

Tittle: Study Customer Churn Analytics for The Banking Industry

Name: Pawan Kanade

Date: 01/02/2024

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 CreditScore and Balance

```
SELECT
  ROUND(AVG(CreditScore),2) AS average_credit_score,
  ROUND(AVG(Balance), 2) AS average_balance
FROM customer_churn;
```

Average CreditScore	650.5288
Average Balance	76485.89

1.2) Summary Statistics for numerical variables using SQL

```
SELECT
'Mean' AS Metric,
ROUND(AVG(CreditScore), 2) AS "Credit Score",
ROUND(AVG(Age), 2) AS "Age",
ROUND(AVG(Tenure), 2) AS "Tenure",
ROUND(AVG(Balance), 2) AS "Balance",
ROUND(AVG(EstimatedSalary), 2) AS "EstimatedSalary"
FROM customer_churn
UNION ALL
SELECT
'Median' AS Metric,
ROUND(PERCENTILE_CONT(0.5) WITHIN GROUP (ORDER BY CreditScore)::numeric, 2) AS
"Credit Score",
ROUND(PERCENTILE_CONT(0.5) WITHIN GROUP (ORDER BY Age)::numeric, 2) AS "Age",
ROUND(PERCENTILE_CONT(0.5) WITHIN GROUP (ORDER BY Tenure)::numeric, 2) AS
"Tenure",
ROUND(PERCENTILE_CONT(0.5) WITHIN GROUP (ORDER BY Balance)::numeric, 2) AS
"Balance",
ROUND(PERCENTILE_CONT(0.5) WITHIN GROUP (ORDER BY EstimatedSalary)::numeric, 2) AS
"EstimatedSalary"
FROM customer_churn
UNION ALL
SELECT
'Min' AS Metric,
MIN(CreditScore) AS "Credit Score",
MIN(Age) AS "Age",
MIN(Tenure) AS "Tenure",
MIN(Balance) AS "Balance",
MIN(EstimatedSalary) AS "EstimatedSalary"
FROM customer_churn
UNION ALL
SELECT
```

```

'Max' AS Metric,
MAX(CreditScore) AS "Credit Score",
MAX(Age) AS "Age",
MAX(Tenure) AS "Tenure",
MAX(Balance) AS "Balance",
MAX(EstimatedSalary) AS "EstimatedSalary"
FROM customer_churn
UNION ALL
SELECT
'STDDEV' AS Metric,
ROUND(STDDEV(CreditScore), 2) AS "Credit Score",
ROUND(STDDEV(Age),2) AS "Age",
ROUND(STDDEV(Tenure),2) AS "Tenure",
ROUND(STDDEV(Balance), 2) AS "Balance",
ROUND(STDDEV(EstimatedSalary), 2) AS "EstimatedSalary"
FROM customer_churn;

```

The entire query provides a comprehensive summary of statistical metrics for different numerical columns in the "customer_churn" 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	EstimatedSalary
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. **IsActiveMember:** 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. Analyzing age groups can reveal patterns in customer behavior.
- 3) **Tenure:** The length of time a customer has been with the company (tenure) is crucial. Longer tenure generally indicates loyalty. Analyzing 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 behavior and financial decisions. Analyzing salary in relation to churn can reveal patterns related to customers with higher or lower incomes.

4) Write the Final Query to identify the Churned/Non-churned Customers

4.1) Query for total customer is churned and not churned

```
select
SUM(CASE WHEN exited = 1 THEN 1 ELSE 0 END) AS Churned,
SUM(CASE WHEN exited = 0 THEN 1 ELSE 0 END) AS Non_Churned
FROM customer_churn;
```

The entire query calculates and returns two counts: the number of churned customers (Churned) and the number of non-churned customers (Non_Churned) from the dataset. This type of query is common in churn analysis to understand the distribution of customers based on their exit status. The results can be useful for visualizations or further analysis to evaluate the proportion of customers who have churned versus those who haven't.

Churned	Non_churned
2037	7963

4.2) Using Window functions, explore the top 5 customers based on certain characteristics and highlight any trends/patterns you see.

```
WITH RankedCustomers AS (
SELECT CustomerId,CreditScore,Balance,Age,Tenure,
NumOfProducts,EstimatedSalary,Exited,
ROW_NUMBER() OVER (PARTITION BY Exited ORDER BY CreditScore DESC) AS
RankByCreditScore,
ROW_NUMBER() OVER (PARTITION BY Exited ORDER BY NumOfProducts DESC) AS
RankByNumOfPrducts,
ROW_NUMBER() OVER (PARTITION BY Exited ORDER BY EstimatedSalary DESC) AS
RankByEstimatedSalary
FROM customer_churn)
SELECT
CustomerId,CreditScore,Balance,Age,Tenure,NumOfProducts,EstimatedSalary,Exited,
CASE WHEN Exited = 1 then 'Churned' else 'NON_Churned' end as Churn_status
FROM RankedCustomers
WHERE RankByCreditScore <= 5 OR RankByNumOfPrducts <=5 OR RankByEstimatedSalary
<=5
ORDER BY Exited DESC, RankByCreditScore, RankByNumOfPrducts,RankByEstimatedSalary;
```

