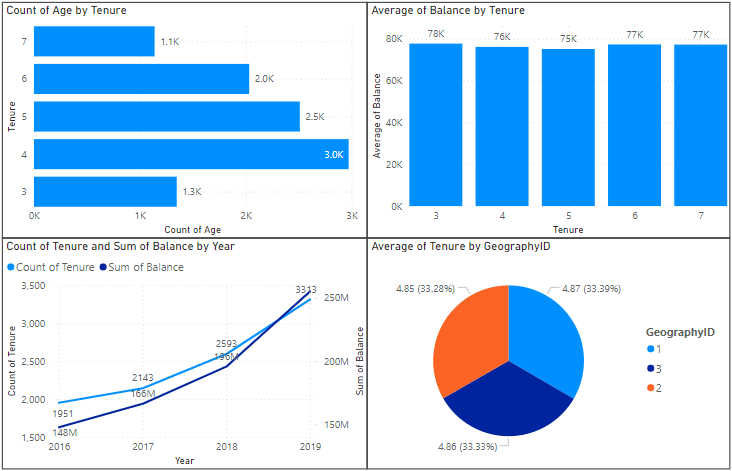


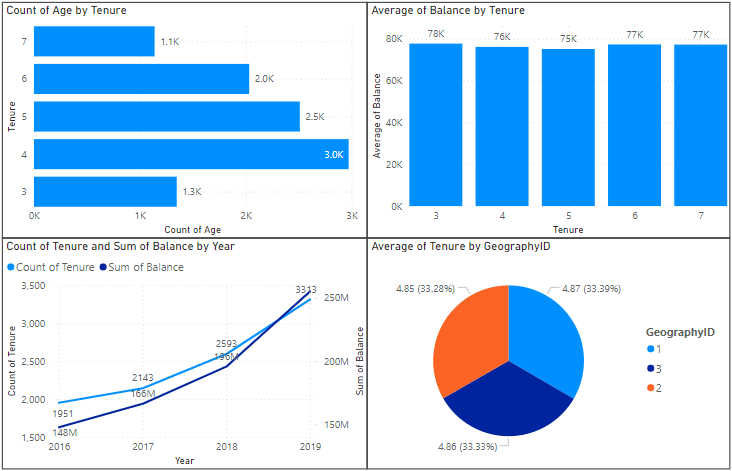
**Question 5. Customer Lifetime Value Forecast: How would you use the available data to model and predict the lifetime (tenure) value of different customer segments?**

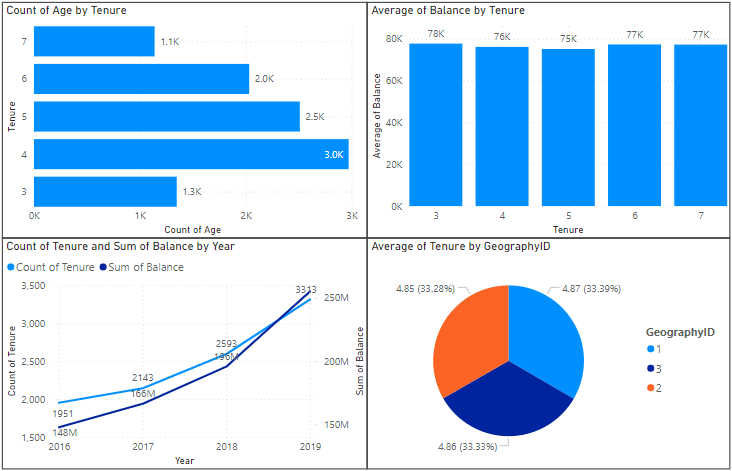
**Solution:-** To predict customer lifetime value (tenure) for different segments in Power BI, we can use different types of charts to highlight key insights:

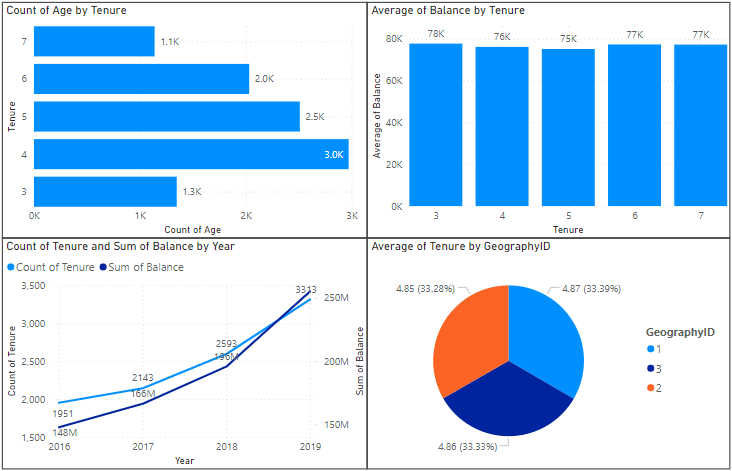
* Line Chart – Helps us track how tenure changes across different customer segments over time.
* Bar Chart – Makes it easy to see which customer segments have longer or shorter predicted tenures.
* Column Chart – Shows how tenure relates to average balance, helping us understand which tenure groups tend to have higher or lower balances.
* Pie Chart – Gives a clear view of how customers are distributed among different segments, with each slice representing a segment's proportion.

By using these visualizations, we can better understand customer behavior, identify trends, and make smarter decisions about customer retention and engagement.









