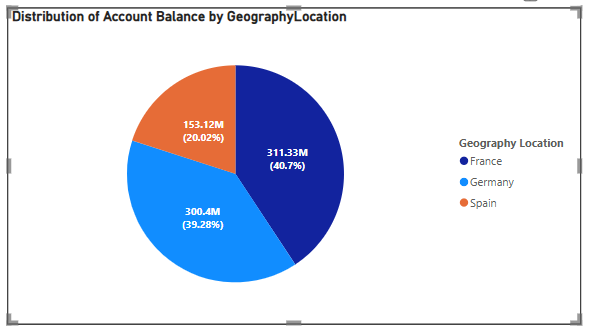
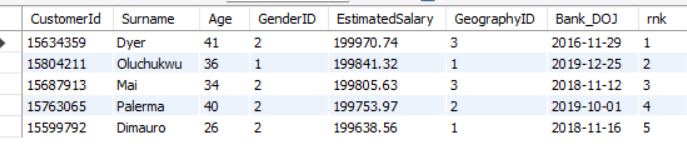
**Bank CRM Analysis**

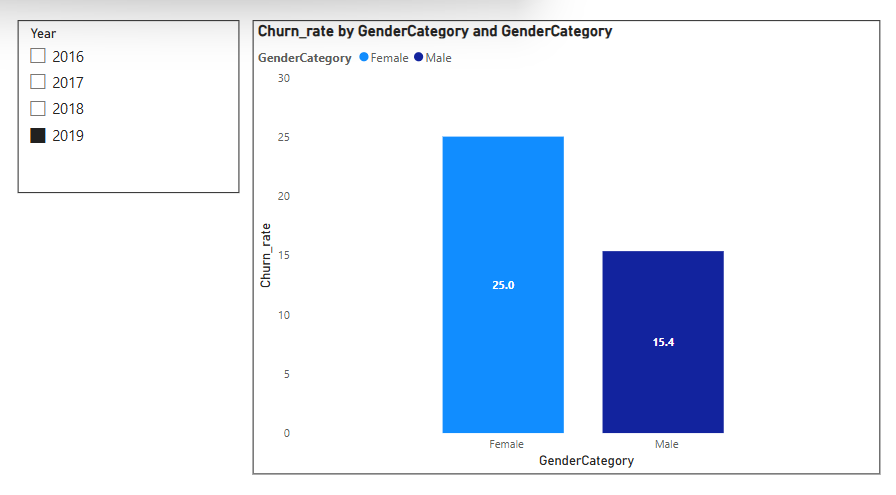
**Objective Questions:**

1. **What is the distribution of account balances across different regions?  
     
   Power BI:** There are currently three regions in the dataset: France, Germany, and Spain. To determine the distribution of account balances, we calculate the Account balances and group them by the regions (Geography).

The distribution of account balance by geography shows that France holds the largest share with 311.33M (40.7%), followed closely by Germany with 300.4M (39.28%), and Spain holds 153.12M (20.02%).

1. **Identify the top 5 customers with the highest Estimated Salary in the last quarter of the year. (SQL)**This query can be used for identifying the customer id and salary of the top 5 customers with the highest estimated salary in the last quarter of the year.

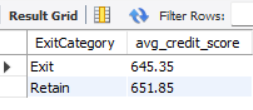
**SQL Query:**  
select \* from(  
select \*,  
dense\_rank() over (order by EstimatedSalary desc) as rnk  
from customerinfo  
where quarter(Bank\_DOJ)=4) a  
where rnk<6;  
  
**Result:**  


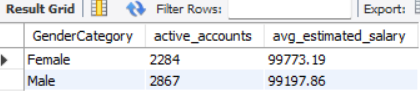
1. **Calculate the average number of products used by customers who have a credit card. (SQL)**  
   **SQL Query:**select   
   floor(avg(NumOfProducts)) as avg\_products   
   from   
   bank\_churn  
   where HasCrCard=1;  
     
   **Result:**
2. **Determine the churn rate by gender for the most recent year in the dataset.  
   Power BI:**The below DAX functions are used to determine the chure rate by gender for the most recent year i.e. 2019. **Churned\_Customers =** CALCULATE(COUNT(Bank\_Churn[CustomerId]),Bank\_Churn[Exited]=1)  
    **Churn\_rate** = DIVIDE([Churned\_Customers],[Total Customers])\*100  
    ****  
   **Conclusion:**

Churn rate for Males in 2019 is 15.4

Churn rate for Females in 2019 is 25.0

1. **Compare the average credit score of customers who have exited and those who remain. (SQL)**  
     
   **SQL Query:**select   
   ExitCategory,  
   round(avg(CreditScore),2) as avg\_credit\_score  
   from   
   bank\_churn bc  
   join   
   exit\_customer e  
   on   
   e.ExitID=bc.Exited  
   group by ExitCategory;  
     
     
   **Result:**



1. **Which gender has a higher average estimated salary, and how does it relate to the number of active accounts? (SQL**)  
   Answer: The query first finds the number of active accounts and average estimated salary for each Gender category.  
     
   **SQL Query:**select  
   g.GenderCategory,   
   count(c.CustomerId) as active\_accounts,  
   round(avg(c.EstimatedSalary),2) as avg\_estimated\_salary  
   from  
   customerinfo c  
   join gender g on g.GenderID=c.GenderID  
   join bank\_churn bc on bc.CustomerId=c.CustomerId  
   where bc.IsActiveMember=1  
   group by g.GenderCategory  
   order by avg\_estimated\_salary desc;  
     
   **Result:**

**Conclusion:** Male have higher active accounts and the average salary of the females who have active accounts is slightly higher than the males.

1. **Segment the customers based on their credit score and identify the segment with the highest exit rate. (SQL)**  
     
   Segment customers by credit score ranges using SQL, computed exit rates for each segment, and identify the segment with the highest exit rate by ordering results in descending order.

**SQL Query:**SELECT

CASE

WHEN bc.CreditScore BETWEEN 300 AND 649 THEN 'Low'

WHEN bc.CreditScore BETWEEN 650 AND 749 THEN 'Medium'

WHEN bc.CreditScore BETWEEN 750 AND 849 THEN 'High'

ELSE 'Excellent'

END AS credit\_segment,

COUNT(CASE WHEN bc.Exited = 1 THEN c.CustomerId END) AS exited\_customers,

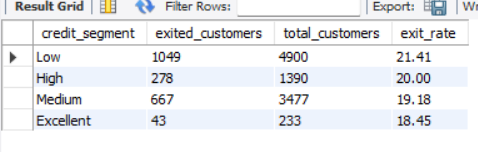
COUNT(c.CustomerId) AS total\_customers,

ROUND((COUNT(CASE WHEN bc.Exited = 1 THEN c.CustomerId END) \* 100.0) / COUNT(c.CustomerId), 2) AS exit\_rate

FROM customerinfo c

JOIN bank\_churn bc ON bc.CustomerId = c.CustomerId

GROUP BY credit\_segment

ORDER BY exit\_rate DESC;  
  
**Result:** **Conclusion:** The segment with Low Credit Score has highest exit rate as compared to other credit segment.

1. **Find out which geographic region has the highest number of active customers with a tenure greater than 5 years. (SQL)**  
   **SQL Query:**  
   select GeographyLocation, count(c.CustomerId) as active\_members

from

customerinfo c

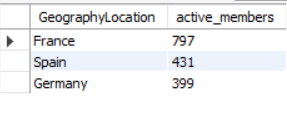
join bank\_churn bc on c.CustomerId=bc.CustomerId

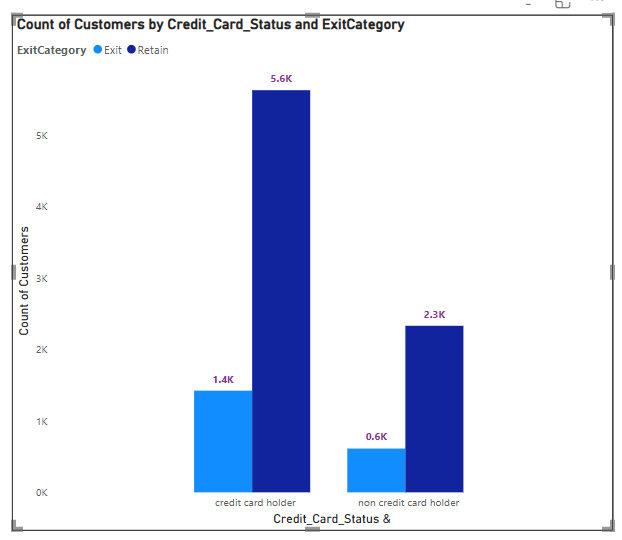
join geography g on g.GeographyID=c.GeographyID

where IsActiveMember=1 and

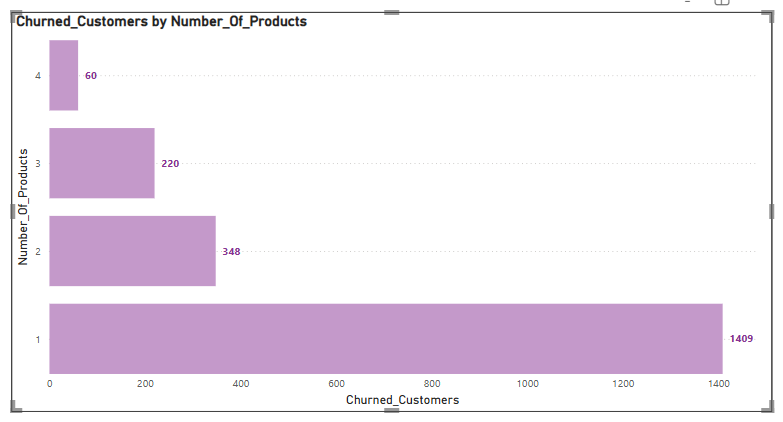
bc.Tenure> 5

group by GeographyLocation

order by active\_members desc;  
  
**Result:** **Conclusion:** France has highest active customers among the 3 region.

1. **What is the impact of having a credit card on customer churn, based on the available data?**  
    **Power BI:**   
   **  
     
   Insights:** It appears that more customers with a credit card (HasCrCard = 1) have churned compared to those without a credit card.

