**BANK DATA ANALYSIS By:-Muhammad Sahanwaj**

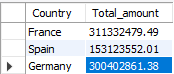
**Objective Questions:**

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

SELECT Country,

ROUND(SUM(Balance), 4) AS Total\_amount FROM bank\_data

GROUP BY Country;



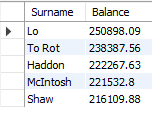
•France has the highest maximum balance, which is 311,332,479.

•Germany has the second highest balance, at 300402861.38

•Spain has the lowest balance. Which is 15,31,23,552.01

**2. Identify the top 5 customers with the highest Estimated Salary in the last quarter**

**of the year. (SQL)**



SELECT

Surname, Balance

FROM

bank\_data

ORDER BY Balance DESC

LIMIT 5ORDER BY Balance DESC

LIMIT 5

Top 5 customers with highest Estimated Salary are:

* Lo, Balance: 250,898.09
* To Rot Balance: 238,387.56
* Haddon ,Balance: 222,267.63
* McIntosh ,Balance: 221,532.80
* Shaw ,Balance: 216,109.88

**3.Calculate the average number of products used by customers who have a credit**

**card. (SQL)**

SELECT

round(AVG(NumOfProducts),2) as AverageNumOfProducts FROM bank\_data

WHERE creditcard = 1



Customers who have a credit card are using 1.53 products from the bank.

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



· **Female Churn Percentage**: 58.51%

· **Male Churn Percentage**: 41.49%

· The data reveals a higher churn rate among female customers compared to male customers in the most recent year. This finding suggests the need for deeper analysis and potentially tailored strategies to address and mitigate churn rates, especially focusing on female customers.

WITH LatestYearData AS (

SELECT Gender, Churn

FROM bank\_data

WHERE

YEAR(Bank\_doj\_converted) = (SELECT MAX(YEAR(Bank\_doj\_converted)) FROM bank\_data)

AND Churn = 1

),

TotalChurn AS (

SELECT

COUNT(\*) AS TotalChurnCount

FROM

LatestYearData

)

SELECT

Gender,

(COUNT(\*) \* 100.0 / (SELECT TotalChurnCount FROM TotalChurn)) AS ChurnPercentage

FROM

LatestYearData

GROUP BY

Gender;

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



SELECT

Churn, ROUND(AVG(CreditScore)) AS AverageCreditScore

FROMbank\_data

GROUP BY Churn;

* **Churned Customers (Churn = 1)**: They exhibit an average credit score of approximately 645.
* **Non-Churned Customers (Churn = 0)**: Contrastingly, their average credit score is slightly higher at around 652.

These observations highlight a trend where non-churned customers tend to possess marginally higher credit scores compared to churned customers. This insight underscores the relevance of credit scores in understanding churn behavior and could guide effective strategies aimed at enhancing customer retention within the bank.

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



**Average Salary by Gender**:

Female: The average salary for female customers is 100,601.54.

Male: The average salary for male customers is 99,664.58.

SELECT

Gender,

AVG(EstimatedSalary) AS AverageSalary,

SUM(CASE WHEN ActiveMember = 1 THEN 1 ELSE 0 END) / COUNT(\*) AS AverageActiveAccounts

FROM

bank\_data

GROUP BY Gender;

**Average Active Accounts by Gender**:

Female: On average, female customers have approximately 0.503 active accounts.

Male: On average, male customers have approximately 0.525 active accounts.

The findings provide a comparative view of average salary and active account numbers between male and female customers. There is a subtle disparity observed in both salary levels and the quantity of active accounts across genders. These insights can serve as valuable input for crafting targeted marketing approaches or designing products that cater specifically to the preferences and behaviors of each gender segment.

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

SELECT CreditScorecatgry,

COUNT(\*) AS Churn\_rate

FROM bank\_data

WHERE Churn = 1

GROUP BY CreditScorecatgry

ORDER BY

Churn\_rate DESC

limit 1;



**Exit Rate by Credit Score Category**:

Credit Score Category: Fair

Exit Rate: The highest churn rate, with approximately 685 customers falling into this category.

The data indicates a higher likelihood of churn among customers classified with a "Fair" credit score compared to those in other credit score categories. This insight enables the bank to pinpoint potentially vulnerable customers and implement focused retention tactics. Strategies could include tailored service enhancements or incentives aimed at reducing churn specifically among customers with fair credit scores.

