

# CUSTOMER CHURN ANALYSIS

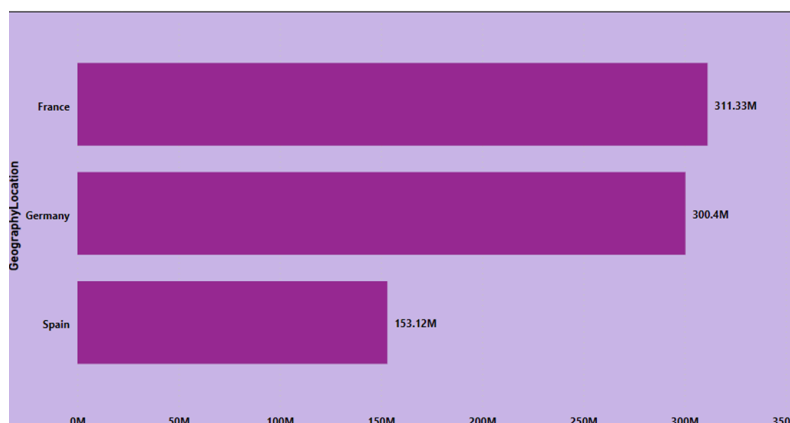
## PROBLEM STATEMENT:

- You are an analytical CRM (Customer Relationship Management) specialist hired by a bank to extract meaningful insights from various customer-related datasets.
- The bank aims to reduce customer churn, improve service delivery, and enhance customer satisfaction. They have provided you with datasets including customer demographics, transaction details, customer exit information, and active customer profiles.

## OBJECTIVE QUESTION

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

- To analyze the distribution of account balance across different regions, we can create a visualization that shows the distribution of account balances for each country represented in our dataset.
- I have created a column chart to show the Account balance graph of customers with the country present in x-axis.
- To visualize better, I have added a year. We can drill up and see the data year-wise as well based on the insights we want to gather from the chart.
- Based on our analysis, France has the highest account balance in all four years. Germany is the second highest and Spain is the lowest in the bracket.



File Name - Power Bi

Sheet Name - Objectives

Visualization used - Bar Chart

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

**SQL Query:**

with bank\_data as (

SELECT CustomerID, EstimatedSalary,(str\_to\_date(Bank\_doj,  
"%m/%d/%Y")) as DoJ\_date

FROM fabankcrm.customerinfo)

select \* , year(DoJ\_date), quarter(DoJ\_date) from bank\_data

where quarter(DoJ\_date)= 4

order by EstimatedSalary desc

limit 5;

	CustomerID	EstimatedSalary	DoJ_date	year(DoJ_date)	quarter(DoJ_date)
▶	15634359	199970.74	2016-11-29	2016	4
	15762331	199929.17	2017-10-07	2017	4
	15804211	199841.32	2019-12-25	2019	4
	15687913	199805.63	2018-12-11	2018	4
	15599792	199638.56	2018-11-16	2018	4

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

**SQL Query:**

select round(avg(NumOfProducts),2) as

Average\_number\_of\_products from bank\_churn

where HasCrCard = 1;

	Average_number_of_products
▶	1.53

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

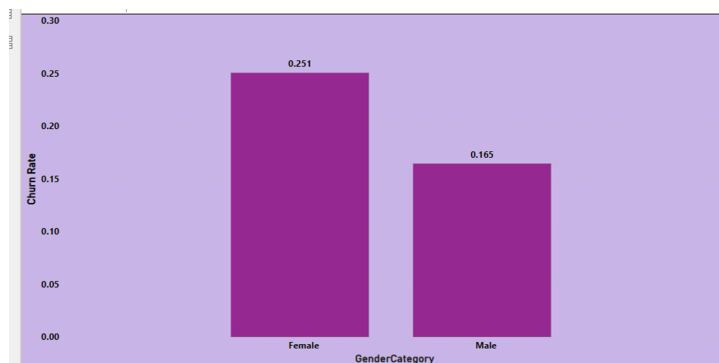
- To determine churn rate, we will first calculate the churn rate of overall customers.
- With the help of DAX, we will create three measures as I have shown below:

**Total Customers** = `COUNT('fabankcrm customerinfo'[CustomerId])`

**Total Exited customers** = `COUNTROWS(FILTER('fabankcrm bank_churn', 'fabankcrm bank_churn'[Exited] = 1))`

**Churn Rate** = `DIVIDE([Total Exited customers], [Total[Customers]])`

- Based on the above measures, we have drawn a bar chart to visualize the gender wise Churn rate. The churn rate for Female customers is 25% which is higher than the males that are 16.5%.



File Name - Power Bi

Sheet Name - Objectives

Visualization used - Column Chart

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

**SQL Query:**

```
select e.ExitCategory, round(avg(bc.CreditScore),2) as  
Average_credit_score from bank_churn bc  
  
join exitcustomer e ON bc.Exited = e.ExitID  
  
group by e.ExitCategory;
```

	ExitCategory	Average_credit_score
	Exit	645.35
	Retain	651.85

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

**SQL Query:**

```
select g.GenderCategory, a.ActiveCategory,  
round(avg(ci.EstimatedSalary),2) as Average_estimated_salary  
  
from customerinfo ci  
  
join bank_churn bc ON ci.CustomerID = bc.CustomerID  
  
Join activecustomer a ON bc.IsActiveMember = a.ActiveID  
  
Join gender g ON ci.GenderID = g.GenderID  
  
group by a.ActiveCategory, g.GenderCategory  
  
order by Average_estimated_salary desc;
```

	GenderCategory	ActiveCategory	Average_estimated_salary
►	Female	Inactive Member	101439.06
	Female		101439.06
	Male	Inactive Member	100181.21
	Male		100181.21
	Female	Active Member	99773.19
	Male	Active Member	99197.86

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

### **SQL Query:**

With count\_of\_churned as

(select count(\*) as churned\_customers from bank\_churn

where Exited = 1)

