**Objective Questions:**

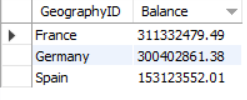
**1. What is the distribution of account balances across different regions?**

**Ans--** SELECT cd.GeographyID, ROUND(SUM(bc.Balance), 2) AS Balance

FROM customerdata cd

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

GROUP BY cd.GeographyID;



**2. Identify the top 5 customers with the highest number of transactions in the last quarter of the year. (SQL)**

**Ans—** Here we will find top5 customers with the highest salary

SELECT CustomerId, EstimatedSalary

FROM customerdata

WHERE YEAR(Bank\_DOJ) = 2019

AND MONTH(Bank\_DOJ) IN (10, 11, 12)

ORDER BY EstimatedSalary Desc

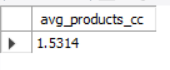
LIMIT 5;



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

**Ans--** Select Avg(NumOfProducts) as avg\_products\_cc

from bank\_churn where HasCrCard = "credit card holder";



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

**Ans--** SELECT

CASE WHEN Exited = 'Exit' THEN 'Exited'

WHEN Exited = 'Retain' THEN 'Retained'

END AS Customer\_Status,

AVG(CreditScore) AS Avg\_CreditScore

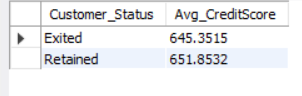
FROMbank\_churn

WHERE

Exited IN ('Exit', 'Retain')

GROUP BY

Customer\_Status;



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

**Ans--** SELECT

GenderID,

AVG(EstimatedSalary) AS avg\_salary,

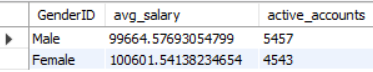
COUNT(\*) AS active\_accounts

FROM

customerdata

GROUP BY

GenderID; ;



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

**Ans--** select CreditScore, count(CustomerId) customer\_count from bank\_churn

where exited= 1

group by CreditScore

order by customer\_count desc limit 1;



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

**Ans--** SELECT cd.GeographyID, COUNT(\*) AS Num\_Active\_Customers

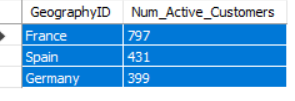
FROM customerdata AS cd

INNER JOIN bank\_churn AS bc ON cd.CustomerId = bc.CustomerId

WHERE bc.Tenure > 5 AND bc.IsActiveMember = 'Active Member'

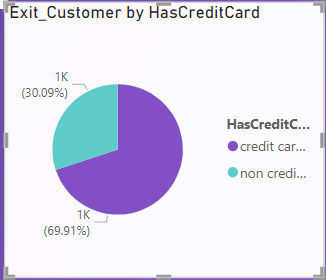
GROUP BY cd.GeographyID

ORDER BY Num\_Active\_Customers DESC;



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

**Ans—**From data it is clear that exit customers are higher in customers having credit card (69.91%) rather than not having credit card (30.09%).



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

**Ans--** SELECT NumOfProducts, COUNT(\*) AS Num\_Customers

FROM bank\_churn

WHERE Exited = 'Exit'

GROUP BY NumOfProducts

ORDER BY Num\_Customers DESC

LIMIT 1;;



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

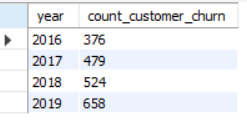
**Ans--** select year(bankDOJ) as year, count(c.CustomerId) as count\_customer\_churn

from bank\_churn b

inner join customerinfo c ON b.CustomerId= c.CustomerId

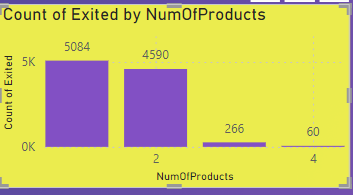
where Exited= 1

group by year(bankdoj);