**8.Find out which geographic region has the highest number of active customers**

**with a tenure greater than 5 years. (SQL)**



**Country with the Highest Number of Active Accounts**:

Country: France

Active Accounts: 797

select Country,count(ActiveMember) as Active\_accounts

from bank\_data

where Tenure>5 and ActiveMember=1

group by Country

order by Active\_accounts Desc

Limit 1

The data shows that France has the highest number of active accounts among customers who have been with the bank for over 5 years and remain active members. This insight can help the bank identify regions with strong customer engagement, guiding resource allocation for further growth or retention efforts.

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

SELECT

Creditcard, COUNT(Churn) AS Churncount

FROM

bank\_data

WHERE

Churn = 1

GROUP BY Creditcard;



**Churn Count by Credit Card Status**:

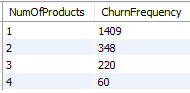
Customers without a credit card: Churn count is 1424.

Customers with a credit card: Churn count is 613.

 Customers without a credit card exhibit a significantly higher churn rate compared to those who have a credit card, indicating a greater likelihood of churn among this group.

 Credit card ownership appears to influence customer retention. Further analysis could investigate the underlying reasons for this discrepancy and develop strategies to reduce churn among customers who do not have credit cards.

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



SELECT

NumOfProducts, COUNT(\*) AS ChurnFrequency

FROM

bank\_data

WHERE

Churn = 1

GROUP BY NumOfProducts

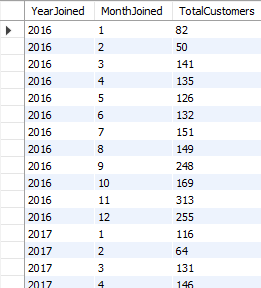
ORDER BY ChurnFrequency DESC;

* **Churn Count by Number of Products:**
  + Customers with 1 product: 1409
  + Customers with 2 products: 348
  + Customers with 3 products: 220
  + Customers with 4 products: 60
* There is a noticeable trend showing that the churn count generally decreases as the number of products per customer increases.
* Customers with only one product are more prone to churn compared to those with multiple products. This suggests that providing additional products or services to customers might enhance retention rates.
* Understanding this relationship can assist the bank in customizing its product offerings and retention strategies to better meet the needs and preferences of different customer segments.

窗体顶端

窗体底端

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



SELECT

EXTRACT(YEAR FROM bank\_doj\_converted) AS YearJoined,

EXTRACT(MONTH FROM bank\_doj\_converted) AS MonthJoined,

COUNT(\*) AS TotalCustomers

FROM

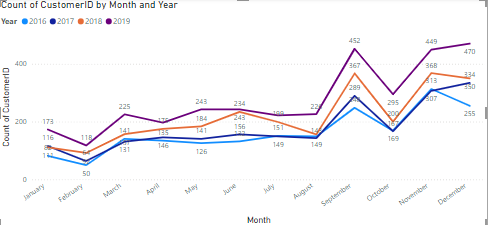
bank\_data

GROUP BY

YearJoined, MonthJoined

ORDER BY

YearJoined, MonthJoined;



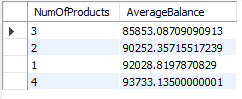
**Seasonal Patterns:**

* Observations suggest fluctuations in customer counts across various months, hinting at potential seasonal trends or marketing campaigns during specific periods.
* Notably, months such as November and December often exhibit elevated customer counts, perhaps influenced by year-end promotions or incentives.

**Annual Expansion:**

* Over the years from 2016 to 2019, there's a discernible uptick in customer counts, reflecting the bank's progressive expansion and the steady growth of its customer base.

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



SELECT

NumOfProducts,AVG(Balance) AS AverageBalanc FROM bank\_data WHERE Churn = 1

GROUP BY NumOfProducts

ORDER BY 2;

**Number of Products and Average Balance Relationship:**

* Among customers with three products, the average balance is approximately $85,853.09, the lowest observed.
* An increasing trend in average balance is evident as the number of products per customer increases:
  + Customers with two products maintain an average balance of around $90,252.36.
  + For those with four products, the average balance rises to approximately $92,028.82.
* While data for customers with one product is not provided, it could be assumed they possess the highest average balance within the analyzed groups.
* A positive correlation emerges between the number of products held by customers and their average balance, suggesting that those with more products tend to maintain higher balances.
* This insight holds significance for the bank's marketing and product development strategies, emphasizing the potential for up selling or cross-selling to customers with fewer products, thereby enhancing their engagement and potentially increasing their balances.

