

BANK CUSTOMER CHURN DATA ANALYSIS

DHANUSHWR K
JUNIOR DATA ANALYST

1. Introduction

Customer churn is a critical issue for banks, as it directly impacts revenue and profitability. By analysing customer data, banks can gain valuable insights into the factors that drive customer attrition and develop strategies to retain their customers.

This report aims to explore bank customer churn by examining a dataset of 10,000 customers from a European bank. The analysis will focus on understanding the characteristics of customers who have churned, identifying the key drivers of churn, and providing recommendations for improving customer retention.

2. Objectives

The primary objectives of this report are:

- **Characterize churned customers:** Identify the demographic, financial, and product-related attributes of customers who have churned.
- **Analyse churn drivers:** Determine the key factors that contribute to customer churn, such as credit score, balance, product usage, and customer service interactions.
- **Develop retention strategies:** Propose recommendations for improving customer retention based on the findings of the analysis.

3. Prepare

This analysis utilizes bank customers churn data. This dataset, publicly available from Kaggle, under [license](#), provides insights into customer usage patterns.

Dataset Field Descriptions

- **CustomerID:** A unique identifier for each customer.
- **Surname:** The customer's last name.
- **CreditScore:** A numerical value representing the customer's creditworthiness.

- **Geography:** The country where the customer resides (France, Spain, or Germany).
- **Gender:** The customer's gender (Male or Female).
- **Age:** The customer's age.
- **Tenure:** The number of years the customer has been with the bank.
- **Balance:** The customer's current account balance.
- **NumOfProducts:** The number of products the customer has with the bank.
- **HasCrCard:** Whether the customer has a credit card (1) or not (0).
- **IsActiveMember:** Whether the customer is an active member of the bank (1) or not (0).
- **EstimatedSalary:** The estimated annual salary of the customer.
- **Exited:** Whether the customer has churned (1) or not (0).

4. Process

- The data was **cleaned using Microsoft excel**.
- The data **didn't contain any duplicate entries nor any blanks**.
- The IsActiveMember, HasCrCard and Exited **columns were modified**. The 0 and 1 entries were changed to No and Yes respectively using **IF condition** formula
- All the **data types were formatted accordingly**.
 - CustomerID: Number
 - Surname: Text
 - CreditScore: Number
 - Geography: Text
 - Gender: Text
 - Age: Number
 - Tenure: Number

- Balance: Number
- NumOfProducts: Number
- HasCrCard: Text
- IsActiveMember: Text
- EstimatedSalary: Number
- Exited: Text

Key variables to focus on:

- **Exited:** This is the target variable, indicating whether a customer has churned.
- **CreditScore:** A higher credit score might indicate a more financially stable customer.
- **Geography:** Different regions might have varying customer behaviors and preferences.
- **Age:** Age could be correlated with customer loyalty and spending patterns.
- **Tenure:** Longer tenure might suggest a stronger customer relationship.
- **Balance:** Higher balances might indicate a more valuable customer.
- **NumOfProducts:** Customers with multiple products might be less likely to churn.
- **HasCrCard:** Credit card usage could be a factor in customer retention.
- **IsActiveMember:** Active members might be more engaged with the bank.
- **EstimatedSalary:** Higher income could indicate a more valuable customer.

Final prepared dataset imported to MySQL:

P1 bank_churn.bank_churn_prepared	
Table Details	
Engine:	InnoDB
Row Format:	Dynamic
Column count:	15
Table rows:	9829
AVG row length:	268
Data length:	2.5 MiB
Index length:	0.0 bytes
Max data length:	0.0 bytes
Data free:	4.0 MiB
Table size (estimate):	2.5 MiB
Update time:	2024-10-14 12:37:35
Create time:	2024-10-14 12:36:38
Auto increment:	
Table collation:	utf8mb4_0900_ai_ci
Information on this page may be outdated. Click Analyze Table to update it.	

5. Analyse

5.1 Data separation:

The prepared data was separated based on Exited status of the customers:

QUERY:

```
CREATE TABLE bank_customers_exited Select * From  
bank_churn.bank_churn_prepared Where Exited = "Yes";
```

```
CREATE TABLE bank_customers_not_exited Select * From  
bank_churn.bank_churn_prepared Where Exited = "No";
```

RESULTS:

- **bank_customers_exited table with 2037 entries were created.**
- **bank_customers_exited table with 7963 entries were created.**

5.2 Age group analysis:

The number of customers churn across age group was analyzed.

QUERY:

Select

```
CASE When Age Between 18 and 20 then "18 - 20"  
      when Age between 21 and 30 then "21 - 30"  
      When Age between 31 and 40 then "31 - 40"  
      When Age between 41 and 50 then "41 - 50"  
      When Age between 51 and 60 then "51 - 60"  
      when Age between 61 and 70 then "61 - 70"  
      when Age between 71 and 80 then "71 - 80"
```

```

        when Age between 81 and 92 then "81 - 92"

        Else 'other'

    END AS Age_Group,

    COUNT(*) AS Count

From bank_churn.bank_customers_exited

GROUP BY Age_Group

ORDER BY Age_Group;

```

RESULTS:

Age_Group	Count
18 - 20	5
21 - 30	143
31 - 40	538
41 - 50	788
51 - 60	448
61 - 70	104
71 - 80	10
81 - 92	1

5.3 Gender analysis:

The number of male and female customers churn was analyzed.

