**Learners have to come up with a Report to support the answers to the following questions and suggestions**

**Objective Questions:**

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

Ans: SELECT geo.GeographyLocation,

ROUND(SUM(bc.Balance),2) AS balances

FROM bank\_churn bc

JOIN customerinfo ci USING (CustomerID)

JOIN geography geo USING (GeographyID)

GROUP BY geo.GeographyLocation ;

Output:

A screenshot of a computer

Description automatically generated

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

ANS:

SELECT \*

FROM customerinfo

WHERE month(BankDOJ) IN (10,11,12)

ORDER BY EstimatedSalary DESC

LIMIT 5;

Output:

A screenshot of a computer

Description automatically generated

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

Ans:

SELECT CustomerId, ROUND(AVG(NumOfProducts),0) As avg\_products

FROM bank\_churn

WHERE HasCrCard = 1

GROUP BY CustomerId;

Output:

A screenshot of a computer

Description automatically generated

1. Determine the churn rate by gender for the most recent year in the dataset.
2. Compare the average credit score of customers who have exited and those who remain. (SQL)

**ANS:**

SELECT (CASE WHEN Exited = 1 THEN 'Exited' ELSE 'Remain'END ) AS exited\_remain,

AVG(CreditScore) AS avg\_creditscore

FROM bank\_churn

GROUP BY exited\_remain

Output:

A screenshot of a computer

Description automatically generated

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

ANS:

WITH active\_avg\_est\_salary AS

(SELECT g.GenderCategory AS gender, ROUND(AVG(c.EstimatedSalary),2) AS active\_avg\_est\_salary

FROM customerinfo c

JOIN bank\_churn b USING (CustomerID)

JOIN gender g USING (genderID)

WHERE IsActiveMember = 1

GROUP BY g.GenderCategory),

inactive\_avg\_est\_salary AS

(SELECT g.GenderCategory AS gender, ROUND(AVG(c.EstimatedSalary),2) AS inactive\_avg\_est\_salary

FROM customerinfo c

JOIN bank\_churn b USING (CustomerID)

JOIN gender g USING (genderID)

WHERE IsActiveMember = 0

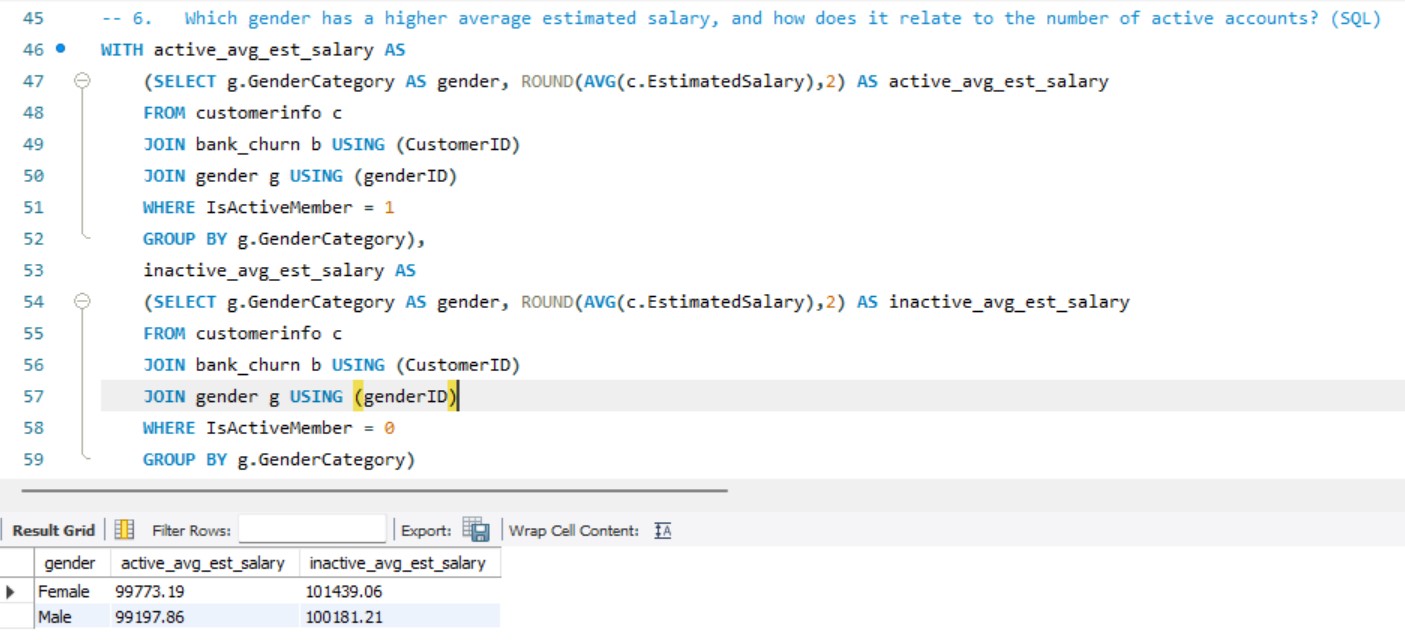
GROUP BY g.GenderCategory)