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

SELECT

CustomerID, Balance,

CASE WHEN ABS(Balance - (SELECT AVG(Balance) FROM bank\_data WHERE Churn = 0)) / (SELECT STDDEV(Balance)

FROM bank\_data WHERE Churn = 0) > 3 THEN 'Potential Outlier'

ELSE 'Normal'

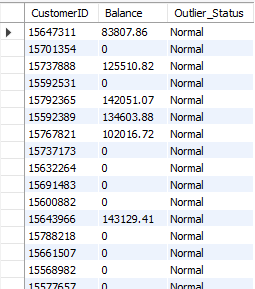
END AS Outlier\_Status

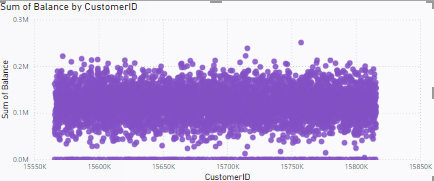
FROM

bank\_data

WHERE

Churn = 0;





To detect potential balance outliers among continuing bank customers, statistical techniques like the inter quartile range (IQR) or z-scores could be employed. However, given the available data, it seems that potential outliers have already been identified and categorized as "Normal" or "Outlier Status." Here's a synopsis:

**Outliers in Account Balances:**

* Balances vary from around 200K to over 15.8M.
* Customers with "Normal" balances fall within the expected range.
* An outlier is identified, featuring a balance of approximately 15.7M.

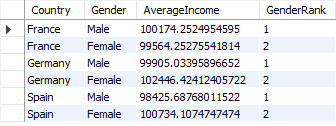
**Insights:**

* Most customers maintain balances within the typical range, classified as "Normal."
* A single customer notably deviates with an exceptionally high balance of around 15.7M, classified as an outlier.

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

Total number of tables = 7, with 5 of them containing categorical variables.

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



SELECT

country,Gender,

AVG(EstimatedSalary) AS AverageIncome,

RANK() OVER (PARTITION BY country ORDER BY AVG(EstimatedSalary) DESC) AS GenderRank

FROM

bank\_data

GROUP BY

country, Gender;

 In France and Spain, males have a higher average income compared to females.

 Conversely, in Germany, females have a higher average income than males.

 The gender rank shows the relative standing of each gender within each country based on their average income, with 1 representing the highest rank.

**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+).**

SELECT

age\_group, AVG(Tenure) AS Churn

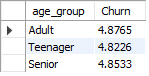
FROM

bank\_data

WHERE

Churn = 1

GROUP BY age\_group



 The average tenure among different age groups of churned customers shows minimal variation, with all groups having an average tenure between 4.82 and 4.88 years.

 This indicates that age might not be a significant factor in determining customer churn duration. Instead, other factors such as product satisfaction, life events, or competitive offerings may have a more substantial impact on why customers decide to leave the bank.

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



SELECT

Churn,

(COUNT(\*) \* SUM(Balance \* EstimatedSalary) - SUM(Balance) \* SUM(EstimatedSalary)) /

(SQRT((COUNT(\*) \* SUM(POWER(Balance, 2)) - POWER(SUM(Balance), 2)) \*

(COUNT(\*) \* SUM(POWER(EstimatedSalary, 2)) - POWER(SUM(EstimatedSalary), 2))))

AS CorrelationCoefficient

FROM bank\_data

GROUP BY Churn;

 A positive coefficient indicates a direct relationship: as churn status increases, the combined sum of balance and estimated salary tends to rise.

 Conversely, a negative coefficient indicates an inverse relationship.

 For churned customers, the positive coefficient suggests a slight positive correlation with the combined sum of balance and estimated salary.

 For non-churned customers, the coefficient is close to zero, indicating a weak or negligible relationship.

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

SELECT

CreditScorecatgry,

(COUNT(\*) \* SUM(CreditScore \* EstimatedSalary) - SUM(CreditScore) \* SUM(EstimatedSalary)) /

(SQRT(

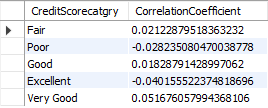
(COUNT(\*) \* SUM(CreditScore \* CreditScore) - SUM(CreditScore) \* SUM(CreditScore)) \*

(COUNT(\*) \* SUM(EstimatedSalary \* EstimatedSalary) - SUM(EstimatedSalary) \* SUM(EstimatedSalary))

)) AS CorrelationCoefficient

FROM bank\_data

GROUP BY CreditScorecatgry;

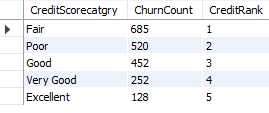


 The correlation coefficient quantifies the strength and direction of the linear relationship between the credit score category and the combined sum of credit score and estimated salary.