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

1. **For customers who have exited, what is the most common number of products they have used?**The below chart provides the insights of customers who have exited by number of products.  
     
     
     
   **Conclusion:** The customers who use 3 or 4 products are less likely to churn.

The customer who uses just 1 product are the customers who have churned the most. So, focusing on those customer base can help to decrease the churn rate.

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.**  
   **SQL Query:**select

year(bank\_doj) as year\_of\_joining,

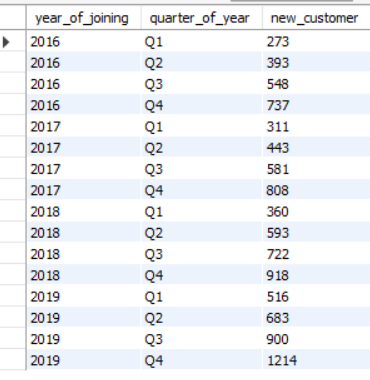
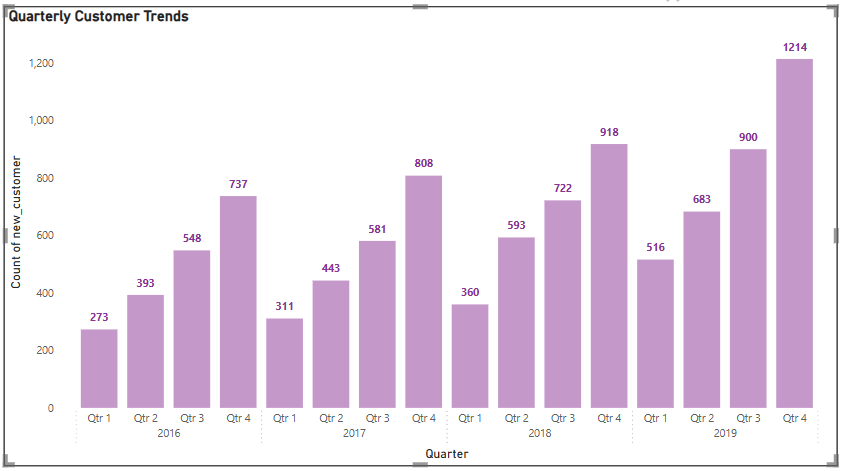
CONCAT('Q', QUARTER(Bank\_DOJ)) AS quarter\_of\_year,

count(\*) as new\_customer

from

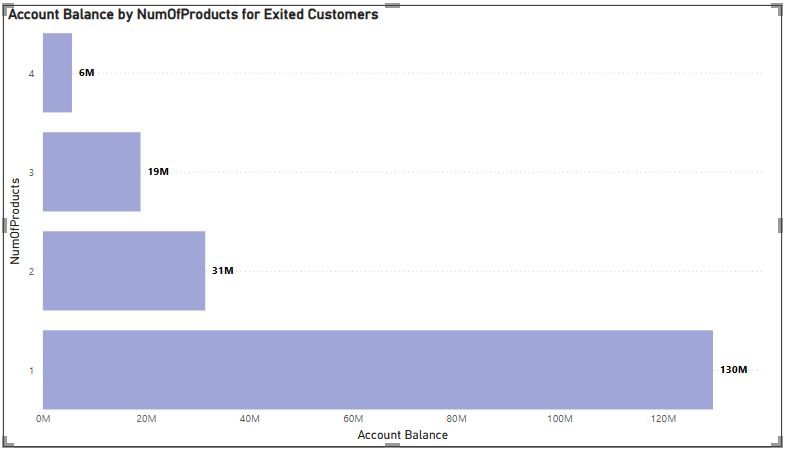
customerinfo

group by year\_of\_joining,quarter\_of\_year

order by year\_of\_joining,quarter\_of\_year;  
   
 **Result:**  
   
   
  
  
**Insights:**

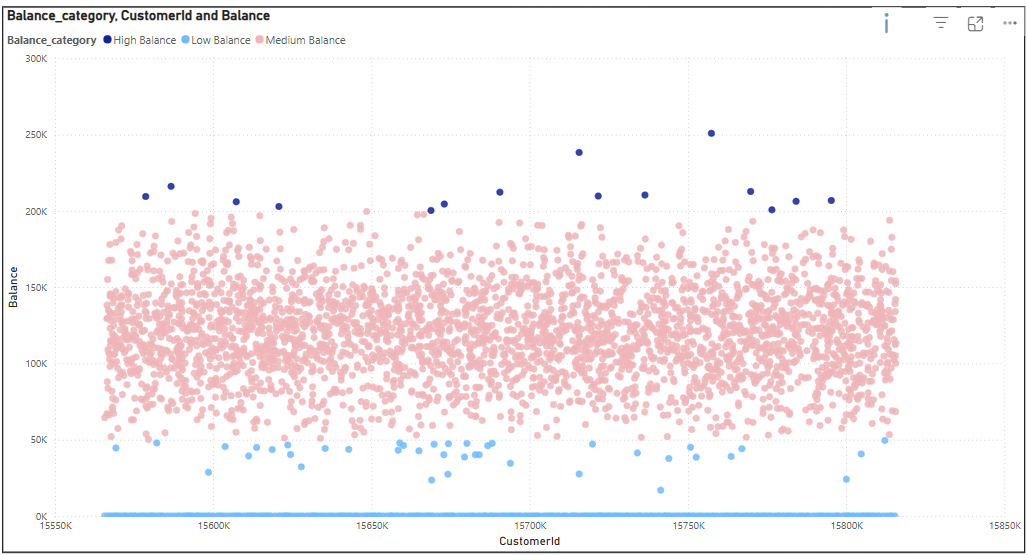
1. New customer acquisition shows a consistent year-on-year increase from 2016 to 2019.
2. Q4 is the strongest quarter every year, with the highest number of new customers.
3. 2019 saw the highest growth, ending with 1,214 new customers in Q4.
4. Each year follows an upward quarterly trend, starting lower in Q1 and peaking in Q4.  
     
   **Conclusion:** The number of customers has increased per year in the bank that is a good sign. Q4 consistently outperforms other quarters, suggesting effective year-end campaigns or seasonal banking behaviour.

It is showing customers satisfaction with the bank services and products. From the data above, we could say that it is almost double in the 4 years of span.

1. **Analyse the relationship between the number of products and the account balance for customers who have exited.**  
     
     
   **Conclusion:** 
   1. The most common number of products used by exiting customers is 1. This suggests that the customers who use least number of products are more likely to churn.
   2. As the number of products used increases. This suggests that customers who churned tend to have fewer products compared to active customers.
   3. The customers who have used 4 products have less balance.
2. **Identify any potential outliers in terms of balance among customers who have remained with the bank.**To identify potential outliers in **Balance** among customers who have **remained** (Retain) used the scatter plot showing **CustomerId vs. Balance** and categorized by Balance\_category:

* **High Balance (dark blue dots):**   
  These are the key outliers. They are customers with balances **greater than 200,000** and they are **visibly fewer** in number compared to medium and low balance categories. These customers stand out above the dense middle cluster and are spaced at the top of the chart — likely in the **220,000–300,000** range.
* **Low Balance (light blue dots):**   
  A few customers have **very low or zero balance**, which could also be considered outliers at the bottom. However, there are many low-balance customers, so only those with **zero** or **close to zero** might be considered outliers.
* **Medium Balance (pink dots):**   
  This is the **densest group**, ranging from **50,001 to 200,000**, and doesn't exhibit obvious outliers since it's the dominant cluster.

### **Conclusion:**

* **High-balance customers (Balance > 200,000)** are the most significant outliers.
* **Zero-balance customers** may also be considered outliers at the lower end.  
  

1. **How many different tables are given in the dataset, out of these tables which table only consists of categorical variables?**  
   **Answer:** There are 7 tables are given in the dataset. They are: Active Customer, Bank Churn, Credit Card, Customer Info, Exit Customer, Gender and Geography.  
     
   Among all the tables the below mentioned tables consists of categorical variables:

* Active Customer: Contains categorical variables like Active Category.
* Credit Card: Contains categorical variables like Category.
* Exit customer: Contains categorical variables like Exit Category.
* Gender: Contains categorical variables like Gender Category.
* Geography: Contains categorical variables like Geography Location.  
    
  **Conclusion:** There is no such table which contains only categorical variables.

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)**This SQL query calculates the average income for males and females within each geographic location and assigns rank based on the average salary.

**SQL Query:**select

GenderCategory,

gy.GeographyLocation,

round(avg(EstimatedSalary),2) as avg\_income,

dense\_rank() over

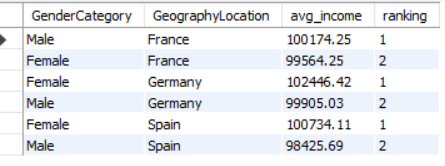
(partition by gy.GeographyLocation order by avg(EstimatedSalary) desc) as ranking

from customerinfo c

join gender g on c.GenderID=g.GenderID

join geography gy on gy.GeographyID=c.GeographyID

group by 1,2;

**Result:**  
Germany female customers has highest average income among all countries and Spain male customers has lowest average income among all countries.

1. **Using SQL, write a query to find out the average tenure of the people who have exited in each age bracket (18-30, 30-50, 50+).**  
   To find the average tenure for each age bracket we first segregate the user base into three segments.
2. People with age from 18-30 as adults.
3. People with age from 31-50 as Middle Aged.
4. People with age above 50 as Old Aged.