**Question 6. 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 evaluate how marketing campaigns affect **customer retention and acquisition**, we need to analyze key metrics and follow a structured approach.

### **1. Identify Key Metrics**

We should focus on important metrics that help measure customer retention and new customer acquisition, such as:

* **Retention Rate** – Percentage of customers who stayed with the company over a certain period.
* **Churn Rate** – Percentage of customers who left within a specific timeframe.
* **Acquisition Rate** – Number of new customers gained during the period.

Additionally, campaign-specific metrics should be considered, such as:-

* **Campaign Reach** – Number of customers exposed to the campaign.
* **Campaign Conversion Rate** – Percentage of customers who engaged or responded positively.

### **2. Prepare the Data**

We need a clean and organized dataset that includes:-

* Customer demographics (from the "customerinfo" dataset).
* Customer tenure and churn status (from the "bank\_churn" dataset).
* Marketing campaign details such as **launch dates, target audience, and outcomes**.

By merging this data, we can create a unified dataset for deeper analysis.

### **3. Segment the Data**

Customers can be grouped based on **demographics, tenure, and behavior**, while marketing campaigns can be categorized by **target audience, channel, and timing**. This helps in understanding which strategies work best for different customer groups.

### **4. Analyze the Impact**

By comparing key metrics **before and after the campaign**, we can see how marketing efforts influenced customer retention and acquisition.

* Calculate retention, churn, and acquisition rates for each segment.
* Use statistical tests or predictive models to see if campaigns significantly impacted customer behavior.

### **5. Additional Information Needed**

To conduct a thorough analysis, we might need:

* Detailed campaign objectives, messaging, and creative assets.
* Customer feedback or survey responses on campaign effectiveness.
* External factors like economic trends or competitor activities that may affect retention and acquisition.

This approach helps businesses optimize their marketing strategies, improve customer engagement, and drive better overall performance.

**Question 7. Customer Exit Reasons Exploration: Can you identify common characteristics or trends among customers who have exited that could explain their reasons for leaving?**

**Solution:-** To understand why customers leave, we can analyze their behavior, compare them with active customers, and identify common patterns.

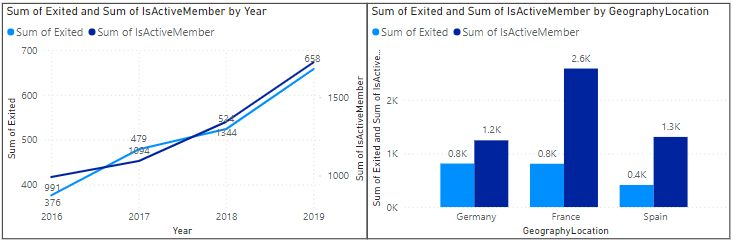
### **Key Steps:**

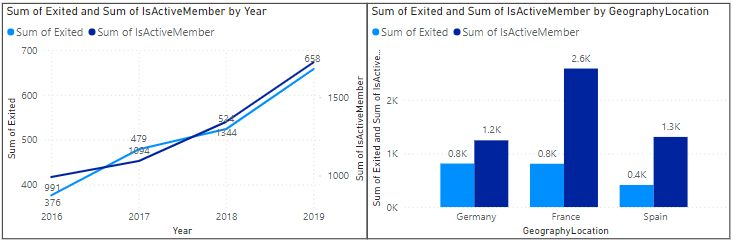
* Use **bar charts** to show exit reasons and **line charts** to track behavior changes before leaving.
* Compare **average behavior metrics** of exiting vs. active customers over time.
* Use **hypothesis testing** or **predictive models** to find key churn factors.

### **Common Reasons for Exit:**

1. Not meeting customer needs.
2. Outdated or slow processes.
3. Poorly integrated banking services.
4. Resistance to change.
5. Bad customer service.

Understanding these trends can help improve retention and customer satisfaction.



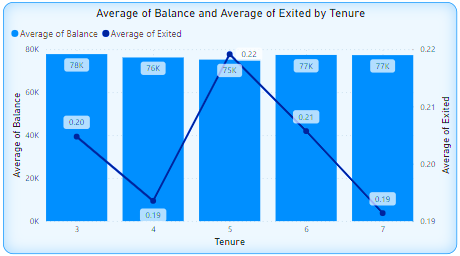


**Question 8. Are 'Tenure', 'NumOfProducts', 'IsActiveMember', and 'EstimatedSalary' important for predicting if a customer will leave the bank?**

**Solution:-**  To predict whether a customer will leave the bank, certain factors play a key role. Let’s evaluate the importance of each:

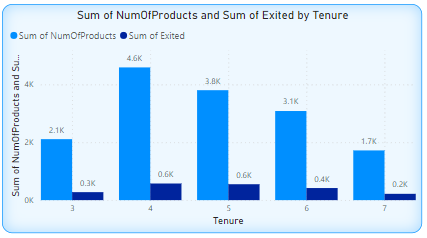
### **Tenure:**

This represents how long a customer has been with the bank. Customers with longer tenure are usually less likely to leave, as they have an established relationship with the bank. Therefore, tenure is an important factor in predicting churn.