SELECT \*

FROM active\_avg\_est\_salary a

JOIN inactive\_avg\_est\_salary i USING(gender);

Output:



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

SELECT CASE WHEN CreditScore < 600 THEN 'Poor(Less Than 600)'

WHEN CreditScore >= 600 AND CreditScore < 700 THEN 'Fair(Between 600 And 700)'

WHEN CreditScore >= 700 AND CreditScore < 800 THEN 'Good(Between 700 And 800)'

ELSE 'Excellent(More than 800)'

END AS segments, Count(Exited) As cnt\_exited

FROM bank\_churn

WHERE Exited = 1

GROUP BY segments

ORDER By cnt\_exited DESC;

Output:

A screenshot of a computer

Description automatically generated

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

ANS: SELECT g.GeographyLocation, Count(c.CustomerId) AS active\_customers

FROM customerinfo c

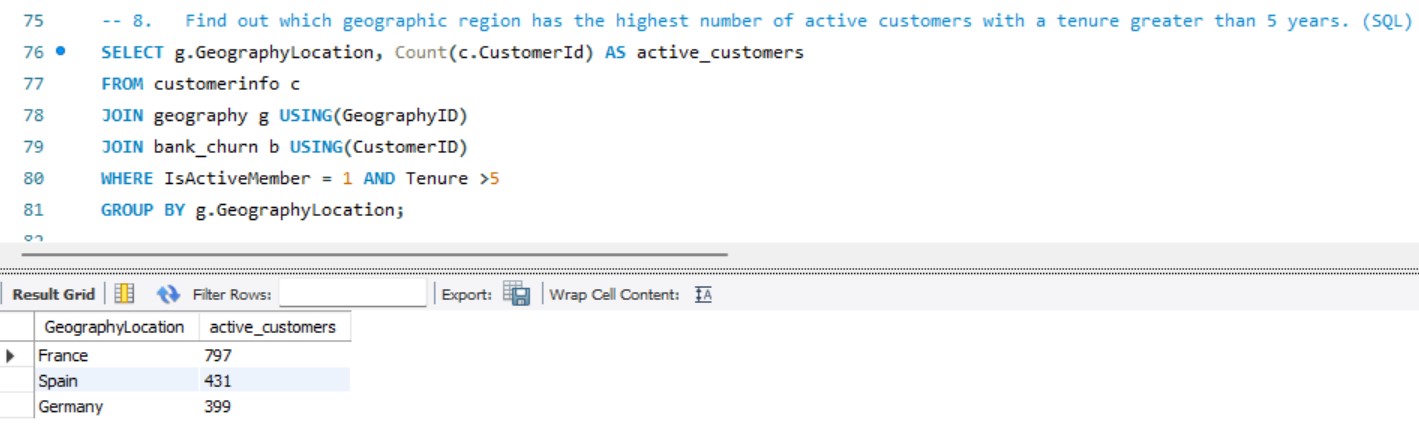
JOIN geography g USING(GeographyID)

JOIN bank\_churn b USING(CustomerID)

WHERE IsActiveMember = 1 AND Tenure >5

GROUP BY g.GeographyLocation

Output:



1. What is the impact of having a credit card on customer churn, based on the available data?
2. For customers who have exited, what is the most common number of products they have used?

