**** Bank CRM-Analysis

**by Yash Kumar**

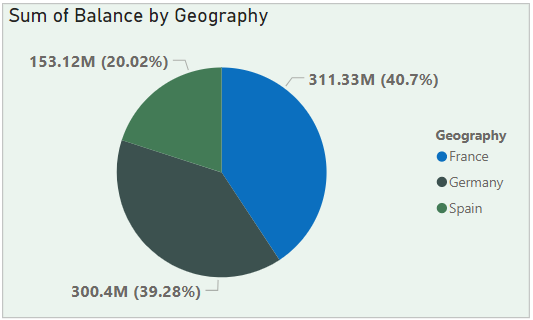
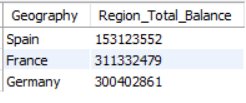
**Objective Questions:**

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

Currently there are three regions in the dataset**:** **France**, **Germany** and **Spain**.

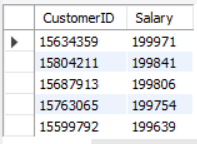
| **Region** | **round Total Balance ($)** |
| --- | --- |
| Spain | 153.12M |
| Germany | 300.4M |
| France | 311.33M |

To find distribution of account balances across regions, we calculate the sum of balance and group by regions (Geography).



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

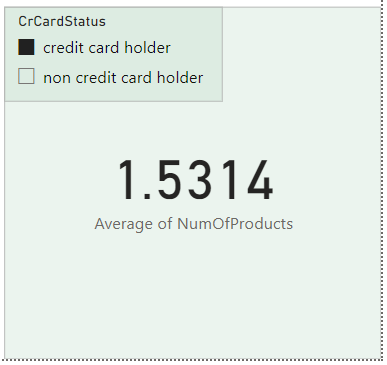
To find top 5 customers with highest Est.Salary, we sort data by Salary in descending order, then limit the rows to 5.



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

**The average number of products used by customers who have a credit card** = **1.53**

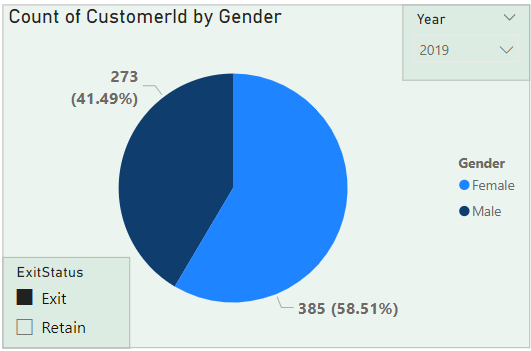
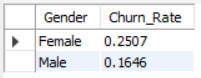
To get this, we calculate the average **NumOfProducts** used by customers, after filtering data by customers who are **CreditCard holders**



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

| **Gender** | Total\_Counts | Exit\_Counts | **Churn\_Rate** |
| --- | --- | --- | --- |
| **Female** | 1537 | 385 | **0.25** |
| **Male** | 1776 | 273 | **0.16** |

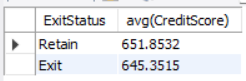
To find this, we calculate the average number of Churns(**Exited**) for each Gender using **group by**.



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

| **Exit Status** | **Avg Credit Score** |
| --- | --- |
| Retain | 652 |
| Exit | 645 |

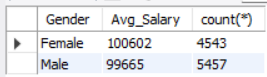
For this, we simply calculate average **CreditScore** of customers while grouping them by **ExitStatus**



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

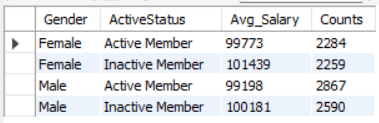
**Female** customers have the higher average estimated salary.

| **Gender** | **Avg Salary** |
| --- | --- |
| Female | 1,00,600 |
| Male | 99,660 |



**No**, there’s no strong correlation with number of Active members and Average  
 estimated salary.

| **Status** | **Gender** | **Active Counts** | **Avg Salary** |
| --- | --- | --- | --- |
| Active | Female | 2284 | 99,770 |
| Male | 2867 | 99,120 |



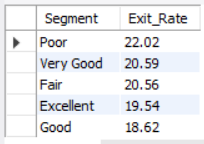
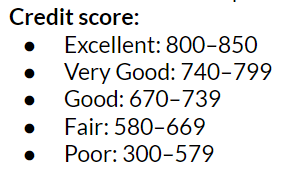
1. Segment the customers based on their credit score and identify the segment with the highest exit rate. (SQL)

Segment with the highest exit rate is **Poor (300-579)** of Credit Scores.

| **Segment** | **Exit Rate** |
| --- | --- |
| Poor | 22.02 |
| Very Good | 20.59 |
| Fair | 20.56 |
| Excellent | 19.54 |
| Good | 18.62 |

For this,

* we group by customers based on credit score
* count total customers of that segment
* count total exited customers of that segment
* and take ratio of it, we get exit rate

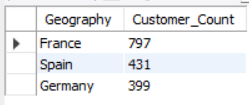


1. Find out which geographic region has the highest number of active customers with a tenure greater than 5 years. (SQL)

**France** has the highest number of active customers (with tenure > 5 year)

| **Geographic Region** | **Customer Count** |
| --- | --- |
| France | 797 |
| Spain | 431 |
| Germany | 399 |

For this, we calculate the count of customers with tenure>5 and group by  
 geography.

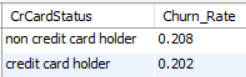


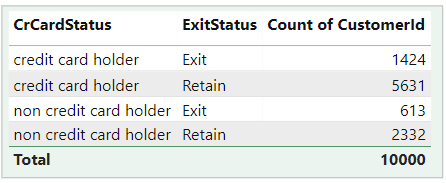
1. What is the impact of having a credit card on customer churn, based on the available data?

Result**:** No significant impact seen.

| **Card Status** | **Churn Rate** |
| --- | --- |
| Non credit card holder | 0.208 |
| Credit card holder | 0.202 |

Based on data we can see having a Credit card doesn’t have a significant impact on churn rates - **It’s just slightly better**.