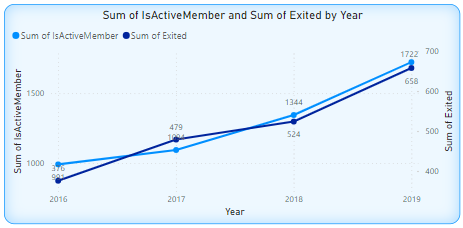
### **Number of Products:**

This shows how many products a customer uses with the bank. Generally, customers with multiple products are more engaged and less likely to leave. However, churn risk may still depend on product satisfaction and overall customer experience.



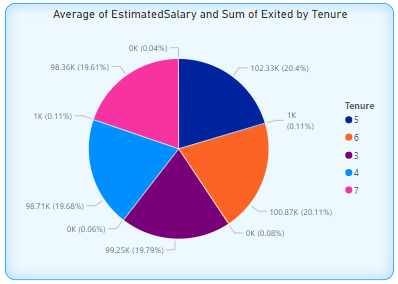
### **Is Active Member:**

This indicates whether a customer actively engages with the bank's services. Active members are typically more involved and loyal, making them less likely to leave compared to inactive customers.



### **Estimated Salary:**

A customer's estimated salary can offer insights into their financial stability. Higher-income customers may be less likely to leave, as they might benefit from premium banking services and find it more convenient to stay.



### **Summary:**

All these factors can play a role in predicting customer churn, but their actual impact depends on the dataset and modeling approach used. Performing data analysis and feature importance evaluation is crucial to identifying the most significant predictors for churn in your specific case.

**Question 9. Utilize SQL queries to segment customers based on demographics and account details.**

**Solution:-**

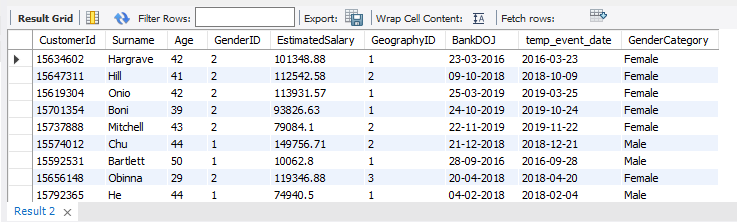
**Gender Segmentation:-**

Begin by analyzing the distribution of customers by gender. This involves linking the **customerinfo** table with the **gender** table using **GenderID**. Examine differences in behavior or preferences across gender categories to identify any notable trends.

SELECT c.\*, g.GenderCategory

FROM customerinfo c

JOIN gender g ON c.GenderID = g.GenderID;



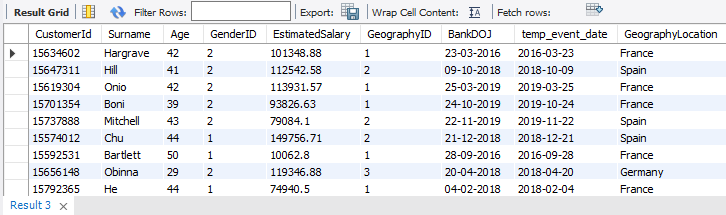
### **Geographic Segmentation:**

Analyze customer locations by linking the **customerinfo** table with the **geography** table using **GeographyID**. Identify regional trends or behavioral differences, such as variations in income, cultural preferences, or responses to marketing campaigns.

SELECT c.\*, geo.GeographyLocation

FROM customerinfo c

JOIN geography geo ON c.GeographyID = geo.GeographyID;



### **Age Group Segmentation:**

Divide customers into different age ranges to understand how their needs and behaviors vary. This can be done by analyzing the **customerinfo** table. Look at the number of customers in each age group and identify patterns. Younger customers might prefer digital banking and online promotions, while older customers may lean towards traditional banking services.

SELECT

CASE

WHEN Age < 18 THEN 'Under 18'

WHEN Age BETWEEN 18 AND 24 THEN '18-24'

WHEN Age BETWEEN 25 AND 34 THEN '25-34'

WHEN Age BETWEEN 35 AND 44 THEN '35-44'

WHEN Age BETWEEN 45 AND 54 THEN '45-54'

WHEN Age BETWEEN 55 AND 64 THEN '55-64'

ELSE '65+'

END AS AgeGroup

COUNT(\*) AS CustomerCount

FROM customerinfo

GROUP BY

CASE

WHEN Age < 18 THEN 'Under 18'

WHEN Age BETWEEN 18 AND 24 THEN '18-24'

WHEN Age BETWEEN 25 AND 34 THEN '25-34'

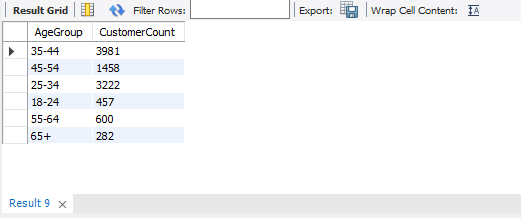
WHEN Age BETWEEN 35 AND 44 THEN '35-44'

WHEN Age BETWEEN 45 AND 54 THEN '45-54'

WHEN Age BETWEEN 55 AND 64 THEN '55-64'

ELSE '65+'

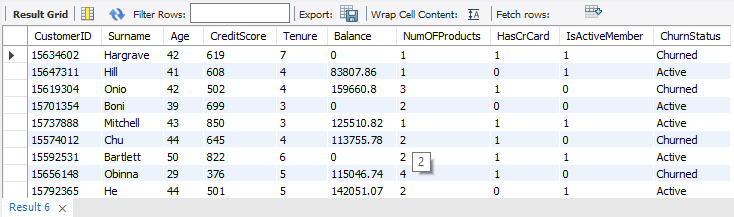
END;



### **Account Details and Churn Analysis:**

Review customer account details and churn status to identify trends in customer retention and attrition. Join the **customerinfo** and **bank\_churn** tables using **CustomerID** to analyze factors like credit score, tenure, balance, product usage, and membership status. Use the **Exited** column to classify customers as active or churned. This segmentation helps in understanding key reasons behind churn, optimizing marketing strategies, and improving customer retention efforts. Regular updates to this analysis ensure it stays relevant as customer behaviors change.

