**Project Title: Bank Customer Churn Prediction Using SQL**

**Introduction**

The objective of this project is to analyze customer data to predict churn rates and identify key factors influencing customer retention in the banking sector. By leveraging SQL for data manipulation and analysis, the project aims to derive actionable insights that can inform marketing strategies and improve customer retention.

**Business Scenario**

The banking industry faces significant challenges in retaining customers amidst increasing competition. Understanding the characteristics of customers who churn versus those who remain is critical for developing targeted retention strategies. This document outlines the key questions, KPIs, findings, and proposed marketing strategies based on the analysis of customer data.

**Key Questions**

1. **Customer Segmentation**: What are the characteristics of customers who churn versus those who remain?
2. **Churn Prediction**: Can we predict which customers are likely to churn based on their attributes?
3. **Product Usage**: How does the number of products a customer has relate to their likelihood of churning?
4. **Demographic Analysis**: What demographic factors (age, gender, country) are associated with higher churn rates?
5. **Balance Impact**: Does a higher account balance correlate with lower churn rates?
6. **Credit Score Influence**: How does a customer's credit score affect their likelihood of remaining an active member?
7. **Tenure Analysis**: Is there a relationship between the length of time a customer has been with the bank and their likelihood to churn?
8. **Salary Insights**: How does estimated salary influence customer retention and product usage?
9. **Active Membership Trends**: What percentage of customers are active members, and how does this relate to churn?
10. **Country Comparison**: Which country has the highest churn rate and what factors contribute to this?
11. **Gender Differences**: Are there significant differences in churn rates between male and female customers?
12. **Product Effectiveness**: Which products are associated with higher retention rates among customers?
13. **Churn Rate Over Time**: How has the churn rate changed over different periods (e.g., by tenure)?
14. **Customer Value Assessment**: What is the average balance of customers who have churned compared to those who have not?
15. **Marketing Strategies**: What targeted marketing strategies could be developed based on customer characteristics to reduce churn?

**Key Performance Indicators (KPIs)**

* **Total Unique Customers**: 10,000
* **Average Balance**: $76,485.88
* **Average Estimated Salary**: $100,000
* **Credit Card Holder Percentage**: 70.55%
* **Average Churn Rate:** 20.37%

**Characteristics of Churned vs. Retained Customers**

* **Age and Balance**: Churned customers have an average age of 44 years and a higher average balance of $91,108, compared to retained customers, who average 37 years in age and have a balance of $73,745. This suggests that older customers with larger balances may feel dissatisfied or have unmet needs, indicating a potential area for improvement in customer service and engagement strategies.

**Churn Prediction Insights**

* **Product Usage**: Customers holding two products show a significantly lower churn rate (0.0758), while those with four products experience a high churn rate (1.00). This indicates that offering a variety of products can enhance customer retention.
* **Demographics**: Female customers exhibit a higher churn rate (0.25) compared to males (0.16), with the average age for females being slightly higher (40.15). Targeting retention strategies towards middle-aged female customers may be beneficial.

**Product Usage and Effectiveness**

* The data suggests that having multiple products correlates with lower churn rates, highlighting the importance for banks to promote bundled services to foster customer loyalty.

**Demographic Analysis**

* **Age**: Customers aged 40-50 are the most likely to churn. Retention efforts should focus on this demographic.
* **Country Comparison**: In Germany, female customers have a notable churn rate of 0.37% despite having an average balance of $119,145 and salary of $102,446, suggesting that financial stability does not guarantee loyalty.

**Balance Impact**

* Interestingly, customers with low balances exhibit lower churn rates (0.19) compared to those with high balances (0.23). This counterintuitive finding may require further investigation into customer satisfaction levels and service quality.

**Credit Score Influence**

* Customers with credit scores below 750 have a 51.28% likelihood of remaining active members, while those above 750 have a slightly higher retention rate at 52%. This underscores the significance of financial health in retention strategies.