**SQL Query:**select

case

when c.age between 18 and 30 then 'Adults'

when c.age between 31 and 50 then 'Middle-Aged'

else 'Old-Aged'

end as age\_bracket,

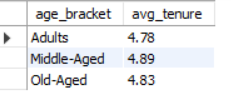
round(avg(Tenure),2) as avg\_tenure

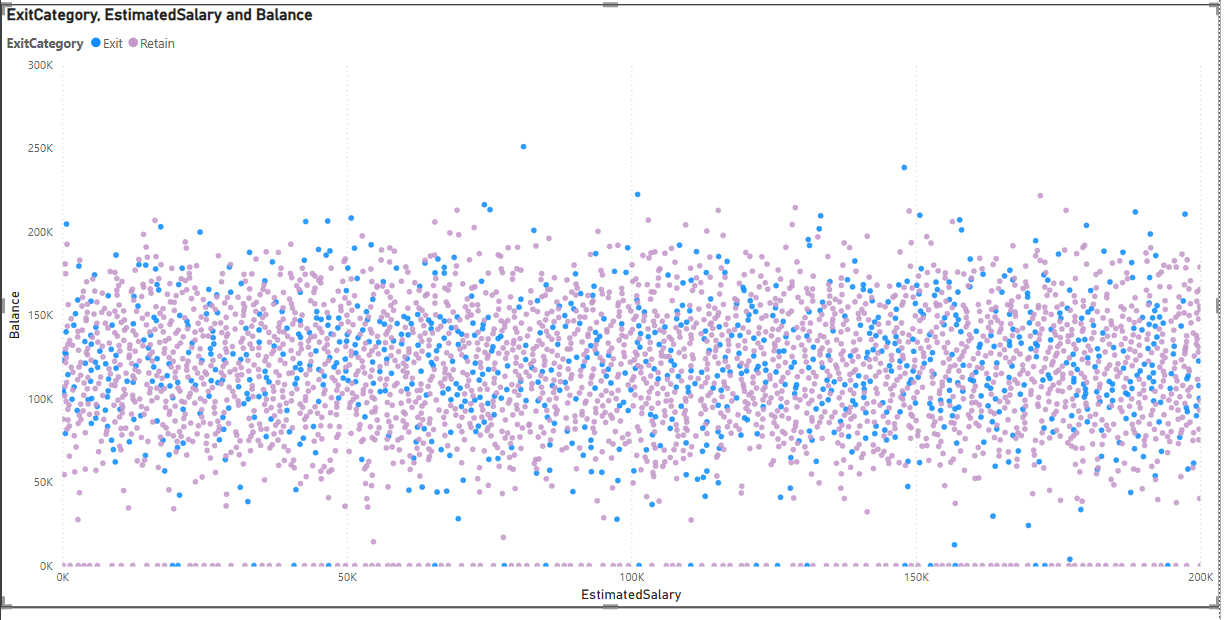
from customerinfo c

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

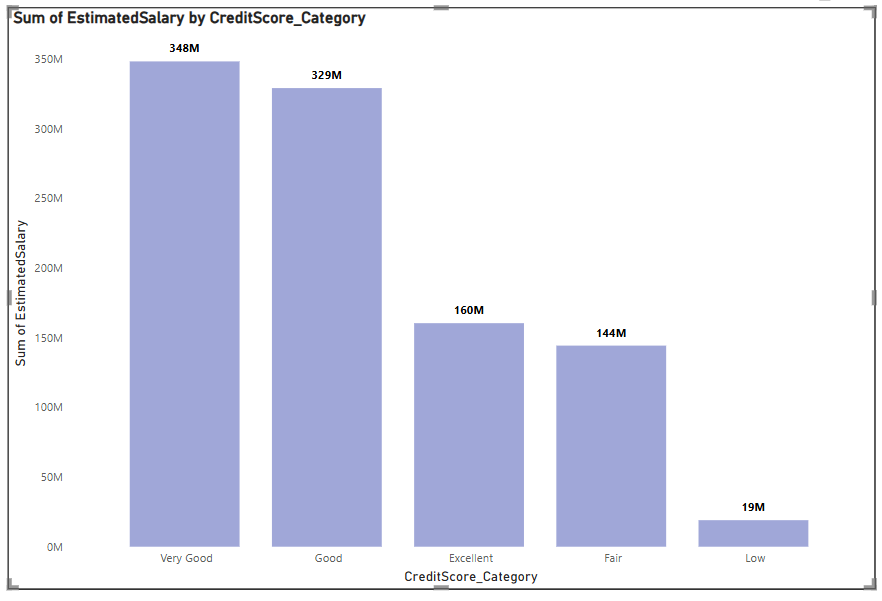
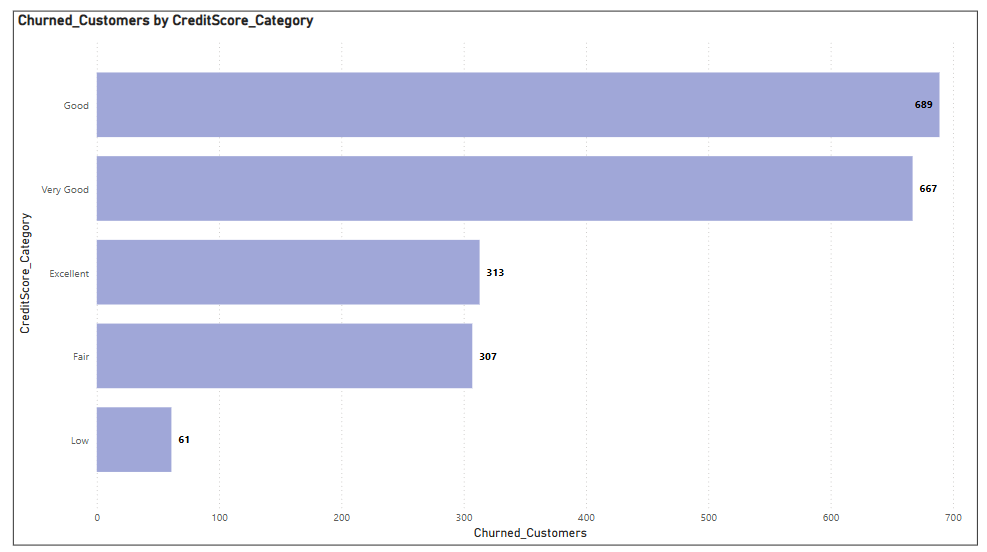
where Exited=1

group by age\_bracket

ORDER BY age\_bracket;  
  
**Result:** 

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 above Scatter plot represents **EstimatedSalary (X-Axis)** vs. **Balance (Y-Axis)**, with **ExitCategory** (Exited vs. Retained) shown in different colors.

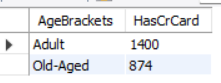
#### **Key Observations**

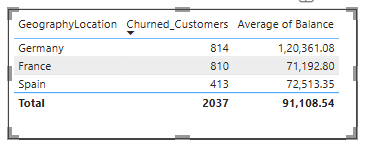
1. **No Clear Correlation**
   * The distribution of points appears **random** with no upward or downward trend.
   * This suggests **no strong correlation** between EstimatedSalary and Balance.
2. **Exited vs. Retained Customers**
   * **Blue (Exited)** and **Pink (Retained)** dots are spread across all salary and balance levels.
   * There is **no specific trend** that indicates that customers with higher/lower salaries have higher/lower balances when they exit.
3. **Low-Balance Customers**
   * Some customers have a **zero or very low balance** across different salary levels.
4. **High-Balance Outliers**
   * Some points are **above 200K balance**, which could be outliers.
5. **Is there any correlation between the salary and the Credit score of customers?**  
   Below DAX function is created to categorize the Credit Score of customers for the available data:  
   CreditScoreCategory = SWITCH(TRUE(),  
   Bank\_Churn[CreditScore]<= 450, "Low",  
   Bank\_Churn[CreditScore] <= 550, "Fair",  
   Bank\_Churn[CreditScore] <= 650, "Good",  
   Bank\_Churn[CreditScore] <= 750, "Very Good",  
   Bank\_Churn[CreditScore] > 750, "Excellent")  
     
     
   There appears to be a **positive correlation** between the estimated salary and the credit score of customers.
6. Higher credit score categories (like Very Good and Good) show significantly higher total estimated salaries.
7. Lower credit score categories (Fair and Low) contribute far less to the total salary.
8. This suggests that customers with better credit scores generally have higher income levels.
9. **Rank each bucket of credit score as per the number of customers who have churned the bank.**  
     
     
     
   Based on the above chart here is the ranking of credit score categories by number of churned customers (from highest to lowest):
10. Good — 689 customers
11. Very Good — 667 customers
12. Excellent — 313 customers
13. Fair — 307 customers
14. Low — 61 customers

So, Good and Very Good categories have the highest churn despite better credit scores, which is an interesting insight.

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.**  
   The below query provides the numbers of customers who have a credit card. We first segregate the customers based on the age into age brackets and then find those buckets that have lesser than average number of credit cards per bucket.  
     
