

|  |
| --- |
|  |
| Analytical CRM development for a bank  Detail analysis on factors responsible for customer churn & identifying customer segments to target them with tailored services. |
| |  |  |  | | --- | --- | --- | | PRATIK NANDANWAR | MARCH’24 | DATA ANALYSIS & SUGGESTIONS | |

**Methodology**

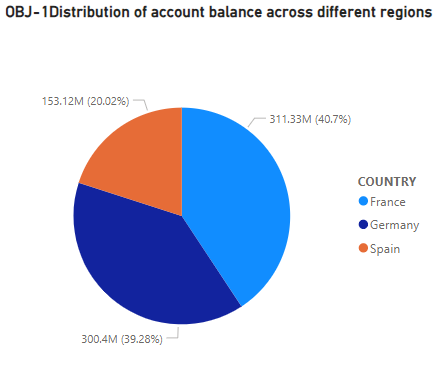
* Data pre-processing & and consolidation were done using **excel**.
  + In which categorical variables were updated in the 2  
    main Sheets named Customer\_info & Bank\_churn   
    from the categorical variable sheets available such as   
    exit customer, active customer, gender & geography.
  + Missing values, duplicates, inconsistencies were checked   
    and cleaned for further analysis.
* These data sheet of “Customer\_info” & “Bank\_churn” were then   
  imported on **MySQL workbench** 
  + Questions related to querying were solved in   
    **MySQL workbench** and remaining were solved & visualized   
    in **Microsoft PowerBI.**

**Objective questions**

**Q-1**

What is the distribution of account balance across different regions?

**Answer**

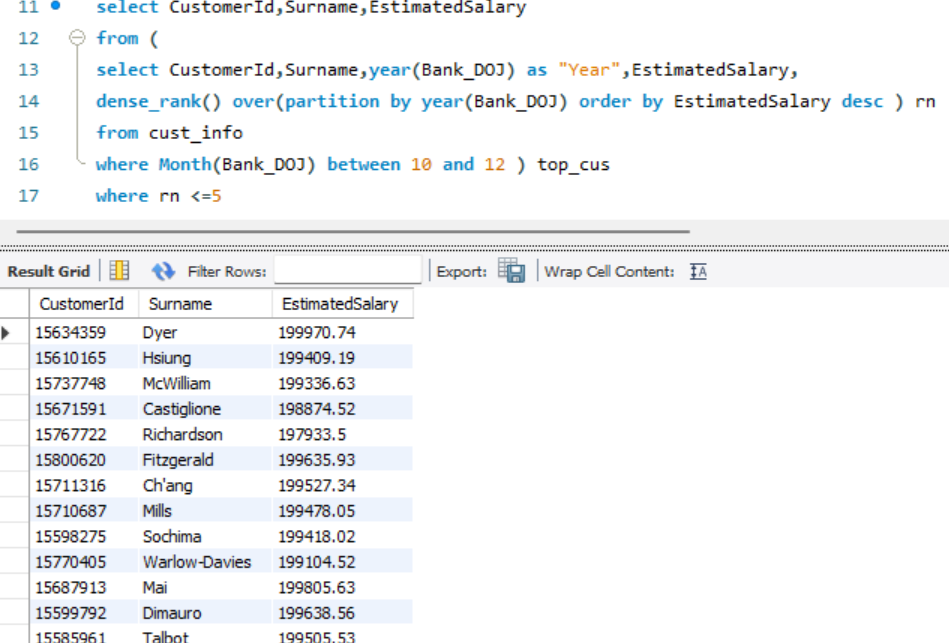


* France – 311.33M
* Germany – 300.4M
* Spain – 153.12M
* This insight of bank balance distribution would help for identifying
  + - Regional Economic Activity.
    - Market Opportunities for bank to plan & then invest.
    - Branch Network Optimization (opening & upgrading new and existing branches).
* The distribution of account balances across different regions provides insights into market dynamics, customer behaviours, and strategic opportunities for the bank to optimize its operations.

**Q-2**

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

**Answer**



* Identifying the top 5 customers with the highest estimated salary in the last quarter of the year can indicate several insights and implications for the bank such as: -
  + - Identification of high value customers
    - Identification of customers who would

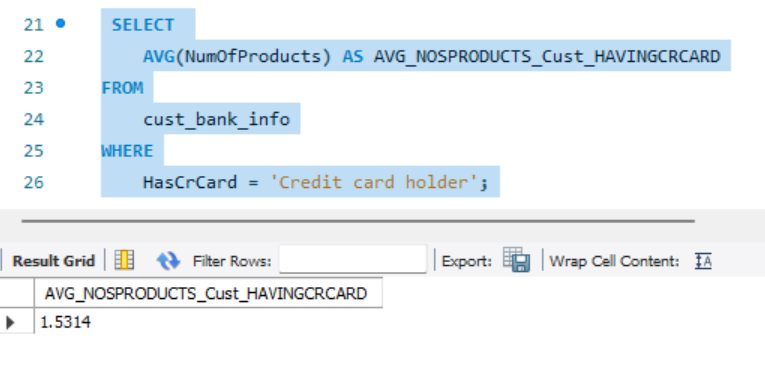
need and afford tailored specific services.

* + - **Future Revenue Potential** by these customers   
      & importance of retaining them.

Q.3

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

Answer



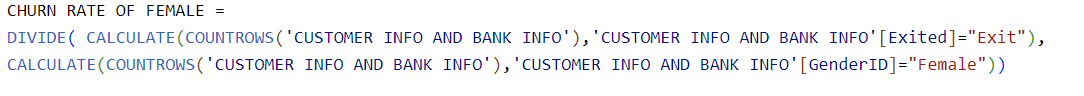
* The Average number of products used by customers who have a credit card is

1.5314

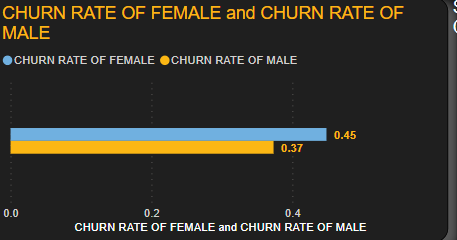
Q.4

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

Answer



Same DAX for Males

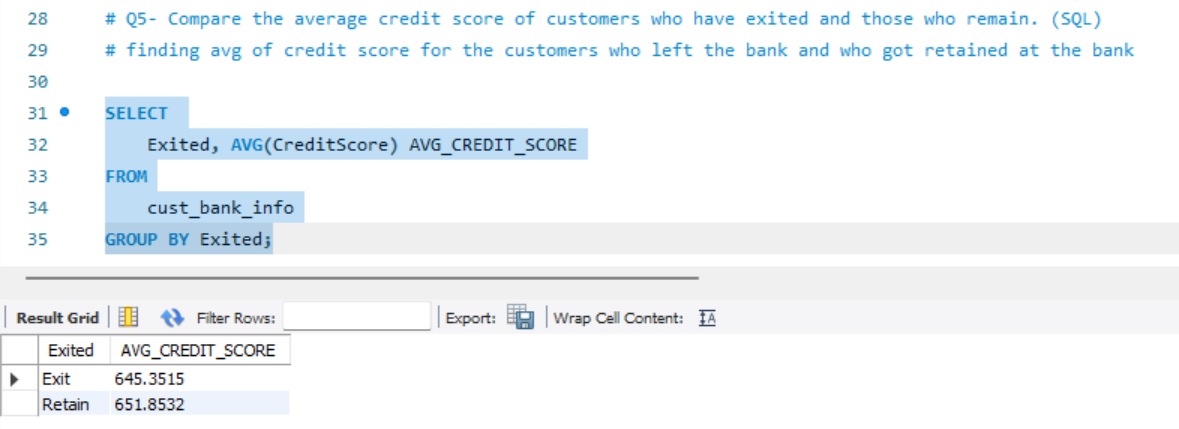


* Churn rate of **female** is **higher** than males.
* Difference of Churn rate by gender can provide several insights such as:
  + - **Gender wise preferences** to bank.
    - **Foresight to bank** to which gender to focus on if the difference if significant.
    - Understanding Gender wise customer feedback.
    - Help in formulating gender specific retention strategies.

**Q.5**

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

Answer



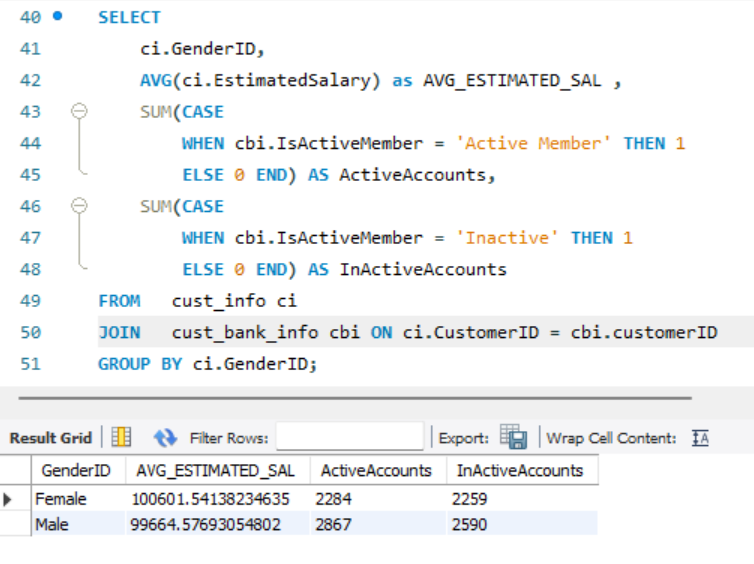
* The average credit score of customers who exited were found to be 645.35

and those who were still with the bank were 651.8.

* However, both the categorical scores comes under the fair category so this would need analysis by further factors for establishing the relation of customers average score and them leaving the bank.
* Credit score can give insights to
  + - * + **Creditworthiness of a customer.**
        + **Credit Risk and Churn.**

**Q.6**

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

**Answer**

**Male**

**99664**

**Active Accounts**

**2867**

**Active Accounts**

**2284**

**Salary**

**Female**

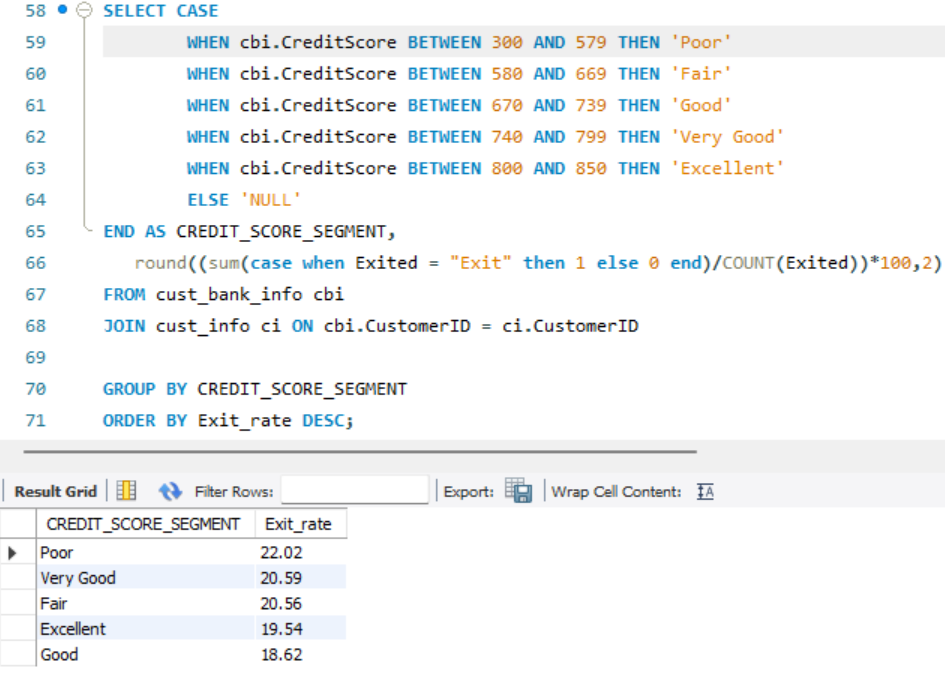
**100601**

* **Female have the higher average estimated salary**  
  having **583 less** active accounts.
* This can indicate that **female category can be the major contributor to the Revenue** to the Bank as their average salary is higher than the male customers.
* This also provides insight to bank on tailoring, manipulating ,and upgrading/Updating the existing and new strategies or policies or business plan.