**Tenure Analysis**

* Churn rates are highest among new customers (0 months) at 0.2300 but decrease to around 0.1722 by the seventh year, indicating that customer loyalty tends to increase with tenure. Early engagement strategies are crucial.

**Salary Insights**

* High-income individuals show a retention rate of 20%, slightly above medium (19.87%) and low-income (19.93%) groups, suggesting that while financial capability influences loyalty, it is not the sole factor.

**Active Membership Trends**

* Active members have a lower churn rate (0.14%) compared to non-active members (0.26%), reinforcing the need for banks to encourage active engagement through personalized services.

Recommendations

1. **Targeted Marketing Strategies**: Focus on middle-aged females in France and Spain and young males in Germany for tailored marketing campaigns.
2. **Product Bundling Promotions**: Encourage customers to hold multiple products by offering incentives or discounts for bundled services.
3. **Enhanced Customer Engagement**: Develop personalized engagement strategies for customers aged 40-50 to address their specific needs and concerns.
4. **Leverage Data Analytics**: Utilize demographic, behavioral, and psychographic data to create more nuanced customer segments for targeted marketing efforts.
5. **Early Customer Relationship Building**: Implement proactive engagement strategies for new customers to foster loyalty from the outset.
6. **Monitor Satisfaction Levels**: Conduct regular surveys and feedback sessions to understand customer satisfaction across different segments, particularly focusing on older customers with higher balances.

**Conclusion**

This document serves as a comprehensive overview for analyzing customer behavior related to churn within the banking sector using SQL-based insights. The findings will guide strategic decisions aimed at improving customer retention and minimizing attrition through targeted marketing efforts and enhanced customer engagement initiatives.

Next Steps

1. Validate findings through further data analysis.
2. Develop predictive models using SQL queries for real-time insights.
3. Implement marketing strategies based on identified customer segments.

This structured approach ensures that all aspects of customer behavior are considered in developing effective retention strategies while leveraging data-driven insights for continuous improvement in service delivery within the banking industry.

**SQL QUERIES USED TO PROVIDE INSIGHTS ON KEY QUESTIONS:**

**Database Creation**

**Create Database:**

CREATE DATABASE BANK\_DB;

USE BANK\_DB;

**Create Table:**

CREATE TABLE CUST\_DT (

CUSTOMER\_ID INT,

CREDIT\_SCORE INT,

COUNTRY VARCHAR(50),

GENDER VARCHAR(7),

AGE INT,

TENURE INT,

BALANCE FLOAT,

PRODUCTS\_NUMBER INT,

CREDIT\_CARD INT,

ACTIVE\_MEMBER INT,

ESTIMATED\_SALARY DECIMAL(8 , 6 ),

CHURN INT

);

**Check for NULL Values in Columns:**

SELECT

CUSTOMER\_ID IS NULL,

CREDIT\_SCORE IS NULL,

COUNTRY IS NULL,

GENDER IS NULL,

AGE IS NULL,

TENURE IS NULL,

BALANCE IS NULL,

PRODUCTS\_NUMBER IS NULL,

CREDIT\_CARD IS NULL,

ACTIVE\_MEMBER IS NULL,

ESTIMATED\_SALARY IS NULL,

CHURN IS NULL

FROM

CUST\_DT;

**Queries**

**Beginner Level:**

**Q1: Total Number of Customers in the Dataset**

SELECT COUNT(\*) AS TOTAL\_CUSTOMERS

FROM CUST\_DT;

Output:



**Q2: Customers with Credit Score Above 700**

SELECT COUNT(\*) AS HIGH\_CREDIT\_SCORE\_CUSTOMER

FROM CUST\_DT

WHERE CREDIT\_SCORE >= 700;

Output:



**Q3: Average Age of Customers**

SELECT AVG(AGE) AS AVERAGE\_AGE

FROM CUST\_DT;

Output:



**Intermediate Level**

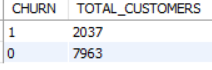
**Q4: Customers Who Have Churned vs. Those Who Have Not**

SELECT CHURN, COUNT(\*) AS TOTAL\_CUSTOMERS

FROM CUST\_DT

GROUP BY CHURN;

Output:



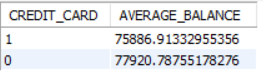
**Q5: Average Balance of Customers with Credit Cards**

SELECT CREDIT\_CARD, AVG(BALANCE) AS AVERAGE\_BALANCE

FROM CUST\_DT

GROUP BY CREDIT\_CARD;

Output:



**Q6: List of Unique Countries**

SELECT DISTINCT(COUNTRY) FROM CUST\_DT;

Output:



**Advanced Level**

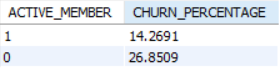
**Q7: Percentage of Active Members Who Have Churned**

SELECT ACTIVE\_MEMBER, AVG(CHURN) \* 100 AS CHURN\_PERCENTAGE

FROM CUST\_DT

GROUP BY ACTIVE\_MEMBER;

Output:



**Q8: Average Estimated Salary for Customers Who Have Churned**

SELECT AVG(ESTIMATED\_SALARY) AS AVG\_ESTIMATED\_SALARY

FROM CUST\_DT

WHERE CHURN = 1;

Output:

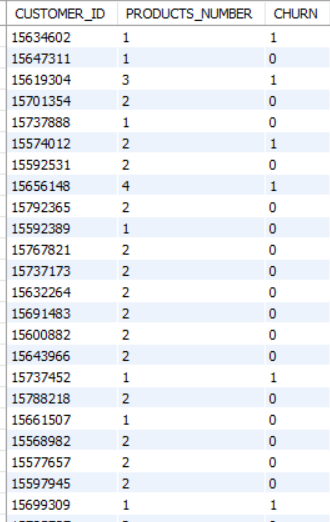


**Q9: Report on Number of Products Held by Each Customer and Their Churn Status**

SELECT CUSTOMER\_ID, PRODUCTS\_NUMBER, CHURN

FROM CUST\_DT;

Output:



**Expert Level**

**Q10: Predict Churn Based on Customer Demographics Using a CASE Statement**

SELECT

CUSTOMER\_ID,

CASE

WHEN AGE < 30 THEN 'YOUNG'

WHEN AGE BETWEEN 30 AND 50 THEN 'MIDDLE-AGED'

ELSE 'SENIOR'

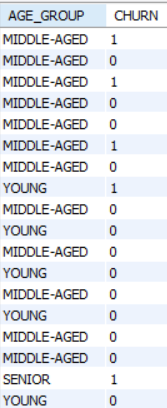
END AS AGE\_GROUP,

CHURN

FROM

CUST\_DT;

Output:



**Q11: Trends in Customer Balances Over Time for Churned Customers**

SELECT TENURE, AVG(BALANCE) AS AVERAGE\_BALANCED\_CHURNED

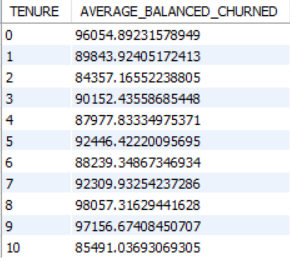
FROM CUST\_DT

WHERE CHURN = 1

GROUP BY TENURE

ORDER BY TENURE;

Output:



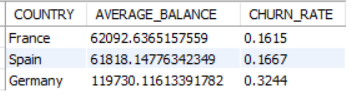
**Q12: Summary Report Grouping Customers by Country (Average Balance and Churn Rate)**

SELECT COUNTRY, AVG(BALANCE) AS AVERAGE\_BALANCE, AVG(CHURN) AS CHURN\_RATE

FROM CUST\_DT

GROUP BY COUNTRY;

Output:



**Q13: Top 10% of Customers Based on Estimated Salary and Their Churn Status**

WITH SALARYRANKED AS (

SELECT \*, NTILE(10) OVER (ORDER BY ESTIMATED\_SALARY DESC) AS SALARY\_DECILE

FROM CUST\_DT

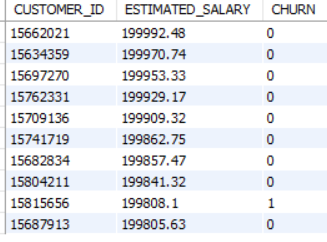
)

SELECT CUSTOMER\_ID, ESTIMATED\_SALARY, CHURN

FROM SALARYRANKED

WHERE SALARY\_DECILE = 1;

Output:



**Q14: Create a View with Tenure Categorized Into Ranges**

CREATE VIEW CUSTOMER\_TENURE\_VIEW AS

SELECT CUSTOMER\_ID,

CASE

WHEN TENURE = 0 THEN 'NEW'

WHEN TENURE BETWEEN 1 AND 3 THEN '1-3 YEARS'

WHEN TENURE BETWEEN 4 AND 6 THEN '4-6 YEARS'

ELSE '7+ YEARS'

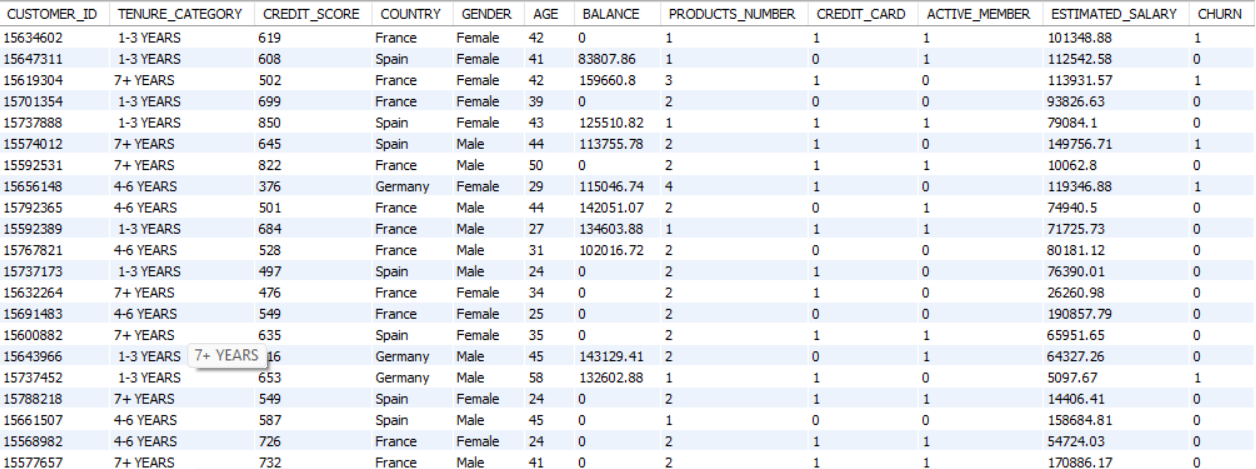
END AS TENURE\_CATEGORY,

CREDIT\_SCORE, COUNTRY, GENDER, AGE, BALANCE, PRODUCTS\_NUMBER, CREDIT\_CARD, ACTIVE\_MEMBER, ESTIMATED\_SALARY, CHURN

FROM CUST\_DT;

SELECT \* FROM CUSTOMER\_TENURE\_VIEW;

Output:



**Business Scenario Questions**

**Customer Segmentation**

**Characteristics of Customers Who Churn vs. Those Who Remain**

SELECT

AVG(AGE) AS AVG\_AGE,

AVG(BALANCE) AS AVG\_BALANCE,

COUNT(\*) AS TOTAL\_CUSTOMERS,

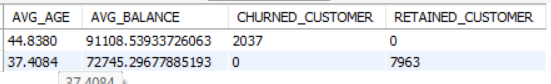
SUM(CASE WHEN CHURN = 1 THEN 1 ELSE 0 END) AS CHURNED\_CUSTOMER,

SUM(CASE WHEN CHURN = 0 THEN 1 ELSE 0 END) AS RETAINED\_CUSTOMER

FROM CUST\_DT

GROUP BY CHURN;