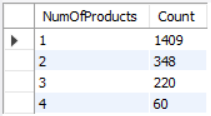


1. For customers who have exited, what is the most common number of products they have used?

For Customers who have **Exited,**

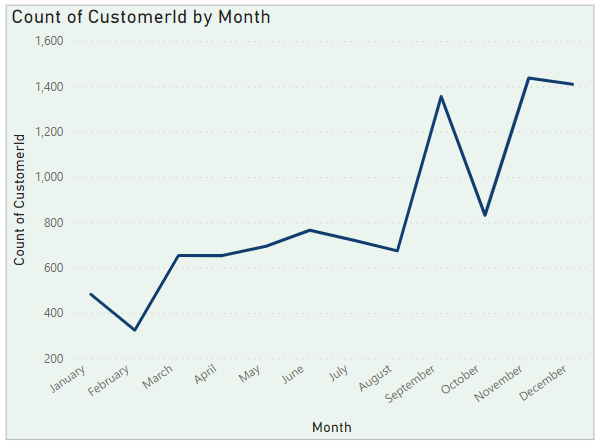
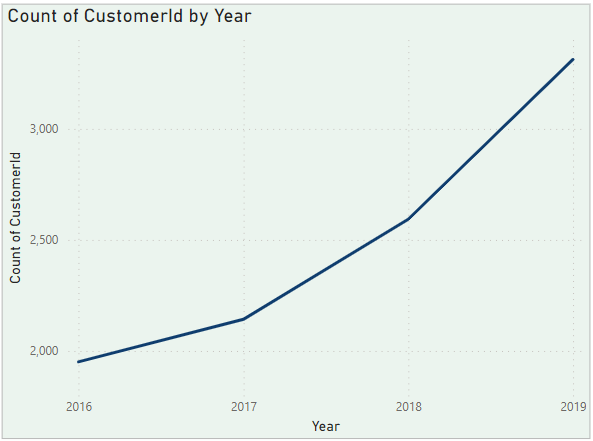
Most common NumOfProduct is **1**

| **Number of Products** | **Count of Customers** |
| --- | --- |
| 1 | 1409 |
| 2 | 348 |
| 3 | 220 |
| 4 | 60 |



1. Examine the trend of customers joining over time and identify any seasonal patterns (yearly or monthly).

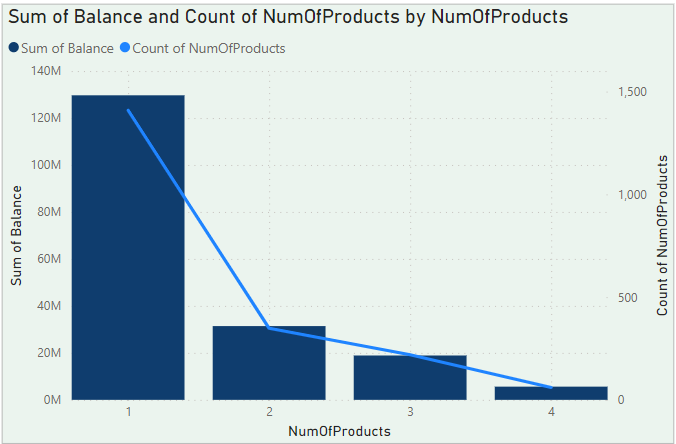
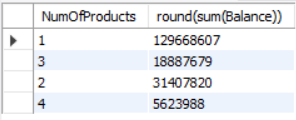
* **Yearly**  
  Count of customers is smoothly increasing over years, without any dramatic fluctuations.
* **Monthly**  
  Count of customers is overall increasing, but with some dramatic exits and the joining of customers during the last quarters of the year(Aug-Dec).



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

Relation that we can see between Num of Products and account balance of customers is:  
  
**Sum of Balance of customers decreases with increase in Number of Products** that indicates more customers use lesser number of products.

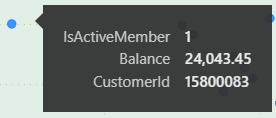
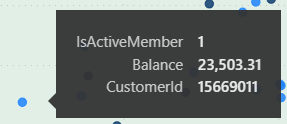
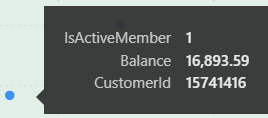
It was analysed on the **Sum** of balance rather than **Average** because average is almost the same for every segment because of similar count of customers.

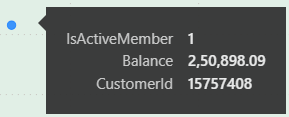
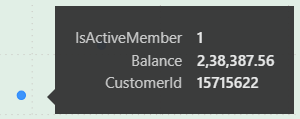


1. Identify any potential outliers in terms of balance among customers who have remained with the bank.

| **Balance Scale** | **Balance ($)** | **CustomerID** |
| --- | --- | --- |
| Lower End | 16,894 | 15741416 |
| 23,503 | 15669011 |
| 24,044 | 15800083 |
| Higher End | 2,38,388 | 15715622 |
| 2,50,898 | 15757408 |

Customers with balance at **Lower End**:

  
  
  
  
  
 Customer with balance at **Higher End**:

  
  
Scatter Plot chart of all customer’s Balances:



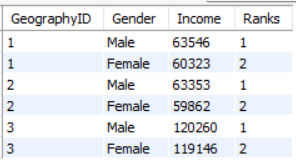
1. How many different tables are given in the dataset, out of these tables which table only consists of categorical variables?

* There are a total of 7 different tables in the **raw** dataset, out of which 5 are categorical ones.

| **Categorical Tables** |
| --- |
| Gender |
| Geography |
| ActiveCustomer |
| ExitCustomer |
| CreditID |

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).

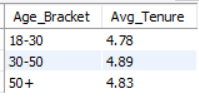
| **GeographyID** | **Gender** | **Avg Income** | **Geographic Rank** |
| --- | --- | --- | --- |
| 1 | Male | 63,546 | 1 |
| Female | 60,323 | 2 |
| 2 | Male | 63,353 | 1 |
| Female | 59,862 | 2 |
| 3 | Male | 120,260 | 1 |
| Female | 119,146 | 2 |



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+).

| **Age Bracket** | **Avg Tenure (years)** |
| --- | --- |
| 18-30 | 4.78 |
| 30-50 | 4.89 |
| 50+ | 4.83 |

For this, we first segment ages in bracket then calculate average tenure and  
 grouping by age bracket



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 correlation coefficient between Estimated Salary and Balance is approximately **0.0128**, as shown in the table below:



Achieved using **CORREL** function in Excel:

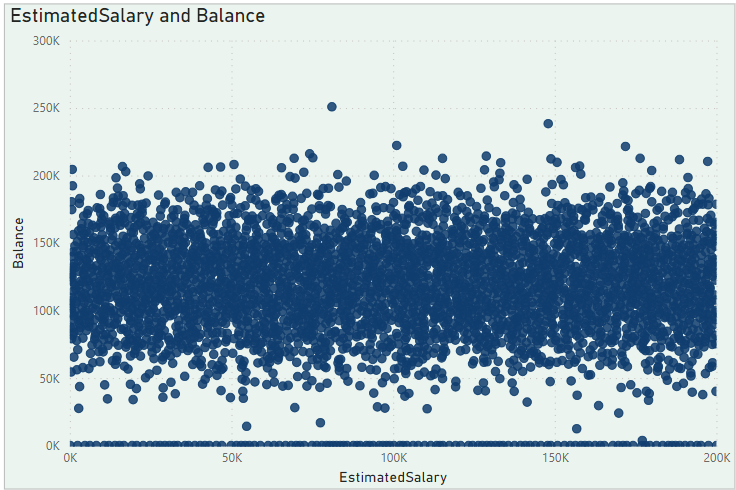




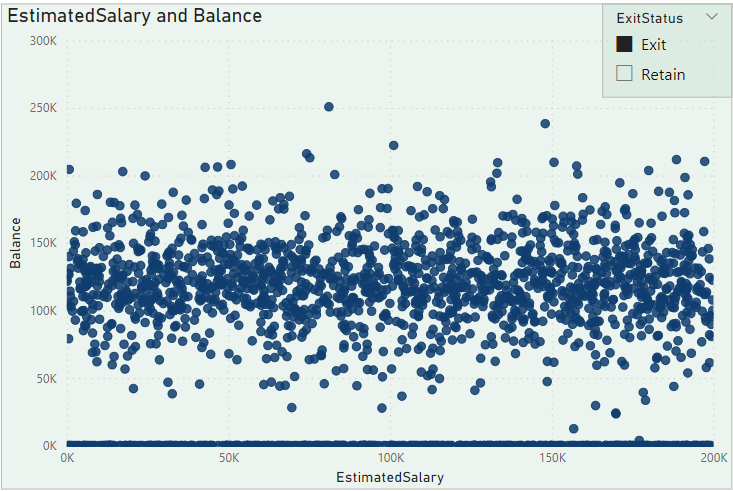
This value indicates a **very weak positive correlation** between the two variables, suggesting that there is no significant direct relationship between a customer's estimated salary and their account balance.

As we can see in the scatter charts below:

* No direct correlation of salary and balance is seen.
* The average balance (of all customers) lies in the range of **70K - 150K**, despite salary difference.



* As for the exited customers, the correlation is still **very weak positive**
* The average balance (of exited customers) lies in the range of **90K - 150K**, despite salary difference.

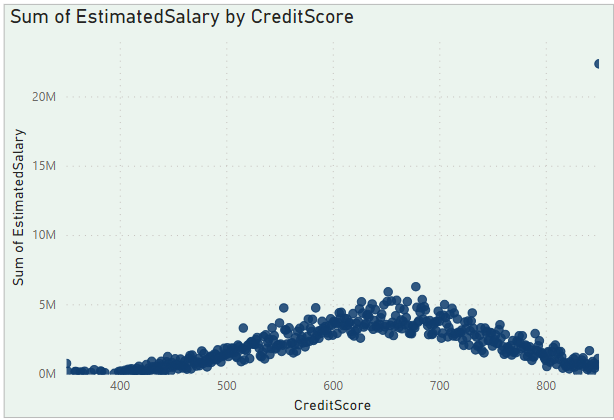


Both customer segments (with & without Exits) have similar average Balance spectrum.  
  
**Exit one is just less dense**.

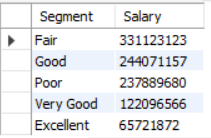
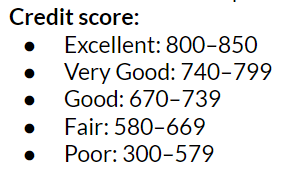
1. Is there any correlation between the salary and the Credit score of customers?  
     
   The correlation coefficient between Estimated Salary and Credit Score is approximately **-0.0014**

****  
 This value indicates a **very weak negative correlation** between the two  
 variables, suggesting that there is **no** significant direct relationship between a  
 customer's estimated salary and their credit score.

but there is **some pattern** we can see forming in the scatter chart below:



| **Segment** | **Sum of Salary ($)** |
| --- | --- |
| Fair | 331M |
| Good | 244M |
| Poor | 237M |
| Very Good | 122M |
| Excellent | 66M |



We can see a rise in sum of salary in middle end of curve, that suggests:

* SumOfSalary increases for mid ranges credit scores
* More customers with high salary lie in mid range of Credit Score, thus higher sum
* We can’t make direct correlation that higher credit score means higher salary because average is similar.

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

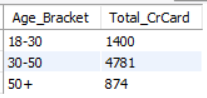
| **Rank** | **Segment** | **Churn Count** |
| --- | --- | --- |
| 1 | Fair | 685 |
| 2 | Poor | 520 |
| 3 | Good | 452 |
| 4 | Very Good | 252 |
| 5 | Excellent | 128 |



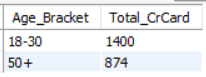
1. According to the age buckets find the number of customers who have a credit card.

| **Age Bracket** | **Total Credit Cards** |
| --- | --- |
| 18 - 30 | 1400 |
| 30 - 50 | 4781 |
| 50+ | 874 |

For this, we calculate the sum of Credit cards and then grouping by Age Brackets.



Also retrieve those buckets that have a lesser than average number of credit cards per bucket.

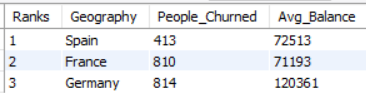


1. Rank the Locations as per the number of people who have churned the bank and average balance of the customers.

| **Rank** | **Geography** | **People Churned** | **Avg Balance** |
| --- | --- | --- | --- |
| 1 | Spain | 413 | 72,513 |
| 2 | France | 810 | 71,193 |
| 3 | Germany | 814 | 120,361 |

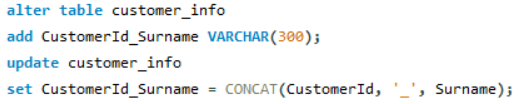
For this, we calculate the number of customers who have exited the bank and average the balance of them by grouping geography and ranking using the window function.

**Ranking is based on ascending order of no. of people churned.**

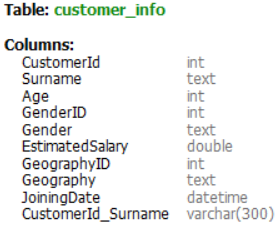


1. As we can see that the “CustomerInfo” table has the CustomerID and Surname, now if we have to join it with a table where the primary key is also a combination of CustomerID and Surname, come up with a column where the format is “CustomerID\_Surname”.