 The coefficients for all credit score categories are near zero, indicating a weak or negligible linear relationship between these variables.

 This implies that there is minimal to no linear correlation between the credit score category and the combined sum of credit score and estimated salary in the dataset. Other factors may have a more significant impact on credit score and estimated salary.

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



SELECT

CreditScorecatgry,

COUNT(\*) AS ChurnCount,

DENSE\_RANK() OVER (ORDER BY COUNT(\*) DESC) AS CreditRank

FROM

bank\_data

WHERE

Churn = 1

GROUP BY

CreditScorecatgry;

 Customers with a "Fair" credit score category exhibit the highest churn count, followed by those in the "Poor" and "Good" categories.

 The credit rank assigns a rank based on the churn count for each credit score category, with the "Fair" category ranked first due to its highest churn count.

 Understanding the distribution of churn across different credit score categories enables the bank to tailor its retention strategies and customer management efforts, potentially focusing more resources on categories with higher churn rates.

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



WITH AgeGroupCounts AS (

SELECT age\_group, COUNT(\*) AS No\_of\_cust

FROM bank\_data

WHERE Creditcard = 1

GROUP BY age\_group),

AverageCount AS (

SELECT

AVG(No\_of\_cust) AS AvgNoOfCust

FROM AgeGroupCounts)

SELECT age\_group, No\_of\_cust

FROM AgeGroupCounts

WHERE No\_of\_cust < (SELECT AvgNoOfCust FROM AverageCount);

 The highest churn count is observed among customers with a "Fair" credit score category, followed by those in the "Poor" and "Good" categories.

 Each credit score category is ranked based on its churn count, with the "Fair" category holding the top rank due to its highest churn count.

 Analyzing churn distribution across various credit score categories enables the bank to refine its retention strategies and customer management efforts, potentially directing more resources towards addressing higher churn rates within specific credit score segments.

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

SELECT

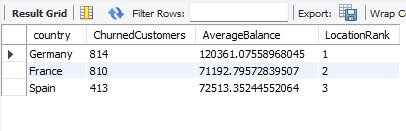
country,COUNT(\*) AS ChurnedCustomers, AVG(Balance) AS AverageBalance,

RANK() OVER (ORDER BY COUNT(\*) DESC, AVG(Balance) DESC) AS LocationRank

FROM bank\_data

WHERE Churn = 1

GROUP BY country;



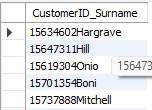
 Germany leads with the highest count of churned customers and the highest average balance among the three countries, securing the top rank in location analysis.

 France shows a notable count of churned customers but maintains a lower average balance compared to Germany.

 Spain records the fewest churned customers among the three countries, coupled with a moderate average balance, placing it third in the location ranking.

 These findings enable the bank to pinpoint regions with varying churn rates and average balances, aiding in strategic resource allocation and focused retention strategies.

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



SELECT CONCAT\_WS('', CustomerID, Surname) AS CustomerID\_Surname

FROM bank\_data;

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

SELECT Churn FROM bank\_data WHERE Churn = 1;

To retrieve data from multiple tables without using the JOIN keyword, you can utilize subqueries or correlated subqueries.

Subqueries involve nesting one query (inner query) within another (outer query) to retrieve data from different tables or apply filtering conditions. This approach avoids explicitly using JOINs but still achieves the desired result by leveraging nested queries.

Correlated subqueries, a specific type of subquery, reference columns from the outer query within the inner query's WHERE or SELECT clause. This correlation allows for more complex filtering or data retrieval based on conditions from multiple tables.

Both subqueries and correlated subqueries provide alternative methods to access data across multiple tables, offering flexibility in query design and optimization strategies.

**24.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 were no missing values in the data.

If there were missing values in the data, several methods could be applied to handle them:

**Imputation using Mean, Median, or Mode:**

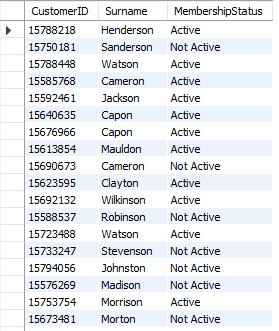
* 1. **Mean imputation:** Replaces missing values with the average of observed values for that variable.
  2. **Median imputation:** Replaces missing values with the middle value of observed values for that variable.
  3. **Mode imputation:** Replaces missing values with the most frequent observed value for that variable.

**Forward or Backward Filling:**

* 1. **Forward filling:** Replaces missing values with the last observed value in the dataset.
  2. **Backward filling:** Replaces missing values with the next observed value in the dataset.