Select case

when CreditScore >=300 and CreditScore<580 then "Poor  
Credit score"

when CreditScore >=580 and CreditScore<670 then "Fair  
Credit score"

when CreditScore >=670 and CreditScore<740 then "Good  
Credit score"

when CreditScore >=740 and CreditScore<800 then "Very  
Good Credit score"

else "Excellent Credit score"

end as CreditScore\_segment, count(\*) as Total\_customers,

ROUND((SELECT churned\_customers FROM  
count\_of\_churned)/count(\*),2) as Exit\_rate

```
from bank_churn
group by CreditScore_segment
order by Exit_rate;
```

	CreditScore_segment	Total_customers	Exit_rate
►	Fair Credit score	3331	0.61
	Good Credit score	2428	0.84
	Poor Credit score	2362	0.86
	Very Good Credit score	1224	1.66
	Excellent Credit score	655	3.11

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

**SQL Query:**

```
select g.GeographyLocation, count(*) As Highest_Active_customers
from customerinfo ci
```

```
Join Bank_churn bc ON ci.CustomerID = bc.CustomerID
```

```
join geography g on ci.GeographyID = g.GeographyID
```

```
where bc.IsActiveMember = 1 and bc.Tenure>5
```

```
group by g.GeographyLocation
```

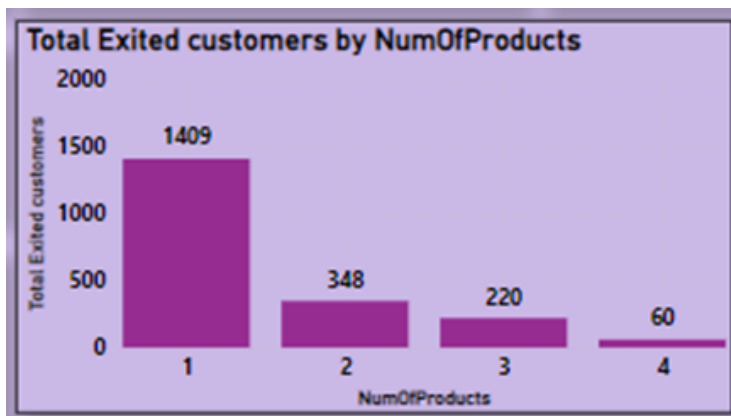
```
order by Highest_Active_customers desc
```

```
limit 3;
```

	GeographyLocation	Highest_Active_customers
►	France	797
	Spain	431
	Germany	399

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

- With the help of the below visual it seems, Customer churned for credit card holders are much higher than non-credit card holders.
- Credit card holders had 1424 customers churned and non-credit card holders had 613.



File Name - Power Bi

Sheet Name - Objectives

Visualization used - Column Chart



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

- To determine the most common products customer have used, is to first create a DAX measure which calculate the number of customers churned.
- We then take a parameter such as Balance to see the number of products vs Account balance so that we can see the relationship of products used by customer having the balance with the Bank.
- Dormant customer seems to have an edge over the customers who are active. We can explore more with the Tenure as well to see if they have bought the products and been with the Bank over 3 years, this implies they are Loyal with the Bank.
- The graph goes down from 1 to 4 with Highest to lowest number of products owned by customers. One being the highest and four being the lowest.

## 11. Examine the trend of customer joining over time and identify any seasonal patterns?

- First, we will extract the relevant data to observe any trend over years or months.
- Through the below SQL result, we can see the Total customers joined by the year and month and I have sorted the result by Total customers and joining year in a descending order.

- From the query we can observe a trend in 2019 the number of customers joining the bank has increased in a great number.
- The cluster is observed of customer joined in the last quarter of the year.
- France contributed the large number of customers coming from country.

### **SQL Query:**

```
select year(Bank_doj) as joining_year, month(Bank_doj) as
joining_month, count(*) as Total_customers,
g.GeographyLocation, Age from customerinfo ci
join geography g on ci.GeographyID = g.GeographyID
group by joining_year, joining_month, g.GeographyLocation, Age
order by Total_customers desc, joining_year desc;
```

### **Trend Observed:**

- Analysis of 27.77% increase in count of customers joining the bank between 2018 and 2019. The increased volume of customers between year 2018 and 2019 is shown in the below Line chart.
- With this, we have observed some key metrics which can lead to a breakthrough of our analysis.

1. France is the largest increase among country with

282 customers.

2. 'Female' accounted for most of the increase among gender with 389 and for 331 for 'male'.

3. '30-50' accounted for most of the increase among other age groups.

4. With 494 customers almost 27% were the credit card holders.



File Name - Power Bi

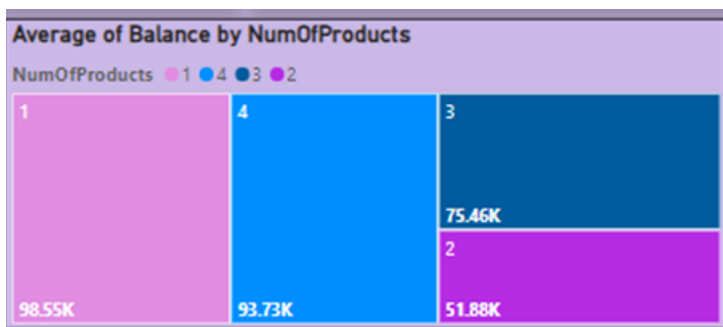
Sheet Name - Objectives

Visualization used - Line Chart

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

- Filter the data for churned customers, and we can see Account balance is drastically decreased between number of products 1 and 2 for the customers who are churned.

- Hence there is a relationship which can be seen where account balance is inversely proportional to the number of products.
- I have used tree map chart to show the visuals.



File Name - Power Bi

Sheet Name – Subjective 1

Visualization used – Tree Map

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

- Firstly, I have filtered out the data who have remained with the bank (non-churned)
- Across Age group, 30-50 had the most interesting recent trend in Account balance and started trending up in 2016, rising by 64.29% (values rising by: 52,249,346.26) in 3 years.