   **SQL Query:**WITH creditinfo AS (

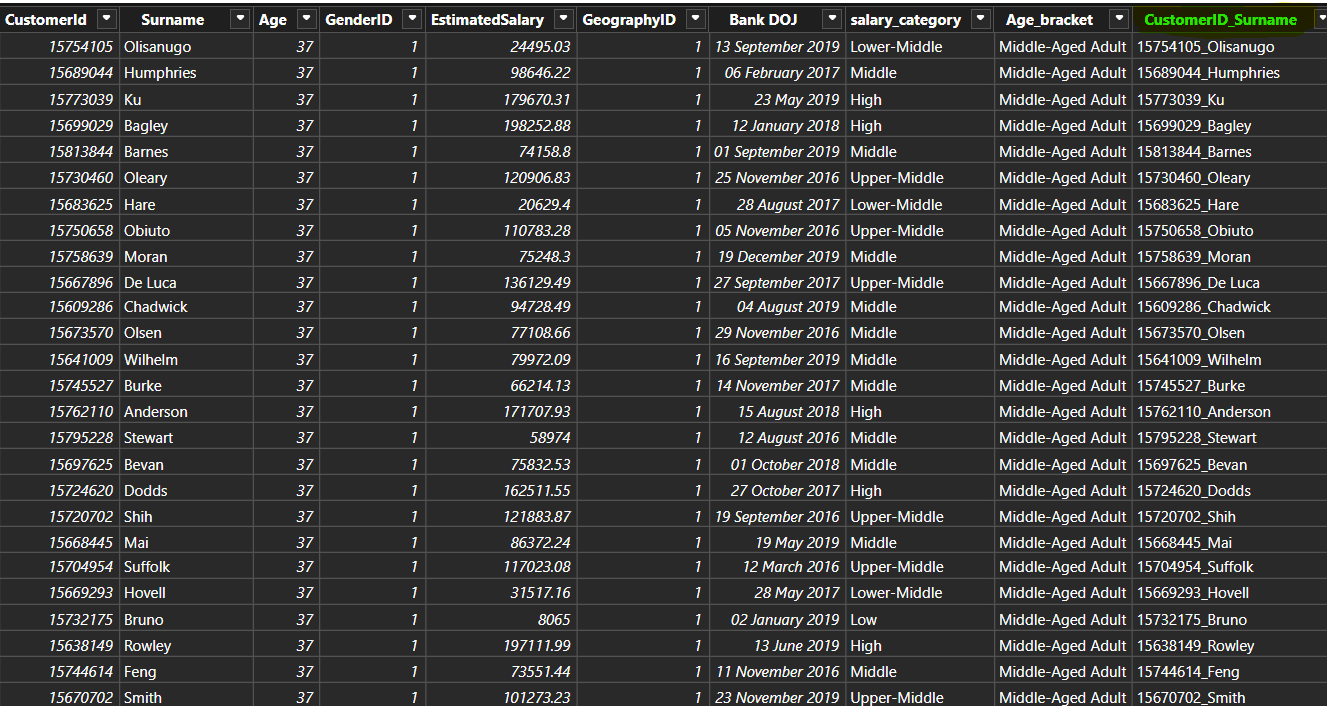
SELECT  
CASE   
WHEN age BETWEEN 18 AND 30 THEN 'Adult'  
WHEN age BETWEEN 31 AND 50 THEN 'Middle-Aged'  
ELSE 'Old-Aged'  
END AS AgeBrackets,  
COUNT(c.CustomerId) AS HasCrCard  
FROM customerinfo c  
JOIN bank\_churn b ON c.CustomerId = b.CustomerId  
WHERE b.Has\_creditcard = 1 -- Ensures filtering is done before counting  
GROUP BY AgeBrackets)

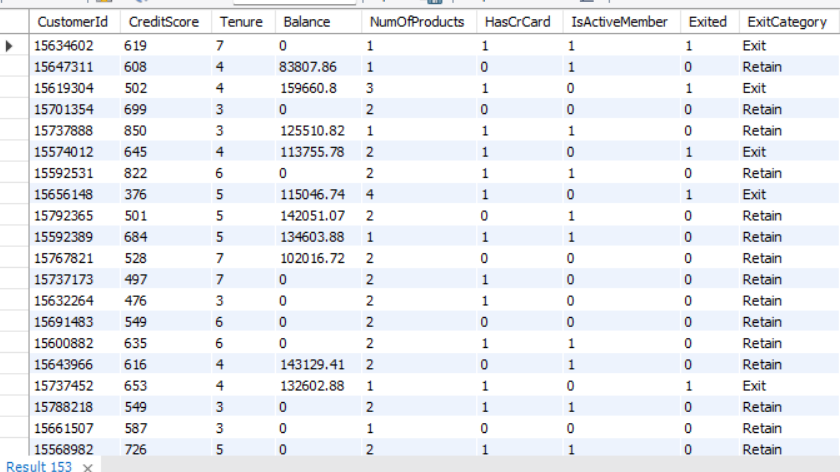
SELECT \* FROM creditinfo  
WHERE HasCrCard < (  
SELECT AVG(HasCrCard)   
FROM creditinfo);  
  
**Result:**

1. **Rank the Locations as per the number of people who have churned the bank and average balance of the customers.  
     
   Power BI:**  
   **  
   Conclusion:**
2. **Germany**: 814 customers exited, with an average balance of approximately 120361.08.
3. **France**: 810 customers exited, with an average balance of approximately 71192.80.
4. **Spain**: 413 customers exited, with an average balance of approximately 72513.35.
5. **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”.**Answer:  
   The column where the format is “CustomerID\_Surname” can be created in POWER BI using the formula:

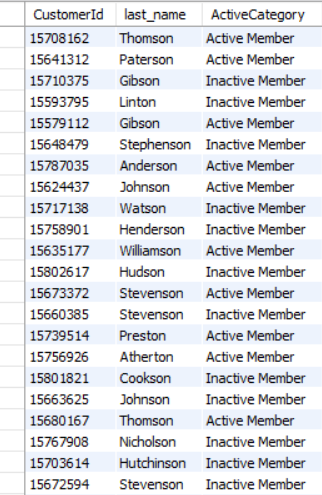
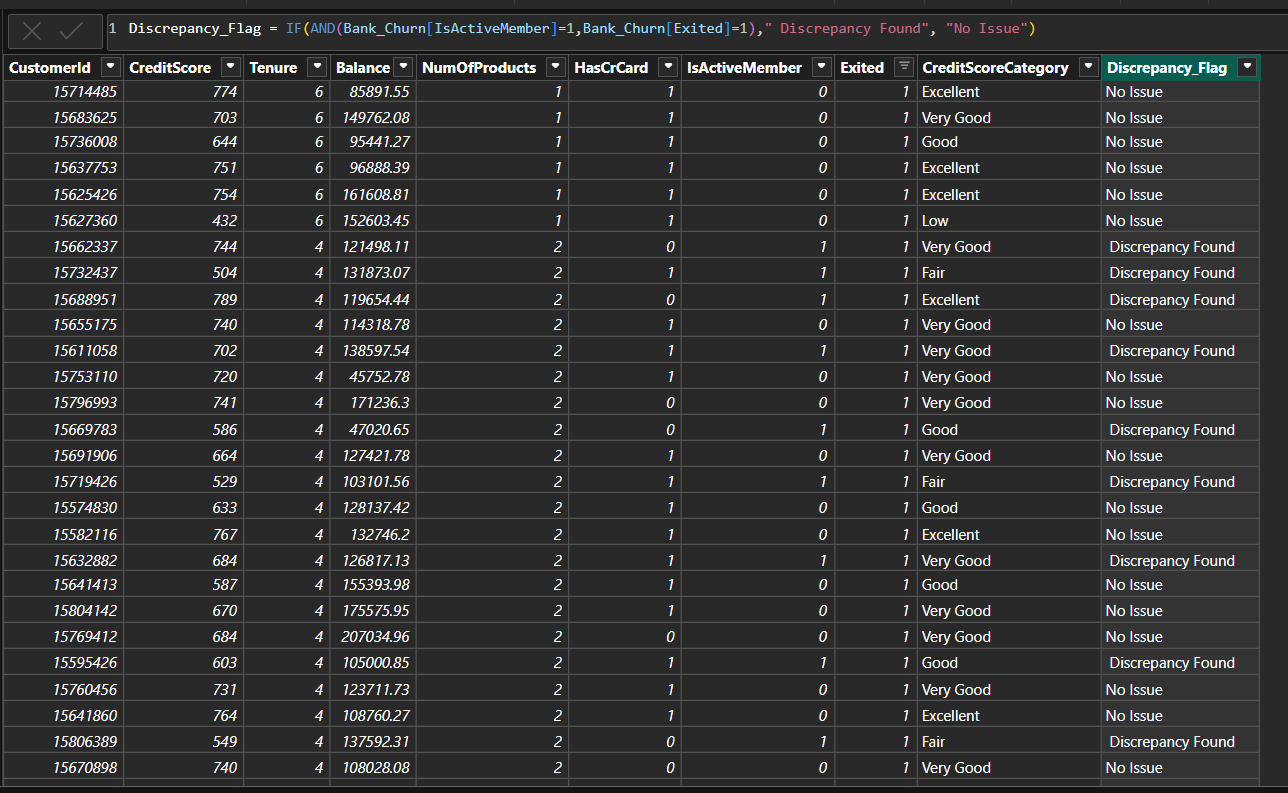
CustomerID\_Surname = CustomerInfo[CustomerId]&”\_“&CustomerInfo[Surname]

Steps: In Table View -> Click on New Column -> Use the Formula.

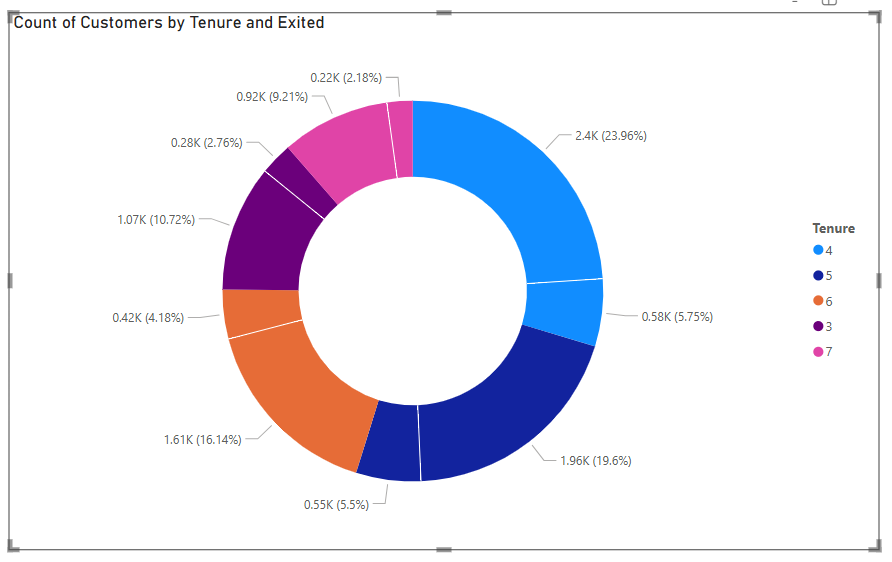
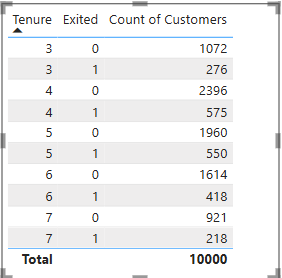
By following the above steps we can create a new column in the prescribed format.  
  