QUERY:

```

SELECT

    CASE

        WHEN Exited = 'Yes' THEN 'Exited'

        WHEN Exited = 'No' THEN 'Not Exited'

        ELSE 'Other'

```

```

        END AS Customer_Status,

        Gender,

        COUNT(*) AS Count

FROM

        bank_churn.bank_churn_prepared

GROUP BY Customer_Status, Gender

ORDER BY Customer_Status, Gender;

```

RESULTS:

Customer_Status	Gender	Count
Exited	Female	1139
Exited	Male	898
Not Exited	Female	3404
Not Exited	Male	4559

5.4 Geography analysis:

The number of customers churn across France, Spain and Germany was analyzed.

QUERY:

```

SELECT

        Geography,

        CASE

                WHEN Exited = 'Yes' THEN 'Exited'

                WHEN Exited = 'No' THEN 'Not Exited'

                ELSE 'Other'

        END AS Customer_Status,

        COUNT(*) AS Count

```

```
FROM bank_churn.bank_churn_prepared

GROUP BY Geography, Customer_Status

ORDER BY Geography, Customer_Status;
```

RESULTS:

Geography	Customer_Status	Count
France	Exited	810
France	Not Exited	4204
Germany	Exited	814
Germany	Not Exited	1695
Spain	Exited	413
Spain	Not Exited	2064

5.5 Credit card analysis:

The number of customers with credit card churn was analyzed.

QUERY:

```
SELECT

HasCrCard,

CASE

    WHEN Exited = 'Yes' THEN 'Exited'

        WHEN Exited = 'No' THEN 'Not Exited'

        ELSE 'Other'

    END AS Customer_Status,

COUNT(*) AS Count

FROM bank_churn.bank_churn_prepared

GROUP BY

    HasCrCard, Customer_Status
```


ORDER BY HasCrCard, Customer_Status;

RESULTS:

HasCrCard	Customer_Status	Count
No	Exited	613
No	Not Exited	2332
Yes	Exited	1424
Yes	Not Exited	5631

5.6 Active member analysis:

The number of active customer churn was analyzed.

QUERY:

```
SELECT
IsActiveMember,
CASE
    WHEN Exited = 'Yes' THEN 'Exited'
        WHEN Exited = 'No' THEN 'Not Exited'
        ELSE 'Other'
    END AS Customer_Status,
COUNT(*) AS Count
FROM bank_churn.bank_churn_prepared
GROUP BY
    IsActiveMember, Customer_Status
ORDER BY IsActiveMember, Customer_Status;
```

RESULTS:

IsActiveMember	Customer_Status	Count
No	Exited	1302
No	Not Exited	3547
Yes	Exited	735
Yes	Not Exited	4416

5.7 Tenure analysis:

The number of customers churn based on tenure was analyzed.

QUERY:

```
SELECT
    Tenure,
    CASE
        WHEN Exited = 'Yes' THEN 'Exited'
        WHEN Exited = 'No' THEN 'Not Exited'
        ELSE 'Other'
    END AS Customer_Status,
    COUNT(*) AS Count
FROM bank_churn.bank_churn_prepared
GROUP BY
    Tenure, Customer_Status
ORDER BY Tenure, Customer_Status;
```

RESULTS:

Tenure	Customer_Status	Count
0	Exited	95
0	Not Exited	318
1	Exited	232
1	Not Exited	803
2	Exited	201
2	Not Exited	847
3	Exited	213
3	Not Exited	796
4	Exited	203
4	Not Exited	786
5	Exited	209
5	Not Exited	803
6	Exited	196
6	Not Exited	771
7	Exited	177
7	Not Exited	851
8	Exited	197
8	Not Exited	828
9	Exited	213
9	Not Exited	771
10	Exited	101
10	Not Exited	389

5.8 Credit score analysis:

The number of customers churn based on their credit score was analyzed.

QUERY:

```
SELECT

    CASE

        WHEN CreditScore BETWEEN 350 AND 400 THEN '350 - 400'

        WHEN CreditScore BETWEEN 401 AND 450 THEN '401 - 450'

        WHEN CreditScore BETWEEN 451 AND 500 THEN '451 - 500'

        WHEN CreditScore BETWEEN 501 AND 550 THEN '501 - 550'

        WHEN CreditScore BETWEEN 551 AND 600 THEN '551 - 600'

        WHEN CreditScore BETWEEN 601 AND 650 THEN '601 - 650'

        WHEN CreditScore BETWEEN 651 AND 700 THEN '651 - 700'

        WHEN CreditScore BETWEEN 701 AND 750 THEN '701 - 750'

        WHEN CreditScore BETWEEN 751 AND 800 THEN '751 - 800'

        WHEN CreditScore BETWEEN 801 AND 850 THEN '801 - 850'

        ELSE 'Other'

    END AS CreditScore_range,

    COUNT(*) AS Count

FROM

    bank_churn.bank_customers_exited WHERE Exited = "Yes"

GROUP BY CreditScore_range

ORDER BY CreditScore_range;
```

RESULTS:

CreditScore_range	Count
350 - 400	19
401 - 450	42
451 - 500	91

501 - 550	216
551 - 600	297
601 - 650	392
651 - 700	361
701 - 750	306
751 - 800	186
801 - 850	127

5.9 Product number analysis:

The number of customer churn based on number of products was analyzed.

QUERY:

```

SELECT

NumOfProducts,

CASE

    WHEN Exited = 'Yes' THEN 'Exited'

        WHEN Exited = 'No' THEN 'Not Exited'

        ELSE 'Other'

    END AS Customer_Status,

COUNT(*) AS Count

FROM bank_churn.bank_churn_prepared

GROUP BY

    NumOfProducts, Customer_Status

ORDER BY NumOfProducts, Customer_Status;
```

RESULTS:

NumOfProducts	Customer_Status	Count
1	Exited	1409
1	Not Exited	3675
2	Exited	348
2	Not Exited	4242
3	Exited	220
3	Not Exited	46
4	Exited	60

5.10 Salary analysis:

The number of customers churn based on the salary was analyzed.