**Q.7**

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

**Answer**

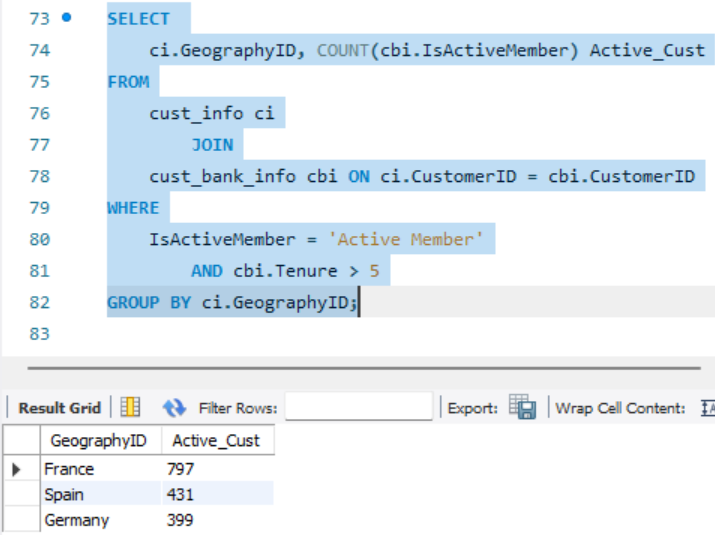
****

* Customers were segmented based on their credit score as per the criteria.
* Highest **customers exited rate** is **from** the segment who had a **“POOR”** credit score having exit rate of 22.02%

**Q.8**

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

**Answer**

****

* **FRANCE** has the highest Active customers with tenure greater than 5 years.
* Identifying the geographic region with the highest number of active customers who have a tenure greater than 5 years can provide several insights for the bank such as:

Customer loyalty

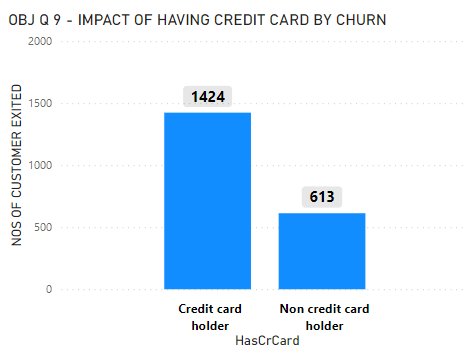
Customer satisfaction level w.r.t region.

Market dynamics.

**Q.9**

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

**Answer**

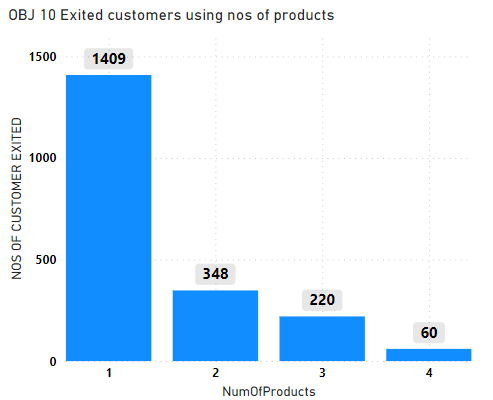
****

* **Customers churned having Credit card – 1424**
* **Customers churned not having Credit card – 613**
* The difference can be clearly seen that when the customers **had credit card more customers left bank** & the customer did not having credit card, the rate of leaving the bank is low i.e. 613
* This establishes the relation that having or not having credit card has corelation with bank churn rates.
* **Opportunities for Credit Card Improvement-**Providing a positive credit card experience to enhance customer retention.

**Q.10**

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

**Answer**

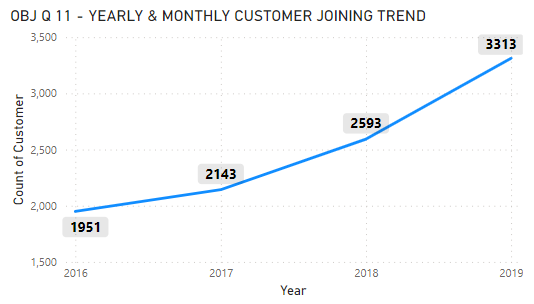
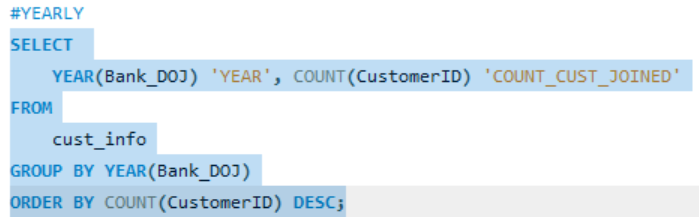
****

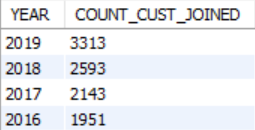
* Most common number of products used by the customers who left the bank was only 1 product.
* Customers using 4 number of products had the minimum number of customers exiting.
* This provides insight so as to strategizing penetrating the customer with more number of products as it results in lesser number of bank churning.

**Q.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.

**Answer**

****



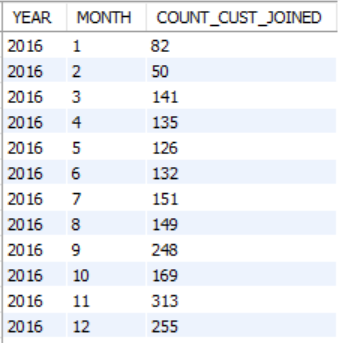
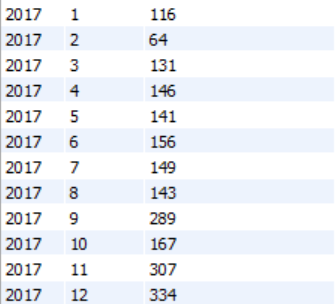
**Yearly -**

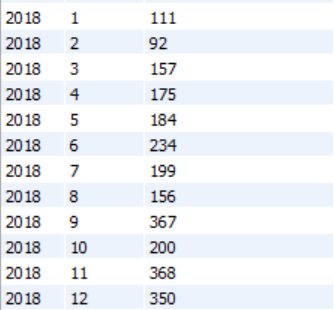
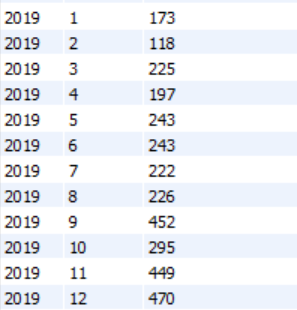
* **Customer strength in - 2016 = 1951**
* **Customer strength in - 2017 = 2143**
* **Customer strength in - 2018 = 2593**
* **Customer strength in - 2019 = 3313**
* The customers are seen to be increasing year by year
* Suggests several positive trends and implications for the bank such as
  + **Growing Customer Base**
  + **Indication if strategies are working or not .**
  + **Customer Acquisition Efficiency**.
  + **Revenue Growth Potential**

**Monthly**

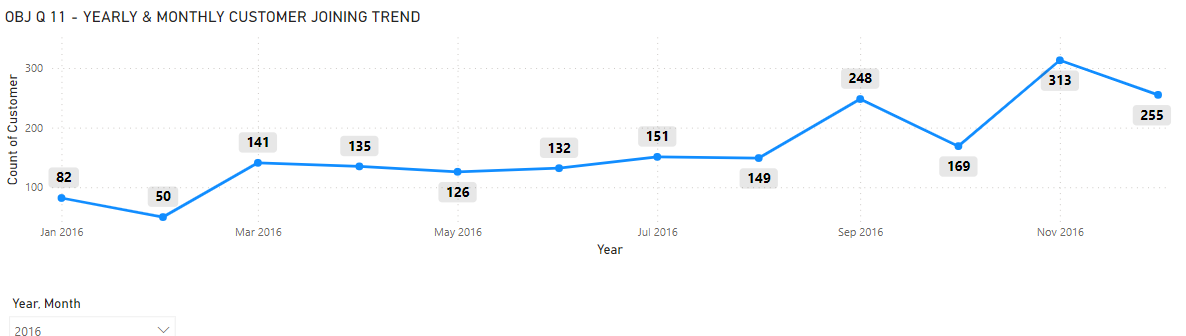
**Data prepared by SQL for every year**

**2016** **2017**

**** ****

**** **2018 2019**

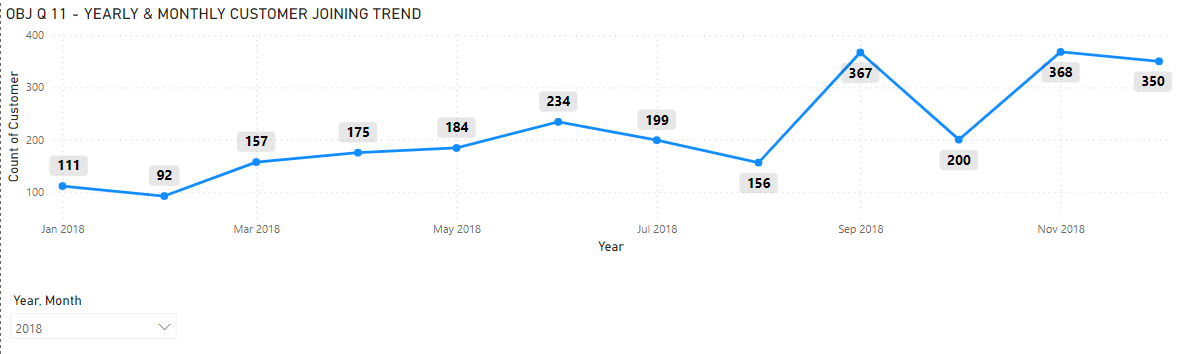
**2016**

****

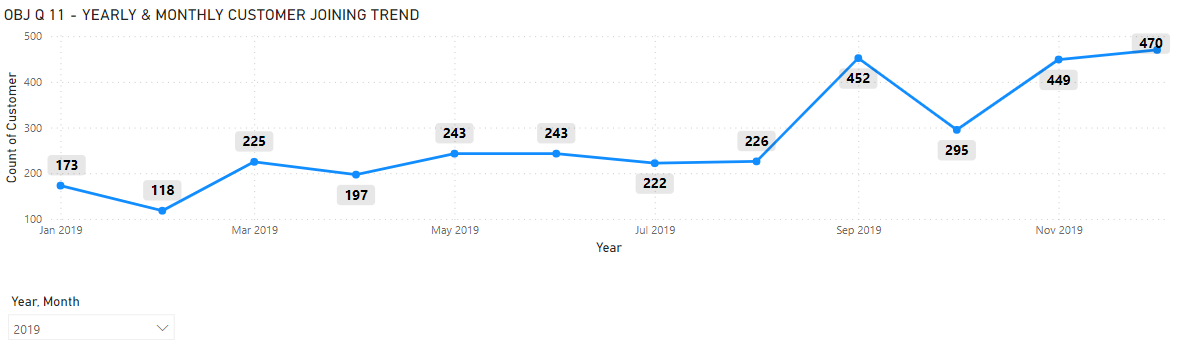
**2017**

****

**2018**

****

**2019**

****

* On looking at the monthly counts of joining over the year a similar pattern can be observed in the fashion of joining in numbers.
* The pattern seen is –

**Nominal start**

**January**

**Dip in numbers**

**February**

**Jump & stagnancy till July**

**March to August**

**Sudden jump and again dip**

**September & october**

**November & December**

**Jump & growth**

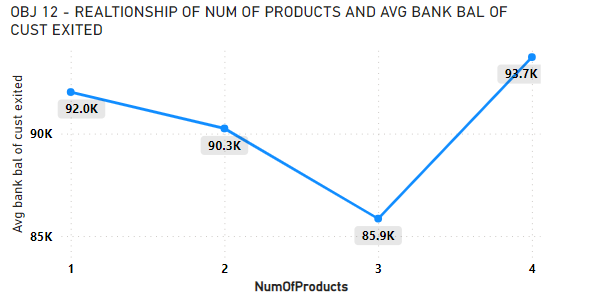
* This pattern repeated itself every year up till 2019 in the same way.