SELECT c.CustomerID, c.Surname,c.Age, b.CreditScore, b.Tenure, b.Balance, b.NumOFProducts, b.HasCrCard, b.IsActiveMember,CASE WHEN b.Exited = 1 THEN 'Churned' ELSE 'Active END AS ChurnStatus FROM customerinfo c JOIN bank\_churn b ON c.CustomerID = b.CustomerID;



### **Average Balance In Different Age Groups:-**

To calculate the average balance for different age groups, join the customerinfo table with the bank\_churn table using the Customer\_ID column. Then, group the data by age categories using a GROUP BY clause and compute the average balance for each age group. A CASE statement is used to classify customers into specific age brackets based on their age.SELECT

CASE

WHEN Age < 18 THEN 'Under 18'

WHEN Age BETWEEN 18 AND 24 THEN '18-24'

WHEN Age BETWEEN 25 AND 34 THEN '25-34'

WHEN Age BETWEEN 35 AND 44 THEN '35-44'

WHEN Age BETWEEN 45 AND 54 THEN '45-54'

WHEN Age BETWEEN 55 AND 64 THEN '55-64'

ELSE '65+'

END AS AgeGroup

Avg(Balance) AS AverageBalance

FROM customerinfo ci

JOIN bank\_churn bc ON ci.CustomerID = bc.CustomerID

GROUP BY CASE

WHEN Age < 18 THEN 'Under 18'

WHEN Age BETWEEN 18 AND 24 THEN '18-24'

WHEN Age BETWEEN 25 AND 34 THEN '25-34'

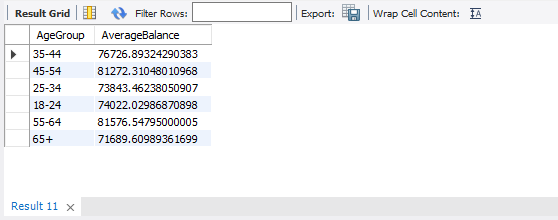
WHEN Age BETWEEN 35 AND 44 THEN '35-44'

WHEN Age BETWEEN 45 AND 54 THEN '45-54'

WHEN Age BETWEEN 55 AND 64 THEN '55-64'

ELSE '65+'

END;



**Question 10. 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?**

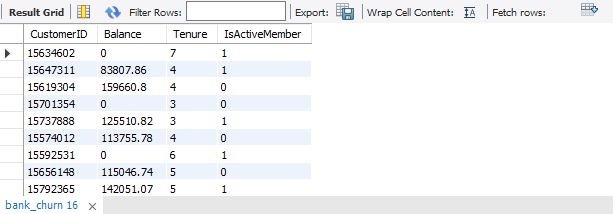
**Solution:-**

**Identifying Customers at Risk of Churn:-** You can identify customers who are likely to leave the bank by setting specific criteria, such as low account balances, inactive memberships, or low engagement levels. The following query selects customers who have already exited (Exited = 1) and either have an account balance below 1000 or are inactive members. This helps in recognizing patterns that contribute to customer churn.:

SELECT CustomerID, Balance, Tenure, IsActiveMember

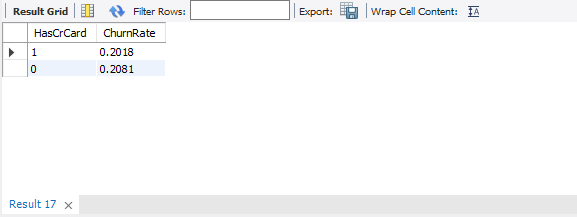
FROM bank\_churn

WHERE Balance < 1000 OR Tenure < 6 OR IsActiveMember = 0;



**Evaluating the Impact of Credit Card Rewards on Customer Retention:-** To measure how credit card rewards affect customer retention, you can compare churn rates between those who have a credit card and those who don’t. This query calculates the average churn rate (Exited) for both groups based on the HasCrCard column. Analyzing this data will help determine whether having a credit card influences a customer's likelihood of staying with the bank.

Select HasCrCard, avg(Exited) as ChurnRate from bank\_churn group by HasCrCard;

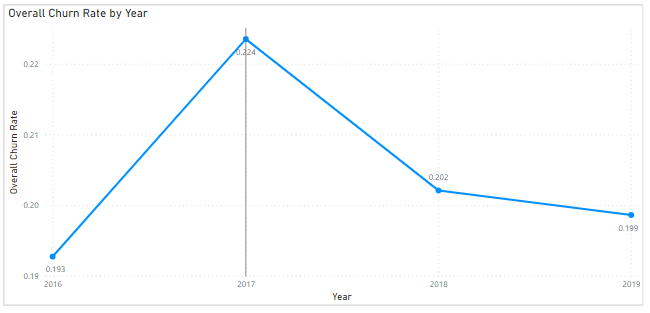


**Question 11. What is the current churn rate per year and overall as well in the bank. Can you suggest some insights to the bank about which kind of customers are more likely to churn and what are the different strategies that can be used to decrease the churn rate.**

To find the churn rate per year first you create a measure that is overall chaurn rate using the DAX fuctions and put that value on a card.