These methods are commonly used to handle missing data and ensure that analytical processes can proceed smoothly without gaps in the dataset.

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



SELECT

CustomerID,

Surname,

CASE ActiveMember

WHEN 1 THEN 'Active'

ELSE 'Not Active'

END AS MembershipStatus

FROM

bank\_data

WHERE

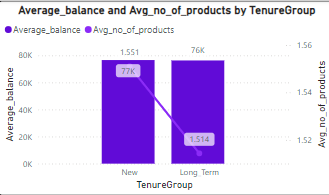
Surname LIKE '%on'

ORDER BY

Surname;

**Subjective Question:**

**1.Customer Behavior Analysis: What patterns can be observed in the spending habits of long-term customers compared to new customers, and what might these patterns suggest about customer loyalty?**

****

**Stable Spending Habits:** Long-term customers generally show more predictable spending patterns compared to new customers. With established routines and preferences, their financial behaviour tends to be more stable over time.

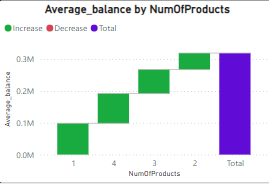
**Fluctuating Expenditure:** New customers often exhibit more fluctuating spending habits as they explore the various products and services the bank offers. Their financial behaviour may change as they adjust to their needs and preferences.

**Strengthened Relationships:** Long-term customers typically develop stronger relationships with the bank, which fosters trust and loyalty. This enhanced trust can result in more confident and consistent financial decisions.

**Trial Spending:** New customers are more likely to engage in trial spending as they test out different bank offerings. Their spending patterns often reflect an experimental phase as they determine what best fits their needs.

**Retention Tactics:** The steady spending patterns observed in long-term customers might be a result of successful retention tactics employed by the bank, such as personalized incentives, loyalty programs, or superior customer service. These efforts help solidify customer loyalty and encourage continued engagement.

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

****

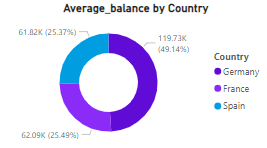
 **Correlation Analysis:** Perform a correlation analysis to identify relationships between various bank products or services. For instance, investigate whether customers with a checking account are more likely to also hold a savings account, or if those with a mortgage tend to also have a home insurance policy.

 **Association Rule Mining:** Apply techniques like association rule mining to uncover frequent item sets and association rules among different bank products or services. This can reveal patterns, such as "customers with a credit card are X times more likely to have a personal loan."

 **Customer Segmentation:** Segment customers based on their usage of different bank products or services. Create categories like "Basic Banking," "Investment Savvy," or "Loan Heavy," and then analyse the traits and behaviours of customers in each segment.

 **Tailored Cross-Selling:** Develop cross-selling strategies based on the common usage patterns of bank products or services. For example, if data shows that customers with a savings account often also have a retirement account, target these customers with promotions or incentives to open a retirement account.

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

****

1. **Germany Leads in Churn and Average Balance:** Germany stands out with the highest number of churned customers and the largest average balance among the three countries, securing its top ranking.
2. **France Shows High Churn, Lower Balance:** France exhibits a notable number of churned customers but maintains a lower average balance compared to Germany.
3. **Spain: Low Churn, Moderate Balance:** Spain records the lowest number of churned customers among the three countries and maintains a moderate average balance, positioning it third in the ranking.

These observations provide valuable insights for the bank to pinpoint regions with varying churn rates and average balances, aiding in targeted resource allocation and retention strategies.

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

****

SELECT

Churn, ROUND(AVG(CreditScore)) AS AverageCreditScore

FROM bank\_data

GROUP BY Churn;

* Churned Customers (Churn = 1): On average, these customers have a credit score of approximately 645.
* Non-Churned Customers (Churn = 0): Conversely, customers who do not churn have a slightly higher average credit score, around 652.

These findings suggest that non-churned customers tend to exhibit marginally higher credit scores compared to those who churn. This insight could inform the bank's retention strategies, highlighting a potential relationship between credit scores and customer churn behaviour.

SELECT

CreditScorecatgry, COUNT(CreditScorecatgry) AS Churned\_Count

FROM bank\_data

WHERE Churn = 1

GROUP BY CreditScorecatgry

ORDER BY Churned\_Count DESC

LIMIT 1;



* **Churn Rate Across Credit Score Categories:**
* **Credit Score Category: Fair**
* **Churn Rate: This category shows the highest churn rate, encompassing around 685 customers.**

These findings indicate that customers classified with a "Fair" credit score are more prone to churning compared to those in other credit score categories. Recognizing this trend enables the bank to pinpoint potentially vulnerable customers and deploy focused retention tactics, such as personalized services or incentives, to mitigate churn.