So strategies can be planned for the months which has a dip in numbers as they are dipping in the same month every year.

**Q.12**

Analyse the relationship between the number of products and the account balance for customers who have exited.?

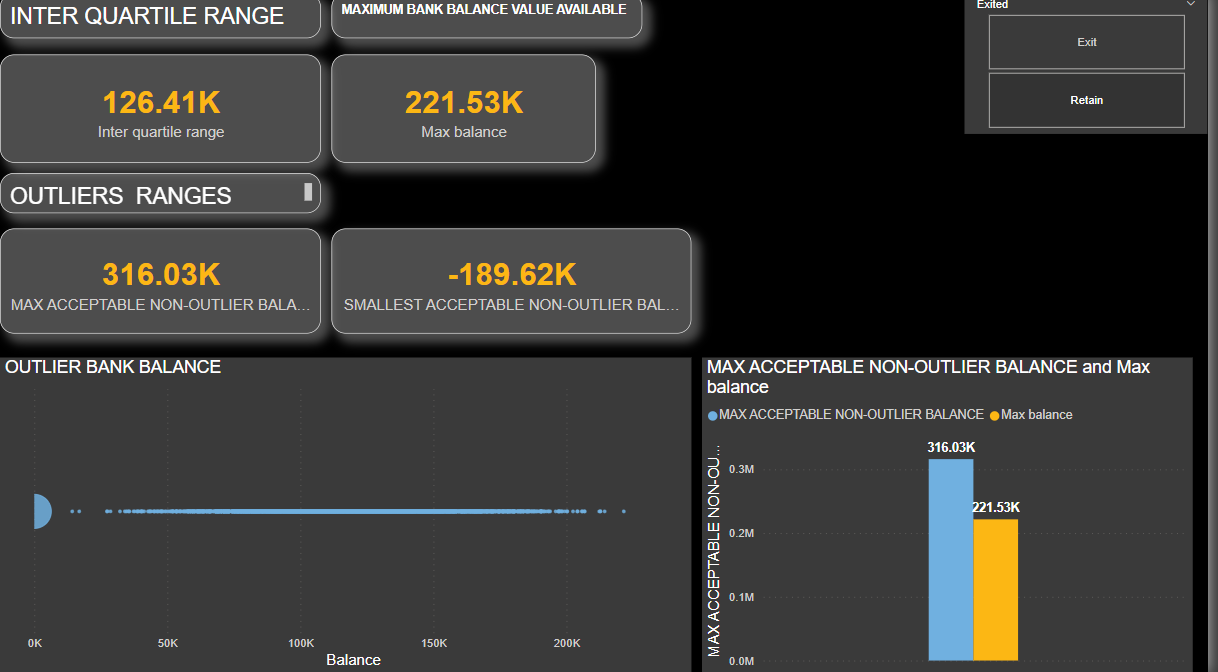
**Answer**

****

* Customers using 4 number of products & services found to be having the highest average bank balance.
* Whereas customers using 3 number of products & services found to be having the lowest average bank balance.
* However the Average bank balance and number of products the customer uses does not vary to very significant large amount.

**Q.13**

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

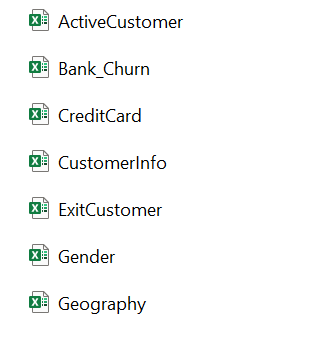
**Answer**

* **No outliers** were found in bank balance of customers who have **remained with the bank.**
* This was calculated by using the concept of interquartile ranges.
  + - * Calculated interquartile ranges
      * Compared the maximum bank balance of customer still with bank to the maximum allowed non outlier balance calculated by interquartile ranges.
* As well in the scatter plot the data points are not significantly distant from the general pattern or cluster of data points hence no outliers was found.

**Q.14**

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

**Answer**

****

* Dataset consisted of7 different tables.
* Out of which 5 tables of:
  + - * + ActiveCustomer
        + CreditCard
        + ExitCustomer
        + Gender
        + Geography

Consisted of only categorical variables tables.

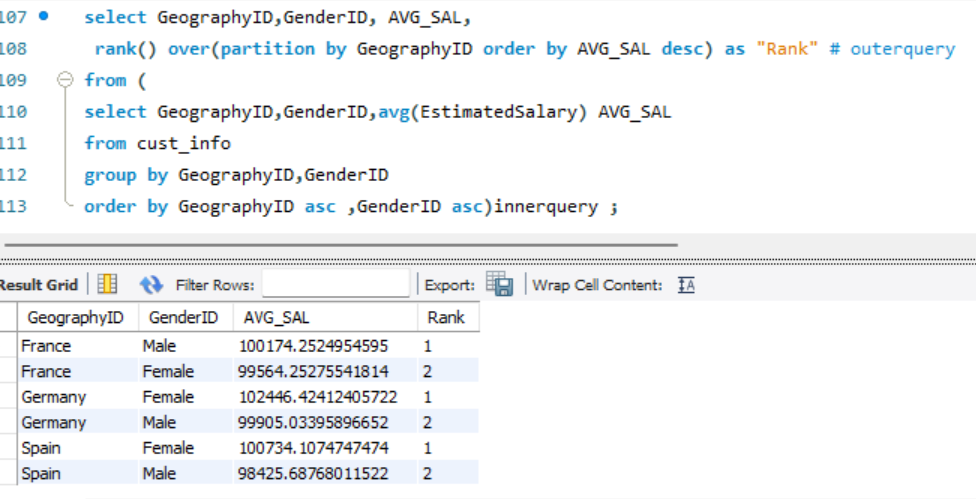
* And 2 tables of:
  + - * Bank\_churn
      * CustomerInfo

Consisted of information regarding customers personal and demographical info and bank related info.

**Q.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)

**Answer**

****

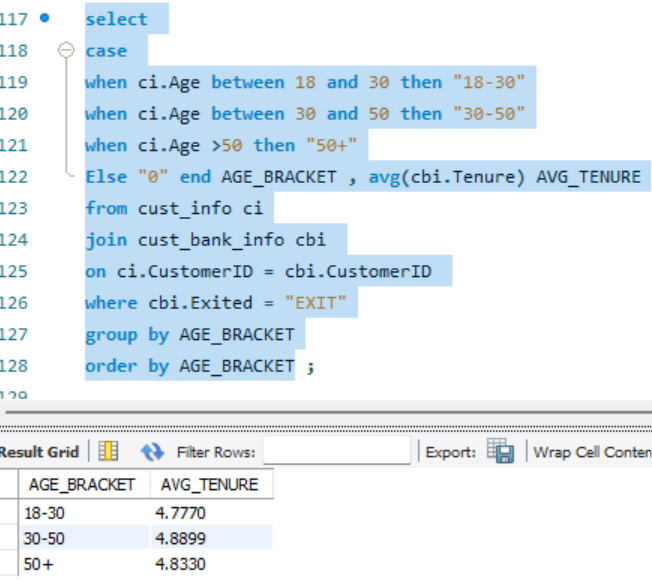
****

* In FRANCE: - **MALE has higher average income** compared to female.
* In GERMANY: - **FEMALE has higher average income** compared to male.
* In SPAIN: - **FEMALE has higher average income** compared to male.

**Q.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+).

**Answer**

****

* The average tenure in each age bracket was calculated and is found to be nearby same.
* Nearby same average tenure across age brackets suggests that customers across all age groups experienced similar levels of satisfaction or dissatisfaction with the bank's service and products.

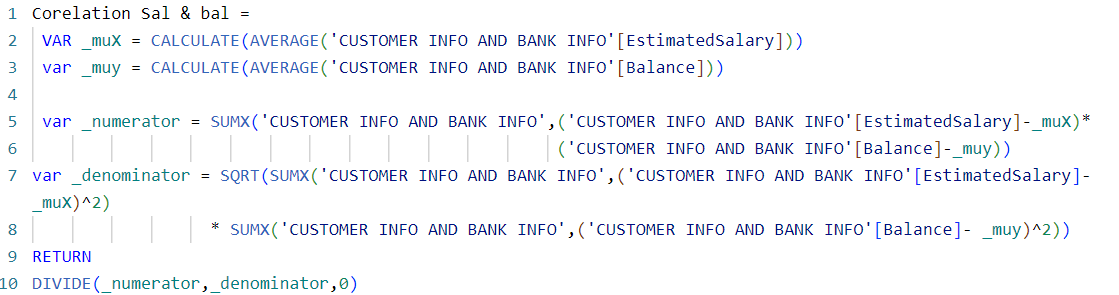
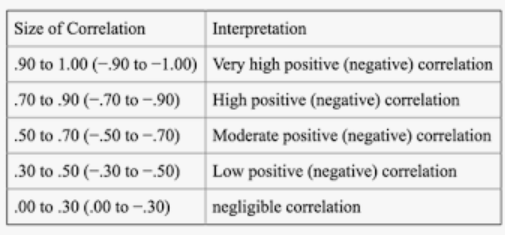
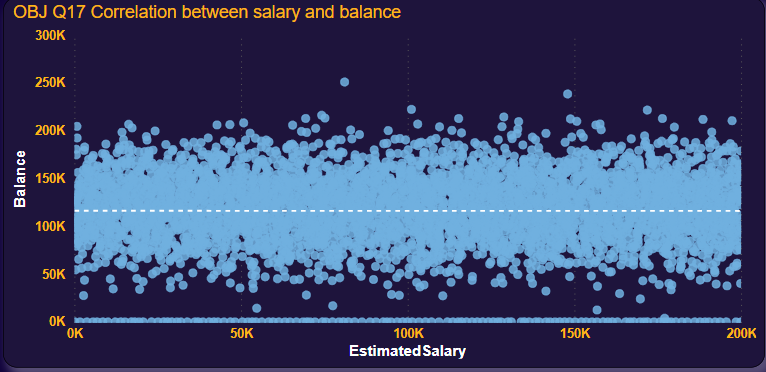
This could reflect both - consistency in the customer experience, or a lack of targeted retention efforts.

* This can also indicate that customers may have faced similar reasons for leaving the bank, regardless of their age demographic.

**Q.17**

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

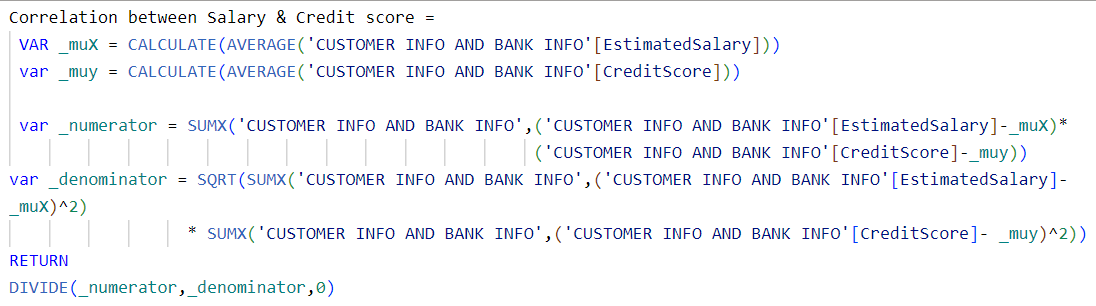
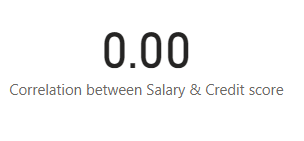
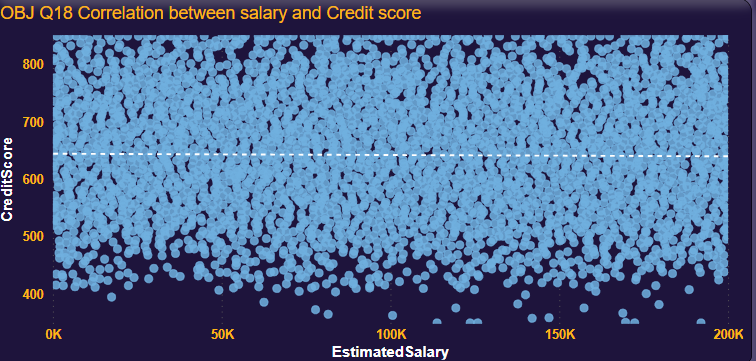
**Answer**

* Correlation was calculated using **POWERBI DAX function** and calculating the corelation coefficient.
* ****
* ****
* ****
* A horizontal pattern with data points spread across the entire range of the x-axis (salary) and y-axis (balance) suggests that there is no linear relationship between the two variables.
* The correlation coefficient (r) would be close to 0, indicating a weak or no correlation between 'salary' and 'balance'.
* The dataset shows **NEGLIGIBLE CORRELATION** between salary & balance of the customers.
* The absence of an upward or downward trend, it means that changes in 'salary' are not associated with predictable changes in 'balance', and vice versa.