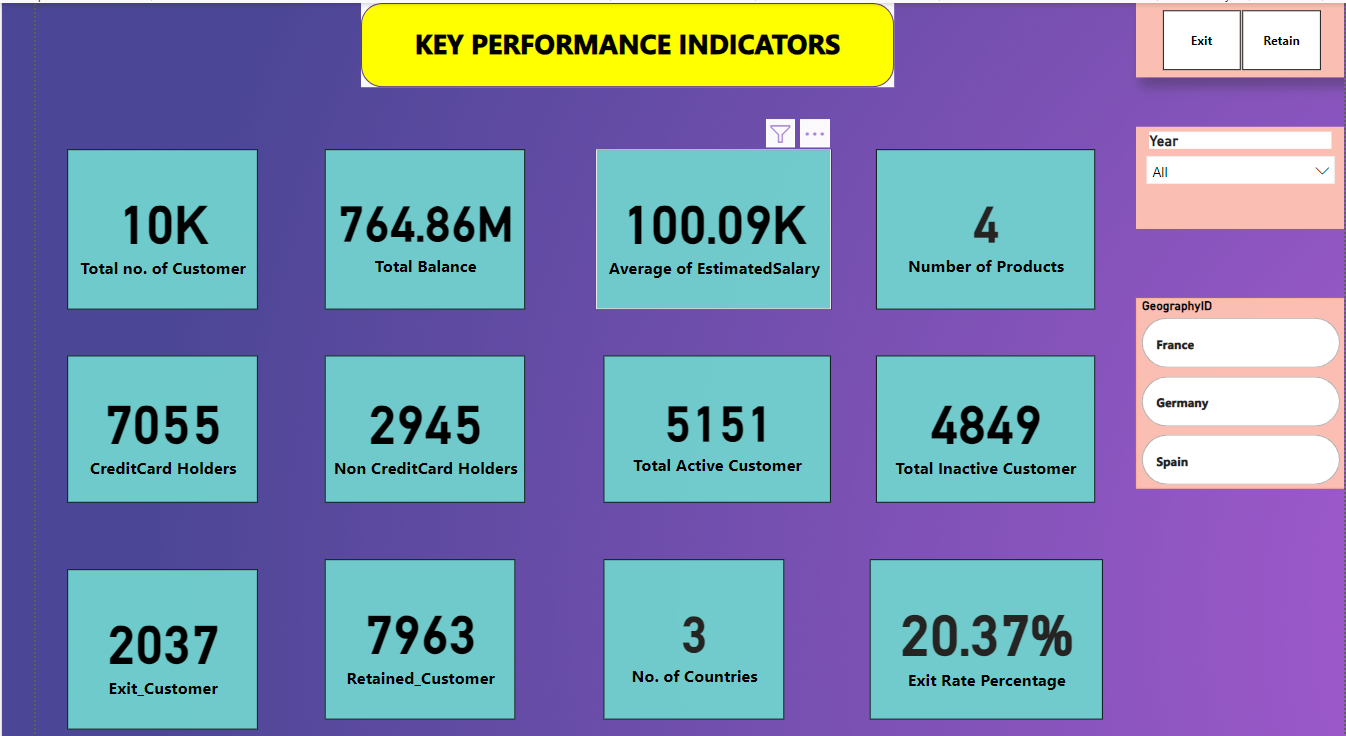
**11. Analyze the relationship between the number of products and the account balance for customers who have exited.**

**Ans—**From data it is clear that customers with less no of products have higher exit rate.



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

**12. Can you create a dashboard incorporating the visuals mentioned above and additionally derive more KPIs if possible?**



**Ans12 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)**

**Ans** SELECT

GeographyID,

GenderID,

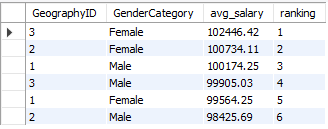
AVG(EstimatedSalary) AS average\_income,

RANK() OVER ( ORDER BY AVG(EstimatedSalary) DESC) AS gender\_rank

FROM

customerdata

GROUP BY GeographyID, GenderID;



**Q13. How many different tables are given in the dataset, out of these tables which table only consist of categorical variables?**

Ans13. Credit ID ,Exit ID , Is Active meber , Gender ,Has CrCard. Are the table which consists pf categorical variables .

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+).

**Ans** SELECT

CASE

WHEN Age BETWEEN 18 AND 30 THEN '18-30'

WHEN Age BETWEEN 31 AND 50 THEN '31-50'

ELSE '50+'

END AS AgeBracket,

Round(AVG(Tenure),2) AS AvgTenure

FROM

bank\_churn bc

JOIN

customerdata cd ON bc.CustomerId = cd.CustomerId

WHERE

bc.Exited = 'Exit'

GROUP BY

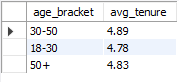
CASE

WHEN Age BETWEEN 18 AND 30 THEN '18-30'

WHEN Age BETWEEN 31 AND 50 THEN '31-50'

ELSE '50+'

END; ;



**Q14Using 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)**

Ans14.SELECT

GeographyID,

GenderID,

AVG(EstimatedSalary) AS average\_income,

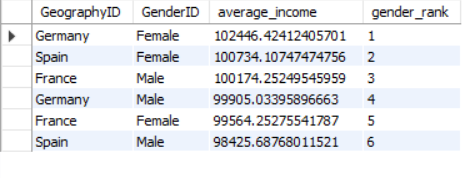
RANK() OVER ( ORDER BY AVG(EstimatedSalary) DESC) AS gender\_rank

FROM

customerdata

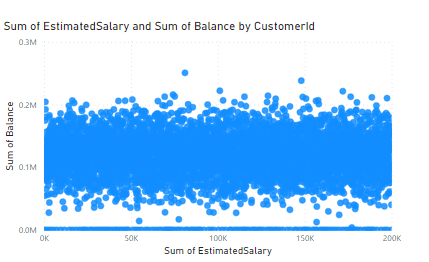
GROUP BY

GeographyID, GenderID;



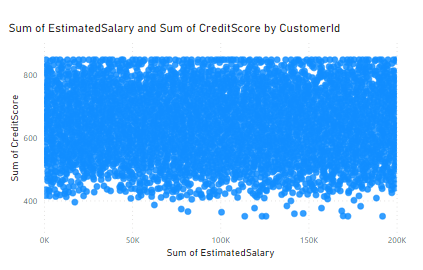
**Q17.Is there any direct correlation between the salary and the balance of the customers? And is it different for people who have exited or not?**

**Ans—**From following scatter plot it is clear that there is no correlation between salary and the balance.



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

**Ans--** From following scatter plot it is clear that there is no correlation between salary and credit score

.

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

**Ans--** with creditbucket as

(

select \*,

case when creditscore between 0 and 579 then 'Poor'

when creditscore between 580 and 669 then 'Fair'

when creditscore between 670 and 739 then 'Good'

when creditscore between 740 and 800 then 'Very Good'

else 'Excellent'

end as creditBucket

from bank\_churn

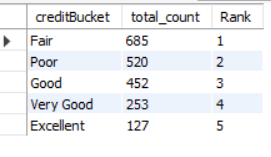
where exited = 'Exit')

select creditbucket, count(CustomerID) as total\_count,

dense\_rank() over(order by count(CustomerID) desc) as ranking

from creditbucket

group by creditbucket; ;



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

**Ans--** WITH AgeBuckets AS (

SELECT

CASE

WHEN c.Age BETWEEN 18 AND 30 THEN '18-30'

WHEN c.Age BETWEEN 31 AND 50 THEN '31-50'

ELSE '50+'

END AS age\_bucket,

CASE

WHEN b.HasCrCard = 'Credit card holder' THEN 1

ELSE 0

END AS has\_credit\_card

FROM

bank\_churn AS b

INNER JOIN

customerdata AS c ON b.CustomerId = c.CustomerId

),

AgeBucketSummary AS (

SELECT

age\_bucket,

COUNT(\*) AS total\_customers,

SUM(has\_credit\_card) AS credit\_card\_customers,

AVG(has\_credit\_card) AS avg\_customers\_with\_credit

FROM

AgeBuckets

GROUP BY

age\_bucket

)

SELECT

age\_bucket,

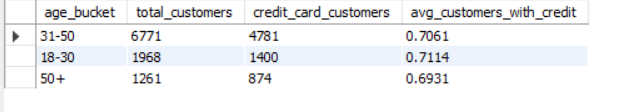
total\_customers,

credit\_card\_customers,

avg\_customers\_with\_credit

FROM

AgeBucketSummary;



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

**Ans--** WITH ChurnStats AS (

SELECT

c.GeographyID,

COUNT(DISTINCT c.CustomerId) AS count\_churn

FROM

customerdata c

JOIN

bank\_churn b ON c.CustomerId = b.CustomerId

WHERE

b.Exited = 'Exit'

GROUP BY

c.GeographyID

)

SELECT

GeographyID,

count\_churn,

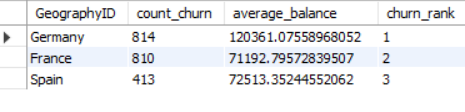
RANK() OVER (ORDER BY count\_churn DESC) AS churn\_rank

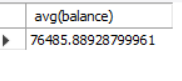
FROM

ChurnStats;--

---------- Average balance

select avg(balance) from bank\_churn;





**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(CustomerId, '\_', Surname) AS CustomerID\_Surname

FROM Customerdata;

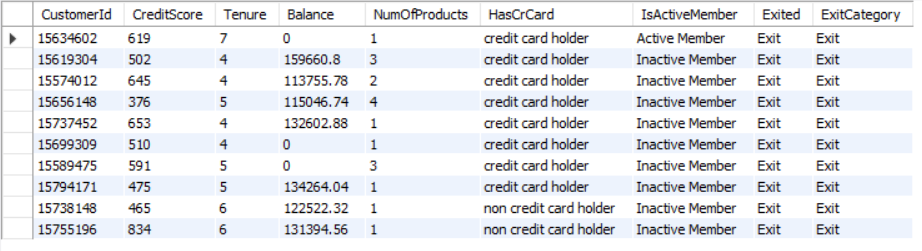


**Q23.Without using “Join”, can we get the “ExitCategory” from ExitCustomers table to Bank\_Churn table? If yes do this using SQL.**

SELECT \*, 'Exit' AS ExitCategory

FROM bank\_churn

WHERE Exited = 'Exit';



**Q24. Were there any missing values in the data, using which tool did you replace them and what are the ways to handle them?**

**SQL**: SQL queries can be used to identify missing values and perform imputation within a database.

**Excel**: Excel's data manipulation features allow users to identify missing values and replace them manually or with built-in functions

**Q25. 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 \*

FROM bank\_churn AS b

INNER JOIN customerdata AS c ON b.CustomerId = c.CustomerId;

SELECT c.CustomerId, c.Surname, b.IsActiveMember

FROM customerdata AS c

INNER JOIN bank\_churn AS b ON c.CustomerId = b.CustomerId

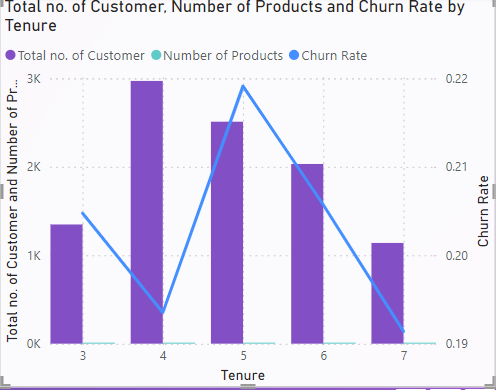
WHERE c.Surname LIKE '%on';



**Subjective Question:**

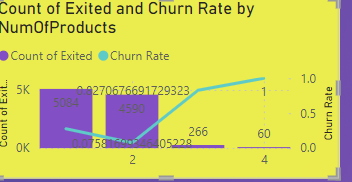
**Customer Behavior Analysis:** What patterns can be observed in the spending habits of long-term customers compared to new customers, and what might these patterns suggest about customer loyalty?

**Ans—** After Analyzing the spending habits of long-term customers compared to new customers can provide valuable insights into customer loyalty and behaviour that Customers with Tenure 4 has spent more and customers with tenure 5 has higher churn rate and loyalty of customers is proportional to tenure, so with higher tenure customers are more loyal.



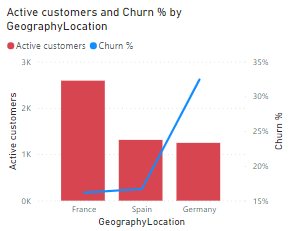
**Product Affinity Study:** Which bank products or services are most commonly used together, and how might this influence cross-selling strategies?