To achieve this, we write and execute the following command in SQL workbench:



And the CustomerID\_Surname will be added as last column in table as shown below:



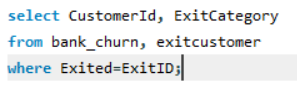
1. Without using “Join”, can we get the “ExitCategory” from ExitCustomers table to Bank\_Churn table? If yes, do this using SQL.

Yes, we can get the **ExitCategory** from ExitCustomers table to Bank\_Churn table without using **Join.**

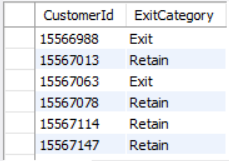
We can achieve this by:

* writing both tables in **from** clause
* using the **where** clause to join by common key

**SQL Query:**

****

**Output:**

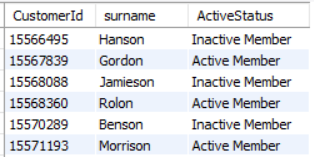
****(limited output shown)

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

* No, there weren’t any missing values in data.
* They were checked through Quality Columns in the View Tab, using Transform Data in PowerBI.

1. Write the SQL query to get the customer IDs, their last name, and whether they are active or not for the customers whose surname ends with “on”.

* For this, we will select customers with a filter (‘where’ clause in SQL) of their surname that ends with “on” and their active status.
* Not all customers are shown.
* Limited sample output has been shown due to large number of output rows (i.e. large number of customers have surname ending with “on”)
* We can see all customers in SQL Workbench after running the query.



**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?

Approach**:**

**Spending** is calculated as **new column** in table by difference of salary and balance of customers

i.e. **Spending = Salary - Balance**

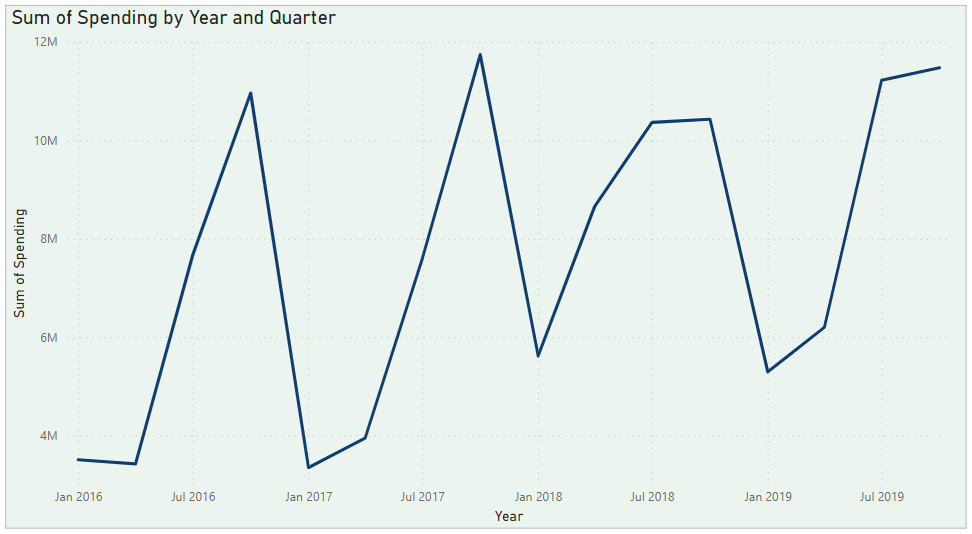
Pattern**:**

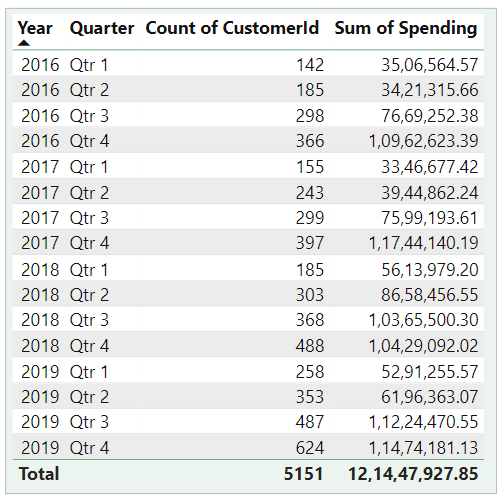
Pattern that can be observed from chart is:

* Sum of Spending of customers group peaks during the later quarter of year and drips down during the early quarter.

Reason**:**

* It’s due to many new customers **joining** the bank during the later quarter of year
* and many of them **exit** in the early quarter of year.





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

While analysing the number of products used by customers can offer a glimpse into product affinity, it's not enough to definitively identify which products are most commonly used together.

Here's why**:**

* **Missing Details:** We lack information on specific products (checking accounts, credit cards, etc.). This makes it impossible to pinpoint clear connections between them.
* **Usage vs Affinity:** Knowing the number of products used doesn't necessarily reveal which ones are used together. A customer might have a checking account, savings account, and safety deposit box, but that doesn't tell us if they use them in a complementary way.

However, this data can still be a stepping stone for further analysis:

* **Data Enrichment:** By incorporating details on specific products held and transaction history, you can conduct a proper market basket analysis, uncovering frequently purchased product bundles.
* **Segmentation:** Customers with a high number of products could be a target segment for deeper analysis to identify commonly used product combinations.

***In conclusion****, while the current data offers limited insights into product affinity, it highlights the importance of richer datasets for crafting effective cross-selling strategies that leverage customer behaviour.*

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

There is NO strong correlation between economic indicators and active accounts & churn rates of different regions.

Parameters taken**:**

* Est.Salary (*Average*)
* Balance (*Average*)
* Spending (*Salary - Balance*)
* Active Counts (*Sum*)
* Exit Counts (*Sum*)

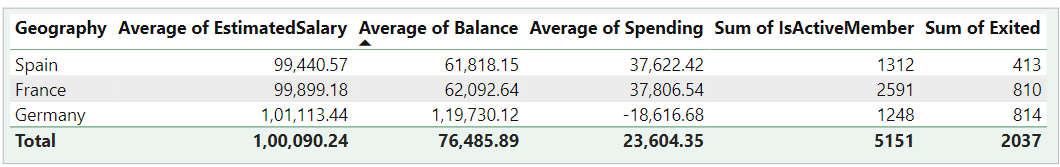
| **Region** | **Avg Salary** | **Avg Balance** | **Avg Spending** | **Active Count** | **Exit Count** |
| --- | --- | --- | --- | --- | --- |
| **Spain** | 99.4K | 61.8K | 37.6K | 1312 | 413 |
| **France** | 99.9K | 62.1K | 37.8K | 2591 | 810 |
| **Germany** | 101.1K | 119.7K | **-**18.6K | 1248 | 814 |

Insights that we get from the table above:

* There is both a higher active count & higher exit ratein **France** despite the fact that avg spending of France and Spain is almost equal.
* Similarly, there is both a lower active count & lower exit rate in **Spain** despite the fact that avg spending of France and Spain is almost equal.
* **Germany** has the lowest active count & highest exit rate.
* Both **France** & **Spain** customers are spending more money than saving
* **Germany** customers seems to be saving more money than spending (*as Avg Spending value is negative)*

| **Region** | **Spending** | **Saving** | **ActiveCounts** | **ExitRate** |
| --- | --- | --- | --- | --- |
| **Spain** | More | Less | Low | Low |
| **France** | More | Less | High | High |
| **Germany** | Less | More | Low | High |

Thus, we can say there’s no correlation of these parameters through an economic indicator.