**Q.18**

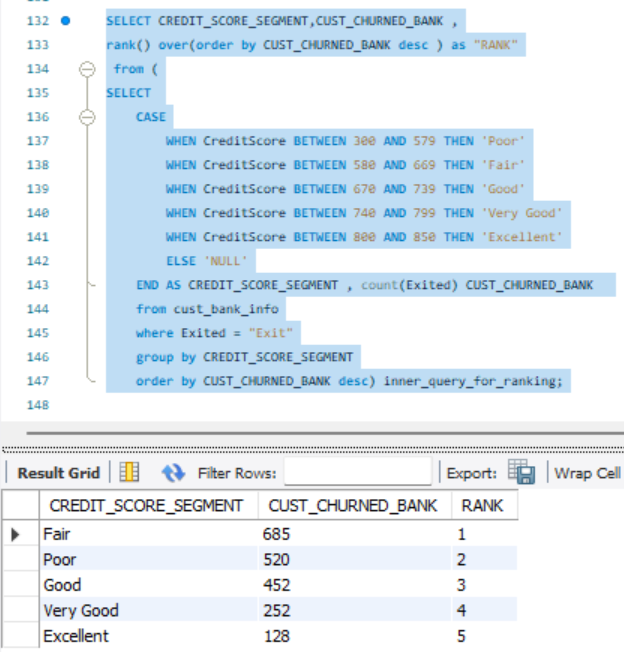
Is there any correlation between salary and Credit score of customers?

**Answer**

* Correlation was calculated using **POWERBI DAX function** and calculating the corelation coefficient.
* 
* 
* 
* The random non-directional pattern shows no association of Salary and credit score of a customer
* T he dataset shows **NO CORRELATION** between Salary & credit score of the customers.

**Q.19**

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

**Answer**

* On ranking each bucket of credit score as per the number of customers who have churned the bank it was observed that most number of people who left the bank had an “Fair” credit score.
* And Customers having Excellent credit score was the last to churn the bank.
* Identifying the relation between credit score and customer churn enables the bank to plan targeted retention strategies tailored to different customer segments.

For example, the bank knows which segment to focus on enhancing satisfaction for customers with lower credit scores by offering personalized services like financial guidance, credit maintenance and its importance.

**Q.20**

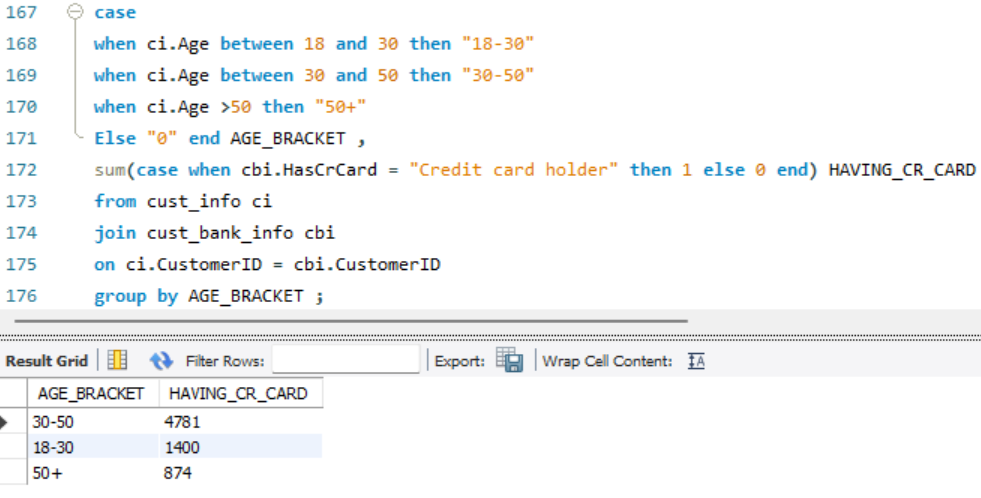
**A-**According to the age buckets find the number of customers who have a credit card.

**B-**Also retrieve those buckets who have lesser than average number of credit cards per bucket.

**Answer**

**A**

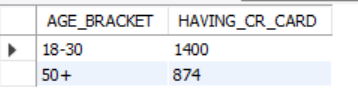
* **Number of people having credit card /age bucket**

****

* It can be observed that Age bracket of 30-50 years has the maximum numbers of credit card holdings.

**B**

Retrieve those buckets who have lesser than average number of credit cards per bucket.

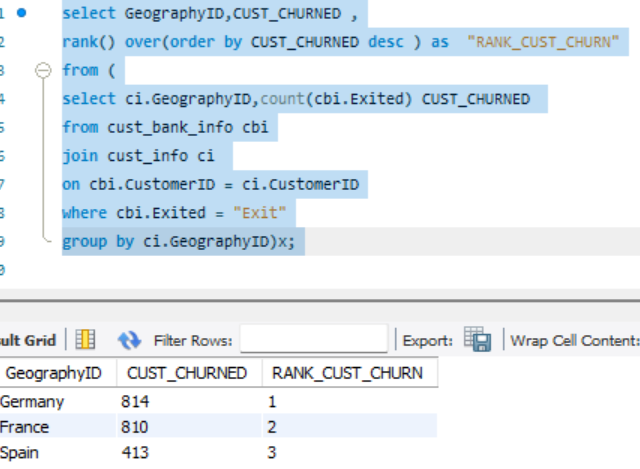
* ****
* Age bracket of 18-30 & 50+ has the credits cards have lesser than average number of credit cards.

**Q.21**

A- Rank the Locations as per the number of people who have churned the bank

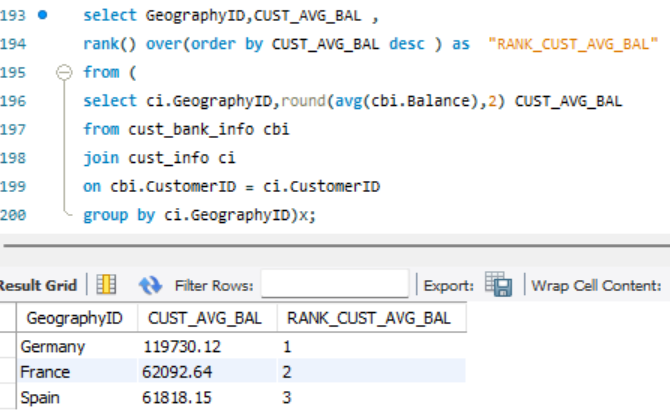
B- and average balance of the customers.

**Answer**



* Germany has the highest customer churn observed & Spain has the lowest customer churn.

**B-** Ranking/**average balance of the customers.**

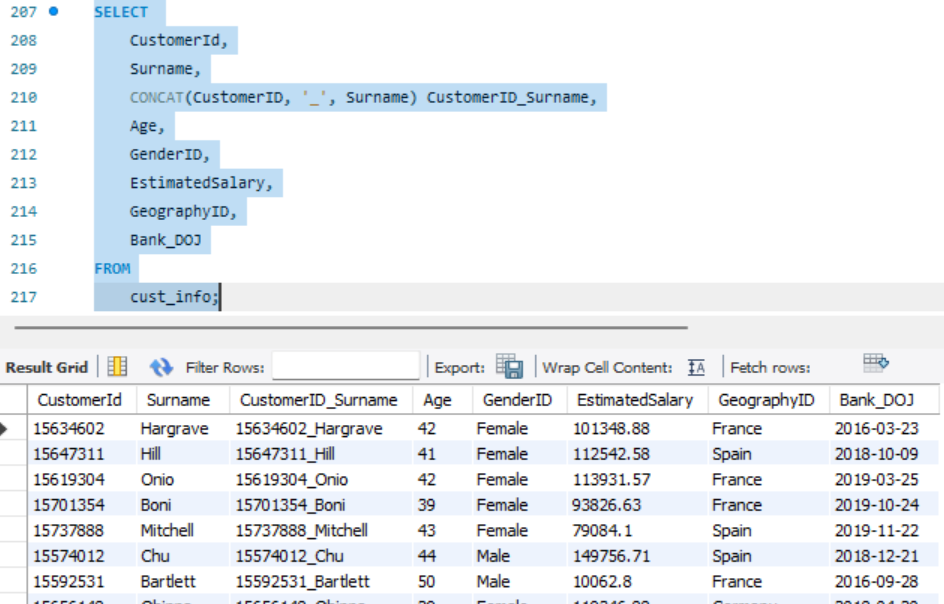
* ****
* **Germany** has the highest average bank balance & Spain has the lowest average bank balance.
* Understanding the customer churn & average bank balance location wise would help monitor and strategize plan for better customer retention and foresight on which location to be focused.

**Q.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”.

**Answer**

* **CONCAT** function can be used for creating a new column where the format is “CustomerID\_Surname”

****

**Q.23**

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

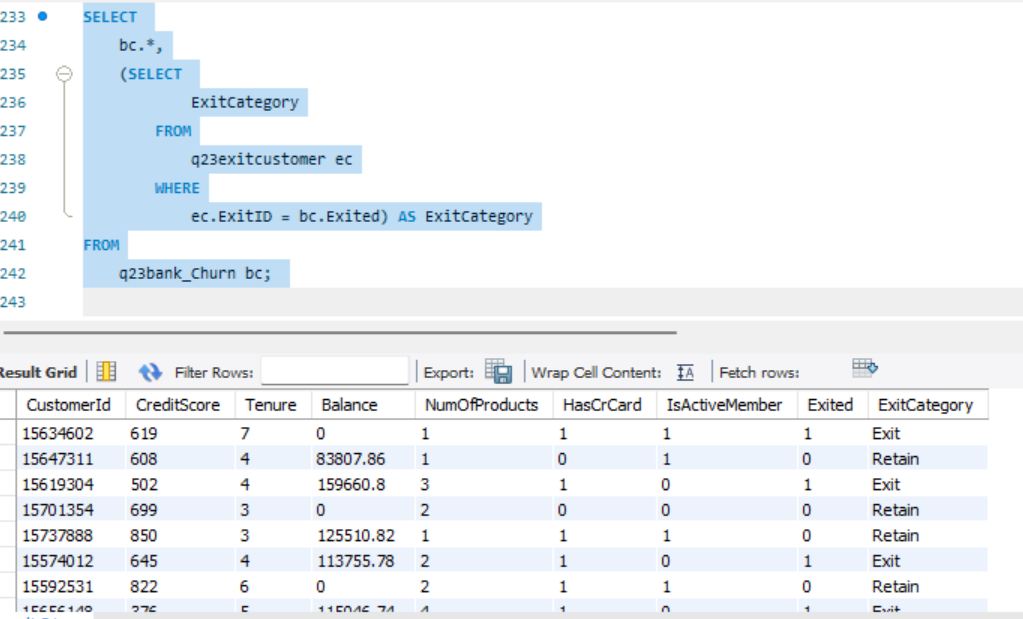
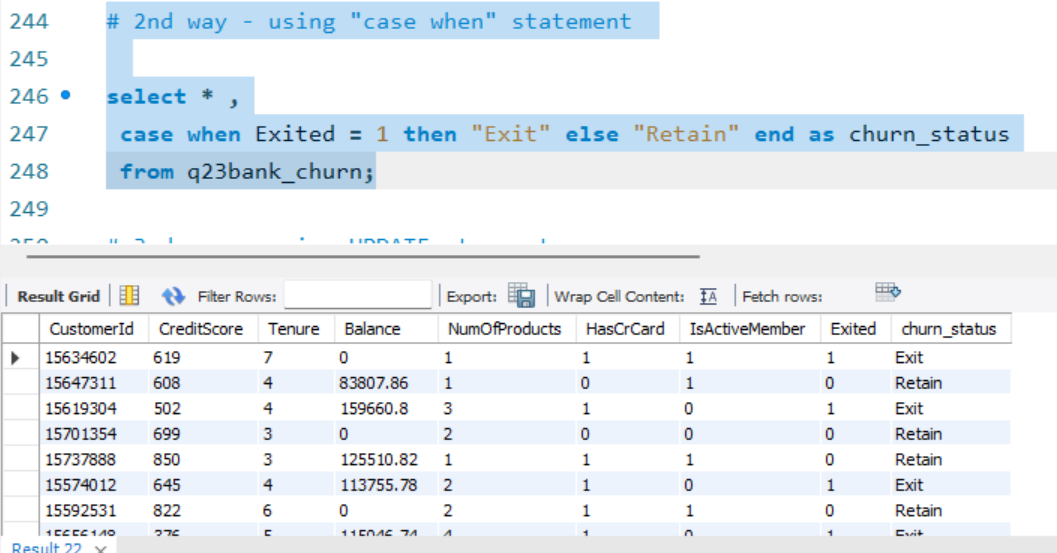
**Answer**

* **YES,** without using “Join” we get the “ExitCategory” from Exit Customers table to Bank Churn table.

**Explanation for data imported**

* I have pre-processed the data in excel before importing it on SQL database -
  + For this as there were 2 main data sheets and remaining 5 were sheets for nominal categorical data i replaced the categorical variables from sheets of "ACTIVE CUSTOMERS", "CREDIT CARD","EXIT CUSTOMER","GENDER" & "GEOGRAPHY" as per their accurate specifications into the main data sheet.
  + Hence for this Question-23 I have imported the needed tables of

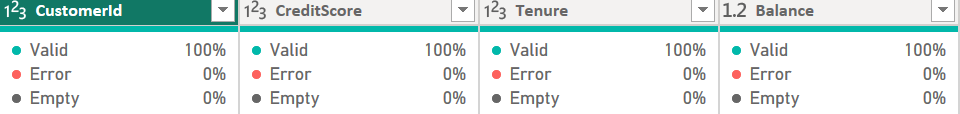
EXIT CUSTOMER & Bank\_churn by the name of "Q23exit\_customers" and "Q23bank\_churn" again in SQL database.

* ANSWER
* **USING CORELATED SUBQUERY**
* 
* **USING CASE WHEN**
* 
* UPDATE & IF state can also be used.

**Q.24**

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

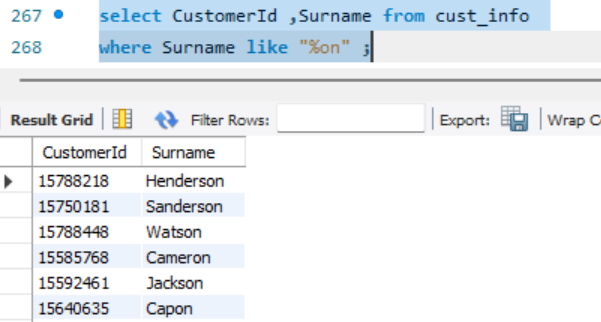
**Answer**

* **No,** there were no missing values in data.
* I used excel & power bi for looking up missing values.
  + - In EXCEL, we can use GoTo(ctrlG)>Specials>Blanks
    - In PowerBi ,in transform data we can observe if any value has error or is missing.
    - 

**Q.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”.

**Answer**

****

* Using wildcard statement the solution was obtained.

OBJECTIVE QUESTIONS END

**Subjective questions**

**Q.1**

Customer Behaviour Analysis:

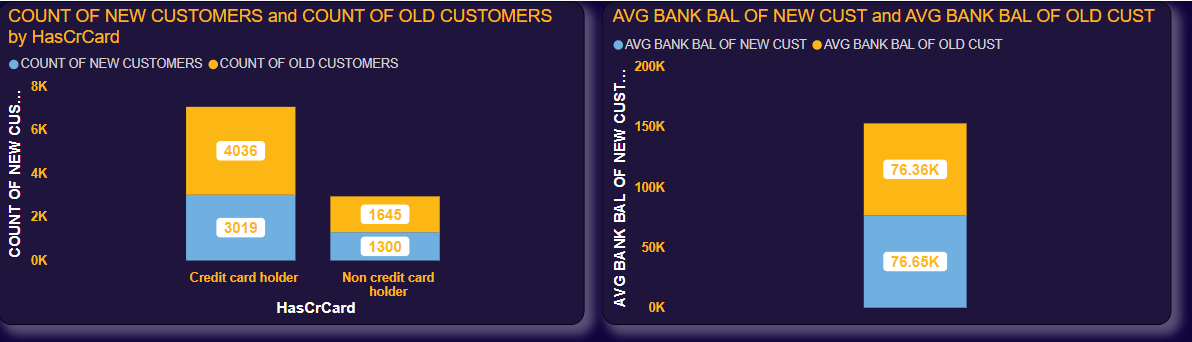
What patterns can be observed in the spending habits of long-term customers compared to new customers,

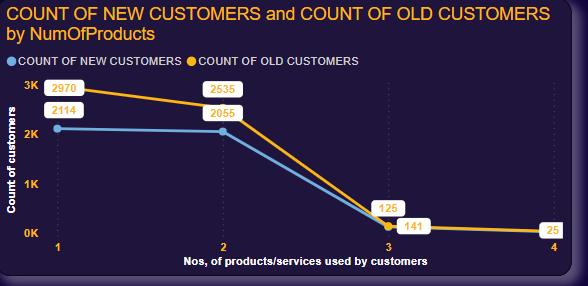
and what might these patterns suggest about customer loyalty?

**Answer**

* Direct information or data regarding the spending habits of customers are not available in the data, it would be beneficial for the analysis and drawing insights or suggestion if any information about spending habits or transaction frequency could have been presented.

Hence the analysis presented would be done pivoting upon the existing data and using it up to benefit.



* 

Spending habits estimation by **AVG BANK BALANCE**

* The average bank balance of the new customers i.e. the customers who has tenure less than equal to 4 years with bank has more or less same average bank balance than the average of balance of the old customers i.e the customers who has tenure greater than 4 years with bank is observed.
* This could indicate that both customer segments maintain similar savings habits, and financial behaviours, spending habits leading to consistent average bank balances.

Spending habits estimation by **COUNT OF CUSTOMERS HAVING OR NOT HAVING CREDIT CARD & ARE NEW OR OLD**

* Major number of old customers (4036) has availed the services of credit card then customers compared to the new customers (3019).
* More **Old customers** having credit card services suggests a **deeper level of engagement and loyalty to the bank** of old customers to bank compared to new customers.
* It may also be due to old customers may understand the benefits, rewards, and perks offered by the bank's credit card programs more deeply than the new customers hence spends and uses the credit card more.  
  **Suggestion**- Efforts and strategies should be derived to make the new customer more aware about the credit card facility from the bank.
* However ,while looking upon the advantages of having credit card ,it also comes with greater responsibility of risk management and appropriate use of credit card for the benefit.

Spending habits estimation by **COUNT OF CUSTOMERS USING NUMBER OF PRODUCTS & ARE NEW OR OLD**

* It can be seen that more old customers use the greater number of bank products & services compared to new customers.  
  The higher utilization of bank products and services by old customers suggests a deeper and more comprehensive financial relationship with the bank and a complex spending habits as they needs more number of products to handle their financial.
* This also indicated that new customers may show more concoious spending habits as they adjust to their financial equation, explore different purchasing behaviours, and know about the bank's products and services**.**

**Loyalty**

* **Old customers** having credit card & other services suggests a **deeper level of loyalty to the bank**.
* As more old customers utilizes multiple bank products and services are likely to have deeper and more established relationships with the bank. They have demonstrated a level of trust and confidence in the bank's offerings.

**Q.2**

**Product Affinity Study:**

Which bank products or services are most commonly used together, and how might this influence cross-selling strategies?

**Answer**

* Services offered by a bank are

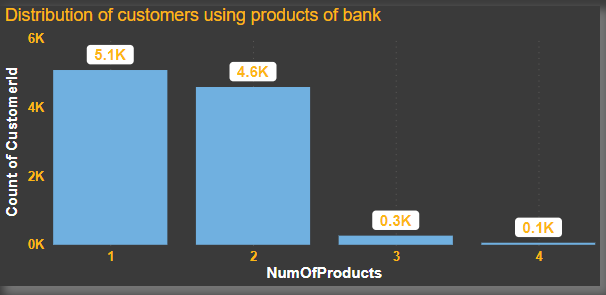
Different Types of Cards for money withdrawal

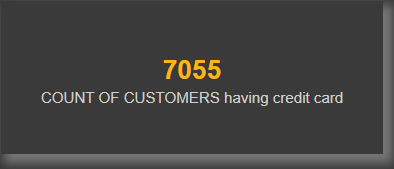
Various types of Loans

Types of accounts

Advisors and advice management

Products offered by the bank

* 
* Top Product that is credit card



* Product-1 that is internet banking – used by 5.1k customers
* Product-2 that is loans used by -4.6k customers
* Product-3 that is Fixed deposits.
* Product-4 that is safe deposit locker service.

**Services that are most commonly used together**

* Recently internet banking due to it’s ease of use and convenience is amongst the top most needed and used services/products by customers of bank.

INTERNET BANKING WITH SEVERAL PRODUCTS

* Combination of several products are used with internet banking like
* **Credit card functionality and maintenance with internet banking.**

Customers often use internet banking to monitor their credit card activity, make payments, and access additional services such as rewards redemption or account management.

* **Loans and Internet Banking**

Loans, such as mortgages, auto loans, personal loans, or business loans, are commonly managed through internet banking platforms. Customers use internet banking to check loan balances, make loan payments, set up automatic payments, and access loan-related documents and information.

* **Fixed Deposits and Internet Banking**

Customers who hold fixed deposit accounts may use internet banking to monitor their deposit balances, view maturity dates, renew or roll over deposits, or request changes to their deposit accounts.

* **Credit Card and Loans**

There are various schemes and offers where using credit cards, customer can take loans and EMI’s on various products in market depending upon the company.For ex. Apple provides EMI’s scheme in INDIA where if you purchase the apple products using particular credit card then you get varied EMI option as well as interest rate options and discounts as well.

* **Influence of these products and combinations of products on cross-selling strategies**
  + - Services and combinations of it offered by a bank and used by a customers can significantly influence cross-selling strategies by creating opportunities to promote additional products or services to existing customers.
    - Financial advisor/advising segment might be used to manage the credit card spending habits by customers and bank can charge the customers for the financial advisors, in turn yielding more revenue.
    - Loans and money lending can be directly done from internet banking with a high interest rate for urgent basis which would yield more profit/revenue.
    - Fixed deposits can be clubbed with other account benefits providing more interest rate to customers which would promote the customers to deposit more money in the bank.

**Conclusion**

* Most used combinations of products

Credit card functionality and maintenance with internet banking.

Loans and Internet Banking

Fixed Deposits and Internet Banking

Credit Card and Loans

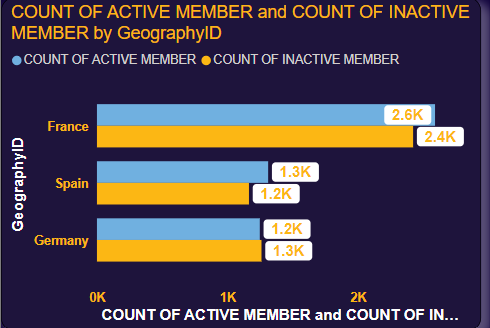
* Services and combinations of it offered by a bank and used by a customers can **significantly influence cross-selling strategies** by creating opportunities to promote additional products or services to existing customers.

**Q.3**

**Geographic Market Trends:**

How do economic indicators in different geographic regions correlate with the number of active accounts and customer churn rates?

**Answer**

* Economic indicators are key statistics that provide insights into the economic performance, health, and stability of countries.
* Statistics like GDP, employments rate, Inflation Rate, Stock Market Indices etc can be considered for checking the correlation with various bank metrics.
* 
* **Gross Domestic Product (GDP):**

GDP growth is often associated with increased economic activity, consumer confidence, and financial transactions. Regions experiencing higher GDP growth rates may see a greater number of active accounts as individuals and businesses engage in banking activities such as savings, investments, and borrowing and lower churn rate. Conversely, regions with slower GDP growth rates may experience stagnation in banking activity and higher rates of account inactivity or higher churn rate.

* **Unemployment Rate:**

The unemployment rate affects the average income, spending habits, and financial stability. Regions with lower unemployment rates may have increases level of bank activity , leading to increased demand for banking services and a higher number of active accounts and lower churn rate.

And , regions with higher unemployment rates may experience lower consumer spending, reduced loan demand, and higher rates of account inactivity or higher churn rate.

* **Inflation rates**

Inflation influences purchasing power, interest rates, and consumer behaviour.

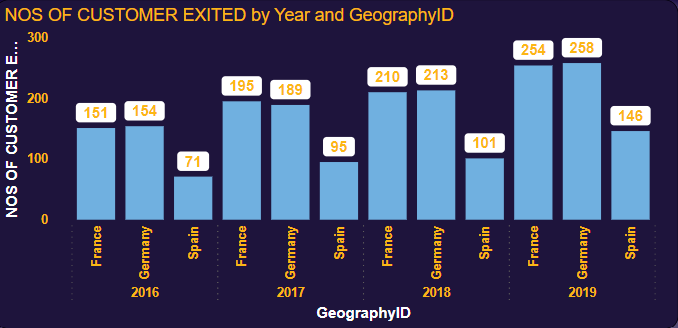
Regions having low inflation rates may see good consumer & customer behaviour, contributing to increased banking activity and the number of active accounts lower churn rate.

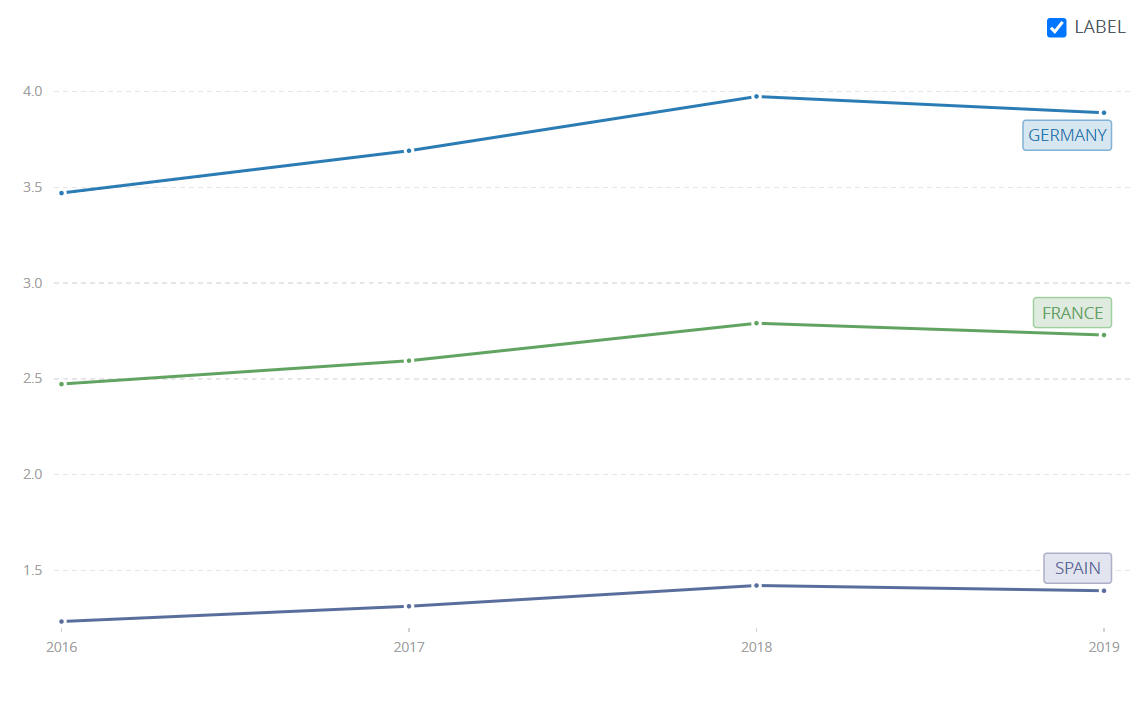
In constrast, regions with high inflation rates may experience lower purchasing power, lower loan demand, and higher account inactivity as customers prioritize essential spending firsts over banking activities.

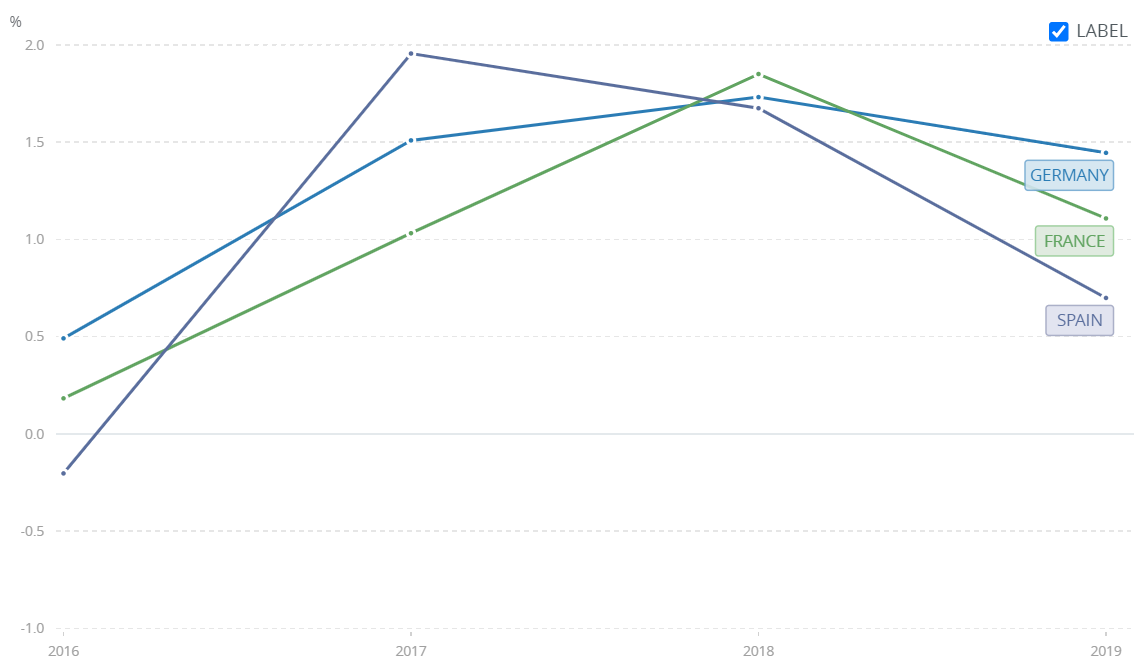
* **New/Updates in Government Policies or rules**

Government policies, and economic development initiatives or new/updates in rules has impact on regional economic growth and banking activity.

When the policies updates or are derived in favour of the customers which benefits customers ,yields greater economic outcomes thus it may increase the bank activity by customers in country.

* 



* Above is the graph for GDP for countries of France,Germany,Spain.  
  In 2019 GDP of France,Germany,Spain lowered than 2018 hence we can see a increase in churn rate.
* Following is the graph of inflation of the countries .
* Due to the inflation rates in the graph we can observe the churn rate being increased in the year where the inflation rate is more and vice versa.

**SUGGESTIONS**

* Some strategies can be followed to manage the Churn rates & accounts activity during economic fluctuations.
* Increased & transparent customer communication.
* Innovative & flexible loans payment options ,low balance accounts allowance can be planned to retain customers.
* Digital banking solutions.
* Providing Risk Management and Security awareness to customers.

**CONCLUSION**

A dependable significant correlation can be seen between churn rate & active accounts to the economic indicators. Which means churn rates & accounts activity are dependable upon the fluctuations of economic indicators in countries.

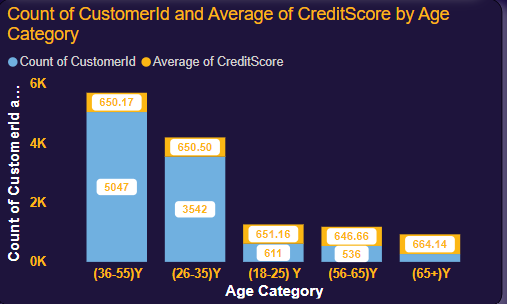
Strategies can be planned for enhanced retaining of the customer base in banks by providing tailored services.

**Q.4**

**Risk Management Assessment:**

Based on customer profiles, which demographic segments appear to pose the highest financial risk to the bank, and why?

**Answer**

* There are certain demographic segments which are correlated to financial risk to bank such as:
  + - * + Lower aged (young customers (<25 years)) & Old aged (elders/retired(>50yrs)) customers.
        + Customers with below average income or unstable/unemployed/freelancing job.
        + Customers with below average credit score.
* 
* Above is the chart for Age category wise numbers of customers as well their average of credit score. However , from our data it looks like the risked age category’s of <25 years & >50 years also has a considerably good average credit score.
* Highest financial risk posers can be
* **Lower aged (young customers (<25 years)) & Old aged (elders/retired(>50yrs)) customers because for** customers <25 years has Limited credit history makes it challenging to assess creditworthiness and predict repayment behaviour.

Lack of financial experience and awareness may lead to higher rates of delinquency or default on loans and credit obligations.

Dependence on student loans or entry-level jobs with lower income levels may result in higher debt-to-income ratios and financial instability.

* **Elderly or Retired Individuals because** majority ofseniors or retirees lives on fixed incomes, such as pensions, Social Security benefits, or retirement savings.
* **Individuals with Poor Credit score** because poor credit scoresindicates a history of late payments, defaults, or bankruptcy, indicating past financial difficulties.
* Customers with below average income or unstable /unemployed/freelancing job because such customers has variable income streams which is not a fixed income to track upon and irregular cash flows which can create irregularities in fulfilling bank requirements on time.

**Suggestions**

* If on the most common demographic segments including above mentioned segments are identified and strategies are made for sustainably tailor segment specific bank experience then it can help in the overall retention rate of customers with a result of risk management and avoiding bad after results and experience for both bank and customer.
* Targeted risk assessment techniques such as
  + - Risk assessment process –

Risk assessment processes involves assessing the potential risks before lending credit to a particular borrower or customer segment.

This assessment would help banks to forecast the chances of repayment and the level of risk exposure they are willing to accept.

* + - Credit scoring models –

These models analyse various factors, such as credit history, payment behaviours, debt levels, length of credit history, and recent inquiries, which generates a credit score that estimates the creditworthiness. Model Ex. FICO Score and Vantage Score.

These models help banks identify low-risk and high-risk borrowers, determine appropriate interest rates and loan terms, and allocate resources efficiently to manage credit risk.

* + - Risk-based pricing mechanisms

Risk-based pricing refers to the practice of setting interest rates and fees for loan products based on the perceived riskiness of the borrower in which high risked customer would be charged high interest rates & vice versa for low risk customer.

* Providing financial education, awareness, and prompt support & services tailored to the specific needs of these segments can help customers to make informed financial decisions and improve or maintain their financial demographics.
* Such techniques can be implemented which can help the bank manage and mitigate financial risks associated with these demographic segments while retaining all the customers.

**Conclusion**

* + - * + Lower aged (young customers (<25 years)) & Old aged (elders/retired(>50yrs)) customers.
        + Customers with below average income or unstable/unemployed/freelancing job.
        + Customers with below average credit score

are the demographic segments appear to pose the highest financial risk to the bank.