1. **Without using “Join”, can we get the “ExitCategory” from ExitCustomers table to Bank\_Churn table? If yes do this using SQL.**  
     
   Answer: This code retrieves customer data from the Bank\_Churn table and adds a new column named "ExitCategory" to classify customers as 'Retain' (not churned) or 'Exit' (churned) based on the value in the Exited column using subquery.  
   **SQL Query:**select   
   bc.\*,  
   (  
   select e.ExitCategory from exit\_customer e where bc.Exited=e.ExitID  
   ) as ExitCategory  
   from bank\_churn bc;  
    **Result:**  
   
2. **Were there any missing values in the data, using which tool did you replace them and what are the ways to handle them?**  
   Answer: **No,** there weren’t any missing values in the data. This eliminates the need for imputation techniques that might introduce assumptions or biases.

If there were missing values in the data, I would use these techniques to handle them:

1. Deletion: This involves removing rows or columns with missing values.
2. Imputing missing values: Fill missing values using statistical methods like mean, median, or values from other related columns.
3. **Replace Missing Values with a Default Value.**
4. **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”.**Answer: This code retrieves the customer id, their last name and active status for all the customers whose surname ends with “on”.  
     
   **SQL Query:**  
   select   
   c.CustomerId,   
   Surname as last\_name,  
   ac.ActiveCategory  
   from customerinfo c  
   join bank\_churn bc on bc.CustomerId=c.CustomerId  
   join active\_customer ac on ac.ActiveID = bc.IsActiveMember  
   where Surname like '%on';  
     
   
5. **Can you observe any data discrepancy in the Customer’s data? As a hint it’s present in the IsActiveMember and Exited columns. One more point to consider is that the data in the Exited Column is absolutely correct and accurate.**Answer: Yes, there are data discrepancy in the Customer’s data. When the customer is Active which mean the customer is still in the bank and hence the data in IsActiveMember column is 1. When the customer exits the bank, the Exited column is 1. But here there is data which represents both the customer Exited and also an active member at the same time. This is due to data discrepancy.  
   To identify this we have created a new column with below DAX function:   
   **Discrepancy\_Flag = IF(AND(Bank\_Churn[IsActiveMember]=1,Bank\_Churn[Exited]=1)," Discrepancy Found", "No Issue")**  
     
   

**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?**  
     
     
     
     
     
   Below is the exit rate of based on the tenure:   
     
   a. The 4 years tenure customers have 5.75% of exit rate.

b. The 5 years tenure customers have 5.5% of exit rate.

c. The 6 years tenure customers have 4.18% of exit rate.

d. The 3 years tenure customers have 2.76% of exit rate.

e. The 7 years tenure customers have 2.18% of exit rate.

Means the oldest customers are loyal to the bank.

### **Recommendations:**

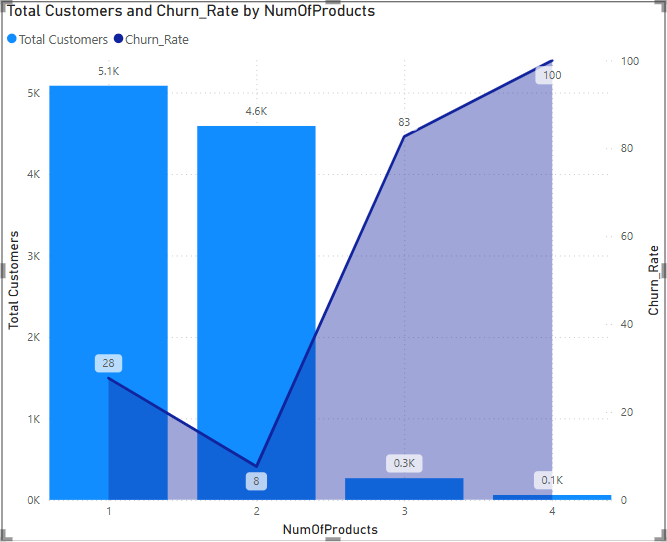
* **Monitor the 4-year tenure group** closely with satisfaction surveys and personalized engagement, as they form a large chunk of the base but show signs of exit.
* **Develop loyalty programs** and incentives for customers approaching 3 and 7 years of tenure to address the observed churn at these points.
* **Segment and analyse spending habits** of long-tenure customers to identify value-driven behaviours and replicate them among newer customers.
* **Implement early engagement strategies** for new customers (within 1–2 years) to increase the likelihood of them reaching mid-tenure, where loyalty is stronger.
* **Use predictive analytics** to detect churn signals in key tenure bands and intervene with tailored retention offers.

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

The data shows that customers with fewer products tend to have a lower churn rate compared to those who have purchased more products.

**Approach:**

Identify commonly paired products and services to uncover usage patterns, noting that fewer product holdings correlate with lower churn rates. Use these insights to refine cross-selling strategies to enhance customer retention.

****

### **Insights:**

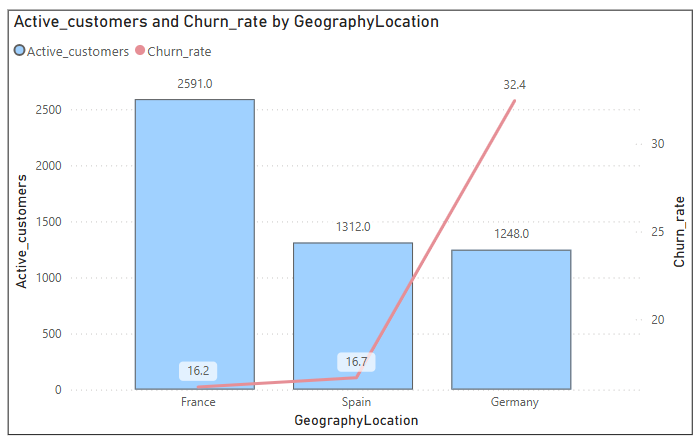
* Most customers have 1 or 2 products (5.1K and 4.6K respectively).
* Churn rate is lowest (8%) for customers with 2 products.
* Churn rate increases significantly for customers with 3 (83%) and 4 (100%) products.
* Very few customers have 3 or more products, indicating deep cross-selling is uncommon.
* Higher product count does not equate to better retention.

### **Recommendations:**

* Target customers with only 1 product to cross-sell a second, focusing on stable combinations.
* Avoid aggressively pushing more than 2 products due to high churn beyond that point.
* Use product affinity analysis to determine common product pairings for smarter cross-sell campaigns.
* Prioritize value and fit of product bundles over quantity when cross-selling.
* Collect feedback from high-product-count customers to identify reasons for churn.
* Build a churn prediction model incorporating product count and engagement metrics.

1. **Geographic Market Trends: How do economic indicators in different geographic regions correlate with the number of active accounts and customer churn rates?  
   Answer:** From the below visual we can easily get to know that the France has high number of active members as compare to other two location but churn rate of Germany is high.

**Approach:**

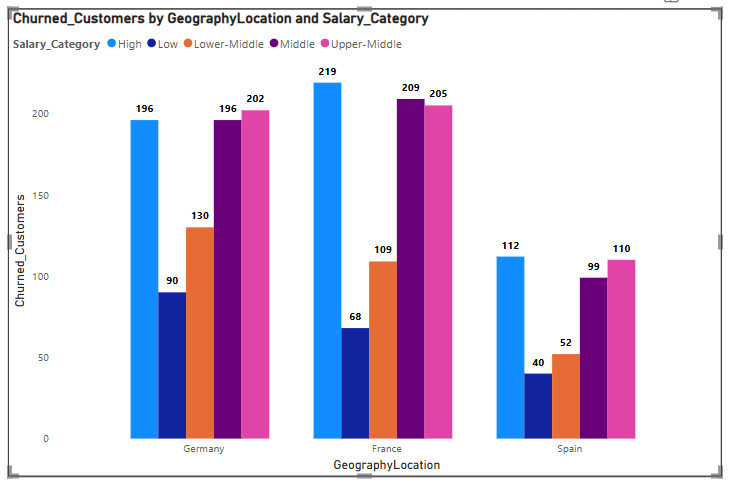
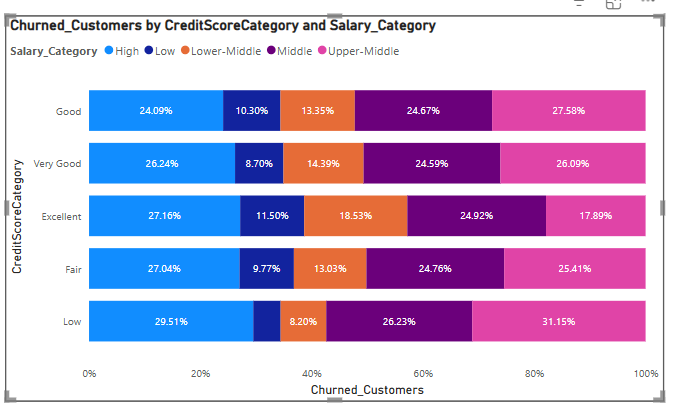
Analyse economic indicators by region to correlate with active account numbers and churn rates, highlighting France's high membership and Germany's high churn. Utilize these findings to tailor regional strategies for retention and growth.