Ans—From the below data it can be observed that Customers with less no of products has less churn rate compare to the customers who bought more no of products.



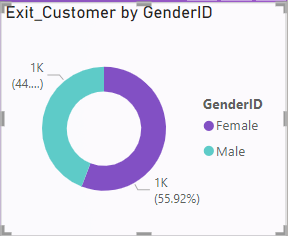
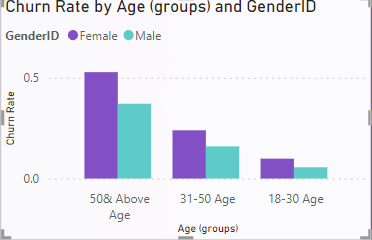
**Geographic Market Trends: How do economic indicators in different geographic regions correlate with the number of active accounts and customer churn rates?**

**Ans**—From graph it is clear that France has more active customers and churn rate is max in Germany.



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

Ans— Ans—From data it is clear that customer churn rate is proportional to the customer age. Female has higher churn rate compared to male.



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

**Ans—** Implementing tailored offers, incentives, and loyalty programs based on customer preferences and behaviour is an effective strategy for encouraging customer retention in the banking industry.

**Marketing Campaign Effectiveness:** How could you assess the impact of marketing campaigns on customer retention and acquisition within the dataset?

**Ans—** Provide personalized assistance and support to customers, addressing their needs and concerns promptly and effectively.

Resolve issues efficiently, demonstrating a commitment to customer satisfaction and loyalty.

Provide exclusive offers and benefits to customers who frequently purchase multiple products, incentivizing continued engagement and loyalty.

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

Ans—Customers with age above 50 are not satisfied with bank terminology or schemes. Customers having credit score below 700 are also leaving. Germany and France are having high churn rates.

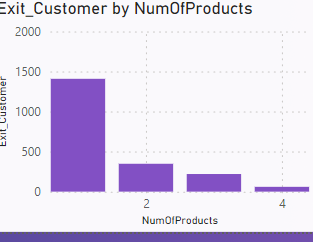
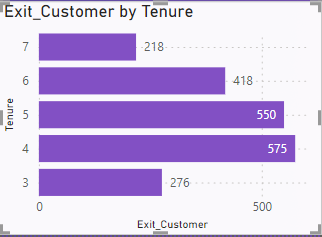
Low levels of customer satisfaction can lead to churn.

Dissatisfaction with customer Service, long wait times, and

Unresolved issues are common reasons why customers churn banks.

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

Ans— From given dataset Customers with tenure 4 and 5 are more likely to exit. Customers with less no of products are more likely to exit and we can not predict it from estimated salary.



**Q7.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?**

Ans—We can create conditional formatting to visually highlight customers at risk on basis of age condition having age >50, and no of products <2 also customers with credit score < 700.

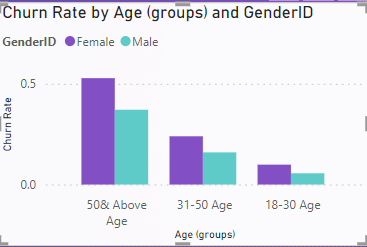
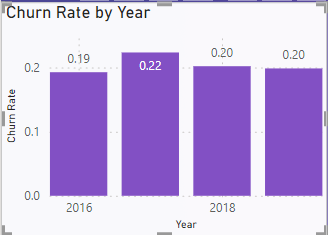
**Q8.What is the current churn rate per year and overall as well in the bank? Can you suggest some insights to the bank about which kind of customers are more likely to churn and what different strategies can be used to decrease the churn rate?**

Ans—In year 2016- 0.189 ,2017- 0.22, 2018- 0.20and in 2019- 0.20

Customers above age 50 and customers having less no of product are more likely to churn.

So if bank provide flexible pricing plans or customizable options that cater to different customer needs. Offering options that align with customers' changing requirements can increase satisfaction and retention.

2.Offer discounts or bonuses for customers who commit to long-term contracts or subscriptions. This can make customers feel valued and less likely to churn.



======================================================================================================================================================