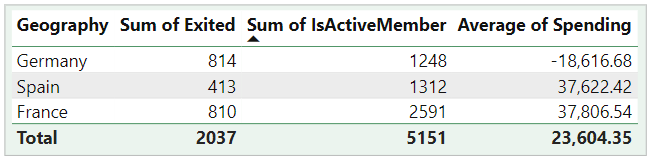


1. **Risk Management Assessment:** Based on customer profiles, which demographic segments appear to pose the highest financial risk to the bank, and why?

**Germany** appears to pose the highest financial risk to the bank**.**

Reasons**:**

* **Highest Churn Rate**Which means more loss and less benefit to the bank in terms of money and loyalty.
* **Lowest Active Counts**Which means lower number of customer activity leading to lesser credibility of bank
* **More Saving than spending**Which means lesser customers would be interested in using Credit Cards and other products of bank



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?

With all the data we have, we can not entirely be correct when predicting Tenure for customers.

Because largely there's no direct correlation of any attribute to Tenure  
and any pattern of attribute of customer falls on every Tenure segment.

But we can get some idea of potential Tenure of customer utilising available data**:**

| **Tenure** | min **Products Used** | **Joining Quarter** | **Joining Months** | **Joining Year** |
| --- | --- | --- | --- | --- |
| **3** | 2 | Qtr3 - Qtr4 | Sept, Oct, Nov, Dec | 2016 |
| **4** | 2 | Qtr1 - Qtr4 | All | 2016 - 2017 |
| **5** | 1 | Qtr1 - Qtr4 | All | 2017 - 2018 |
| **6** | 1 | Qtr1 - Qtr4 | All | 2018 -2019 |
| **7** | 1 | Qtr1 - Qtr3 | All excluding Oct, Nov, Dec | 2019 |

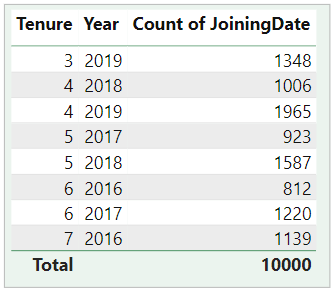
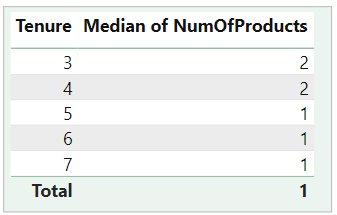
From the table above, we can predict the Tenure of customers based on attributes (columns) they fall in - **The more attributes matches, the greater the chances of Tenure match.**

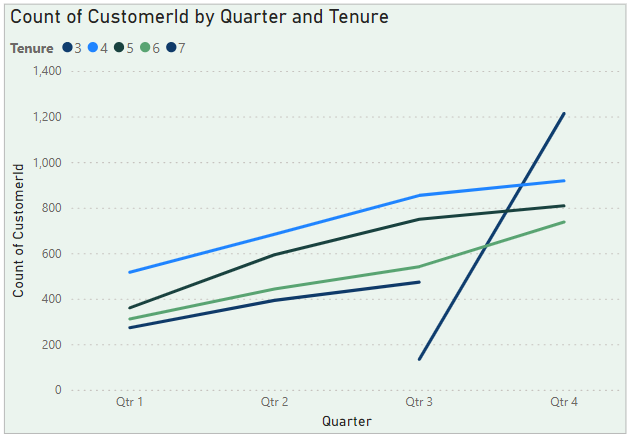
**For example:**

* If a customer has joined between Quarter 3 and Quarter 4 (i.e. in months of Sept, Oct, Nov or Dec)
* has used 2 Products of bank already
* and may/maynot has joined in year 2016
* **Tenure is likely to be 3 years.**

*(Above Table has been created from data of charts below, created using PowerBI visualisations)*

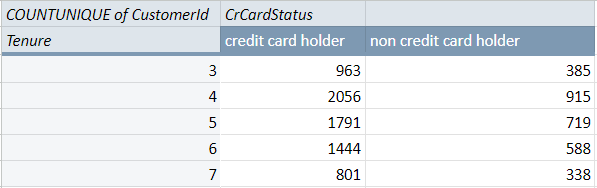
Detailed view of data can be seen from charts below:

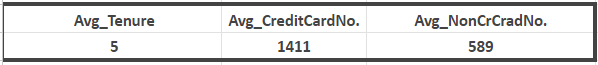




We also have an extra data table - **CreditCardStatus** - which doesn’ty have any direct correlation  
 with Tenure.

But might be of some help in understanding **count of credit/non credit card holders that fall in what  
 range of Tenure**.





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?

Existing data consists of **ActiveCounts**, **ExitRate**, **NumOfProducts** & **Geography** attributes that we can use to assess the impact of marketing campaign:

| **Attributes** | **Impact** |
| --- | --- |
| Active Counts | To assess the count of active customers increment with marketing campaigns. |
| Exit Rate | To assess different exit rates of customers for different campaigns. |
| Num of Products | To assess the increment or decrement number of products used by customers for different campaigns. |
| Geography | To have a demographic centred marketing campaign for different regions. |

While the available data (active counts, exit rates, number of products, geography, and age) offers valuable insights, **a truly comprehensive assessment of marketing campaign effectiveness requires additional information**.

Here's how we can leverage the existing data and identify what's missing for a clear picture:

Utilising Existing Data:

* **Customer Segmentation:** Segment customers based on age, geography, and number of products used. This allows for targeted analysis of campaign performance within specific groups.
* **Trend Analysis:** Analyse changes in active counts, exit rates, and product usage over time. This can reveal trends potentially influenced by marketing campaigns.
* **Control Groups:** If possible, identify control groups not exposed to specific campaigns. Comparing their behaviour with exposed groups helps isolate campaign impact.