Ans: SELECT NumOfProducts,count(CustomerId) AS no\_of\_customers

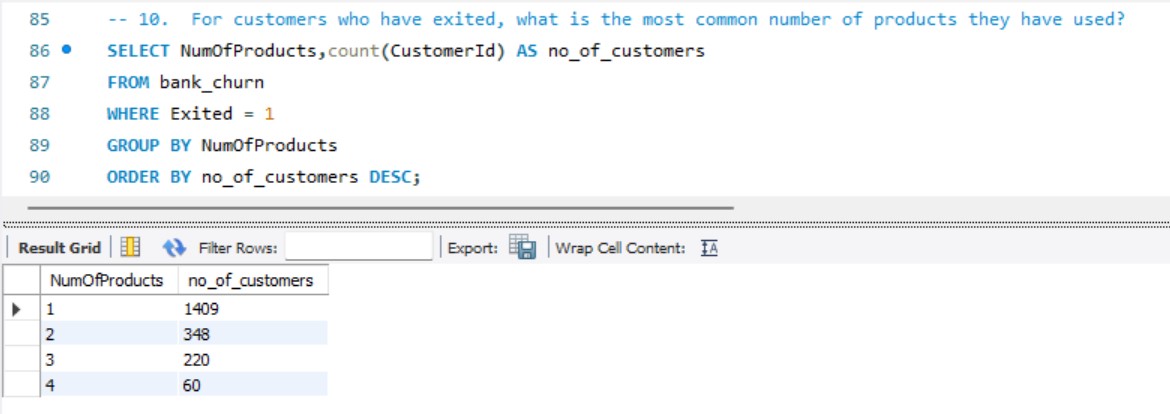
FROM bank\_churn

WHERE Exited = 1

GROUP BY NumOfProducts

ORDER BY no\_of\_customers DESC;

Output:



1. 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.

Ans: SELECT YEAR(BankDOJ) AS join\_year,

MONTHNAME(BankDOJ) AS join\_month,

COUNT(CustomerID) AS Customers

FROM customerinfo

GROUP BY join\_year,join\_month

ORDER BY join\_year DESC,join\_month;

Output:

A screenshot of a computer

Description automatically generated

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

Ans: SELECT NumOfProducts,

Round(AVG(Balance),2) AS avg\_balance

FROM bank\_churn

WHERE Exited =1

GROUP BY NumOfProducts

Order BY NumOfProducts ASC;

Output:

A screenshot of a computer

Description automatically generated

1. Identify any potential outliers in terms of balance among customers who have remained with the bank.
2. How many different tables are given in the dataset, out of these tables which table only consists of categorical variables?

ANS:

1.Activecustomers

2.Bank\_churn

3.CreditCard

4.CustomerInfo

5.ExitCustomer

6.Gender

7.Geography

Out of these tables Bank\_Churn having the categorical Variables.

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

WITH avg\_income AS

(SELECT gg.GeographyLocation,g.GenderCategory,ROUND(AVG(c.EstimatedSalary),2) AS average\_income

FROM customerinfo c

JOIN gender g USING (GenderID)

JOIN geography gg USING (GeographyID)

GROUP BY gg.GeographyLocation,g.GenderCategory

order by gg.GeographyLocation,g.GenderCategory)

SELECT \*,RANK() OVER(PARTITION BY GenderCategory ORDER BY average\_income DESC) AS rn

FROM avg\_income

Output:

A screenshot of a computer

Description automatically generated

1. Using SQL, write a query to find out the average tenure of the people who have exited in each age bracket (18-30, 30-50, 50+).

SELECT CASE WHEN c.age BETWEEN 18 and 30 THEN '18-30'

WHEN c.age BETWEEN 30 AND 50 THEN '30-50'

ELSE '50+'

END AS age\_brackets,

AVG(bc.Tenure) AS avg\_tenure

FROM customerinfo c

JOIN bank\_churn bc USING (CustomerID)

WHERE bc.Exited = 1

GROUP BY age\_brackets

Output:

A screenshot of a computer

Description automatically generated

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

Ans: Customers having Fair credit score having more estimated salary 330 M of total salary Good and poor credit score customers having 244 M and 237 M total salary. Who are having Top credit score those are very few.A graph of blue rectangular objects

Description automatically generated with medium confidence

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

A graph with blue rectangles

Description automatically generated

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

ANS:

1.Activecustomers

2.Bank\_churn

3.CreditCard

4.CustomerInfo

5.ExitCustomer

6.Gender

7.Geography

Out of these tables Bank\_Churn having the categorical Variables.

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

WITH avg\_income\_location\_gender AS

(SELECT ci.GeographyID,

g.GenderCategory,

ROUND(AVG(ci.EstimatedSalary),2) As avg\_income

FROM customerinfo ci

JOIN gender g USING(GenderID)

GROUP BY ci.GeographyID,

g.GenderCategory

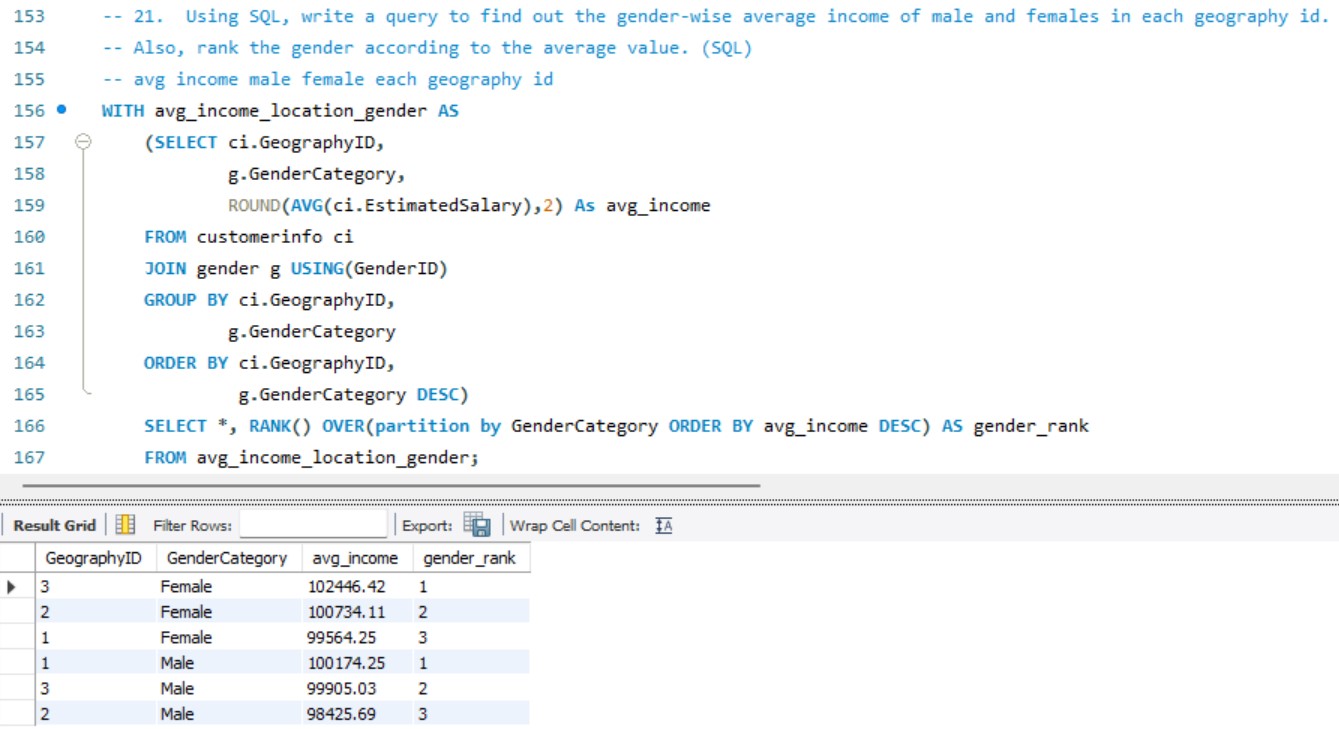
ORDER BY ci.GeographyID,

g.GenderCategory DESC)

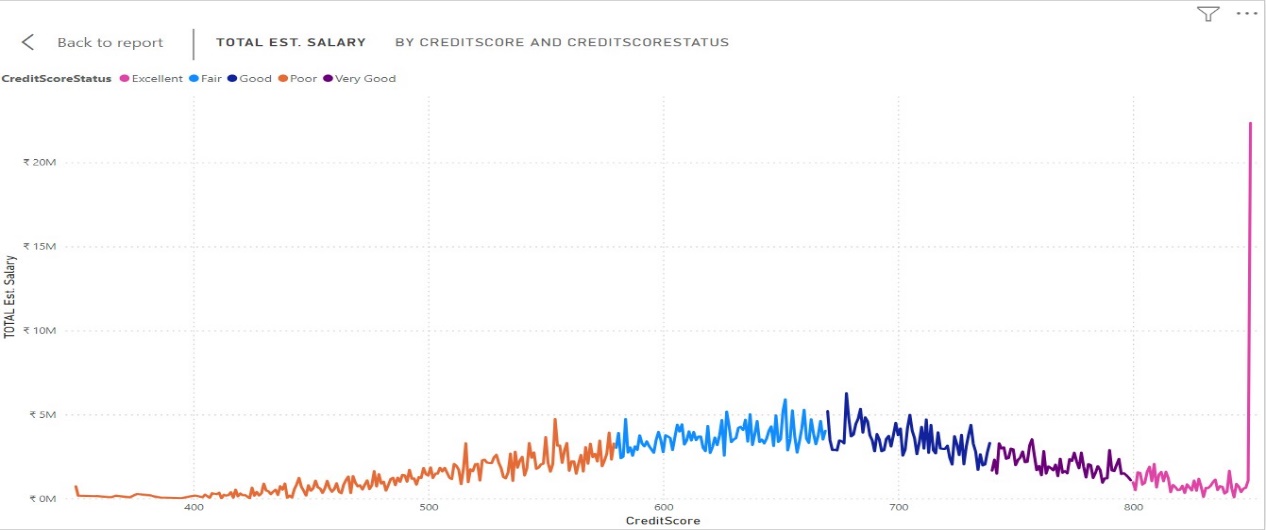
SELECT \*, RANK() OVER(partition by GenderCategory ORDER BY avg\_income DESC) AS gender\_rank

FROM avg\_income\_location\_gender;

Output:



1. Using SQL, write a query to find out the average tenure of the people who have exited in each age bracket (18-30, 30-50, 50+).
2. 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?
3. Is there any correlation between the salary and Credit score of customers?



1. 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”.

Ans:

SELECT DISTINCT b.CustomerID,c.Surname As last\_name, a.ActiveCategory

FROM customerinfo c

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

JOIN activecustomer a ON b.IsActiveMember = a.ActiveID

WHERE c.Surname LIKE "%on" AND a.ActiveCategory IS NOT NULL

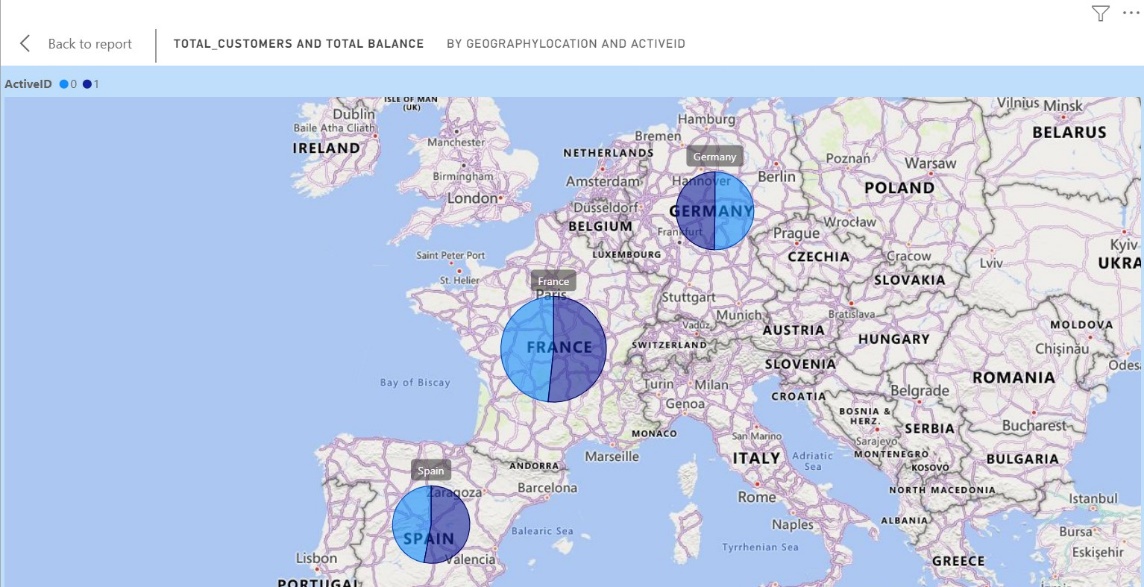
Output:

A screenshot of a computer

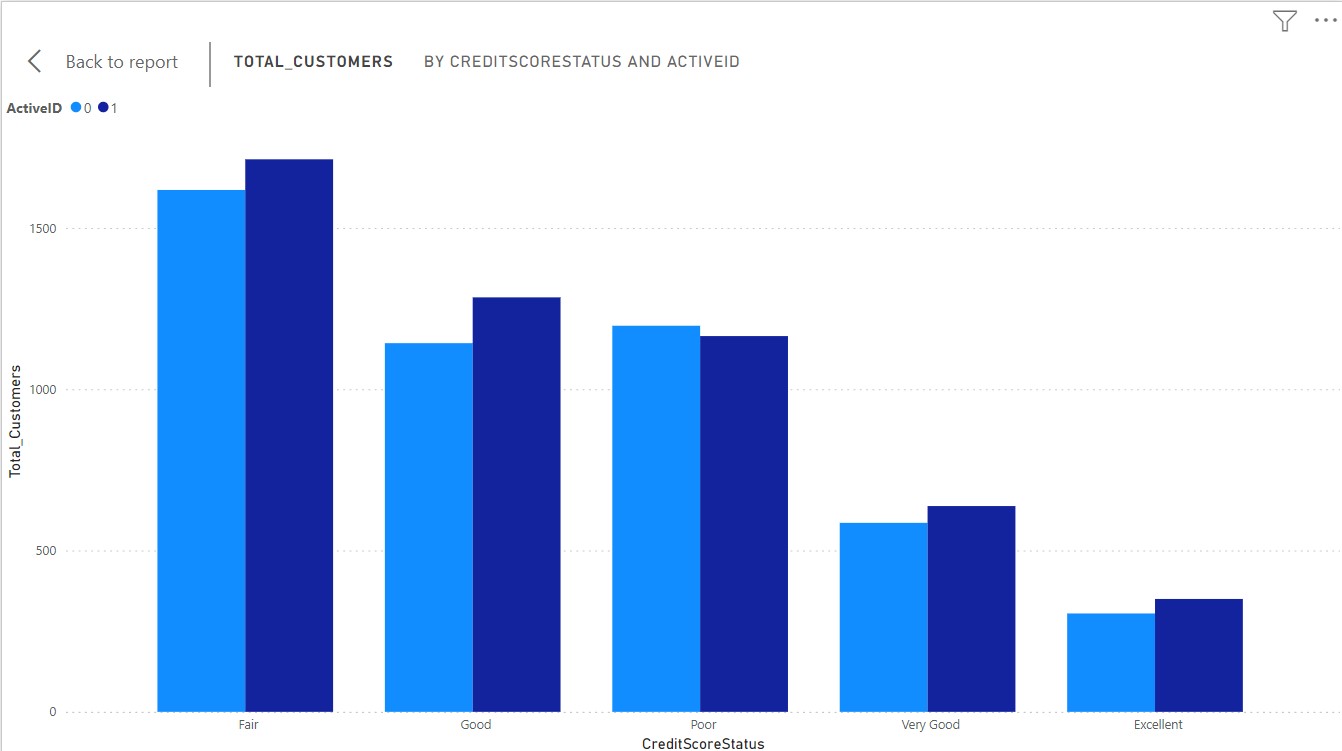
Description automatically generated

**Subjective Question:**

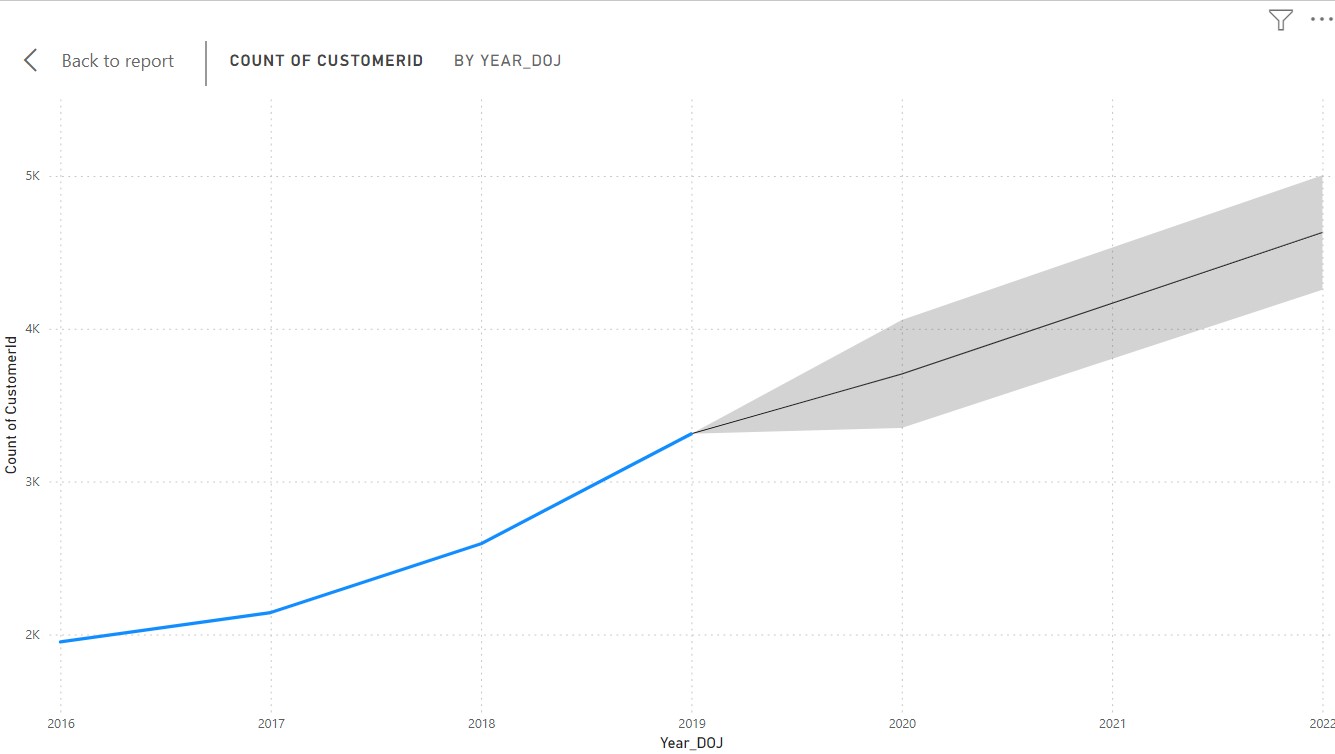
1. Customer Behavior Analysis: What patterns can be observed in the spending habits of long-term customers compared to new customers, and what might these patterns suggest about customer loyalty?
2. Product Affinity Study: Which bank products or services are most commonly used together, and how might this influence cross-selling strategies?
3. Geographic Market Trends: How do economic indicators in different geographic regions correlate with the number of active accounts and customer churn rates?



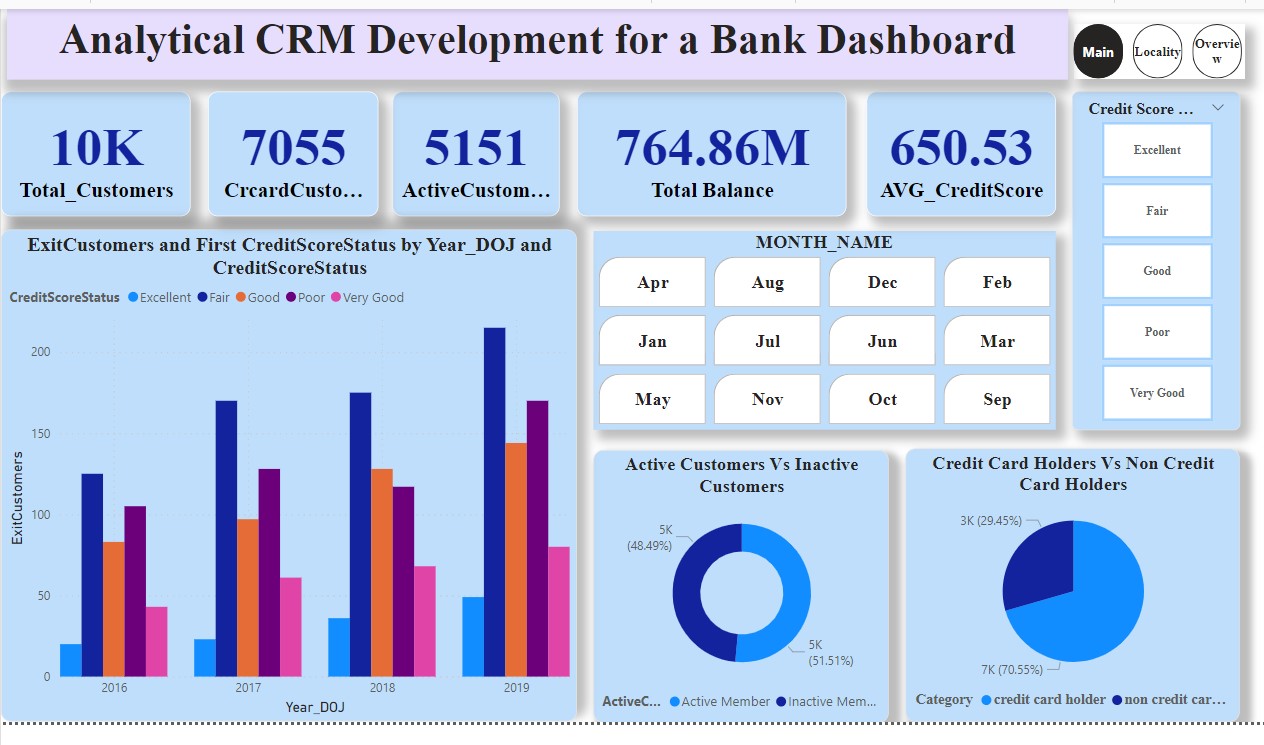
1. Risk Management Assessment: Based on customer profiles, which demographic segments appear to pose the highest financial risk to the bank, and why?



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



1. Marketing Campaign Effectiveness: How could you assess the impact of marketing campaigns on customer retention and acquisition within the dataset? What extra information would you need to solve this?
2. Customer Exit Reasons Exploration: Can you identify common characteristics or trends among customers who have exited that could explain their reasons for leaving?
3. Are 'Tenure', 'NumOfProducts', 'IsActiveMember', and 'EstimatedSalary' important for predicting if a customer will leave the bank?
4. Utilize SQL queries to segment customers based on demographics and account details.
5. 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?
6. 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?
7. Create a dashboard incorporating all the KPIs and visualization-related metrics. Use a slicer in order to assist in selection in the dashboard.



A screenshot of a computer

Description automatically generated

A close-up of a graph

Description automatically generated

1. How would you approach this problem, if the objective and subjective questions weren't given?
2. In the “Bank\_Churn” table how can you modify the name of the “HasCrCard” column to “Has\_creditcard”?