Overall Churn Rate = DIVIDE(COUNTROWS(FILTER(bank\_churn, bank\_churn[Exited] = 1)), COUNTROWS(bank\_churn))



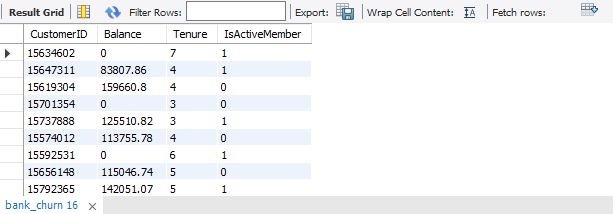
And then use the line chart to visualize churn rate per year. That will show the in 2017 the value of churn is maximum.

**Customer are more likely to churn:-** To spot customers at risk of leaving, you can look at key factors like low account balances, inactive memberships, or minimal engagement. This query identifies customers who have already churned (Exited = 1) and either have a balance below 1000 or are not active members:

SELECT CustomerID, Balance, Tenure, IsActiveMember

FROM bank\_churn

WHERE Balance < 1000 OR Tenure < 6 OR IsActiveMember = 0;



**Strategies that can be used to decrease the churn rate:-**

1. Improve Customer Experience
2. Personalization
3. Enhance Product Value
4. Customer Engagement
5. Retention Programs
6. Proactive Outreach
7. Customer Feedback And Satisfaction Surveys
8. Educational Resources
9. Exit Interviews
10. Continuous Monitoring and Analysis

**Question 12. Create a dashboard incorporating all the KPIs and visualization related metrics. Use a slicer in order to assist in selection in the dashboard.**

### **Data Preparation and Dashboard Features in Power BI:-**

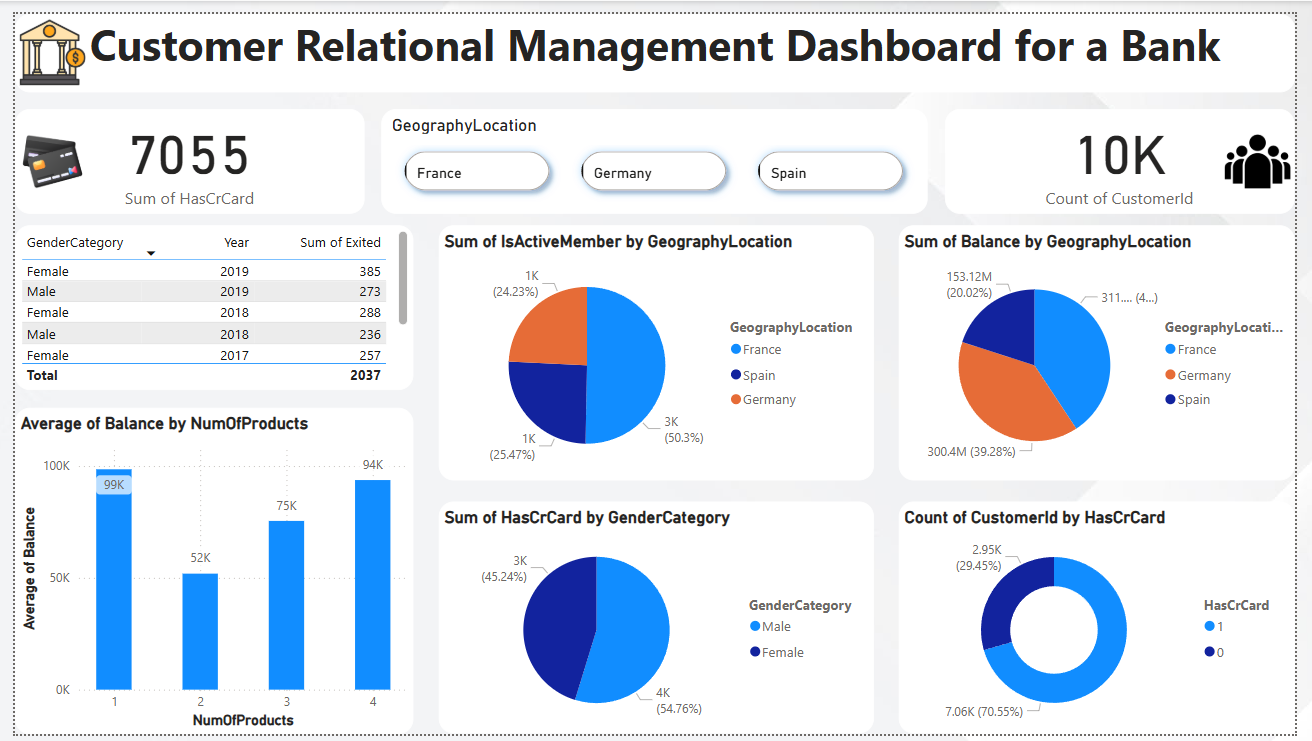
Microsoft Power BI allows users to connect to various data sources, including Excel, CSV, MySQL, SQL Server, Oracle, and SAP Business Warehouse. It enables the creation of interactive reports with visually rich graphics that can be shared via websites or blogs.

#### **Customizable Dashboards:-**

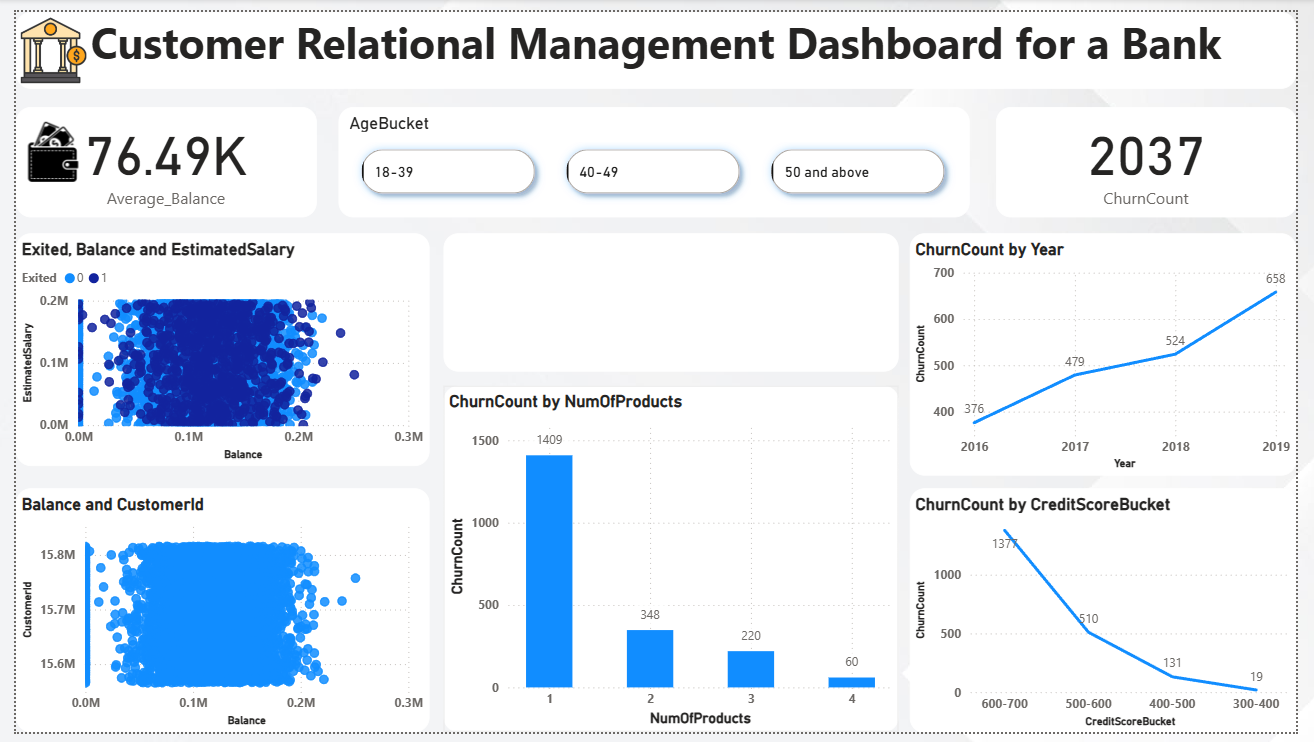
Power BI provides the flexibility to design dashboards tailored to specific needs. A Power BI dashboard is a single-page visual representation of key insights, summarizing crucial information in a concise manner.

For our dashboard, we will create a three-page layout:

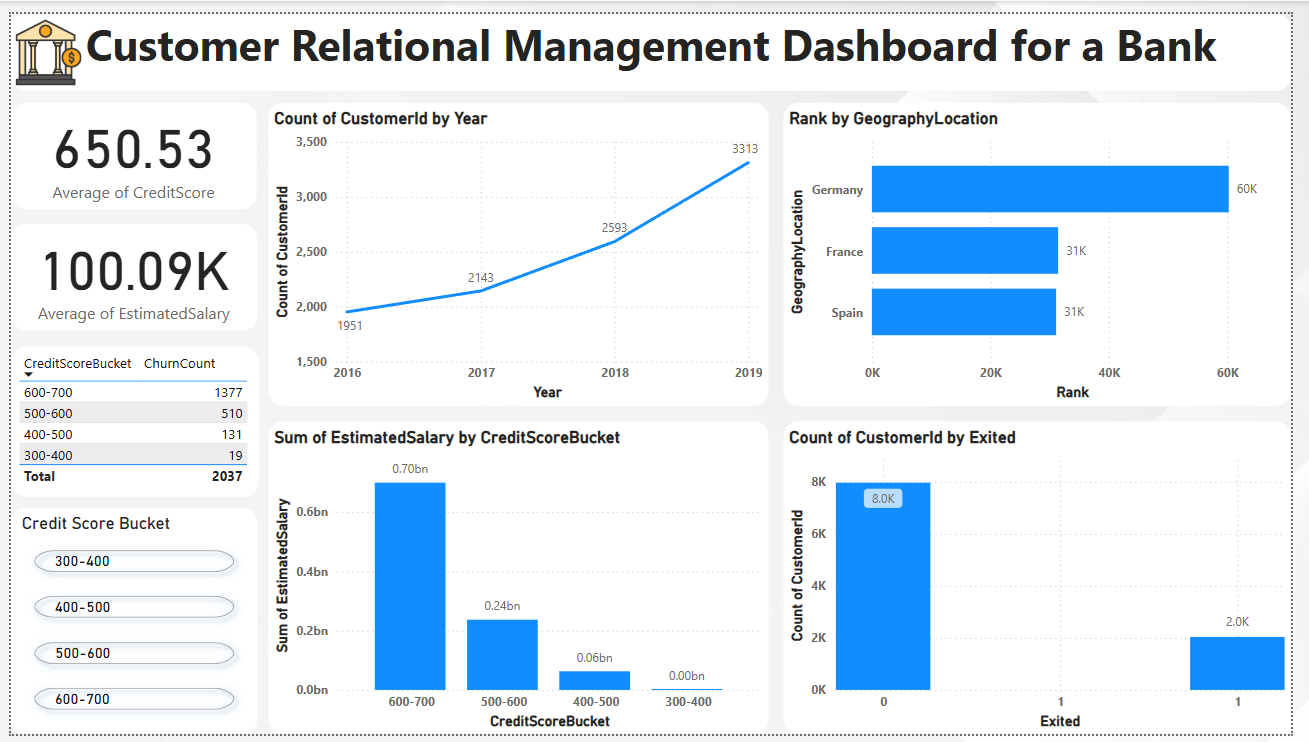
* **Page 1**: Includes a **Stacked Column Chart, Donut Chart, Pie Chart, Table, Card, and a Slicer** to effectively display key metrics and trends.
* **Page 2 & 3**: Can focus on deeper analysis, filtering options, and customer segmentation insights.



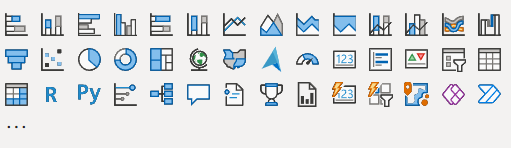
On **Page 2**, we will include a **Stacked Column Chart** to analyze **Churn Count by Number of Products**. Two **Scatter Charts** will be used one to examine **Customer Balances** and another to show the relationship between **Exited Customers, Balance, and Estimated Salary**. A **Line Chart** will track **Churn Count by Year**, while a **Table** will categorize **Age Buckets for Credit Card Ownership**. Additionally, a **Card** will display the **Average Balance**, and a **Slicer** will allow dynamic filtering for better insights.



On Page3, we will include a BarChart to display RankbyGeographyLocation. A StackedAreaChart will show the SumofEstimatedSalarybyCreditScoreBucket, while another StackedAreaChart will track the CountofExitedCustomers. A LineChart will visualize the CustomerCountbyYear, and a Table will organize CreditScoreBucketsandChurnCount. Additionally, a Card will highlight the Rank, and a Slicer will enable dynamic filtering for better analysis.



**Rich graphical Visualizations:-** Power BI offers a variety of visual elements for building and editing reports. These visuals are accessible in the **Visualizations** pane. When using **Power BI Desktop** or the **Power BI Service (app.powerbi.com)**, a default set of pre-installed visuals is readily available for use.

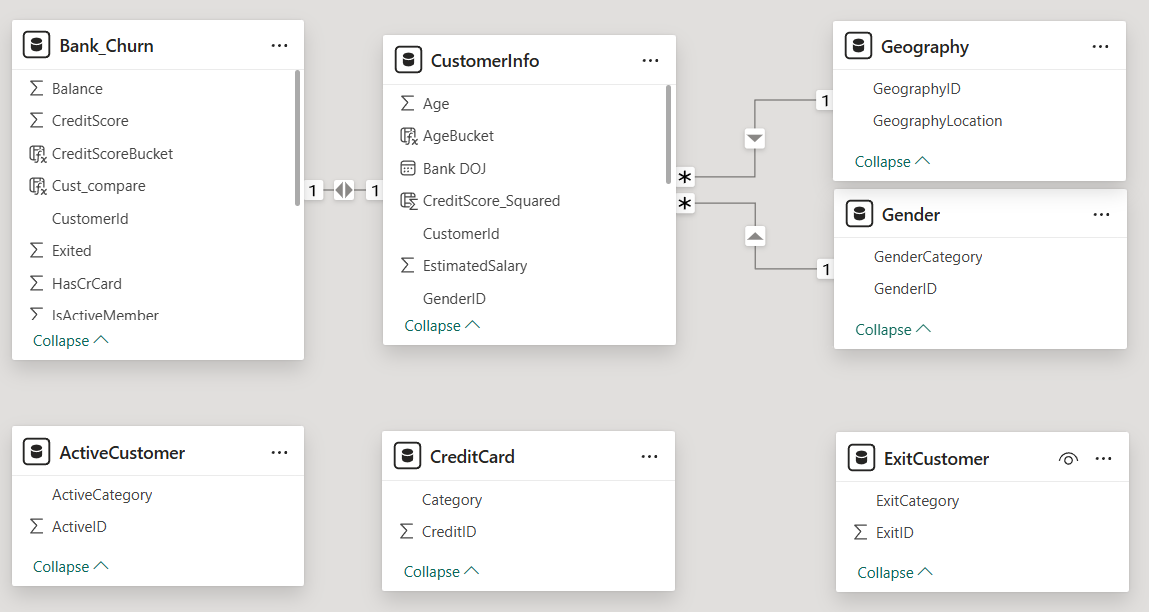


However, we are not restricted to these built-in visuals. Clicking the ellipses in the **Visualizations** pane provides access to additional report visuals, including **custom visuals**. These are developed using the **Custom Visuals SDK**, allowing businesses to tailor data representation to their specific needs.