**Q.5**

Customer Lifetime Value Forecast:

How would you use the available data to model and predict the lifetime (tenure) value of different customer segments?

**Answer**

* Lifetime Customer Value (LCV), is a metric used by businesses to estimate the total revenue or profit generated by a customer or a segment over the entire duration of their relationship with the company.
* Lifetime customer value is a crucial metric for understanding the long-term value of customers to the business and guiding strategic decision-making in areas such as marketing, customer acquisition, retention, and customer relationship management.
* Predicting the lifetime (tenure) value of different customer segments is to predict the revenue generated by various customer segments like customers segments of poor to excellent credit score ,or age segments, or tenure segment i.e how much revenue can be /is generated by each segment.
* For this we can use the formulae of segment lifetime value i.e   
  Segment life time value = Segment purchase value \* segment purchase Frequency \*

Segment lifetime

* + Segment life time value- is the total expected revenue or profit generated by a customer segment over the entire duration of their relationship with the bank.
  + Average purchase value (APV) for a specific segment would be the profit generated from each segment. As we don’t have any data related to profit metrics we would assume the Average profit generated by a customer in a typical bank which is 13.9% which would be the total revenue generated by the segment and divde it by total numbers of deposits or transactions made by the customers in that particular segment.
  + Average purchase frequency(APF) for a specific segment would be the average number of transactions made by each customer within the segment during the specified time period. But we don’t have the transaction data in the dataset.
  + Segment Lifetime refers to the average duration of the customer relationship within the segment.

**Conclusion**

Calculating Customer Lifetime Value (CLV) would offer several benefits  
by helping in **Strategic Decision-Making,Resource Allocation,Retention Strategies,Cross-Selling and Upselling Opportunities,Risk Management**

Which would help the bank better manage and lower the bank churn rate and identify profitable customer segments and target them with tailored services.

**Q.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?

**Answer**

Various logical & analytical tools can be used for assessing the marketing campaigns such as:

Extra information would you need to solve this

**Case 1 (Before-and-After Comparison)**

* Marketing campaign to be assessed is already implemented
* In this case I would need additional data such as:
* Transactions data,
* Duration & time of the campaign start & end.
* Customer data post half time past the campaign start date.
* Deposits frequency & amounts of deposits.
* Data from loan departments (Pre campaign & post campaign both)
* Data from investment departments (Pre campaign & post campaign both)
* Data from IT departments regarding LOGINS, NEW USER ACCOUNTS CREATE COUNTS, CLICK RATES.
* Data of feedbacks & customers surveys.
* For assessing the impact:
  + Average of deposits ,if increased or decreased or are stagnant can be assessed on the time duration of data of before implementation of marketing campaign , if the average of deposits increases which would mean that the marketing was effective and on the % of increase of deposits % of effectiveness could be calculated.
  + Customer acquisition could be a good estimator for looking up of assessment if the marketing campaign worked. If the customer joining the bank increases after the release of campaign or there is a sudden jump or increase in numbers of customers then it can be said that marketing campaign proved beneficial.
  + Like-wise if the loan applications, FD’s accounts, online banking accounts increases then the effectiveness of the marketing campaign can be predicted.

**Case 2** (**Control group comparison**)  
If the impact on marketing campaign to be assessed is planned to be implemented and prediction has to be made upon the impacts that a type of campaign would make

* Divide customers into two groups: one exposed to the marketing campaign (treatment group) and one not exposed (control group).
* Compare retention and acquisition metrics between the two groups over a specified time period to measure the incremental impact of the marketing campaign.
* **Segmentation Analysis**:

Segmenting customers based on demographic, behavioural, or age wise.

Assess the impact of the campaign on different customer segments by comparing retention and acquisition metrics within each segment.

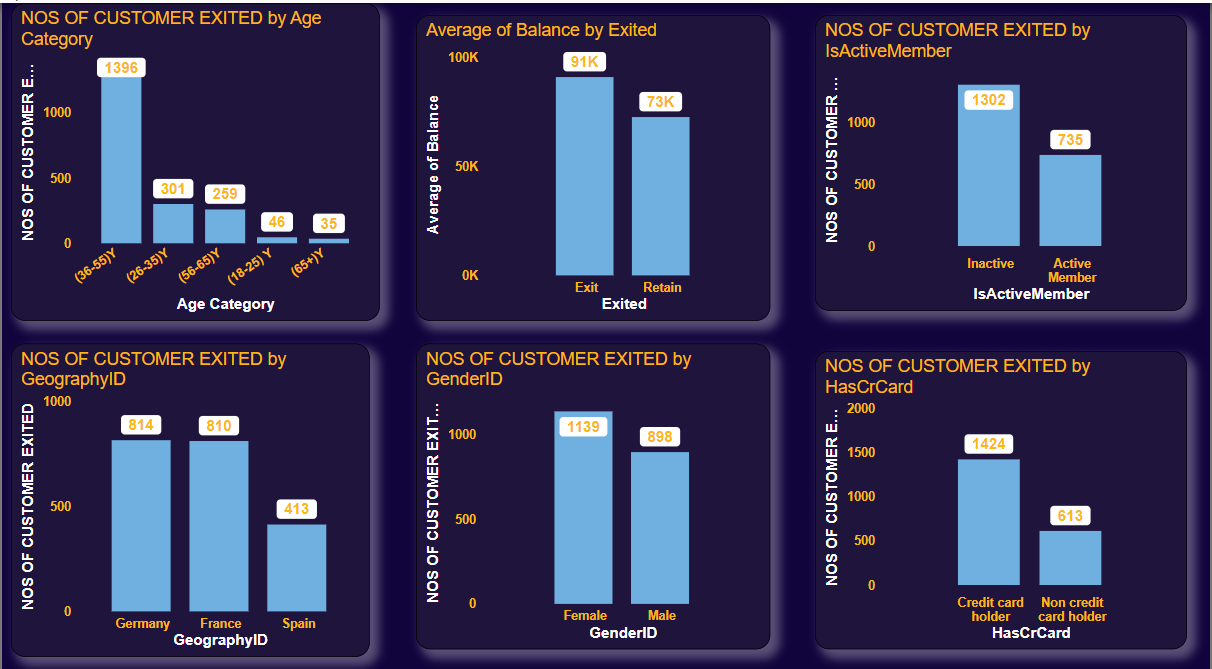
Identify segments that show the most significant improvements in retention or acquisition rates because of the campaign.

**Q.7**

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

**Answer**

* Common characteristics or trends among customers who have exited  
  can be identified through pivoting the numbers of customers left and analysing the characteristics they belonged to.
* This would help us exploring reasons for leaving by the means of the identifying the characteristics they belonged to.
* Below are the charts for of characteristics and trends of exited customers.

****

* Large number of customers exited belongs to the age groups of 36-55 years. This segment comprises individuals who are in the prime of their careers, raising families, or planning for retirement.

Key characteristics: Established credit history, stable income levels, home ownership, retirement savings, family financial responsibilities.  
This can be due to the lack of prompt services from bank or lack of additional benefits from bank end as this age group generally wants higher benefits and rewards as well prompt service management for their financial needs.

* Average bank balance of the customers who have exited is greater than who has retained.

This can indicate that the customers had higher financial services needs which could not be fulfilled by the bank.

* Germany & France has the greater numbers of customer leaving the bank than Spain. This can be the indicator of lack of services at the offices and services at Germany & France.
* Female has the greater numbers leaving the bank, which can indicate that there can be lack in female specific services offered through the bank.
* Customers having Credit card has the greater numbers of customers leaving the bank than customers not having credit card. This can be due to higher charges on credit card transactions, or higher interest rates, or lack of offers on credit cards which other competitive banks may be providing.

**Suggestions**

* Services specific to age group of 36-55 should be tailored more and more benefits shall be tried to offer.
* Strategy regarding managing complex financial needs shall be planned and tailored.
* Monitoring and enhancement shall be focused more on locations of Germany & France.
* Female centric services shall be updated and offered more.
* Credit card scheme shall be managed and restructured which would provide more benefits to customers.

**Conclusion**

Common characteristics such as female gender ,age groups was identified and trends such as customers having credit card and from Germany & France was identified among customers who have exited.

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

**Q.8**

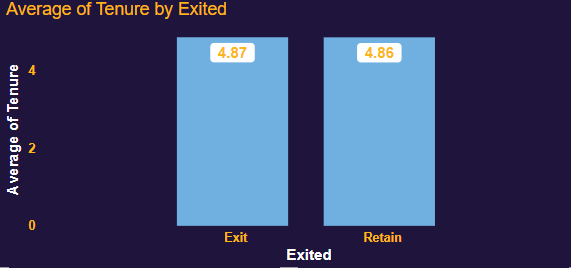
Are 'Tenure', 'NumOfProducts', 'IsActiveMember', and 'EstimatedSalary' important for predicting if a customer will leave the bank?

**Answer**

* **Tenure**

Yes, tenure of a customer is amongst important factors for predicting whether a customer will leave the bank because Customers with longer tenure may have established relationships with the bank, making them less likely to leave.

Shorter tenure may indicate a higher likelihood of churn, as new customers may not have developed strong ties to the bank yet.

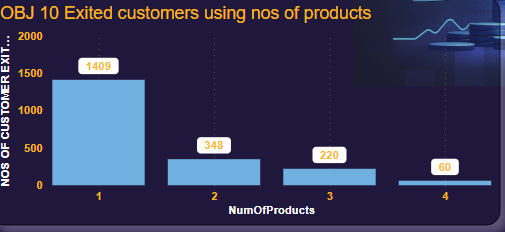
****

However, the average tenure of customers leaving & still with bank is same which indicates there can be gaps and lack in services provided by the bank.

* **Number Of Products used by customers**

Customers who use multiple services, such as checking accounts, savings accounts, loans, credit cards, and investment products, may be more deeply integrated into the bank's ecosystem.

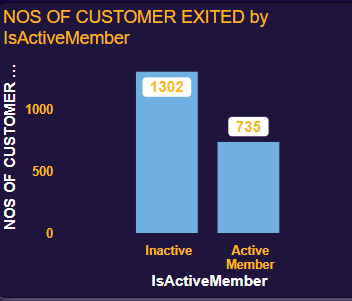
Higher engagement with multiple services could indicate stronger loyalty and lower churn risk.



Customers using more numbers of products are amongst the least to leave the bank & vice versa can be seen in the dataset.

* **Is customer Active or not**

Active members are generally less likely to churn compared to inactive members, as they continue to derive value from their relationship with the bank.



From the graph it can be seen that less numbers of Active members are there leaving the bank compared to inactive members.

* **Salary**

Higher salary levels may indicate greater financial stability and ability to maintain banking relationships.

Customers with higher salaries may be less sensitive to fees or pricing changes and less likely to switch banks for financial reasons.

However in the available bank dataset the average salary of customers left and still with the bank are more or less the same which may be the indicator of lack of services from bank and lower customers satisfaction.

* **Suggestions**
* Innovative strategy and marketing shall be implemented to keep the customers engaged with the bank for long tenure.
* Marketing shall be done & customer shall be educated so as to provide & as much as benefits and rewards possible which would promote the customer to use more numbers of products.

**Conclusion**

* **Yes,** Tenure, Number Of Products customer is using, is customer active or not, and Salary of customers are important and have correlation for predicting if a customer will leave the bank.

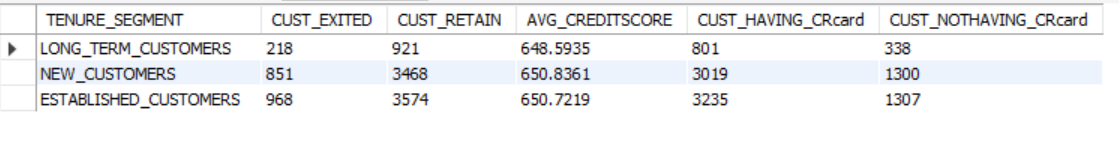
Are 'Tenure', 'NumOfProducts', 'IsActiveMember', and 'EstimatedSalary' important for predicting if a customer will leave the bank

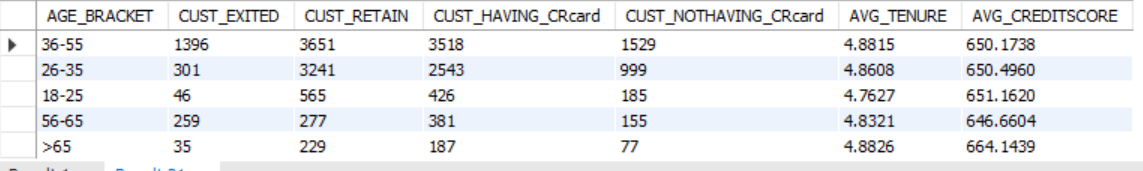
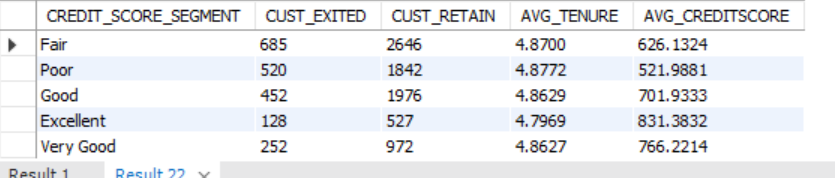
**Q.9**

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

**Answer**

* **Segmentation as per tenure**

****

* **Segmentation as per tenure**
* **Segmentation as per tenure**

**Conclusion**

Customers are segmented based on demographics and account details.

**Queries are present in SQL script file.**

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

**Q.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?

**Answer**

* Conditional formatting setup to visually highlight customers at risk of churn and to evaluate the impact of credit card rewards on customer retention can be done in EXCEL as well as POWERBI.
* 
* However more data needed on credit card rewards to figure out the retention rates.

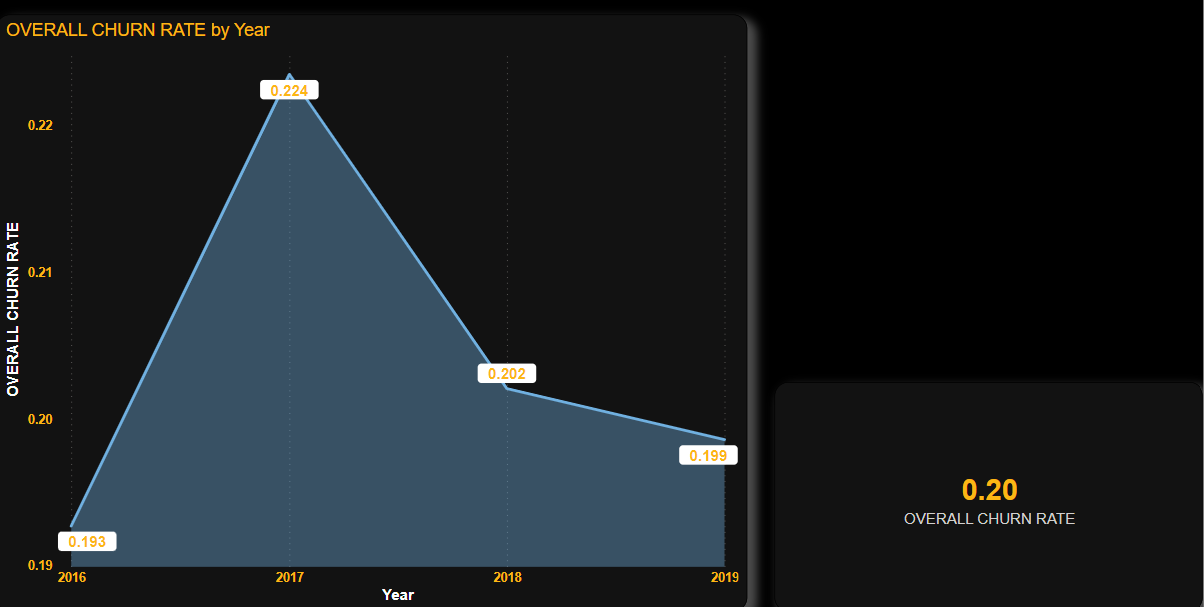
Are 'Tenure', 'NumOfProducts', 'IsActiveMember', and 'EstimatedSalary' important for predicting if a customer will leave the bank

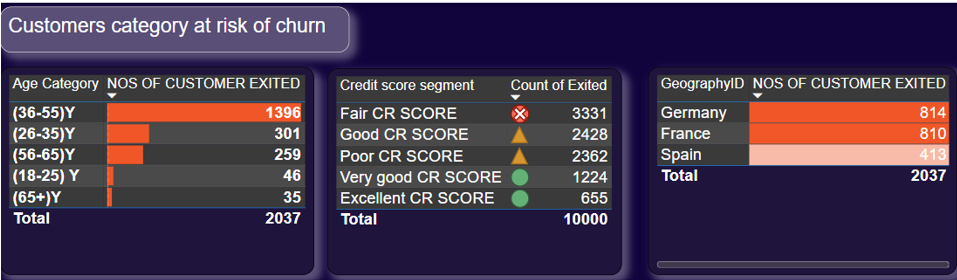
**Q.11**

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.

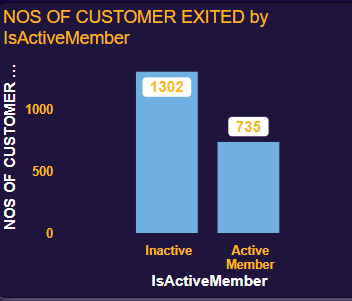
**Answer**

The churn rate/year is as follows



* 2017 is observed to be having a highest churn rate and 2016 to be having the lowest churn rate.
* Overall churn rate is found to be 20.37
* 
* Customers within the age range of 30-50 ,when they are not given optimal support & services tends to quickly shift to other entity churning the bank.

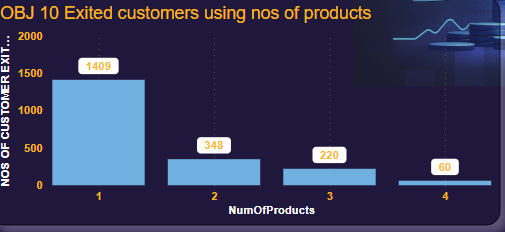
Active members are generally less likely to churn compared to inactive members, as they continue to derive value from their relationship with the bank.



From the graph it can be seen that less numbers of Active members are there leaving the bank compared to inactive members.

* Customers who use multiple services, such as checking accounts, savings accounts, loans, credit cards, and investment products, may be more deeply integrated into the bank's ecosystem.

Higher engagement with multiple services could indicate stronger loyalty and lower churn risk.



Customers using more numbers of products are amongst the least to leave the bank & vice versa can be seen in the dataset. Hence the customers using just 1 product from bank are more likely to leave the bank.

**Strategy for reducing churn rate**

* Reducing churn rate is very important for a well-functioning bank, customer base and sustaining business growth.

Here are **several strategies** that banks can implement to decrease churn rate:

* + **Tailored customer service –**

Provide exceptional customer service in all mediums including digital as well physical

Training of staff to be attentive, empathetic, and responsive to customer needs and concerns.

Resolve customer issues promptly and efficiently to enhance satisfaction and loyalty.

* + Planned product offerings and rewards which would benefit the customer as well as the bank profitability.
  + Offer competitive interest rates, fees, and features to attract and retain customers
  + **Monitor Customer Satisfaction**:

Feedbacks & surveys shall be taken very seriously and if any issue arises it should be resolved quickly.

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

**Q.12**

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

**Answer**

****

* **KPIs and visualization related metrics & slicers were included in dashboards**

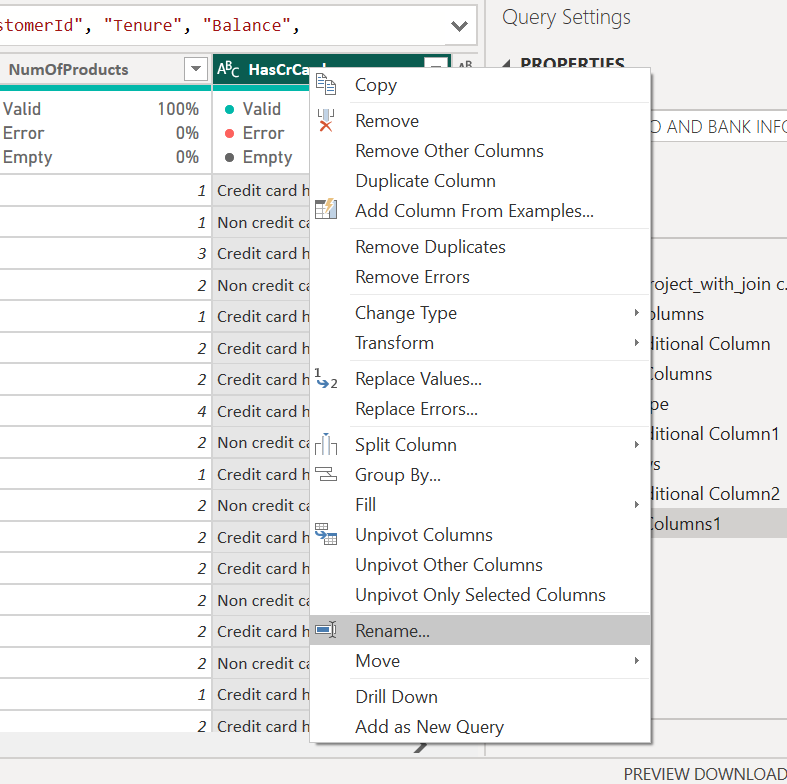
Are 'Tenure', 'NumOfProducts', 'IsActiveMember', and 'EstimatedSalary' important for predicting if a customer will leave the bank

**Q.14**

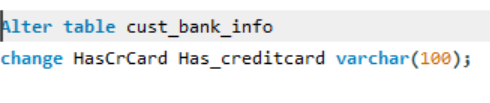
In the “Bank\_Churn” table how can you modify the name of “HasCrCard” column to “Has\_creditcard”?

**Answer**

* It can be changed in power bi power query editor directly right clicking upon the column and renaming the column.

****

* As well it can be changed through SQL query pasted below

****

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

**Q.13**

How would you approach this problem, if the objective and subjective questions were not given?

**Answer**

**Data reading**

* I would first thoroughly read the unclean data and understand the data. This helps to get a general idea about how the data structure is.

**Data cleaning**

* Post this I would gain a rough idea upon what design and structure and design in a overall manner the data has i.e what type of information is present in the data ,what are the data types present in the data and how manipulating them can make the data more readable and be made more information giver by changing the design of the data for ex First name and Last name is given is given in two different columns so I would concatenate their full names to make it more readable, if date and times are present in a single columns I would make it separate for the later convenience to use it in analysis.
* I would have merged the data in power bi and then query in SQL so as to get a cleaner and readable query for the further team and eliminate unnecessary join statements in SQL query’s.

**Approach for attempting this problem without questions**

**Getting the understandings of data by extracting, pivoting, relating it to available metrics.**

* After than I would get all the key metrics such as how much countries are there which is needed to be analysed, how much customer base are there, how much revenue is being generated in total etc.
* Then I would segregate the data earlier identified and relate it with numerous segments that is the total revenue generated, distributions of bank deposits, divisions of males and females, churning rate, customer satisfaction present and available by the countries present in the data set.
* Like this I would pivot all the data according to various metrics and segments which can be already available or derived like segments such as age groups, credit scores etc and again relating them with aspects of banks such as churn rates, low or high satisfactions score etc.

Also I would try to relate the available data to the metrics which are not available in the data like GDP’s ,inflation of countries etc.

* **Estimation of the probable problem statements & issues banks can face.**
* **Problems statements can be seen derived from the pivoting of data.**

Common and global entities have interrelated problems which are point of concerns such as low cash flows, customer churnings stagnant growths etc.

Also, many patterns would be visible while pivoting the data

For e.g. Females could be seen churning more compared to males   
any specific countries would have low average bank balances compared to other cities.  
Through this problems and segment wise issues could be identified which would be leveraged to extract meaningful conclusions and suggestions to tackle the problems in a structured manner in order to find solutions on the objective that is to be achieved.

SUBJECTIVE QUESTIONS END