### **Insights:**

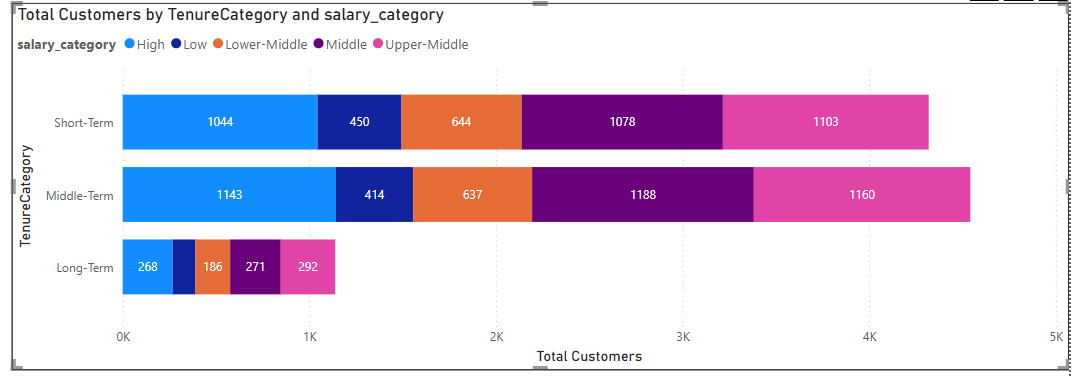
* **France** has the **highest number of active customers** (2,591) with a **moderate churn rate** (16.2%).
* **Spain** has **moderate active customers** (1,312) with a **slightly higher churn rate** (16.7%).
* **Germany** has the **lowest number of active customers** (1,248) but a **significantly high churn rate** (32.4%).
* High churn in Germany suggests dissatisfaction or weaker customer engagement, despite comparable active customer count with Spain.
* France’s larger customer base and relatively low churn imply stronger customer retention and potentially better economic engagement.

### **Recommendations:**

* **Germany:** Investigate key causes of churn through customer feedback and satisfaction surveys. Improve service offerings, communication, or localized support.
* **France:** Leverage successful strategies in France (product mix, communication, or incentives) and consider applying them in other regions.
* **Spain:** Monitor closely — churn rate is rising slightly and may need proactive retention efforts.
* **Localized strategies:** Develop region-specific retention plans, possibly tied to regional economic indicators, product preferences, or competition.
* **Customer engagement:** Launch campaigns to strengthen loyalty in Germany, such as rewards programs or personalized outreach.

1. **Risk Management Assessment: Based on customer profiles, which demographic segments appear to pose the highest financial risk to the bank, and why?  
     
     
     
     
     
   Insights:**a)Germany has the highest churned customers count with 814, especially in High, Middle and Upper-Middlesalary categories. France with 810 and Spain with 413 churned customers.  
     
   b) Across Credit Score categories, Low scores have high proportions in Middle and Upper-Middle salary ranges.  
     
   c) Even Good and Very Good credit score categories show churn in Upper-Middle and Middle salary brackets.  
     
   c) High salary customers are not immune — churn is notable across all credit categories.  
     
   e) Middle-income customers appear across all risk categories — showing potential hidden risks.

Conclusion:   
While low salary + low credit score customers are risky, higher salary segments with fair to good credit scores also demand attention due to their high volume.   
Bank should monitor not just low salaries, but also higher salary groups with declining credit profiles to prevent unexpected losses.

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?**Answer:   
     
     
     
   **Insights:** Customer Tenure Value Forecast

* Middle-Term & Short-Term customers dominate. Focus retention here.
* Long-Term customers are fewer but valuable — nurture them.
* Combine Salary, Credit Score, and Tenure for predictive modelling.
* Use models (like Logistic Regression) to predict tenure and churn.
* Focus on Upper-Middle salary segments for higher lifetime value.
* Prioritize retention and upselling for Middle-Term customers.

**Recommendation:**

* Focus retention efforts on Short-Term, high-potential salary groups.
* Upsell/cross-sell strategies for Middle-Term to extend tenure.
* Incentivize Long-Term loyalty programs for sustainable revenue.

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?**To evaluate how marketing campaigns affect customer retention and acquisition, we need to consider several key strategies and additional data points:

**Enhanced Customer Service:** Improve customer service by offering personalized assistance, resolving issues quickly, and addressing customer feedback efficiently.

**Targeted Offers:** The marketing team should create special offers and provide additional security features for customers over the age of 50, as well as offer incentives to customers who purchase multiple products. Special promotions for credit card holders should also be considered.

### Key Strategies:

**Customer Segmentation:** Group customers by age, location, and the number of products they use. This segmentation allows for a more focused analysis of how different campaigns perform within specific groups.

**Trend Analysis:** Study changes in active customer numbers, exit rates, and product usage over time. This helps identify trends that might be influenced by marketing efforts.

**Control Groups:** If possible, establish control groups that haven't been exposed to certain campaigns. Comparing these groups to those that have been exposed can help determine the true impact of the campaigns.

**Campaign Details & Timing:** Detailed information about campaign content, channels used (e.g., online, offline), and launch dates. This data is essential for linking campaign exposure to changes in customer behavior.

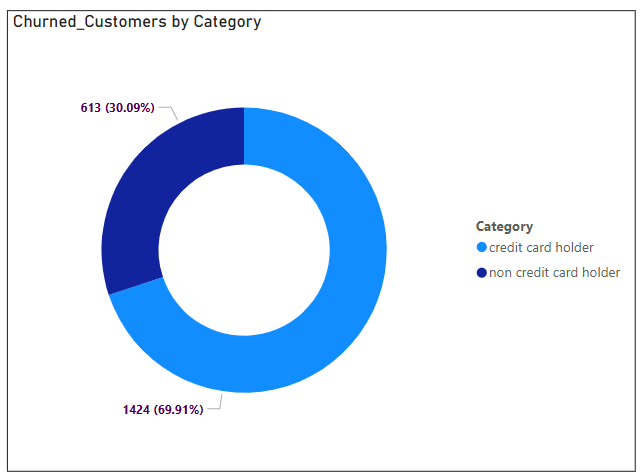
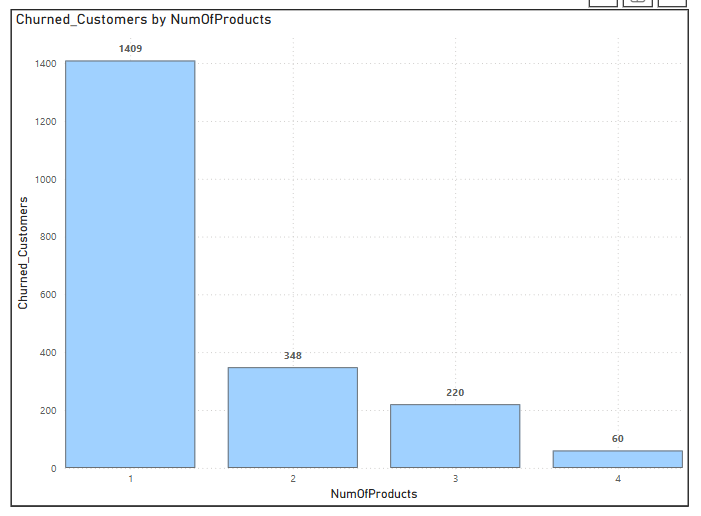
**Customer Acquisition Channel:** Information on how customers were originally acquired (e.g., referral, online advertisement). This helps assess whether campaigns are effective in retaining customers acquired through different channels.

**Customer Lifetime Value (CLV):** Calculate the CLV to understand the long-term impact of marketing campaigns on customer retention and revenue.

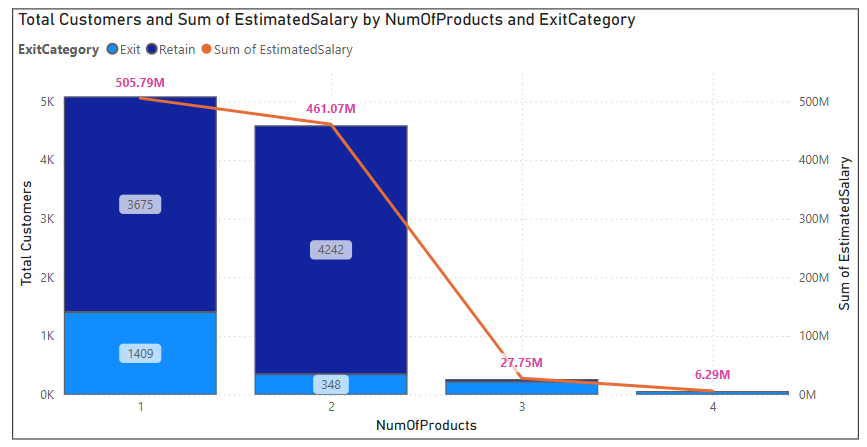
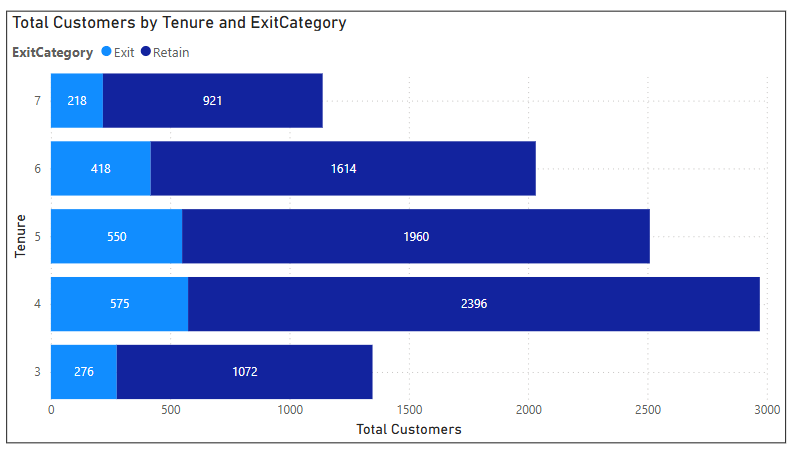
1. **Customer Exit Reasons Exploration: Can you identify common characteristics or trends among customers who have exited that could explain their reasons for leaving?**To find the common characteristics or trends among customers who have exited that could explain their reasons for leaving we can use the below chart analysis.