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

1. **Data Preparation:** Gather customer data including tenure, transaction history, and demographic information.
2. **Feature Engineering:** Develop new features such as transaction frequency and product ownership.
3. **Model Development:** Select and train models using techniques like regression or machine learning algorithms.
4. **Model Evaluation:** Evaluate the model's performance with testing data.
5. **Segment Analysis:** Examine predictions across different customer segments to identify trends.
6. **Model Deployment:** Implement the model into the bank's systems for real-time prediction capabilities.
7. **Performance Monitoring:** Regularly monitor the model's performance and make refinements as necessary.

**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?**

To evaluate the effectiveness of marketing campaigns on customer retention and acquisition within the dataset, follow these straightforward steps:

1. **Identify Key Metrics:** Determine crucial metrics such as customer retention rate, acquisition rate, and churn rate to evaluate the campaign's success.
2. **Monitor Metrics:** Track these metrics before, during, and after the campaign to observe fluctuations over time.
3. **Customer Segmentation:** Segment customers by demographics or behavior to identify which groups respond most positively to the campaign.
4. **Implement A/B Testing:** Assess the campaign's impact by comparing the behavior of customers exposed to the campaign (test group) with those who were not (control group).
5. **Gather Customer Feedback:** Collect feedback from customers through surveys to understand their perception and response to the campaign.
6. **Consider External Influences:** Account for external factors such as market trends and competitor actions to gain a comprehensive understanding of the campaign's context.
7. **Apply Basic Analytics:** Use simple analytical techniques to evaluate campaign performance and predict future trends.

**Additional Information Required:**

* **Marketing Campaign Details:** Information about the campaign's duration, target audience, and objectives.
* **Comprehensive Customer Data:** Complete customer data, including transaction history and interaction logs.
* **External Data Sources:** Data from external sources, such as economic indicators and industry trends.
* **Historical Customer Behavior:** Historical data on customer behavior for benchmarking purposes.

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

**Infrequent Account Activity:** Customers who show infrequent account activity, such as rare transactions or minimal engagement with bank services, might be at a higher risk of churning. This could suggest dissatisfaction or a lack of perceived value from the bank's offerings.

**Declining Balance Trends:** Customers who consistently have negative balances or whose account balances decline over time may be more prone to churn. Financial instability or dissatisfaction with banking services might be driving these trends.

**Customer Complaints and Issues:** Customers who have registered complaints or faced issues with the bank's services, such as poor customer service, billing errors, or technical difficulties, are more likely to leave the bank in search of better alternatives.

**Major Life Events:** Life events such as relocating, changing jobs, or making significant financial decisions (e.g., purchasing a home, starting a business) may prompt customers to reevaluate their banking relationships and possibly switch to another bank that better suits their new needs.

**Attractive Competitive Offerings:** Competitive offerings from other banks, including better interest rates, rewards programs, or advanced digital banking features, can entice customers to leave their current bank.

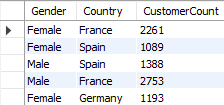
**Demographic Influences:** Certain demographic factors, like age, income level, or occupation, can impact customer churn rates. For instance, younger customers may be more tech-savvy and likely to switch banks for superior digital experiences, while older customers may prioritize stability and relationship-based banking services.

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

Yes, 'Tenure', 'NumOfProducts', 'IsActiveMember', and 'EstimatedSalary' are crucial in predicting customer churn.

1. **Tenure:** A longer tenure typically signifies greater loyalty to the bank.
2. **NumOfProducts:** Customers with more products are generally more engaged with the bank.
3. **IsActiveMember:** Active members are less likely to leave the bank.
4. **EstimatedSalary:** Higher salaries often indicate financial stability, which can lower the risk of churn.