Extra Information for Absolute Assessment:

* **Campaign Details & Timing:** Details on campaign content, channels used (online, offline), and launch dates are crucial. This allows for direct correlation between campaign exposure and changes in customer behaviour.
* **Customer Acquisition Channel:** Understanding how customers were initially acquired helps assess if campaigns are effective at retaining customers acquired through different channels (referral, online ad, etc).
* **Customer Lifetime Value:** Calculate customer lifetime value to assess the long-term impact of campaigns on customer retention and overall revenue generation.

By enriching your data with these additional elements, you can move  
 beyond basic correlations and establish causation.

This empowers you to:

* **Optimise Campaigns:** Identify which campaigns resonate best with specific customer segments and adjust strategies for maximum impact.
* **Measure ROI:** Calculate the return on investment for marketing campaigns by comparing acquisition costs to customer lifetime value.

**Conclusion:**

The existing data provides a decent foundation for assessing marketing campaign effectiveness.

However, incorporating campaign details, customer acquisition channels, and customer lifetime value unlocks a deeper understanding of campaign impact.

By segmenting your data, conducting trend analysis, and utilising control groups when possible, you can gain valuable insights.

Ultimately, a data-driven approach that combines existing data with additional information will empower optimise campaigns, maximise ROI, and drive long-term business growth.

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

For customers who have exited, we have identified some common trends between them summarised in table below

| **Status** | **Num of Products (median)** | **CreditScore (median)** | **Avg Balance ($)** |
| --- | --- | --- | --- |
| **Exit** | 1 | 646 | 91,100 |
| **Retain** | 2 | 653 | 72,750 |

With help of table above, we can identify which customers is likely to  
 Exit or get Retained.

For Example:

Customer is likely to **Exit** if:

* Number of Products used **<2**
* CreditScore is **<650**
* Average Balance in their bank account **<= 75,000**

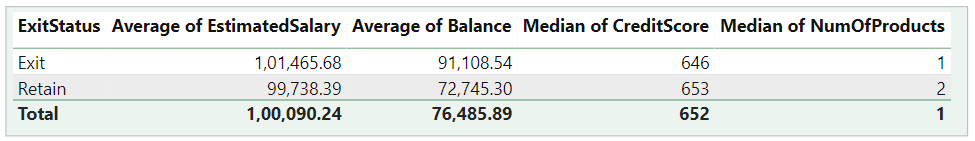
Customer is likely to get **Retain** if:

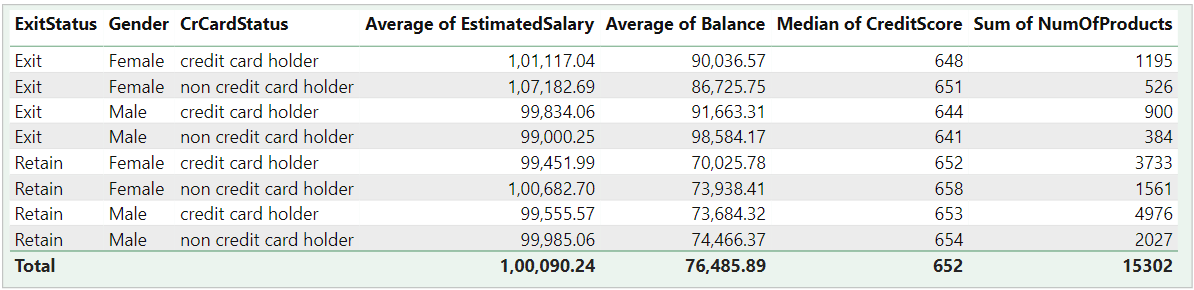
* Number of Products used **>=2**
* CreditScore is **>=650**
* Average Balance in their bank account **>= 85,000**

*Thus we can say a customer is* ***:***

| **Likely to** | **if Num of Products** | **if CreditScore** | **if Avg Balance ($)** |
| --- | --- | --- | --- |
| **Exit** | <2 | <650 | >85,000 |
| **Retain** | >=2 | >=650 | <75,000 |

After experimentation of permutation and combination of attributes with Exit and Retain status of customer, we came up with the trend found in visualisations below:



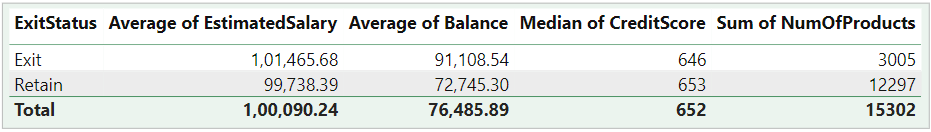


1. Is 'Tenure', 'NumOfProducts', 'IsActiveMember', and 'EstimatedSalary' important for predicting if a customer will leave the bank?

Among the given options, just **NumOfProducts** and **IsActiveMember** is enough for predicting if a customer will leave the bank.

As seen in the previous problem (Question No. 7), we do not need, or rather can not determine by **Estimated Salary,** as they are close enough (for a salary) to be similar.

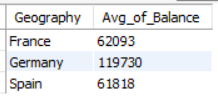
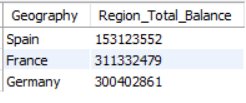
With the help of a solution in the previous problem (Ques.7), we can use Credit Score & Num of Products to determine likelihood of customers leaving the bank. **(refer Ques. 7)**



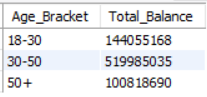
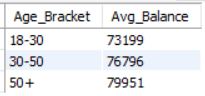
1. Utilise SQL queries to segment customers based on demographics and account details.
2. Demographics considered from dataset: **Geography**, **Gender** & **Age (**Age Bracket**)**
3. Account Detail considered**: Balance**

**Individual Segmentation**:

* Geography (Total Balance & Average Balance)

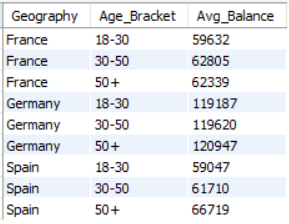
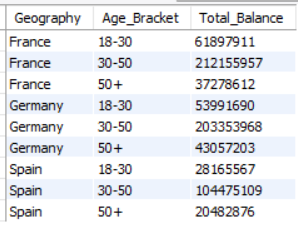


* Age(Total & Average Balance)

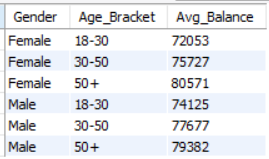


**Mixed Segmentation**:

* Geography & Age (Total & Average Balance)

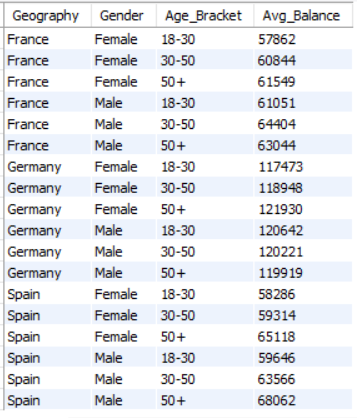


* Gender & Age (Average Balance)



**All Mix Segmentation**:

Average Balance of all permutation & combination of demographics



For the SQL queries, we can see that in the sql file in the subjective section as what queries we have utilised to segment customers.  
  
For additional account details we can change attributes we are choosing to respective queries and achieve desired results.

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?

**Created Churn Risk Score:** Created a calculated column on basis of  
 currently 3 reasonable columns, i.e.

| Churn Risk Score (CRS) = Balance **/** **(**Credit Score **\*** NumOfProducts**)** |
| --- |

that implies Churn Risk is directly proportional to Balance   
 and inversely proportional to Credit Score & Number of Products

Note**:** Customer with Balance 0 will have CRS equal to 0

**Created the table visual:** Build a table incorporating  
 factors like:

* Number of products
* Credit scores
* Balance
* Tenure
* other account activity

**Conditional Formatting:**

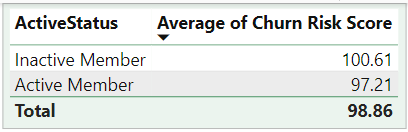
* Select the churn risk score column in your table
* Right-click and choose **Conditional formatting**
* Highlight risk levels using background colours:
  + **Green:** Low risk (score below a defined threshold)
  + **Yellow:** Moderate risk (score within a specific range)
  + **Red:** High risk (score exceeding a defined threshold).
  + **Gray:** Too Low value/ Null to be judged

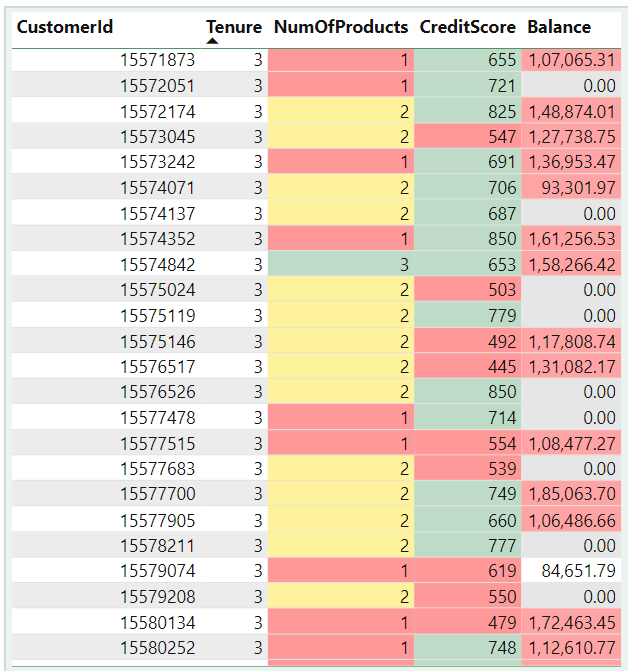
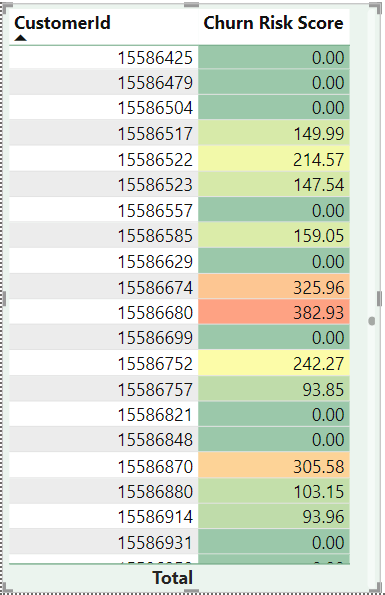
**Benefits:**

This colour-coded table instantly flags at-risk customers, allowing you  
 to prioritise retention efforts. Think of it as a heatmap within your table,  
 guiding your focus towards potential churners.

**Remember:**

* **Judging must be for Active Customers.**
* Adjustment of thresholds and colours based on data and risk assessment should be done periodically.
* We can consider exploring additional formatting options for even more visual cues as our data would grow.



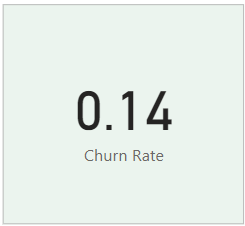
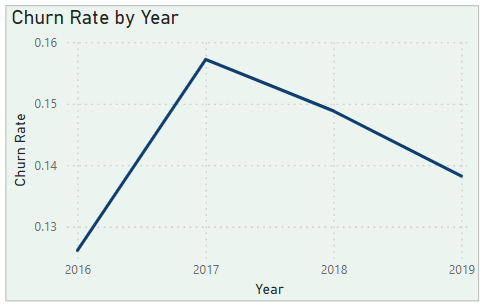
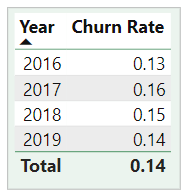


1. What is the current churn rate per year and overall, as well in the bank? Can you suggest some insights to the bank about which kind of customers are more likely to churn and what different strategies can be used to decrease the churn rate?

| Current Churn Rate (Overall) **= 0.14** |
| --- |

Current Churn Rate (Yearly)**:**

| **Year** | **Churn Rate** |
| --- | --- |
| 2016 | 0.13 |
| 2017 | 0.16 |
| 2018 | 0.15 |
| 2019 | 0.14 |

Customers more likely to churn:

**(**Refer to solution of Question 7**)**

Here are key, effective strategies to combat churn:

**Data-Driven Churn Prediction:**

* Leverage machine learning to analyse customer data and predict churn risk.
* Proactively target high-risk segments with personalised retention efforts.

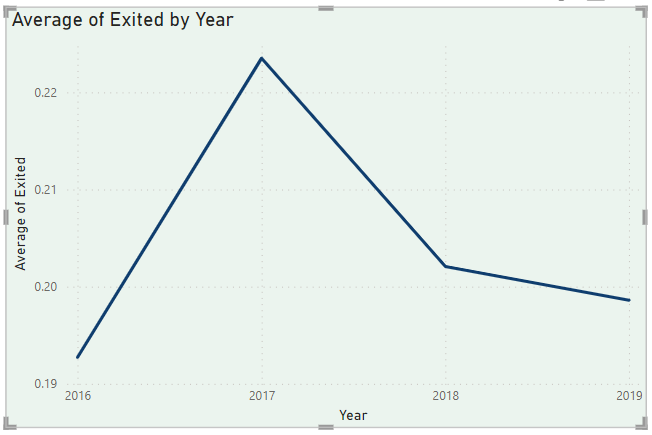
**Personalized Incentives & Rewards:**

Analyse customer behaviour and offer targeted rewards **–** cashback for high spenders, financial consultations for low product users **–** to increase engagement

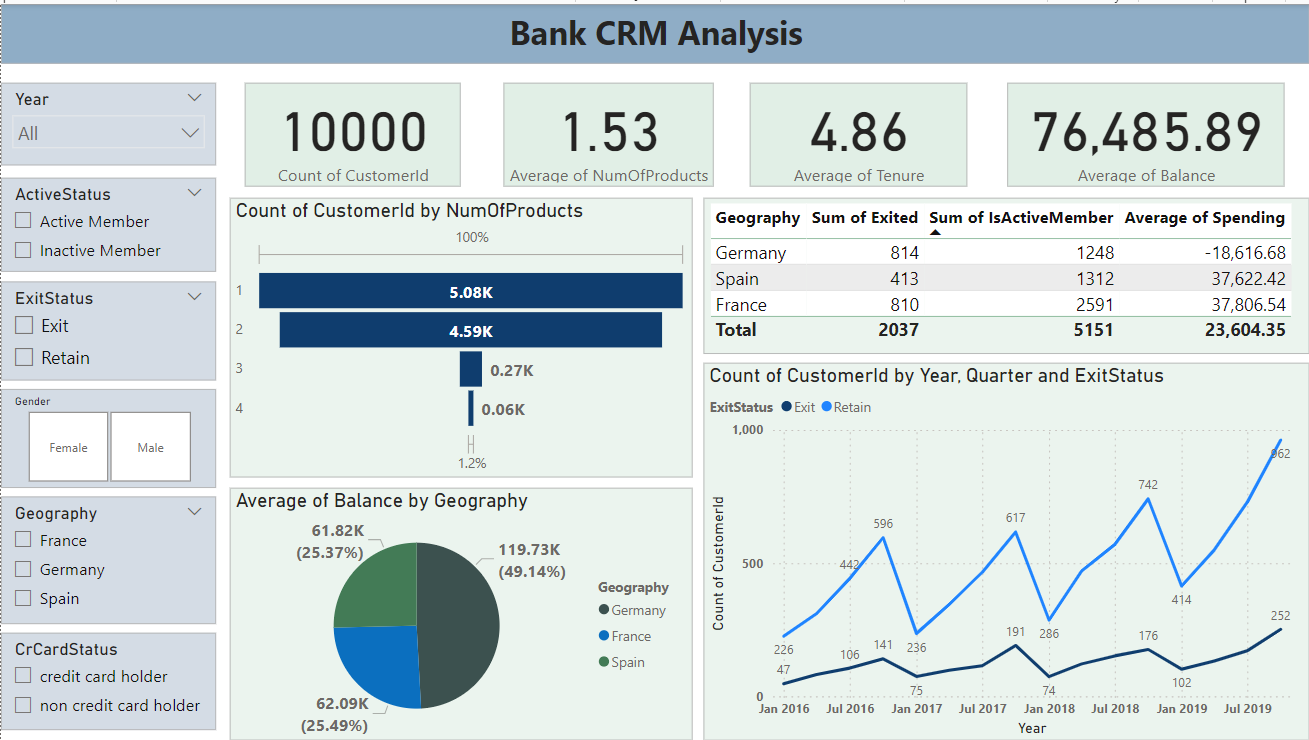
**Frictionless Customer Journeys:**

Simplify processes like account opening and online banking to enhance user experience and reduce churn caused by frustration.

*By focusing on these core strategies, banks can gain a data-driven understanding of churn risk, personalise customer experiences, and  
ultimately build stronger, more loyal customer relationships.*



1. Create a dashboard incorporating all the KPIs and visualisation-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?

Approaching a data analysis problem without specific objective and subjective questions requires a flexible and exploratory approach. Here's a solution tailored to this scenario:

**1. Understand the Context:**Begin by gaining a thorough understanding of the context and purpose of the data analysis. This might involve consulting with stakeholders or reviewing project documentation to identify any implicit objectives or expectations.

**2. Explore the Data:**Conduct exploratory data analysis to familiarise yourself with the dataset. Examine the structure, distributions, and relationships between variables. Look for any patterns, anomalies, or interesting trends that emerge.

**3. Identify Potential Insights:**Based on your exploration of the data, brainstorm potential insights or hypotheses that could be explored further. Consider both quantitative patterns and qualitative aspects that may be relevant to the problem domain.

**4. Generate Hypotheses:**Formulate hypotheses or conjectures based on your initial observations. These hypotheses can serve as guiding principles for further analysis and experimentation.

**5. Iterative Analysis:**Iteratively apply various analytical techniques to test and refine your hypotheses. This could include statistical analysis, machine learning algorithms, or qualitative methods depending on the nature of the data and the problem, but in this project we would run multiple different queries and play with data in Excel, PowerBI & SQL.

**6. Visualisation and Interpretation:**Use data visualisation techniques to communicate your findings effectively. Visualisations can help uncover patterns, highlight relationships, and convey insights to stakeholders in a meaningful way.

**7. Synthesise Findings:**Synthesise your findings into coherent narratives or themes that provide a holistic understanding of the data. Look for overarching patterns or trends that emerge across different aspects of the analysis.

**8. Seek Validation and Feedback**  
Validate your findings through peer review, domain experts, or comparison with external sources if possible. Seek feedback from stakeholders to ensure that your analysis resonates with their expectations and contributes to their understanding of the problem.\

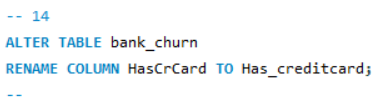
**9. Iterate and Refine:**Iterate on your analysis as needed, refining your approach based on new insights or feedback from stakeholders. Be open to exploring alternative hypotheses or perspectives to deepen your understanding of the problem.

**10. Document and Communicate:**  
Document your analysis process, methodologies, and findings in a clear and transparent manner. Communicate your results effectively to stakeholders, highlighting key insights, implications, and recommendations for further action.

*By following this approach, we can systematically explore and analyse the data even in the absence of specific objective and subjective questions, ultimately deriving valuable insights that can inform decision-making and drive positive outcomes for the project.*

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

We can write following SQL query for that:



This will modify the name of the “**HasCrCard**” column to “**Has\_creditcard**” column.