**Insights:**

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

****   
  
**Conclusion:** The main reasons for exit are:

* Credit Card Ownership: The chart displays the number of customers who exited the bank (churned) categorized by whether they own a credit card (HasCrCard) or not. According to the chart, a significantly higher number of customers with credit cards exited (1,424) compared to those without credit cards (613).
* Less Products Purchased: The number of customers who exited the bank is highest for those who purchased 1 product (1,409), and the number of exits decreases as the number of products purchased count increases. Customers who purchased four or more products have the lowest exit count (approximately 60). This suggests that customers who engage more with the bank by purchasing multiple products are less likely to churn.

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

* **Tenure:** Majority of customers with moderate tenure (4–6 years) are retained, but exits are still visible — tenure helps, but not a sole factor.
* **Number of Products:** Customers with fewer products (1 or 2) form the majority of exits. As products increase, customers tend to stay (though customer count drops).
* **Estimated Salary:** Salary doesn’t show a direct correlation — customers exit across salary ranges. Lower correlation.
* **IsActiveMember (from analysis context):** Usually, active members tend to stay longer, reducing churn. It’s an influential feature for churn prediction.

**Conclusion:**   
Tenure, number of products, and activity status are strong predictors of churn. Estimated salary alone is less significant.

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

g.GeographyLocation,

CASE

WHEN bc.CreditScore BETWEEN 300 AND 649 THEN 'Low'

WHEN bc.CreditScore BETWEEN 650 AND 749 THEN 'Medium'

WHEN bc.CreditScore BETWEEN 750 AND 849 THEN 'High'

ELSE 'Excellent'

END AS credit\_score\_segment,

CASE

WHEN bc.Tenure BETWEEN 3 AND 5 THEN 'Short-Term'

WHEN bc.Tenure BETWEEN 6 AND 7 THEN 'Long-Term'

else 'Other'

END AS TenureCategory,

cr.Category as credit\_card\_Category,

ac.ActiveCategory,

e.ExitCategory,

round(avg(c.EstimatedSalary),2) as avg\_estimated\_salary,

round(avg(bc.balance),2) as avg\_balance

from customerinfo c

join geography g on g.GeographyID=c.GeographyID

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

join active\_customer ac on ac.ActiveID = bc.IsActiveMember

join credit\_card cr on cr.CreditID = bc.HasCrCard

join exit\_customer e on e.ExitID = bc.exited

group by

GeographyLocation,

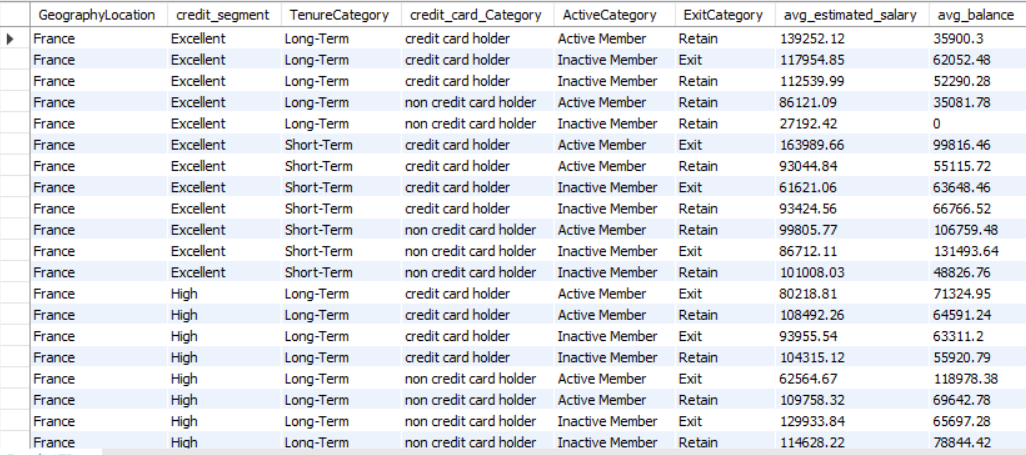
credit\_segment,

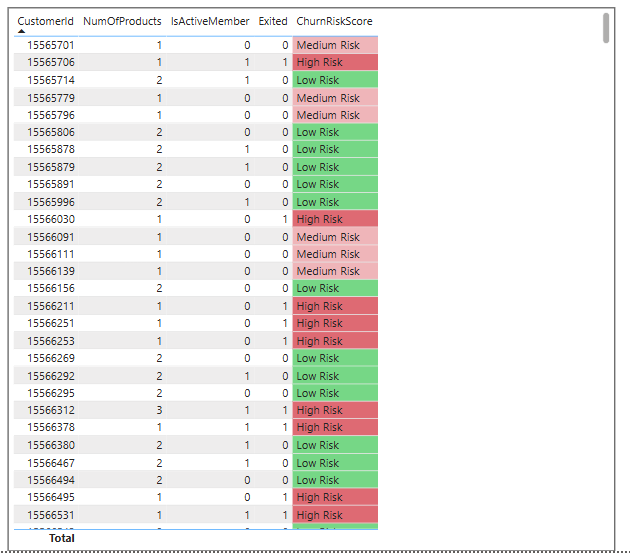
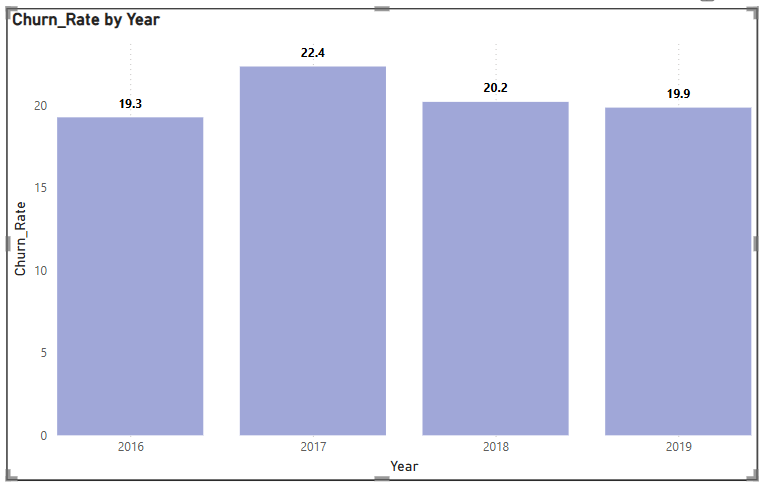
TenureCategory,

cr.Category,

ac.ActiveCategory,

.ExitCategory

order by g.GeographyLocation, credit\_score\_segment; **Result:  
 **

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?  
   Answer:** To create a conditional formatting for highlighting the customers at risk of churn, created a new Column which categorize the customers into 3 segments based on the no of products, exit status and active status.  
     
   **Power BI:**  
   ChurnRiskScore = IF(  
   Bank\_Churn[Exited] = 1, "High Risk",  
   IF(Bank\_Churn[NumOfProducts] = 1 && Bank\_Churn[IsActiveMember] = 0, "Medium Risk",   
   "Low Risk")  
   )  
    ****
2. **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?**The Bank’s overall churn rate is 20.37% with year on year fluctuations. **Insights:** Year wise Churn rate:  
    a) 2016: 19.3%(Lowest) b)2017: 22.4%(Highest) c)2017: 20.2% d)2017: 19.9%  **  
   Customer Segments Prone to Churn:**

Analysis reveals several customer segments with higher likelihoods of churn:

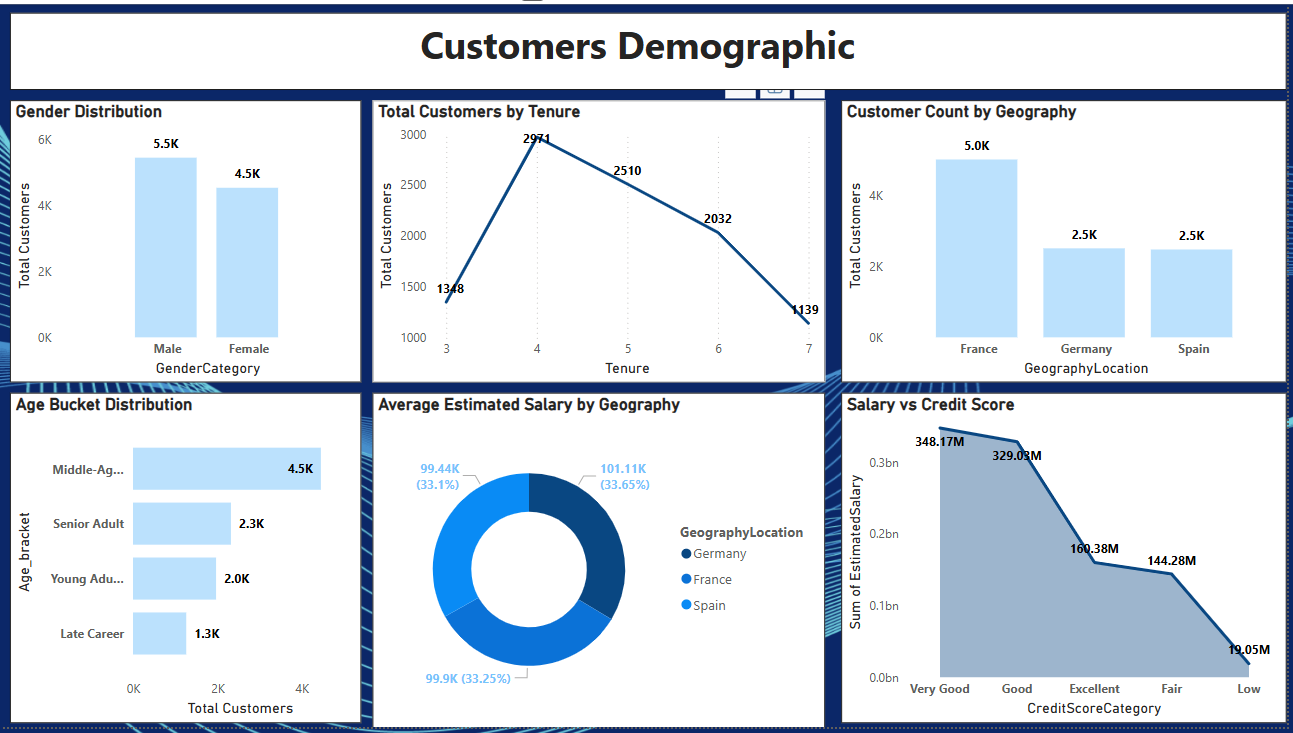
* **Single Product Users**: Customers using only one product may not see enough value in the bank's offerings, particularly when competitors provide broader or more integrated services.
* **Credit Card Holders**: Potential factors for credit card holders churning include:
  + Insufficient credit limits.
  + Lack of appealing rewards programs.
  + High fees associated with the card.
* **Tenure of 4-5 Years**: These customers might be coming off from promotional offers or discounts, making them susceptible to competitors offering more attractive rates or features.
* **High Salary Customers**: High earners may have more financial options and are more likely to switch banks for slightly better benefits or interest rates elsewhere.

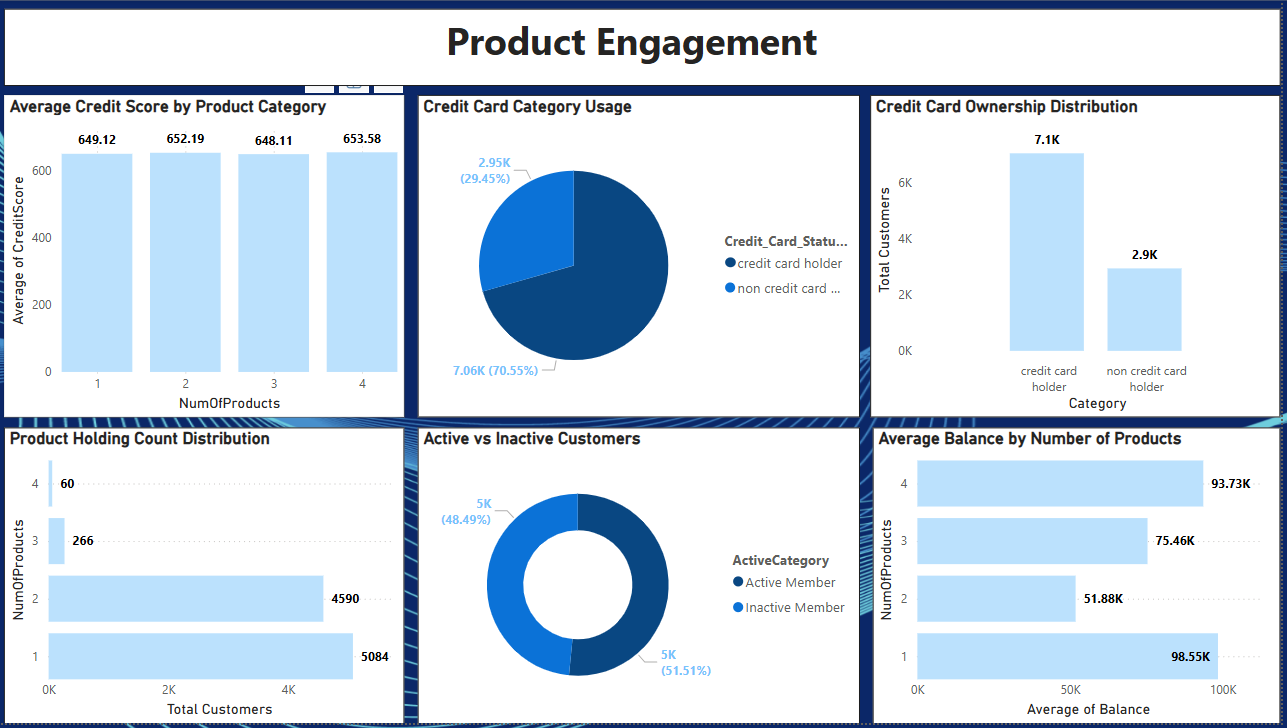
**Recommendations to Reduce Churn:**

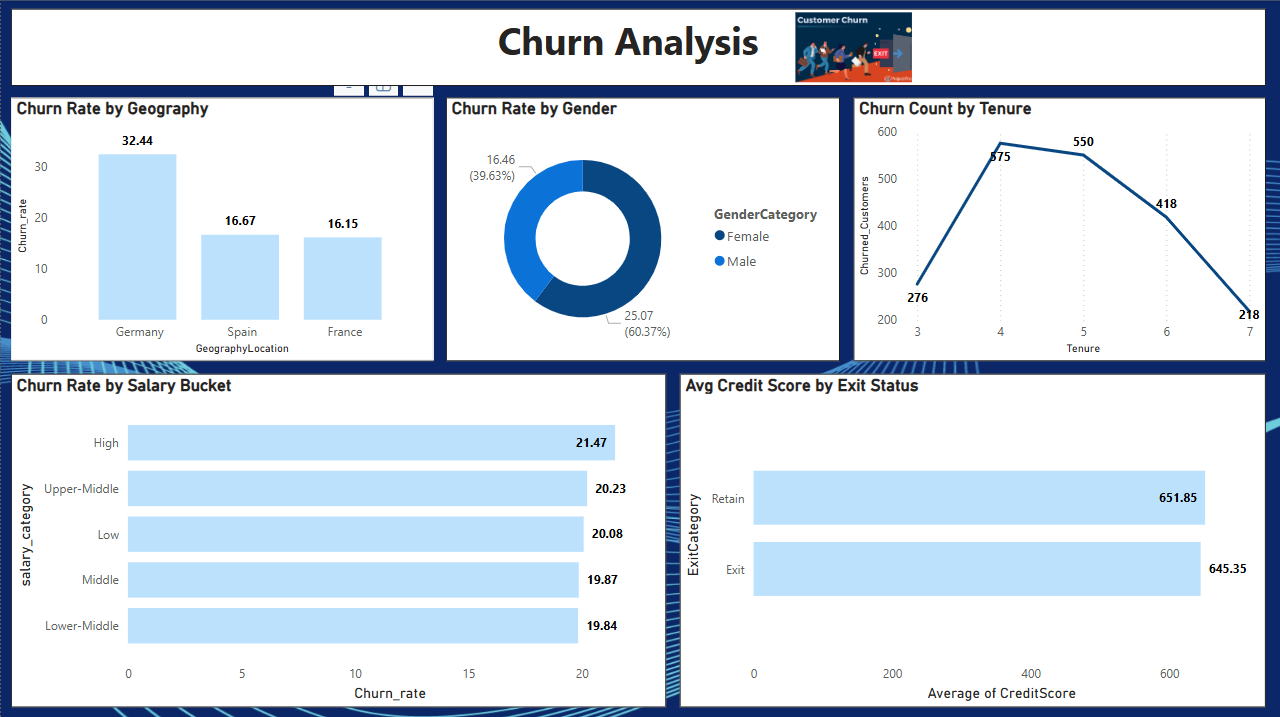
* **Targeted Product Bundles**:
  + Develop tailored product bundles to address the specific needs of customers who use only one product. Highlight the added value and potential cost savings of using multiple services.
* **Enhanced Credit Card Rewards**:
  + Improve the credit card offering by:
    - Increasing credit limits based on customer behaviour and creditworthiness.
    - Aligning rewards programs with customer preferences (e.g., travel, cashback for certain categories).
    - Reducing or eliminating annual fees, especially for high-value customers.
* **Retention Offers for Existing Customers**:
  + For customers nearing the end of their introductory offers, proactively offer personalized retention deals, such as:
    - Extending the promotional rates.
    - Providing discounts on other products or services.
* **Customer Satisfaction Surveys**:
  + Regularly survey customers to gather insights into why they leave. Use this feedback to fine-tune retention strategies and address customer pain points.
* **Relationship Management for High-Value Customers**:
  + Assign dedicated relationship managers to high-value customers, offering them personalized services and exclusive benefits to strengthen loyalty and meet their individual needs.

By focusing on these strategies, the bank can reduce churn, retain high-value customers, and increase overall customer satisfaction.

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?  
     
    Understanding the Context:**

* Define the goal of the analysis: customer churn prediction, marketing effectiveness, tenure forecasting, etc.
* Identify key stakeholders and their expectations.
* Determine how insights will be used for decision-making.

**Data Exploration & Cleaning:**

* Examine dataset structure (columns, data types, missing values).
* Perform summary statistics & distributions (e.g., customer tenure trends, product usage).
* Identify potential discrepancies (e.g., IsActiveMember vs. Exited inconsistency).

**Hypothesis Formation:**

* Do customers with longer tenure have lower churn rates?
* Does NumOfProducts influence customer retention?
* Are high-income customers more likely to stay?

**Data Analysis & Insights:**

* Use DAX measures for calculations (e.g., churn rate, tenure impact and customer segmentation).
* Apply Power BI visualizations (conditional formatting, secondary axes, and trend lines).
* Compare trends across different groups (e.g., geographic churn rates, high-risk customer segments).

**Testing & Validation:**

* Cross-check insights with additional calculations.
* Validate findings using historical data or industry benchmarks.
* Seek feedback from stakeholders to ensure relevance.

**Visualization & Communication:**

* Present insights through interactive dashboards.
* Apply conditional formatting to highlight churn risk levels.
* Use sorting, filtering, and tooltips for better data interpretation.

**Actionable Recommendations:**

* + Improve credit card rewards to boost retention.
  + Focus marketing efforts on high-churn customer segments.
  + Identify at-risk customers and implement retention measures.

1. **In the “Bank\_Churn” table how can you modify the name of the “HasCrCard” column to “Has\_creditcard”?  
   Answer:** This SQL query changes the name of the “HasCrCard” column to “Has\_creditcard” in the Bank\_churn table. It is a more descriptive name which provides more clarity. **SQL Query:**   
   Alter table bank\_churn rename column HasCrCard to Has\_creditcard;  
     
   **Result:** The output shows that the column name “HasCrCard” has been changed to “Has\_creditcard**”.**  
   