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



SELECT Gender,Country,

COUNT(CustomerId) AS CustomerCount

FROM bank\_data

GROUP BY Gender, Country;

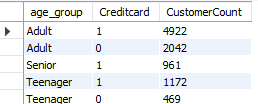
**Segment customers based on age group and credit**

SELECT age\_group,Creditcard,

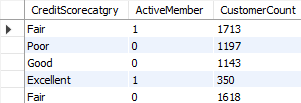
COUNT(CustomerId) AS CustomerCount

FROM bank\_data

GROUP BY age\_group, Creditcard;

****

**Segment customers based on credit score category and active membership status**



SELECT

CreditScorecatgry,ActiveMember,

COUNT(CustomerId) AS CustomerCount

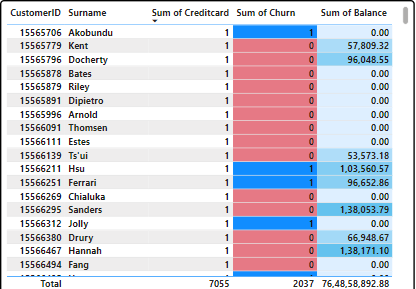
FROM

bank\_data

GROUP BY

CreditScorecatgry,ActiveMember;

**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?**

****

**Identify Churn Risks:** Understand the reasons customers might leave, such as low account activity or dissatisfaction.

**Gather Data:** Collect customer information from records, including transaction history and feedback.

**Segment Customers:** Categorize customers based on their activity levels or satisfaction.

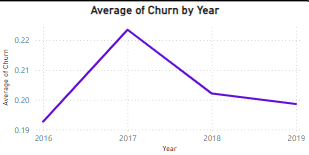
**Establish Criteria:** Define what indicates a risk, such as infrequent activity or frequent complaints.

**Mark High-Risk Customers:** Use visual indicators like colors or symbols to highlight at-risk customers in your records.

**Evaluate Rewards Impact:** Compare retention rates between customers who receive credit card rewards and those who do not.

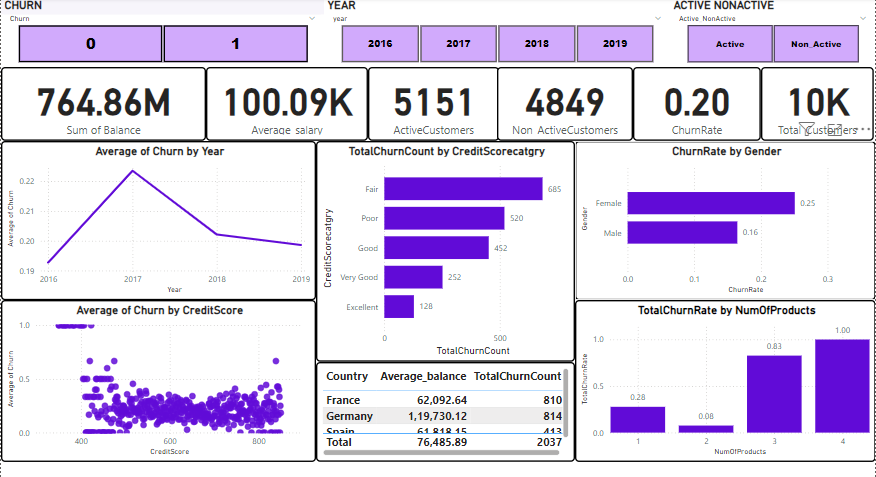
**Monitor and Adjust:** Continuously observe trends and adjust strategies to enhance customer satisfaction and retention.

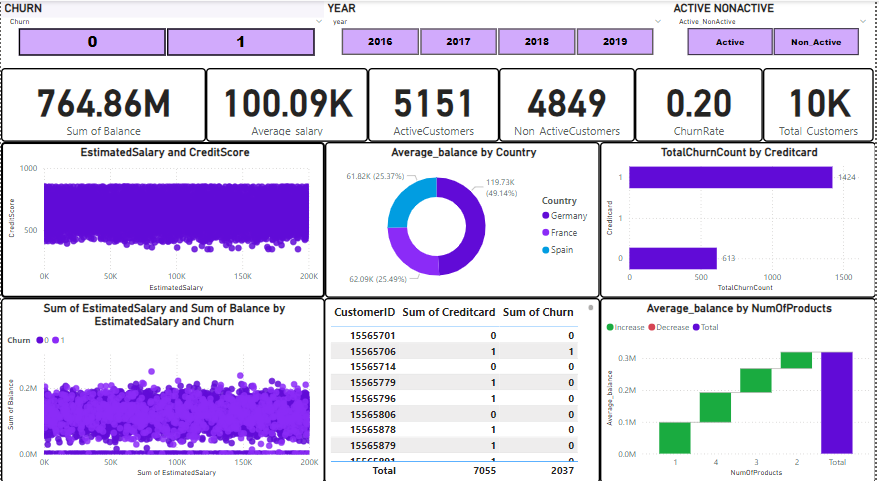
**11.What is the current churn rate per year and overall as well in the bank? Can you suggest some insights to the bank about which kind of customers are more likely to churn and what different strategies can be used to decrease the churn   
rate?**