- Likewise, customer base for the Age-group, 50+ had the lowest account balance observe which can further investigated by the team.

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

- Columns that consist entirely of categorical variables are:

1. Country
2. Gender
3. Credit\_card\_status
4. Activity\_status
5. Churn\_status

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

**SQL Query:**

with ranked\_salaries as (

```
    select g.GenderID, geo.GeographyID, g.GenderCategory,
    geo.GeographyLocation, avg(ci.EstimatedSalary) as
    Average_Salary,
```

```
    dense_rank() over(partition by g.GenderID order by
    avg(ci.EstimatedSalary) DESC) As ranked_salary
```

```
    from customerinfo ci join gender g ON ci.GenderID =
    g.GenderID
```

```

Join geography geo ON ci.GeographyID = geo.GeographyID

group by g.GenderID, geo.GeographyID, g.GenderCategory,
geo.GeographyLocation)

```

```

select rs.GenderCategory as Gender, rs.GeographyLocation as
Country, round(rs.Average_salary,2) as

```

```

Average_Salary, rs.ranked_salary from ranked_salaries rs;

```

	Gender	Country	Average_Salary	ranked_salary
►	Male	France	100174.25	1
	Male	Germany	99905.03	2
	Male	Spain	98425.69	3
	Female	Germany	102446.42	1
	Female	Spain	100734.11	2
	Female	France	99564.25	3

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+).

### SQL Query:

```

select

```

```

case

```

```

    when ci.Age>17 and ci.Age<=30 then '18-30'

```

```

    when ci.Age>29 and ci.Age<=50 then '30-50'

```

```

else '50+'

```

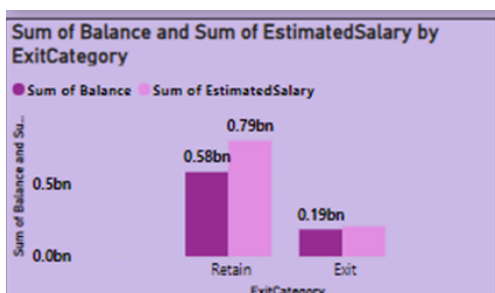
```
end as Age_group,  
round(avg(bc.Tenure),2) as Average_Tenure from bank_churn bc  
join customerinfo ci ON bc.CustomerId = ci.CustomerId  
  
where bc.Exited = 1    -- exited=1 defines the customers who are  
churned.  
  
group by Age_group;
```

	Age_group	Average_Tenure
▶	30-50	4.89
	18-30	4.78
	50+	4.83

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

- To analyze the correlation between salary and balance for customers who have exited or retained with the bank, we can use scatter plots and add trendlines to visualize the relationship as you can see in the below snapshots.
- The trendline for churned customers is slightly below the 100k account balance mark, while the trendline for non-churned customers is slightly upwards over 100k mark. This indicates that, in general, non-churned customers tend to have higher account balances for a given salary level compared to churned customers.

- The positive slope of both trendlines suggests a positive correlation between salary and account balance, meaning higher salaries are associated with higher account balances, for both customer segments.
- Hence, it is confirmed that there is a direct correlation between these two variables.



File Name - Power Bi

Sheet Name – Subjective 1

Visualization used – Clustered Column Chart

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

- To analyze the correlation between salary and credit score for the customers, we can use scatter plots and add trendlines to visualize the relationship as you can see in the below snapshots.
- Based on our scatter plot between average credit score and estimated salary the correlation is slightly decreasing, it



suggests a weak negative correlation between these two variables.

- A weak negative correlation implies that as the estimated salary increases, the average credit score tends to decrease slightly, or vice versa. However, it's important to note that the correlation is weak, meaning the relationship is not very strong or pronounced.



File Name - Power Bi

Sheet Name – Subjective 1

Visualization used – ScatterPlot Chart

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

**SQL Query:**

```
with cte_credit_score_buckets as (  
  select bc.CustomerId,  
  CASE
```

```

        when bc.CreditScore between 800 and 850 then 'Excellent'
    when bc.CreditScore between 740 and 799 then 'Very Good'
    when bc.CreditScore between 670 and 739 then 'Good'
    when bc.CreditScore between 580 and 669 then 'Fair'
    when bc.CreditScore between 300 and 579 then 'Poor'
    else 'Unknown'
END as credit_score_bucket, bc.Exited
from bank_churn bc)

select credit_score_bucket, count(*) as num_churned_customers,
rank() over(order by count(*) desc) as rank_by_churned
from cte_credit_score_buckets
where Exited = 1
group by credit_score_bucket
order by rank_by_churned;

```

credit_score_bucket	num_churned_customers	rank_by_churned
Fair	685	1
Poor	520	2
Good	452	3
Very Good	252	4
Excellent	128	5

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

**SQL Query:**

```
with cte_age_buckets as (  
    select ci.CustomerId,  
    case  
        when ci.Age between 18 and 30 then '18-30'  
        when ci.Age between 31 and 50 then '31-43'  
        when ci.Age between 44 and 56 then '44-56'  
        when ci.Age between 57 and 66 then '57-66'  
        else '66+'  
    end as Age_group, bc.HasCrCard from customerinfo ci inner join  
    bank_churn bc ON ci.CustomerId = bc.CustomerId),  
cte_having_credit_card as (  
    select Age_group, count(case when HasCrCard =1 then 1  
    end) as num_credit_card_holders  
    from cte_age_buckets group by Age_group)  
select Age_group, num_credit_card_holders from  
cte_having_credit_card  
where num_credit_card_holders < (select  
avg(num_credit_card_holders) from cte_having_credit_card)  
order by num_credit_card_holders;
```