QUERY:

SELECT

CASE

WHEN EstimatedSalary BETWEEN 0 AND 20000 THEN '0 - 20k'

WHEN EstimatedSalary BETWEEN 20001 AND 50000 THEN '20k - 50k'

WHEN EstimatedSalary BETWEEN 50001 AND 100000 THEN '50k - 100k'

WHEN EstimatedSalary BETWEEN 100001 AND 150000 THEN '100k - 150k'

WHEN EstimatedSalary BETWEEN 150001 AND 200000 THEN '150k - 200k'

ELSE 'Other'

END AS Salary_Range,

Exited,

COUNT(*) AS Count

FROM

```

bank_churn.bank_churn_prepared

GROUP BY

Salary_Range, Exited

ORDER BY Salary_Range;

```

RESULTS:

Salary_Range	Exited	Count
0 - 20k	No	788
0 - 20k	Yes	198
100k - 150k	No	2038
100k - 150k	Yes	517
150k - 200k	No	1928
150k - 200k	Yes	527
20k - 50k	No	1176
20k - 50k	Yes	291
50k - 100k	No	2033
50k - 100k	Yes	504

5.11 Bank balance analysis:

The number of customers churn based on their bank balance was analyzed.

QUERY:

```

SELECT

CASE

    WHEN Balance BETWEEN 0 AND 20000 THEN '0 - 20k'

    WHEN Balance BETWEEN 20001 AND 50000 THEN '20k - 50k'

    WHEN Balance BETWEEN 50001 AND 100000 THEN '50k - 100k'

    WHEN Balance BETWEEN 100001 AND 150000 THEN '100k - 150k'

```

```

        WHEN Balance BETWEEN 150001 AND 200000 THEN '150k - 200k'

        WHEN Balance BETWEEN 2000001 AND 2500001 THEN '200K - 250K'

        ELSE 'Other'

    END AS Balance_Range,

    Exited,

    COUNT(*) AS Count

FROM

    bank_churn.bank_churn_prepared

GROUP BY

    Balance_Range, Exited

ORDER BY Balance_Range;

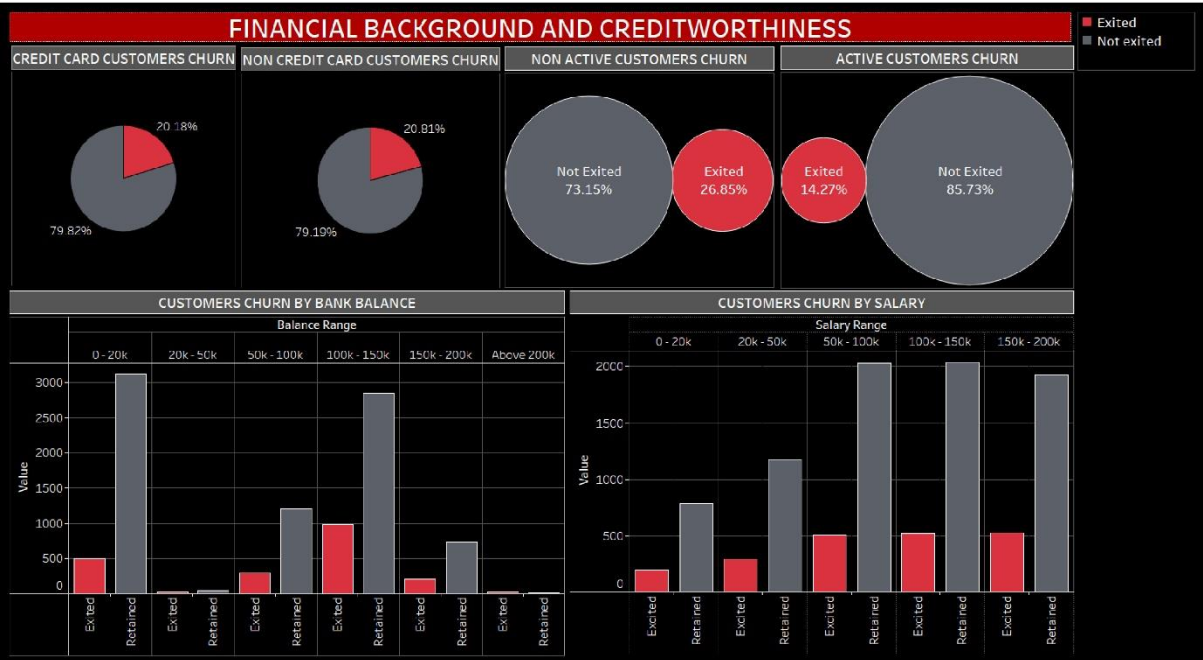
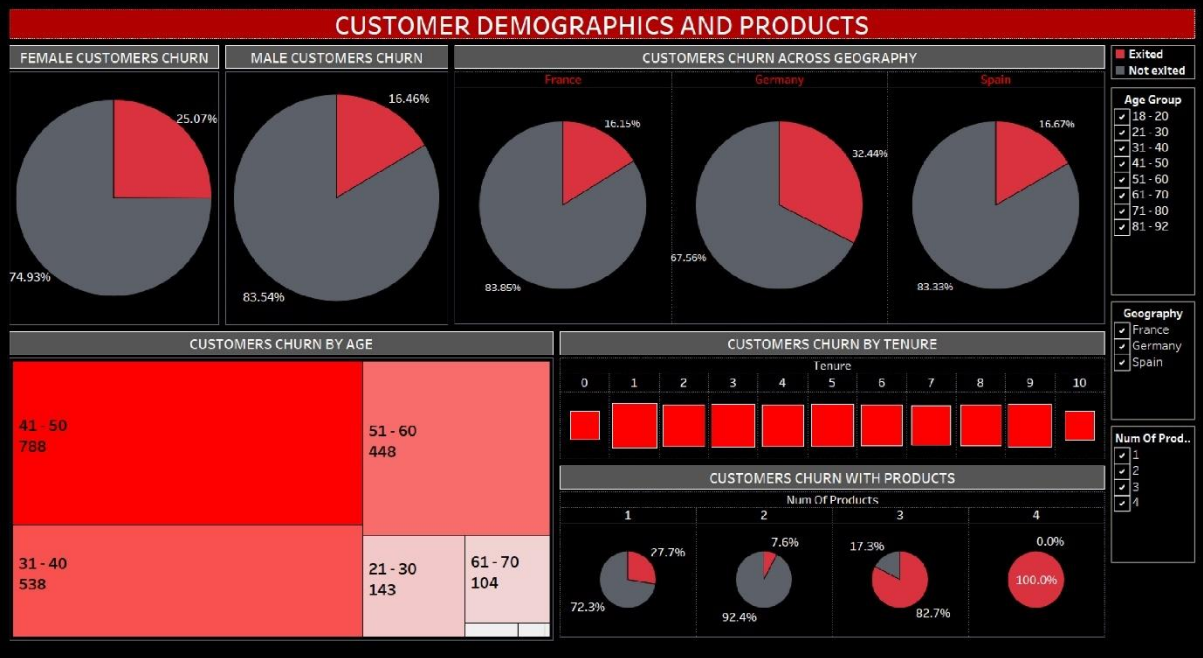
```