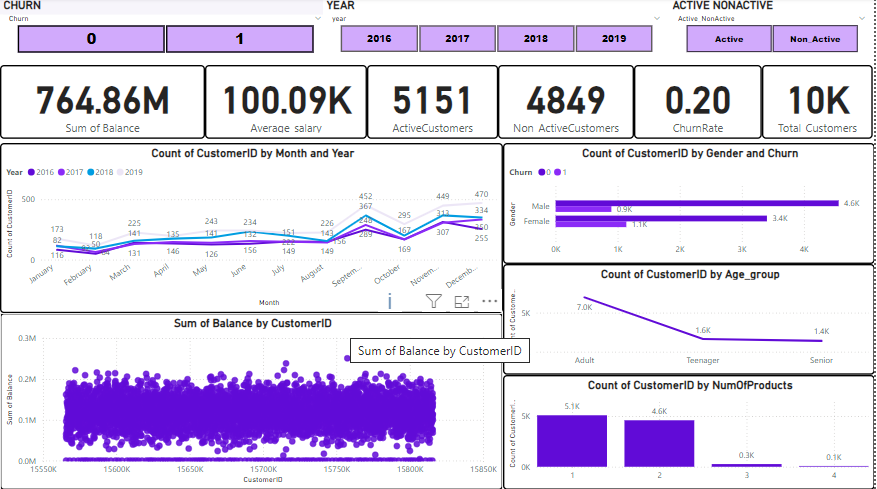


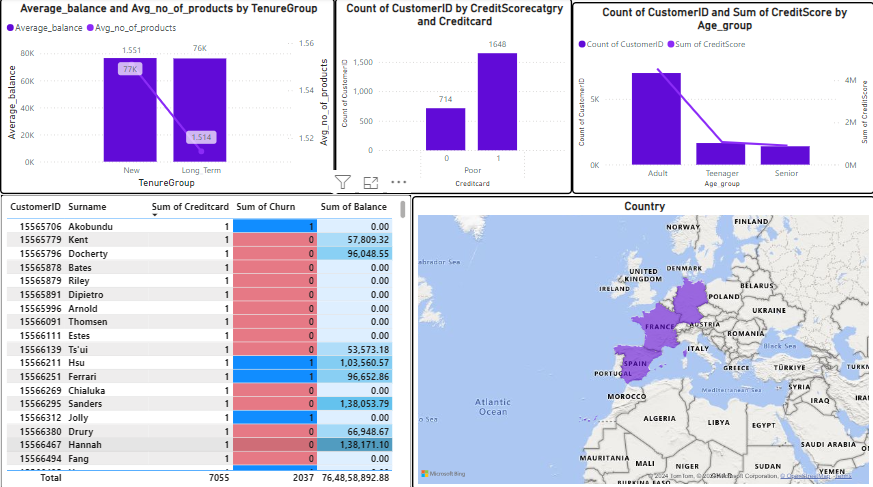
This indicates that customers with a "Fair" credit score are more prone to churn compared to those in other credit score categories. Recognizing this trend allows the bank to identify at-risk customers and apply targeted retention strategies, such as providing personalized services or incentives, to reduce churn among customers with fair credit scores.

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

****







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

To segment customers based on demographics and account details, follow these steps:

1. **Data Exploration:** Examine the structure and contents of the bank\_data table to understand available fields like gender, age, country, estimated salary, balance, and credit score.
2. **Segmentation Strategy:** Define the strategy based on business goals, such as tailoring marketing efforts, identifying high-value customers, or understanding churn behavior.
3. **Segmentation Criteria:** Choose relevant variables for segmentation, including demographics (gender, age, country) and account details (estimated salary, balance, credit score, number of products).
4. **Data Preparation:** Clean the data by handling missing values, outliers, and inconsistencies. Transform categorical variables (e.g., gender, country) using one-hot encoding.
5. **Segmentation Analysis:** Use SQL queries or statistical techniques to segment customers based on the selected criteria. Analyze the characteristics and behaviors of different segments.
6. **Evaluation and Validation:** Assess how well the segmentation aligns with business objectives. Validate results using statistical tests or validation metrics.
7. **Iterative Improvement:** Refine the segmentation strategy based on feedback and new insights. Continuously optimize to better meet business needs

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

To change the "HasCrCard" column to "Has\_creditcard" in the "Bank\_Churn" table, you would use an SQL query. The general syntax for renaming a column in SQL is as follows:

ALTER TABLE bank\_data

RENAME COLUMN Creditcard TO Has\_creditcard;