Age_group	num_credit_card_holders
66+	160
57-66	357
44-56	357
18-30	1400

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

**SQL Query:**

```
with cte_churn_info as (
select ci.GeographyID, bc.Exited, bc.Balance from customerinfo ci
INNER join bank_churn bc on ci.CustomerId = bc.CustomerId),
```

```
cte_churn_count as (
select GeographyID, COUNT(CASE WHEN Exited = 1 THEN 1
END) as churn_count from cte_churn_info
GROUP by GeographyID),
```

```
cte_avg_balance as (
select GeographyID, AVG(Balance) as avg_balance from
cte_churn_info
GROUP BY GeographyID)
```

```
select g.GeographyLocation,
```

```

cc.churn_count,

round(ab.avg_balance,2) as Average_balance,

RANK() OVER (ORDER BY cc.churn_count DESC,
ab.avg_balance DESC) as location_rank

from geography g

LEFT join cte_churn_count cc on g.GeographyID = cc.GeographyID

LEFT join cte_avg_balance ab on g.GeographyID =
ab.GeographyID

order by location_rank;

```

	GeographyLocation	churn_count	Average_balance	location_rank
▶	Germany	814	119730.12	1
	France	810	62092.64	2
	Spain	413	61818.15	3

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”.

### **SQL Query:**

```

select concat(CAST(CustomerId as CHAR), '_', Surname) as
CustomerID_Surname

from customerinfo;

```

	CustomerID_Surname
▶	15634602_Hargrave
	15647311_Hill
	15619304_Onio
	15701354_Boni
	15737888_Mitchell

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

- Yes, we can retrieve the "ExitCategory" from the "ExitCustomers" table and merge it into the "Bank\_Churn" table without using a JOIN by using a subquery. Here's the Sql query:

### **SQL Query:**

```
SELECT CustomerId,Exited
FROM bank_churn,exitcustomer
WHERE Exited=Exited;
```

- In this query, the subquery (SELECT ExitCategory FROM ExitCustomers ec WHERE ec.CustomerId = bc.CustomerId) is used to fetch the "ExitCategory" from the "ExitCustomers" table for each customer in the "Bank\_Churn" table. The subquery is correlated to the outer query by matching the "CustomerId" from both tables.



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”.

- 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.

### **SQL Query:**

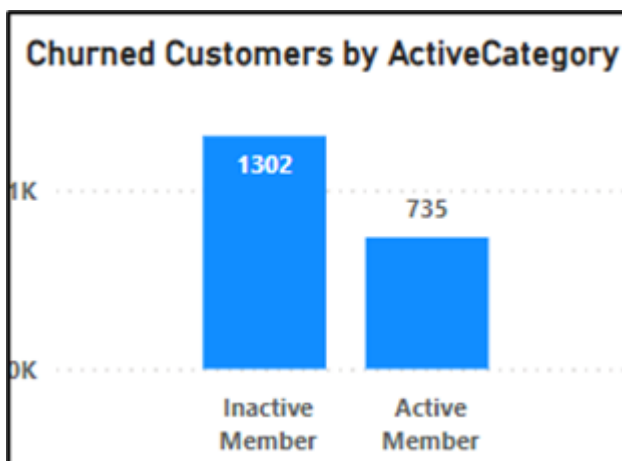
```
select ci.CustomerId, ci.Surname,  
  
CASE when bc.IsActiveMember = 1 then 'Active' else 'Inactive'  
END as Activity_status  
  
from customerinfo ci  
  
join bank_churn bc on ci.CustomerId = bc.CustomerId  
  
where ci.Surname like '%on';
```

	CustomerId	Surname	Activity_status
▶	15788218	Henderson	Active
	15750181	Sanderson	Inactive
	15788448	Watson	Active
	15585768	Cameron	Active
	15592461	Jackson	Active
	15640635	Cannon	Active



26. Can you observe any data disruption in the Customer's data? As a hint it's present in the IsActiveMember and Exited columns. One more point to consider is that the data in the Exited Column is absolutely correct and accurate.

- There is some noticeable discrepancy in the Customers data as we can see from the below graph, among the customers who have already exited the bank are still showing as “Active Member” and the number for that is quite large as 735 which should not be there because as per logic:
- Churned customers = Inactive customers



# 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?

- By analyzing spending habits of long-term customers compared to new customers, we can gain valuable insights into customer loyalty.
- Created a calculated column named "Customer Category" to segment customers:

1. "Long-Term": Tenure > 3 years
2. "New": Tenure <= 3 years

- Created a Stacked bar chart showing the number of customers within two customer segments as defined.

- Within each category, I have also included a different category which is credit card information. Credit card status shows the customer with the credit score and the average amount spent in both segments.

- Some key insights to summarize from our analysis:

1. Customer under 'New' segment in Excellent category made up 11.52% in Average account balance.
2. Germany had the highest Average of account balance at 119.5K followed by France at 62K and Spain 61.8K.

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

- The most common services used by customers are Credit card and NumOfProducts are the number of products customers has purchased.
- I have created a matrix showing the less number if products tend to have more number of credit cards.
- if customers who have purchased the more products are also likely to go for Credit card service, So this can impact the

cross-selling strategy where bank wants to sell its service where customer has already the three or four number of products

Credit Card Category vs Tenure			
Tenure	credit card holder	non credit card holder	Total
3	963	385	1348
4	2056	915	2971
5	1791	719	2510
6	1444	588	2032
Total	7055	2945	10000

File Name - Power Bi

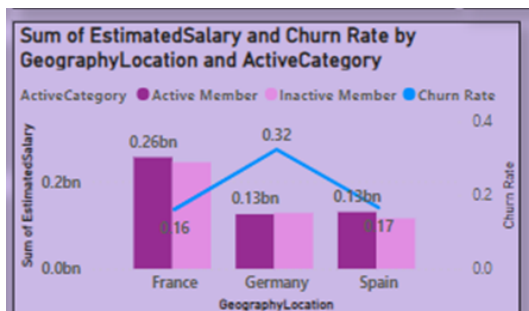
Sheet Name – Subjective 2

Visualization used – Matrix Table

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

- Talking about economic factors affecting churn rate in different geographic regions such as Germany, Spain and France

- We can see Active customer in Germany rose up to 51.60% of Churn Rate and have the lowest number of products bought by the customers.
- France being the second region contributing 11.5% of churn rate and lastly Spain with the 10.7% of churn rate.