Query Explanation:

This query aims to rank customers based on different criteria (credit score, number of products, estimated salary) within their exit status (churned or non-churned). It then selects the top 5 customers for each criterion and presents information about these customers

along with their churn status. The ordering in the final result allows for a quick comparison of churned and non-churned customers based on the specified criteria.

customerid	creditscore	balance	age	tenure	numofproducts	estimatedsalary	exited	churnstatus
15739578	850	128663.9	49	6	1	65769.3	1	Churned
15750335	850	108508.82	43	0	3	184044.8	1	Churned
15676571	850	0	55	6	1	944.41	1	Churned
15681007	850	128548.49	35	2	4	75478.95	1	Churned
15678497	850	0	48	2	1	169425.3	1	Churned
15672152	850	122506.38	37	9	1	199693.84	1	Churned
15603851	704	127785.17	32	7	4	184464.7	1	Churned
15661903	699	80764.03	43	3	1	199378.58	1	Churned
15664270	692	142084.04	45	6	4	188305.85	1	Churned
15689341	655	0	50	10	4	179267.94	1	Churned
15724161	644	137285.26	40	9	4	77063.63	1	Churned
15755262	608	89763.84	41	3	1	199304.74	1	Churned
15815656	541	100116.67	39	9	1	199808.1	1	Churned
15661670	524	107818.63	31	8	1	199725.39	1	Churned
15812888	447	0	41	3	4	197490.39	1	Churned
15762793	850	136980.23	36	0	2	99019.65	0	Non-churned
15588219	850	106871.81	38	1	2	29333.01	0	Non-churned
15719958	850	124548.99	39	3	2	120380.12	0	Non-churned
15760427	850	124788.18	40	6	1	65612.12	0	Non-churned
15643438	850	0	20	7	2	31288.77	0	Non-churned
15664808	790	0	37	3	3	98897.32	0	Non-churned
15732299	756	0	67	4	3	93081.87	0	Non-churned
15798200	707	0	35	2	3	94148.3	0	Non-churned
15662021	685	0	42	2	2	199992.48	0	Non-churned
15709643	675	0	32	1	3	85901.09	0	Non-churned
15634359	639	98635.77	41	5	1	199970.74	0	Non-churned
15709136	620	0	28	8	2	199909.32	0	Non-churned
15697270	608	153325.1	27	4	1	199953.33	0	Non-churned
15649183	598	0	35	8	3	88658.73	0	Non-churned
15762331	569	178755.84	37	9	1	199929.17	0	Non-churned

4.3) If you were to choose the top 5 customers based on any criteria of your choice to offer a reduced interest rate, how would you choose them? What criteria would you use? What do you expect the outcome to be?

1) Top 5 customer based on credit score

```
WITH TopCustomers AS (  
SELECT  
CustomerId,Surname,CreditScore,Exited,  
ROW_NUMBER() OVER (ORDER BY CreditScore DESC) AS RankByCreditScore  
FROM customer_churn)  
SELECT CustomerId,Surname,CreditScore,  
CASE WHEN Exited = 1 then 'Churned' else 'NON_Churned' end as Churn_status  
FROM TopCustomers  
WHERE RankByCreditScore <= 5;
```

The query identifies the top 5 customers based on their credit scores from the "customer_churn" table. Customers are ranked based on their credit scores, and only those with a rank of 5 or less are included in the final result.

2) Output

customerid	surname	creditscore	churn_status
15785798	Uchechukwu	850	NON_Churned
15762793	Jones	850	NON_Churned
15672152	Grant	850	Churned
15612776	McKay	850	NON_Churned
15607275	Ch'ang	850	Churned

5) Identify Top Three Focus Areas

5.1) Credit Score: (Customer Satisfaction and Loyalty)

Identify and implement strategies to enhance satisfaction for customers with lower credit scores. Provide targeted promotions or services to improve the creditworthiness of these customers. Establish personalized credit management plans for customers with varying credit scores.

5.2) Tenure: (Enhanced Customer Retention)

Develop strategies to reward and retain customers with longer tenures. Introduce loyalty programs or exclusive benefits for customers who have been with the institution for an extended period. Implement targeted communication to express appreciation for customers' long-term commitment.

5.3) NumOfProducts: (Promotion of Multiple Products)

Encourage customers to explore and utilize additional financial products. Design promotions or bundled offerings to incentivize the adoption of multiple products. Provide personalized recommendations based on the customer's financial needs and preferences to increase the number of products they use.

6) Conclusion

These focus areas specifically target aspects related to CreditScore, Tenure, and NumOfProducts to enhance customer satisfaction, increase loyalty, and promote the utilization of multiple financial products. The success of these strategies can be measured through metrics such as customer retention rates, increased product adoption, and overall customer satisfaction scores.