1)Summary Statistics

The banking company wants to understand what is the **Average Credit Score** and **Balance** of the customers in the Dataset.

1.1 For Average Credit Score and Balance

```
SELECT
AVG (CreditScore) AS average_credit_score,
AVG (Balance) AS average_balance
FROM CUSTOMER;
```

Average CreditScore	650.5288
Average Balance	76485.89

1.2 Summary Statistics for numerical variables using SQL

```
SELECT
AVG(CreditScore) AS average_credit_score,
AVG(Age) AS average_age,
AVG(Tenure) AS average_tenure,
AVG(Balance) AS average_balance,
AVG(EstimatedSalary) AS average_estimated_salary,
PERCENTILE_CONT(0.5) WITHIN GROUP (ORDER BY CreditScore) AS Median_Credit_Score,
PERCENTILE_CONT(0.5) WITHIN GROUP (ORDER BY Age) AS Median_Age,
PERCENTILE_CONT(0.5) WITHIN GROUP (ORDER BY Tenure) AS Median_Tenure,
PERCENTILE_CONT(0.5) WITHIN GROUP (ORDER BY Balance) AS Median_Balance,
PERCENTILE_CONT(0.5) WITHIN GROUP (ORDER BY EstimatedSalary) AS
Median_EstimatedSalary,
MIN(CreditScore) AS Min_Credit_Score,
MIN(Age) AS Min_Age,
MIN(Tenure) AS Min_Tenure,
MIN(Balance) AS Min_Balance,
MIN(EstimatedSalary) AS Min_Estimated_Salary,
MAX(CreditScore) AS Max_Credit_Score,
MAX(Age) AS Max_Age,
MAX(Tenure) AS Max_Tenure,
MAX(Balance) AS Max_Balance,
MAX(EstimatedSalary) AS Max_Estimated_Salary,
STDDEV(CreditScore)AS Std_Credit_Score,
```

```
STDDEV(Age) AS Std_Age,
STDDEV(Tenure)AS Std_Tenure,
STDDEV(Balance)AS Std_Balance,
STDDEV(EstimatedSalary) AS Std_Estimated_Salary
FROM CUSTOMER;
```

The entire query provides a comprehensive summary of statistical metrics for different numerical columns in the "CUSTOMER" table, including mean, median, minimum, maximum, and standard deviation. The results are organized in a way that facilitates comparison across these metrics for each column.

Output:

Metric	Credit Score	Age	Tenure	Balance	Estimated Salary
Mean	650.53	38.92	5.01	76485.89	100090.24
Median	652	37	5	97198.54	100193.92
Min	350	18	0	0	11.58
Max	850	92	10	250898.09	199992.48
STDDEV	96.65	10.49	2.89	62397.41	57510.49

This table appears to represent descriptive statistics for several variables, likely from a dataset related to credit scoring or financial information. This table provides a comprehensive overview of the central tendency, variability, and range of each variable in the dataset. It offers insights into the distribution of credit scores, ages, tenures, balances, and estimated salaries among the individuals in the dataset.

• There is considerable variability in credit scores, ages, tenure, account balances, and estimated salaries within the dataset.

• Account balances show a right-skewed distribution, with some individuals having high balances.

• The distribution of age is slightly right-skewed, indicating a higher concentration of younger individuals.

• The distribution of tenure is relatively spread out, with some individuals having shorter tenures.

• Estimated salary has a relatively symmetric distribution with a moderate spread

2) Determine the metrics to understand Customer Churn

To understand and analyse customer churn in the dataset, the following metrics are crucial.

1. Credit Score: Credit score may impact a customer's financial stability and likelihood to churn. Lower credit scores might indicate financial struggles, affecting their decision to continue with a service.

2. Age: Age can be a significant factor in churn analysis. Younger customers may be more tech-savvy and open to switching services, while older customers might prefer stability. Analysing age groups can reveal patterns in customer behaviour.

3. Tenure: The length of time a customer has been with the company (tenure) is crucial. Longer tenure generally indicates loyalty. Analysing tenure in relation to churn can provide insights into whether newer customers are more prone to leaving.

4. Balance: Customer balances can reflect their financial engagement. Higher balances might indicate financial commitment, while low or negative balances may suggest financial strain, influencing the decision to churn.

5. Number of Products: The number of products a customer uses can indicate their level of engagement with the company. Customers with multiple products are usually more deeply integrated into the ecosystem, making them less likely to churn.

6. Is Active Member: Active membership status indicates customer engagement. Inactive members might be more prone to churn. Analysing the behaviour of active vs. inactive members provides insights into the correlation and churn.

7. Estimated Salary: Customer salary can influence their spending behaviour and financial decisions. Analysing salary in relation to churn can reveal patterns related to customers with higher or lower incomes.

8. **Geography**: The geographical location of customers may impact their behaviour. Cultural, economic, or regional factors can contribute to churn patterns. Analysing churn across different geographies can uncover location-specific trends.

9. Gender: Gender-based analysis can provide insights into whether there are differences in churn behaviour between male and female customers.

10. **Has Credit Card**: Whether a customer has a credit card with the company may influence their commitment. Customers with credit cards may have a higher level of financial integration, affecting their likelihood to churn.

These metrics offer a comprehensive view of customer behaviour and engagement, helping to identify patterns and factors influencing churn in the dataset.

3)The top metrics that will help us best understand the churn pattern.

1) **Credit Score:** Credit score may impact a customer's financial stability and likelihood to churn. Lower credit scores might indicate financial struggles, affecting their decision to continue with a service.

2) **Age:** Age can be a significant factor in churn analysis. Younger customers may be more tech-savvy and open to switching services, while older customers might prefer stability. Analysing age groups can reveal patterns in customer behaviour.

3) **Tenure:** The length of time a customer has been with the company (tenure) is crucial. Longer tenure generally indicates loyalty. Analysing tenure in relation to churn can provide insights into whether newer customers are more prone to leaving.

4) **Balance:** Customer balances can reflect their financial engagement. Higher balances might indicate financial commitment, while low or negative balances may suggest financial strain, influencing the decision to churn.

5) **Estimated Salary:** Customer salary can influence their spending behaviour and financial decisions. Analysing salary in relation to churn can reveal patterns related to customers with higher or lower incomes.