File Name - Power Bi

Sheet Name – Subjective 2

Visualization used – Line And Clustered Chart

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

- To assess the financial risk posed by different demographic segments based on customer profiles, we can create a measure which calculates the financial risk score by summing the product of Account Balance and CreditScore for churned

customers (where Exited = 1). Higher scores indicate higher financial risk.

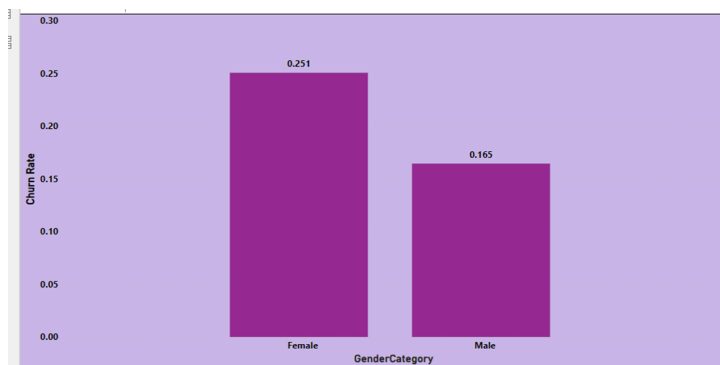
➤ DAX Expression:

Financial Risk Score = SUMX (FILTER('bank\_churn', [Exited] = 1),  
[Balance] \* [CreditScore])

- We have created a scatter plot visualization for Financial Risk Score by Age.
- I have also taken gender into consideration and created a Column chart visualization for Financial Risk Score by Gender.
- Scatter plot of financial risk scores against age shows a slightly decreasing trendline, lower than a straight line, it could indicate the following:
  - Lower risk with increasing age: This could imply that older customers generally pose a lower financial risk to the bank compared to younger customers.
  - Stability and financial maturity: With increasing age, individuals typically have more stable income sources, better

financial management skills, and a lower propensity for defaulting on loans or engaging in risky financial behaviour.

- Column chart shows the positive increase of financial risk score for Female customers.



File Name - Power Bi

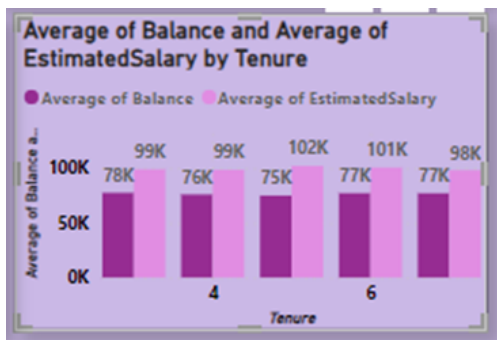
Sheet Name – Objective

Visualization used – Column Chart

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

- To predict the lifetime (tenure) value of different customer segments, I have Calculated Customer Lifetime Value (CLV):

- Create a new calculated column in my dataset to calculate CLV for each customer, with the help of below formula.
- $CLV = \text{Balance} * \text{Tenure} * (1 - \text{Churn\_Rate})$
- Where Churn\_Rate is the rate at which customers are leaving the bank. This rate is based on the churn\_status column (churned/not churned).
- Tenure of six accounted for 25.00% of Sum of Customer Lifetime Value.



File Name - Power Bi

Sheet Name - Subjective 2

Visualization used - Clustered Column Chart



## 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?

- Firstly, we should identify the marketing campaign data and for that we would need tables that contain information about marketing campaigns, such as campaign name, start date, end date, target audience, and channels used.
- Secondly, we will link campaign data with customer data by joining the marketing campaign data with the customer data using a common identifier (e.g., CustomerId) to analyze the impact on customer behavior.
- Later we can calculate customer retention and acquisition rates by defining retention as the percentage of customers who remained with the bank after the campaign compared to before, and acquisition as the percentage of new customers acquired during the campaign period.
- For visualizations we can create Line charts or Bar charts to compare retention and acquisition rates before, during, and after the campaigns. This can help identify the impact of each campaign on customer behavior.

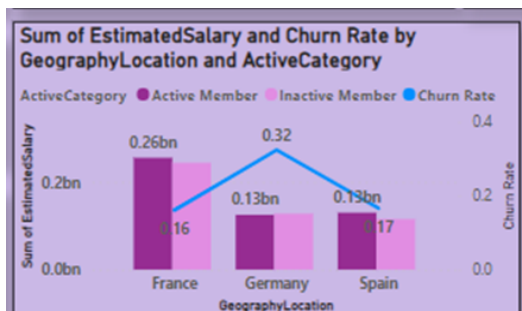
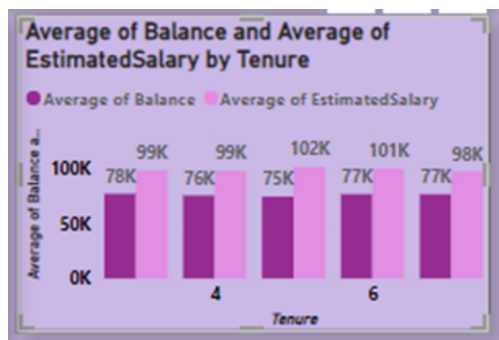
## 7. Customer Exit Reasons Exploration: Can you identify common characteristics or trends among customers who have exited that could explain their reasons for leaving?