RESULTS:

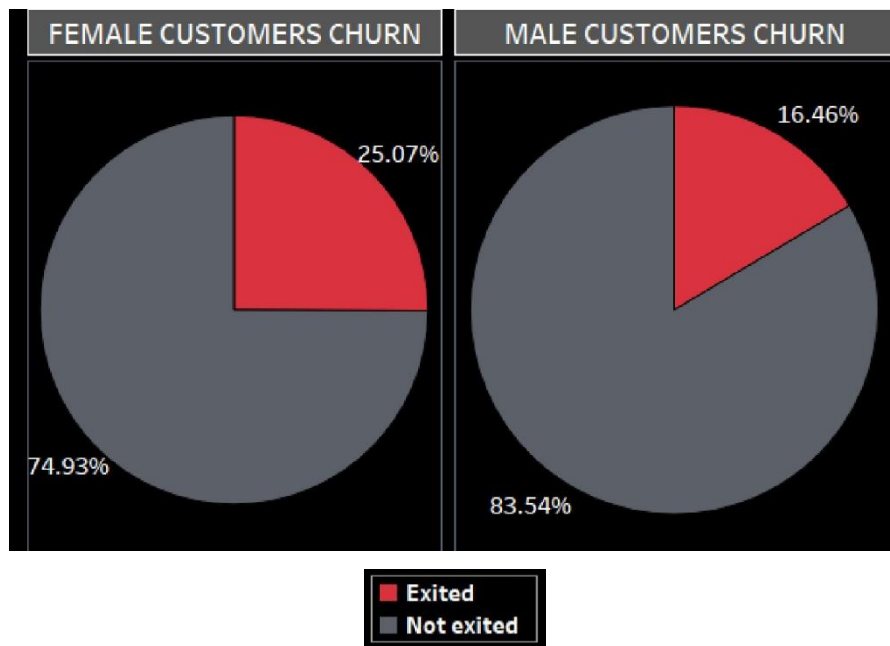
Balance_Range	Exited	Count
0 - 20k	No	3119
0 - 20k	Yes	502
100k - 150k	No	2843
100k - 150k	Yes	987
150k - 200k	No	730
150k - 200k	Yes	205
20k - 50k	No	47
20k - 50k	Yes	24
50k - 100k	No	1209
50k - 100k	Yes	300
Other	No	15
Other	Yes	19