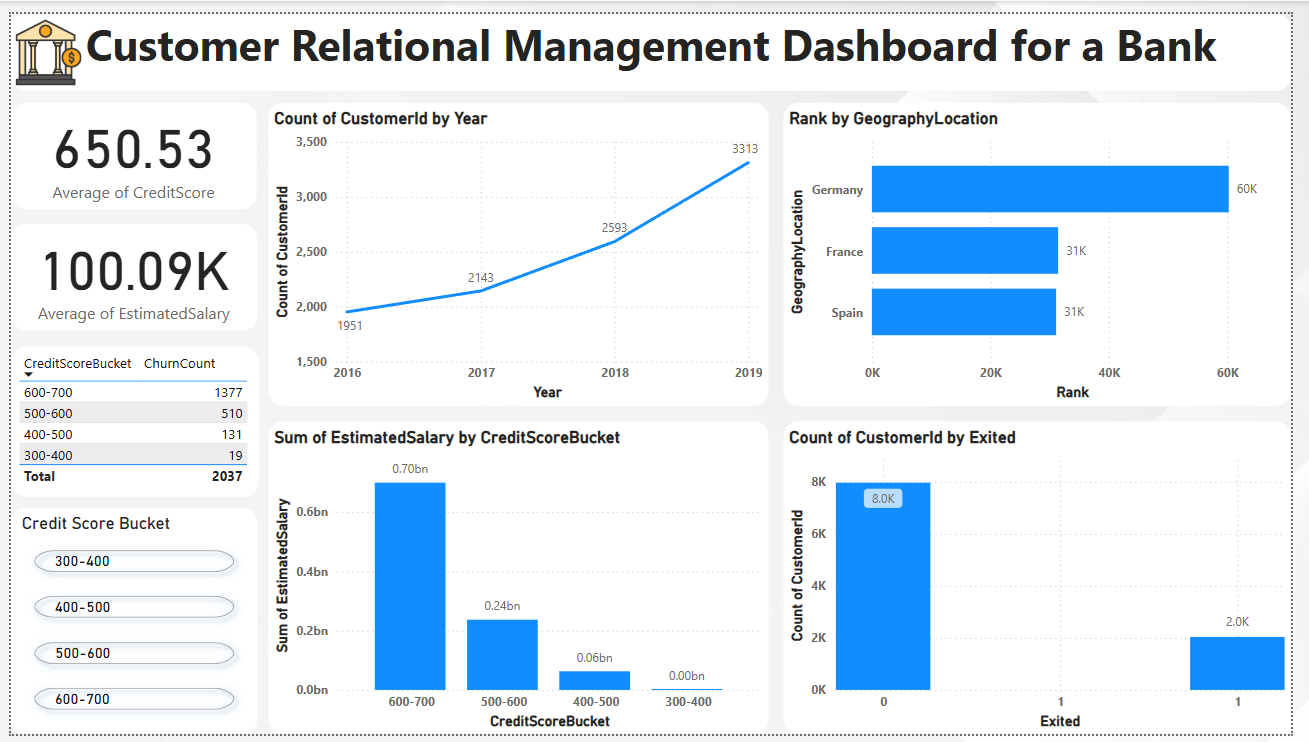
**Question 13. How would you approach this problem, if the objective and subjective questions weren't given?**

**Solution:-** Without defined objectives or questions, a broad approach to analyzing and utilizing this database schema can follow these steps:

1. **Schema Analysis** – Examine the structure of each table, identifying column names, data types, and relationships to understand how the database is organized.



1. **Identifying Relationships** – Determine how tables are connected, such as foreign keys linking records across tables. For example, fields like CustomerID, GenderID, and GeographyID likely reference corresponding records in the **customerinfo** table. The schema diagram with arrows typically illustrates these relationships.
2. **Data Exploration** – Analyze the contents of each table to understand the stored information and its structure. This could involve running sample queries or checking data distributions to identify trends and inconsistencies.



1. **Querying and Analysis** – Perform queries to extract insights, such as calculating summary statistics, joining tables for a comprehensive view, or filtering data based on specific conditions.
2. **Data Integrity Check** – Validate the accuracy and consistency of the data by checking for anomalies, verifying foreign key constraints, ensuring correct data types, and identifying duplicate or missing records.
3. **Documentation and Collaboration** – Maintain clear documentation of the database schema, assumptions, and analyses performed. This aids in understanding the structure and supports teamwork.
4. **Security Measures** – Implement necessary security protocols to protect sensitive information and prevent unauthorized access.

These steps provide a structured approach to exploring and analyzing the database. Additional actions may be required depending on specific goals and use cases.

**Question 14. In the “Bank\_Churn” table how can you modify the name of “HasCrCard” column to “Has\_creditcard”?**

To rename the column **"**HasCrCard**"** to **"**Has**\_**creditcard**"** in the **"**Bank**\_**Churn**"** table, you would generally use the ALTER TABLE statement in SQL. Here’s how you can achieve this:

ALTER TABLE Bank\_Churn

RENAME COLUMN HasCrCard TO Has\_creditcard;