- To identify common characteristics or trends among customers who have exited the bank, we can perform a detailed analysis of the data focusing on churned customers.
- We will start by filtering the dataset to include only customers who have churned (Exited).
- We will explore the Demographics by looking at the demographic information (gender, age, country) of churned customers to see if there are any patterns. For example, are there certain countries or age groups more likely to churn?
- I found age-group 18-30+ has the lowest churn rate comparing to the age group of 30-50
- Females have the highest churn rate on all three age groups.
- France is the leading country with the highest estimated salary of customers with 84M, with that Germany and Spain are the other two having estimated salary 80M and 43M
- Customer with Credit card holder – Country wise and Year wise, the number of highest products brought are with the France whereas Germany and Spain are the rest two countries.

- We can check the tenure as well, whether new customers are more likely to get churned. I have created a column chart to see the number of customers with the Tenure to get the high-level scenario of tenure and customers relationship.

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

- To determine if 'Tenure', 'NumOfProducts', 'IsActiveMember', and 'EstimatedSalary' are important for predicting if a customer will leave the bank, we can create visualizations and perform analysis using Power BI.
- Created a stacked column chart for number of products by churn status and Activity status. Customers with the active status tends to stay with the Bank rather than non-active status.
- Also, created a donut chart to visualize the churn status by Activity status and as we have seen customer churn rate is 20% but the active members don't make difference.
- For active members, number of customers churned are 4849.
- For inactive members, number of customers churned are 5151.



File Name - Power Bi

Sheet Name – Subjective 2

Visualization used – Clustered Column Chart and Line  
and Clustered Column Chart

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

➤ To segment customers based on demographics and account details using SQL, I have used the following queries:

➤ Segmentation by Gender:

**SQL Query:**

```

SELECT ci.GenderID, COUNT(*) AS TotalCustomers,
round(AVG(ci.Age)) AS AvgAge, round(AVG(bc.Balance),2) AS
AvgBalance,

round(AVG(bc.NumOfProducts),2) AS AvgNumOfProducts,
round(AVG(ci.EstimatedSalary),2) AS AvgSalary

FROM customerInfo ci join bank_churn bc ON ci.CustomerId =
bc.CustomerId

GROUP BY GenderID;

```

	GenderID	TotalCustomers	AvgAge	AvgBalance	AvgNumOfProducts	AvgSalary
▶	2	4543	39	75659.37	1.54	100601.54
	1	5457	39	77173.97	1.52	99664.58

### **SQL Query:**

#### ➤ Segmentation by Country:

```

SELECT g.GeographyLocation, COUNT(*) AS TotalCustomers,
round(AVG(ci.Age)) AS AvgAge, round(AVG(bc.Balance),2) AS
AvgBalance,

round(AVG(bc.NumOfProducts),2) AS AvgNumOfProducts,
round(AVG(ci.EstimatedSalary),2) AS AvgSalary

FROM customerInfo ci join bank_churn bc ON ci.CustomerId =
bc.CustomerId

join geography g ON ci.GeographyID = g.GeographyID

GROUP BY g.GeographyLocation;

```

	GeographyLocation	TotalCustomers	AvgAge	AvgBalance	AvgNumOfProducts	AvgSalary
▶	France	5014	39	62092.64	1.53	99899.18
	Spain	2477	39	61818.15	1.54	99440.57
	Germany	2509	40	119730.12	1.52	101113.44

### SQL Query:

#### ➤ Segmentation by Credit Card Status:

```
SELECT cc.category, COUNT(*) AS TotalCustomers,
round(AVG(ci.Age)) AS AvgAge, round(AVG(bc.Balance),2) AS
AvgBalance,
```

```
round(AVG(bc.NumOfProducts),2) AS AvgNumOfProducts,
round(AVG(ci.EstimatedSalary),2) AS AvgSalary
```

```
FROM customerInfo ci join bank_churn bc ON ci.CustomerId =
bc.CustomerId
```

```
join creditcard cc ON bc.HasCrCard = cc.CreditID
```

```
GROUP BY cc.category;
```

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?

- To 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, follow these steps:

➤ First, identify at-risk customers by using the churn\_status column to know who have churned (exit) and those who have not (retain). We can create a measure to calculate the churn rate.

➤ To apply conditional formatting for at-risk customers:

1. Go to the visualization where you want to highlight at-risk customers (a chart or table).
2. Select the column representing customerID or names.
3. Go to the 'Conditional Formatting' option in the formatting pane.
4. Choose 'Font Color' or 'Background Color' based on your preference.
5. Define the condition based on the churn rate measure. For example, if the churn rate is above a certain threshold, set the formatting to highlight those customers.
6. To evaluate the impact of credit card rewards:
7. We can create a new measure to calculate the retention rate for customers with credit cards and without credit cards separately.
8. Use a visualization (bar/column chart) to compare the retention rates of these two groups.
9. Lastly, we can apply the same conditional formatting as above to visually highlight any differences in retention rates between credit card and non-credit card holders.

CustomerId	Tenure	NumOfProducts	HasCrCard	Churn Risk
15565706	4	1	1	Medium
15565714	3	2	0	Low
15565878	4	2	1	Low
15565879	5	2	1	Low
15565996	4	2	1	Low
15566292	7	2	0	Low
15566312	3	3	1	Low
15566378	6	1	0	Medium
15566380	5	2	1	Low
15566467	3	2	1	Low
15566531	4	1	0	Medium
15566633	5	2	1	Low
15566689	5	2	1	Low
15566735	5	2	1	Low

### Evaluating Credit Card Rewards Impact

1. **Create a new measure** called "Retained Customers" that counts customers who have not exited:

Retained Customers =

`COUNTROWS(FILTER(Bank_Churn, Bank_Churn[Exited] = 0))`

2. **Create another measure** called "Retained Credit Card Customers" that counts customers with credit cards who have not exited:

Retained Credit Card Customers = `CALCULATE([Retained Customers], Bank_Churn[HasCrCard] = 1)`

3. **Create a KPI visual** to track the "Retained Credit Card Customers" percentage out of total customers:

Retained Credit Card Customers % = `DIVIDE([Retained Credit Card Customers], [Retained Customers])`