Output:



**Churn Prediction**

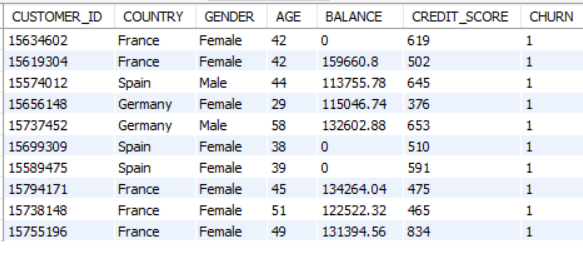
**Predict Which Customers Are Likely to Churn**

SELECT CUSTOMER\_ID, COUNTRY, GENDER, AGE, BALANCE, CREDIT\_SCORE, CHURN

FROM CUST\_DT

WHERE CHURN = 1;

Output:



**Product Usage**

**Relation Between Number of Products and Likelihood of Churning**

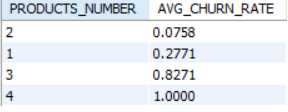
SELECT

PRODUCTS\_NUMBER, AVG(CHURN) AS AVG\_CHURN\_RATE

FROM CUST\_DT

GROUP BY PRODUCTS\_NUMBER;

Output:



**Demographic Analysis**

**Demographic Factors Associated With Higher Churn Rates**

SELECT

COUNTRY,

GENDER,

AVG(AGE) AS AVG\_AGE,

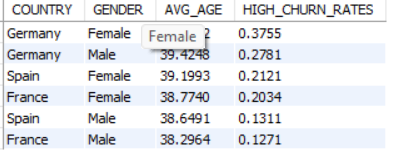
AVG(CHURN) AS HIGH\_CHURN\_RATES

FROM CUST\_DT

GROUP BY COUNTRY, GENDER

ORDER BY HIGH\_CHURN\_RATES DESC;

Output:



**Balance Impact**

**Correlation Between Higher Account Balance and Lower Churn Rates**

SELECT

CASE

WHEN BALANCE >= 125000 THEN 'HIGH BALANCE'

ELSE 'LOW BALANCE'

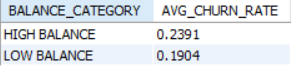
END AS BALANCE\_CATEGORY,

AVG(CHURN) AS AVG\_CHURN\_RATE

FROM CUST\_DT

GROUP BY BALANCE\_CATEGORY;

Output:



**Tenure Analysis**

**Relationship Between Tenure and Likelihood to Churn**

SELECT

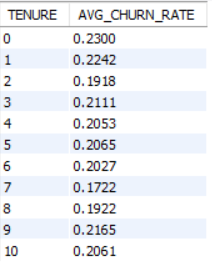
TENURE, AVG(CHURN) AS AVG\_CHURN\_RATE

FROM CUST\_DT

GROUP BY TENURE

ORDER BY TENURE;

Output:



**Marketing Strategies**

**Targeted Marketing Strategies to Reduce Churn**

SELECT

COUNTRY, GENDER, AGE\_GROUP, COUNT(\*) AS POTENTIAL\_TARGETS

FROM

(SELECT

GENDER, COUNTRY,

CASE

WHEN AGE BETWEEN 18 AND 30 THEN 'YOUNG'

WHEN AGE BETWEEN 30 AND 55 THEN 'MIDDLE\_AGE'

ELSE 'SENIOR'

END AS AGE\_GROUP

FROM CUST\_DT

WHERE CHURN = 1) AS SUBQUERY

GROUP BY GENDER, COUNTRY, AGE\_GROUP;

Output:

