**Capstone Project Objective Questions**

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

select

distinct c.geographyid,

g.GeographyLocation,

round(sum(b.balance) over(partition by GeographyLocation),2) as balance

from customer\_id c

left join bank\_churn b on c.CustomerId = b.CustomerId

left join sql\_project\_geography g on c.GeographyID = g.GeographyID

order by GeographyID;

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

select

surname as Customer\_surName,

max(estimatedsalary) as Salary

from customer\_id

group by 1

order by Salary desc

limit 5;

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

SELECT DISTINCT ROUND(AVG(bc.NumOfProducts),2) AS AvgProducts

FROM customer\_id c

LEFT JOIN bank\_churn bc

ON c.CustomerId = bc.CustomerId

where bc.HasCrCard = 1;

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

Unable to change date format inside sql !

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

select exited, avg(creditscore) as avg\_creditscore from bank\_churn

group by exited;

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

select c.gender,

round(avg(c.estimatedsalary),2) avg\_estimated\_salary,

IsActiveMember

from customer\_id c

join bank\_churn bc on c.CustomerId=bc.CustomerId

group by c.Gender, IsActiveMember;

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

with cte1 as(

select

CustomerId,

creditscore,

exited,

case

when creditscore >= 800 and creditscore <= 850 then 'excellent'

when creditscore >= 740 and creditscore <= 799 then 'verygood'

when creditscore >= 670 and creditscore <= 739 then 'good'

when creditscore >= 580 and creditscore <= 669 then 'fair'

else 'poor'

end as credit\_type

from bank\_churn)

select CustomerId, creditscore, exited, credit\_type,

count(exited) over(partition by credit\_type) as max\_exit\_rate from cte1

order by max\_exit\_rate desc;

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

select distinct country,

count(c.customerid) as active\_customers

from customer\_id c

join bank\_churn bc on c.customerid=bc.CustomerId

where IsActiveMember = 1 and Tenure > 5

group by Country limit 1;

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

with churndata as (

select

HasCrCard,

count(\*) as total\_customers,

sum(case when exited = '1' then 1 else 0 end) as churned\_customers

from bank\_churn

group by HasCrCard

)

select

HasCrCard,

total\_customers,

churned\_customers,

churned\_customers/total\_customers as churn\_rate

from churndata;

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

select Numofproducts, count(NumOfProducts) as common\_num\_products from bank\_churn

where exited = 1

group by Numofproducts;

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

select

c.CustomerId,

c.EstimatedSalary,

c.Country,

c.Gender,

c.BankDOJ,

bc.tenure,

bc.HasCrCard

from customer\_id c

join bank\_churn bc

on c.CustomerId=bc.CustomerId

where exited = '1' and tenure in (4,5,6,7)

order by tenure desc;

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

select

NumOfProducts,

round(avg(Balance),2) as AvgBalance,

count(customerID) as CustomerCount

from bank\_churn

where Exited = 1

group by NumOfProducts

order by NumOfProducts;

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

select customerID, NumOfProducts as products\_bought

from bank\_churn

where exited = 0 and NumOfProducts>1

order by NumOfProducts;

1. **Can you create a dashboard incorporating the visuals mentioned above and additionally derive more KPIs if possible?**
2. **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)**

select country, gender,

round(avg(estimatedsalary),2) as avg\_salary,

rank() over (partition by country order by avg(estimatedsalary) desc) as gender\_rank

from customer\_id

group by gender, Country

order by Country, gender\_rank;

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 31 AND 50 THEN '31-50'

WHEN c.Age >= 51 THEN '50+'

END AS age\_bracket,

AVG(bc.Tenure) AS avg\_tenure

FROM

customer\_id c

JOIN

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

WHERE

bc.exited = 1

GROUP BY

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

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

WHEN c.Age >= 51 THEN '50+'

END

ORDER BY age\_bracket;

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

No corelation was found.

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

No corelation was found.

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

with credit\_score\_bucket 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 credit\_score\_bucket from bank\_churn

where exited =1)

select

credit\_score\_bucket, count(CustomerId) as total\_count,

dense\_rank() over (order by count(CustomerId) desc) as ranks

from credit\_score\_bucket

group by credit\_score\_bucket;

**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.**

with agebucket as (

select

case

when c.age between 18 and 30 then '18-30'

when c.age between 31 and 50 then '31-50'

when c.age >= 51 then '50+'

end as agebucket,

count(distinct c.customerid) as num\_customers,

avg(case when b.HasCrCard = 1 then 1 else 0 end) as avg\_credit\_cards

from customer\_id c

join bank\_churn b

on c.CustomerId = b.CustomerId

group by agebucket

)

select

agebucket,

num\_customers,

avg\_credit\_cards

from agebucket

where avg\_credit\_cards < (select avg(avg\_credit\_cards) from agebucket);

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

with locations as (

select

c.Country,

count(c.customerid) as customer\_count,

avg(b.balance) as avg\_balance

from customer\_id c

join bank\_churn b

on c.CustomerId = b.CustomerId

where Exited = 1

group by Country

)

select

country,

customer\_count,

round(avg\_balance,0) as average\_balance,

rank() over(order by customer\_count desc) as ranks

from locations;