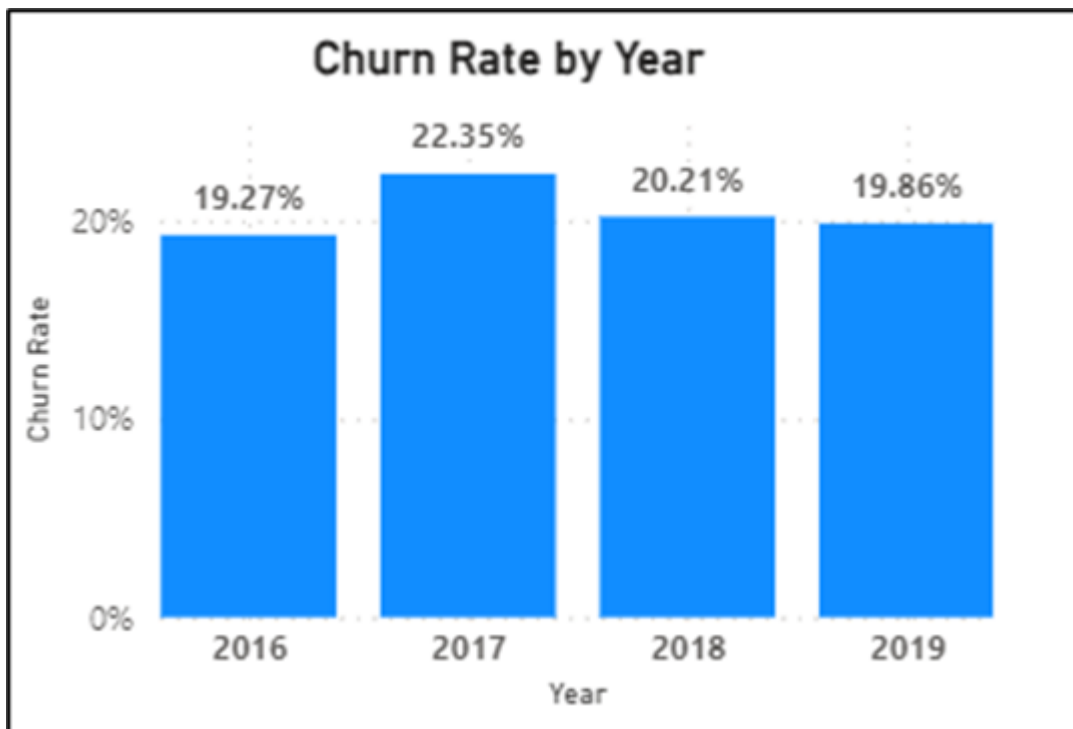
4. Monitor the KPI over time to evaluate the impact of credit card rewards on customer retention. An increase in the percentage indicates that credit card rewards are effectively reducing churn among credit card holders.

By combining conditional formatting to identify high-risk customers and measures to track retained credit card customers, you can gain valuable insights into churn patterns and the effectiveness of credit card rewards in improving customer 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 are the different strategies that can be used to decrease the churn rate.

- To determine the current churn rate per year and overall, in the bank, as well as to provide insights and strategies to decrease the churn rate.
- I used DAX measure Churn rate which I previously calculated.
- Created visualizations to analyze the churn rate based on demographics (age, gender, country), credit score, tenure, number of products, credit card status and activity status.
- Use slicers and filters to drill down into specific segments of customers who are more likely to churn.
- Insights and strategies:

1. Based on my analysis, The customers with lower credit scores, lower balances, shorter tenures, fewer products, and lower salaries might be more likely to churn.
2. Strategies to decrease the churn rate:
3. Improve customer engagement through personalized interactions and targeted marketing campaigns.
4. Offer loyalty programs or incentives to retain customers.
5. Improve service delivery and address customer complaints promptly.

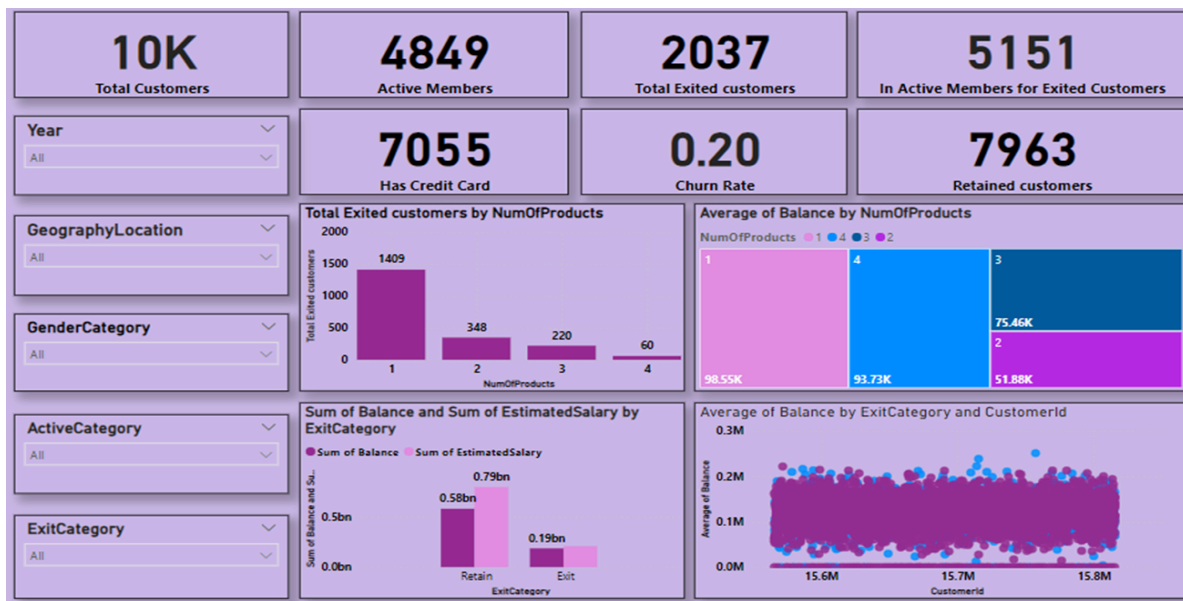
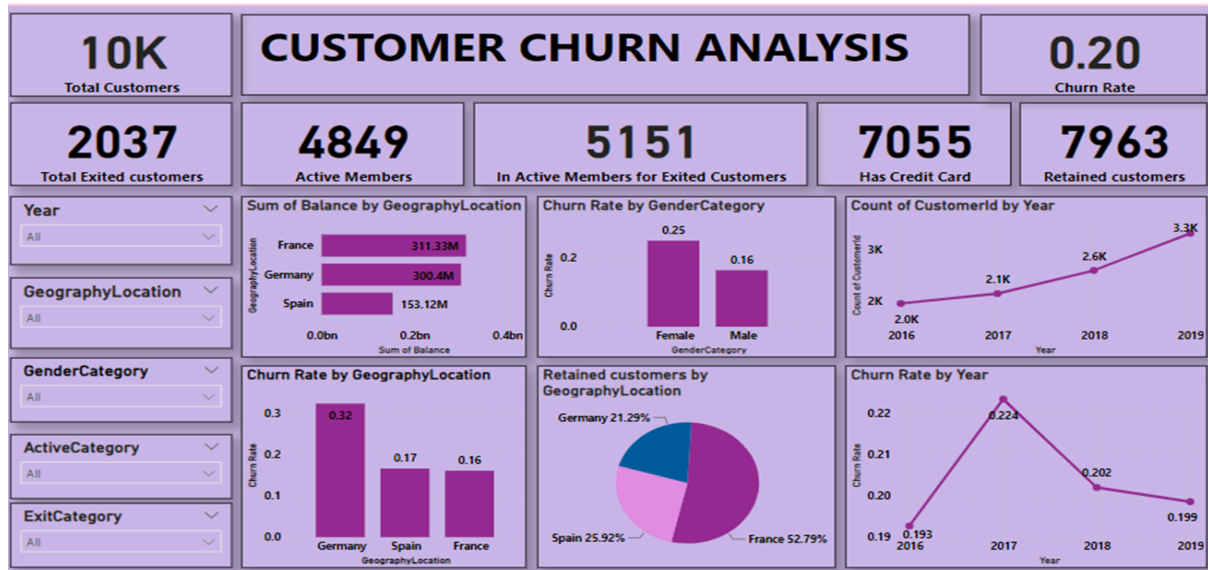


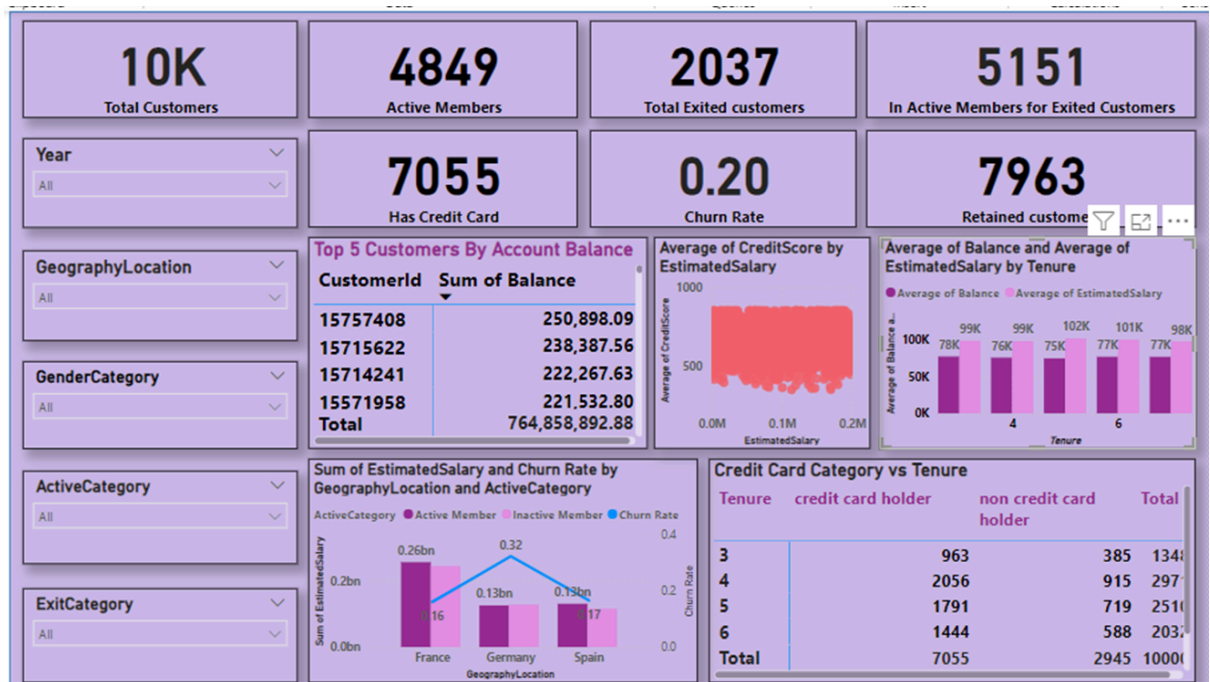
12. Create a dashboard incorporating all the KPIs and visualization related metrics. Use a slicer in order to assist in selection in the dashboard.

- Created multiple reports to answer the Subjective problems.

➤ Please check all the reports attached below.

# DASHBOARD





13. 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.

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

### **SQL Query:**

```
ALTER TABLE bank_churn  
RENAME COLUMN HASCrCard To Has_creditcard;
```

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

## **Conclusion**

- Effective use of CRM analytics can significantly reduce customer churn, improve service delivery, and enhance customer satisfaction.
- The insights derived from the provided datasets offer valuable guidance for strategic decision-making.
- Implementing the recommended actions will enable the bank to meet customer needs better and foster long-term loyalty.