6. VISUALIZATION:

DASHBOARDS



Customer Demographics and Products - Dashboard Analysis



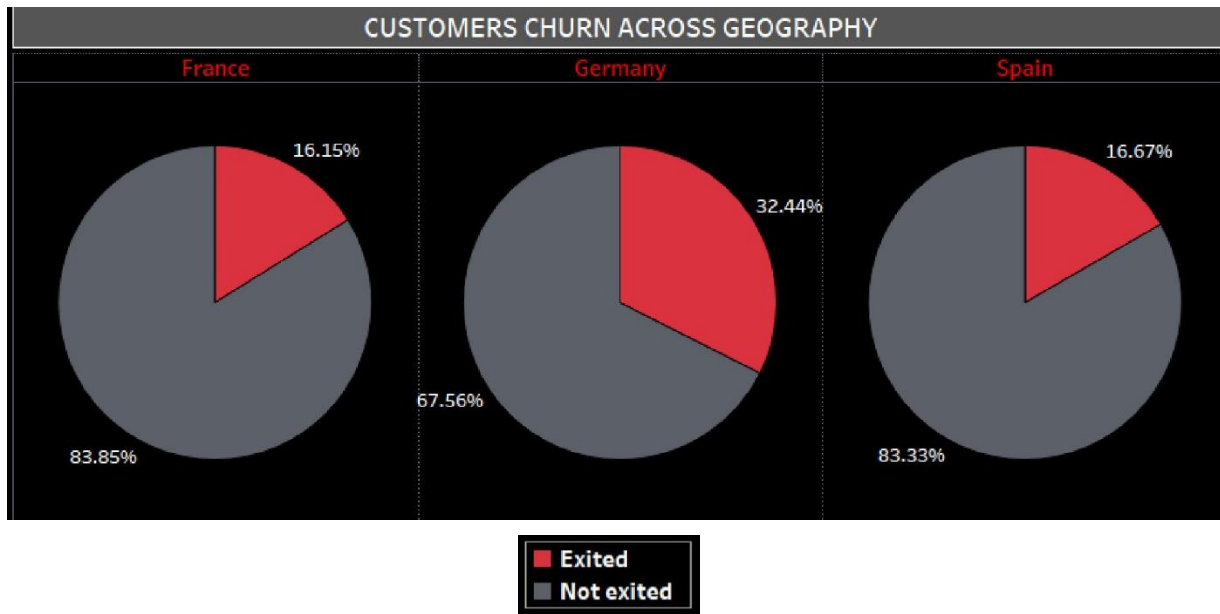
1. Female vs. Male Customer Churn (Pie Charts)

- **Female Customers:**
 - 25.07% churned (exited).
 - 74.93% retained (not exited).
- **Male Customers:**
 - 16.46% churned.
 - 83.54% retained.

Inference:

- Female customers are more likely to leave than male customers, with a churn rate of 25.07% compared to 16.46%. This suggests that gender may play a role in customer loyalty, and targeted retention strategies could be useful for female customers.
-

2. Customer Churn by Geography (France, Germany, Spain - Pie Charts)

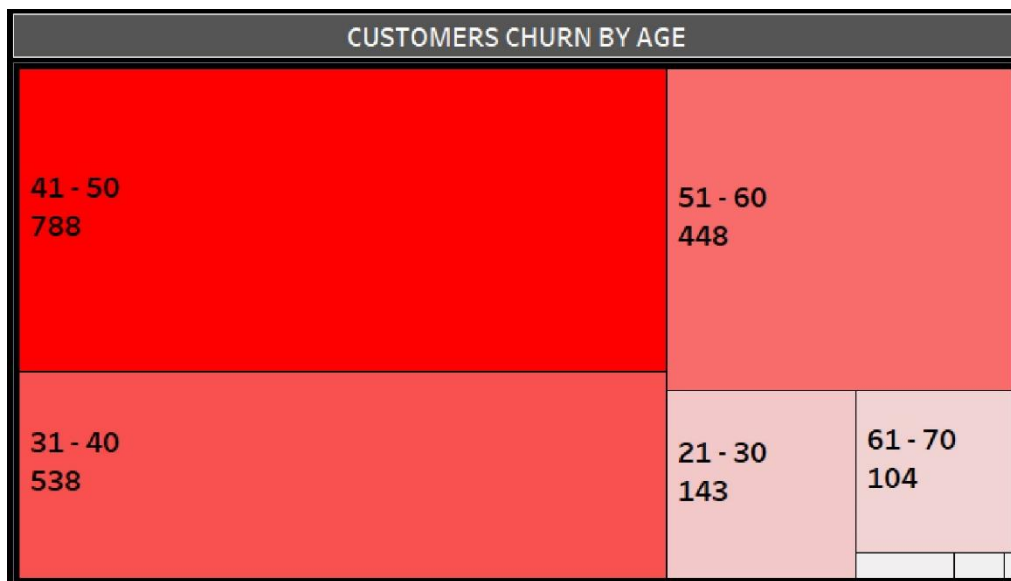


- **France:**
 - 16.15% churned, 83.85% retained.
- **Germany:**
 - 32.44% churned, 67.56% retained.
- **Spain:**
 - 16.67% churned, 83.33% retained.

Inference:

- Germany exhibits the highest churn rate (32.44%), indicating customer dissatisfaction or competitive pressure in that region. In contrast, France and Spain show similar and much lower churn rates (~16%), suggesting better customer retention in those countries.
-

3. Customer Churn by Age (Tree Map)

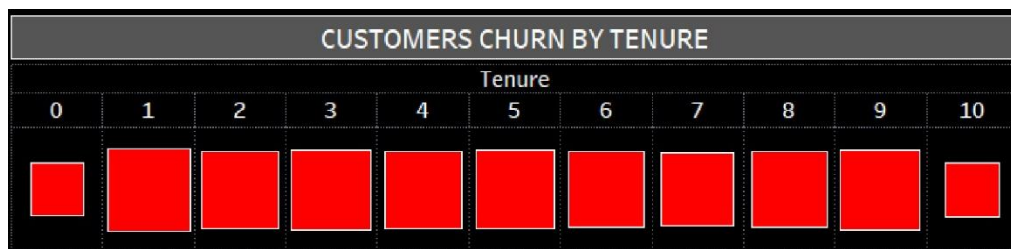


- **Age Group 41-50:** 788 customers churned (largest segment).
- **Age Group 31-40:** 538 churned.
- **Age Group 51-60:** 448 churned.
- **Age Group 21-30:** 143 churned.
- **Age Group 61-70:** 104 churned.

Inference:

- The **41-50 age group** has the highest number of churned customers, indicating potential challenges in engaging or retaining customers in this segment. Younger customers (21-30) and older customers (61-70) show relatively lower churn.
-

4. Customer Churn by Tenure



1. High Churn Period (Tenure 1-3 years)

- **Largest boxes appear for Tenure 1, 2, and 3.**
- **Churn peaks** during this early period, especially in the **first and third years**, suggesting that customers tend to leave **within the first three years** of their relationship.
- This indicates that **the first three years are critical** for retention efforts, especially after the first year (churn in year 1 is the highest).

2. Moderate Churn Period (Tenure 4-6 years)

- **Slight reduction in box sizes** appears for Tenure 4, 5, and 6.
- Though churn is still happening, it becomes **more stable** during this period, with values showing **only slight variation year to year**.
- This suggests that **mid-tenure customers** are somewhat more engaged and less likely to leave than early-tenure customers.

3. Decline in Churn (Tenure 7-10 years)

- **Smaller box sizes** are noticeable for **years 7 through 10**, with **year 10 showing the least churn**.
- Customers who have remained this long are **more loyal** and less likely to churn.
- **Churn gradually decreases** beyond year 6, emphasizing the importance of customer satisfaction and loyalty-building for long-term retention.

4. Outliers: Year 0 and Year 10

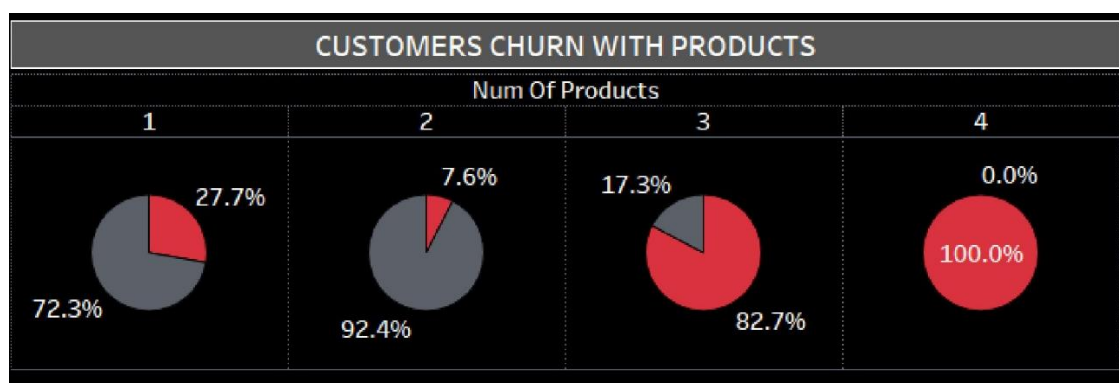
- **Year 0:** A **medium-sized box** indicates **some churn happening very early**, likely from customers leaving soon after joining.
 - **Year 10:** The **smallest box size**, reflecting **very low churn**, indicates that customers who reach this stage are highly loyal.
-

Summary of Key Insights from Tenure Heatmap:

1. **First three years are the most critical** for retention, with **churn peaking in year 1**. Retention efforts during this period should focus on **onboarding, customer engagement, and addressing early dissatisfaction**.
2. **Mid-tenure customers (4-6 years)** show **stable churn patterns**, signaling moderate engagement but with room for improvement.
3. **Long-tenure customers (7-10 years)** are **highly loyal**, with **minimal churn** by year 10. Retention efforts beyond year 6 should shift toward **maintaining loyalty and providing rewards**.

This heatmap highlights the importance of **early interventions** in the first 3 years to reduce churn, followed by **sustained customer satisfaction efforts** to build long-term loyalty.

5. Customer Churn by Number of Products (Pie Charts)



- **1 Product:** 27.7% churned, 72.3% retained.
- **2 Products:** 7.6% churned, 92.4% retained.
- **3 Products:** 17.3% churned, 82.7% retained.
- **4 Products:** 100% churned.



Inference:

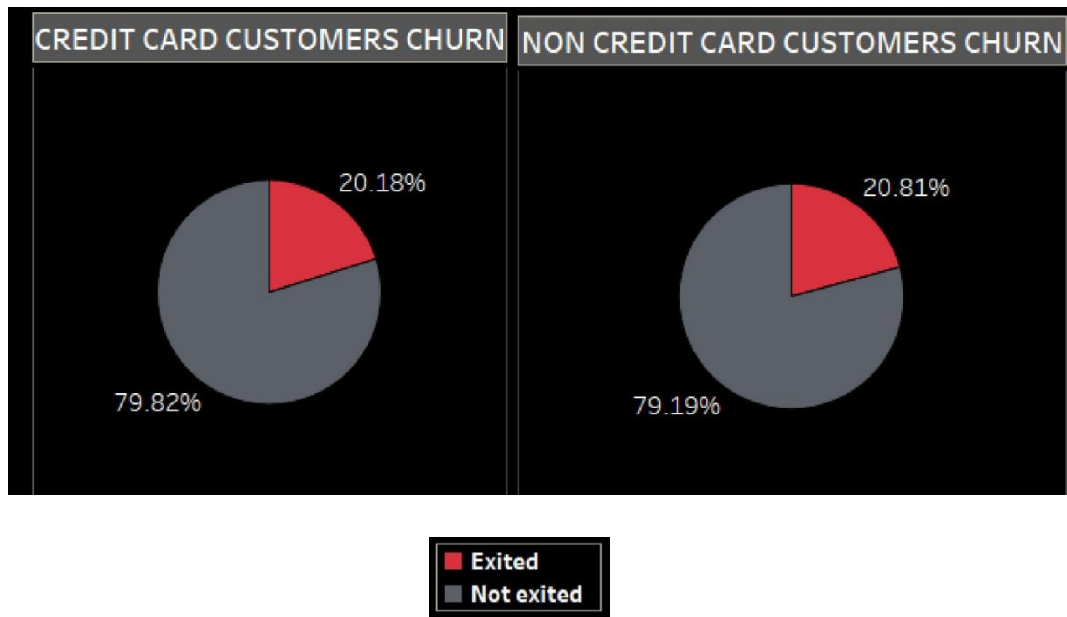
- Customers with **4 products** have the highest churn rate after those with 4 products. This suggests that **product engagement** plays a key role in retention—encouraging customers to adopt multiple products could reduce churn.
 - The 100% churn for customers with 4 products might indicate a **data anomaly** or a unique case.
-

Overall Insights:

- **Gender:** Females churn more than males (25.07% vs. 16.46%).
- **Geography:** Germany has the highest churn (32.44%), suggesting localized retention efforts may be needed.
- **Age:** The 41-50 age group has the largest churn, highlighting it as a key focus area for customer retention.
- **Products:** Encouraging customers to hold at least 2-3 products may improve retention.
- **Tenure:** Churn is evenly distributed across different tenures, indicating that **loyalty programs targeting long-term customers may not be as effective** as engagement strategies for newer customers.

Financial Background and Creditworthiness - Dashboard Analysis

Credit Card vs. Non-Credit Card Customers Churn (Pie Charts)

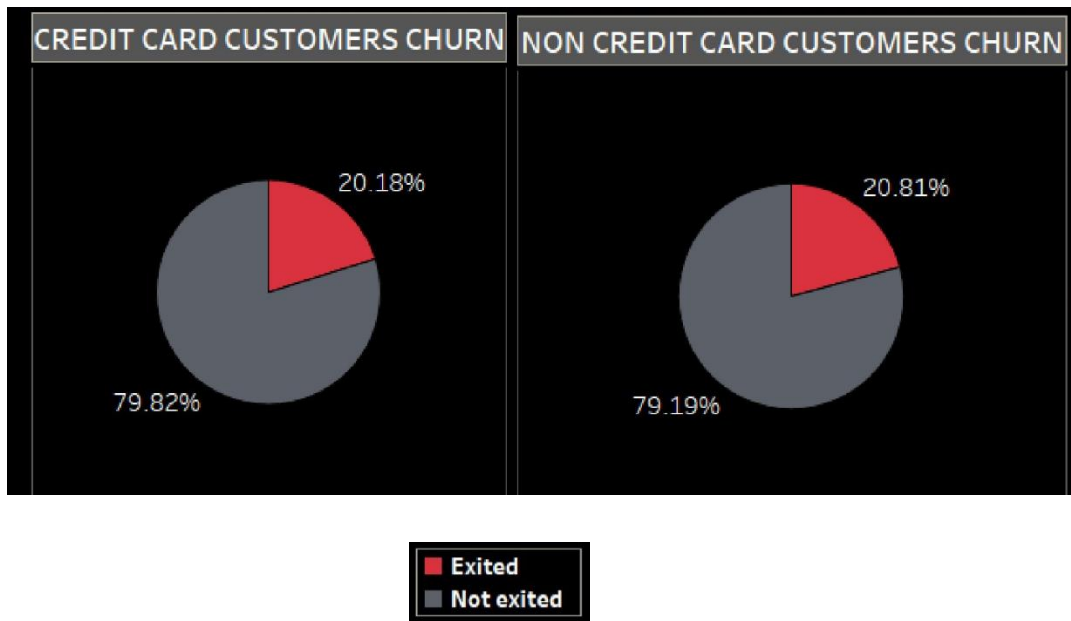


- **Credit Card Customers:**
 - 20.18% exited (churned).
 - 79.82% retained (not exited).
- **Non-Credit Card Customers:**
 - 20.81% exited.
 - 79.19% retained.

Inference:

- Both **credit card holders** and **non-credit card holders** exhibit nearly identical churn rates (~20%). This suggests that **credit card ownership alone does not have a major impact** on customer churn.
-

2. Non-Active vs. Active Customers Churn (Pie Charts)



- **Non-Active Customers:**

- 26.85% exited.
- 73.15% retained.

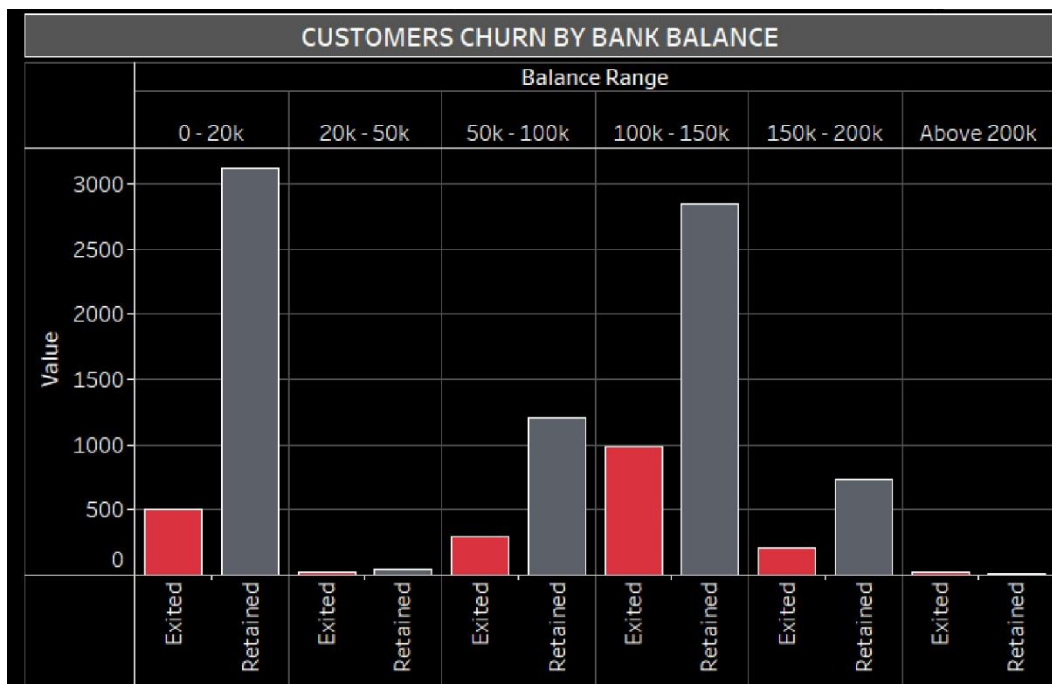
- **Active Customers:**

- 14.27% exited.
- 85.73% retained.

Inference:

- **Non-active customers** are almost **twice as likely to churn (26.85%)** compared to active customers (14.27%). This indicates that customer activity plays a crucial role in retention. Efforts to **engage inactive customers** could significantly reduce churn.
-

3. Customers Churn by Bank Balance (Bar Chart)



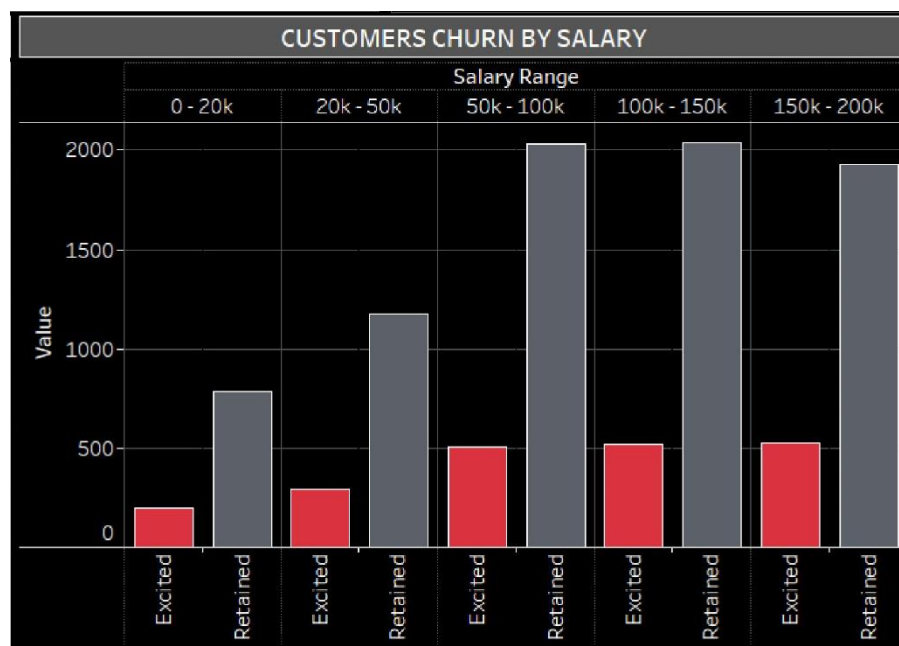
- **Balance Range and Churn:** There appears to be a correlation between balance range and customer churn.
- **Lower Balance Ranges:** Customers in lower balance ranges (0-20k and 20k-50k) have significantly higher churn rates. This might suggest that financial constraints or affordability issues could be contributing factors to customer attrition.
- **Higher Balance Ranges:** Customers in higher balance ranges (50k-100k, 100k-150k, 150k-200k, Above 200k) have varying churn rates.

Potential Explanations:

- **Financial Stress:** Customers in lower balance ranges might be more susceptible to financial stress or economic hardship, which could lead to churn.
- **Product Suitability:** The bank's products and services might not be well-suited to the financial needs of customers in lower balance ranges.

- **Competitive Offers:** Customers in lower balance ranges might be more attracted to competitive offers from other banks that cater specifically to their financial situation.
- **Value Perception:** Customers with higher balances might perceive the bank's products and services as more valuable or relevant to their financial goals.

4. **Customers Churn by Salary (Bar Chart)**



- **Salary Range and Churn:** There appears to be a correlation between salary range and customer churn.
- **Lower Salary Ranges:** Customers in lower salary ranges (0-20k and 20k-50k) have lower churn rates.
- **Higher Salary Ranges:** Customers in higher salary ranges (50k-100k, 100k-150k, 150k-200k) have higher churn rates. This could indicate that customers with higher incomes might be more constrained with the bank's services or have stronger financial stability.

Potential Explanations:

- **Financial Stability:** Customers in higher salary ranges might not be susceptible to financial stress or economic hardship, which could lead to churn.
 - **Product Suitability:** The bank's products and services might not be well-suited to the financial needs of customers in lower salary ranges.
 - **Competitive Offers:** Customers in lower salary ranges might be more attracted to competitive offers from other banks that cater specifically to their financial situation.
 - **Value Perception:** Customers in higher salary ranges might perceive the bank's products and services as more valuable or relevant to their financial goals.
-

Overall Insights:

- **Activity Levels:** Non-active customers churn almost twice as much as active ones, indicating the need for proactive engagement.
- **Bank Balances:** High churn among **customers with balances of 100k-150k** suggests that even high-value customers can feel dissatisfied. Focused efforts might be required to **retain high-balance clients**.
- **Salary Levels:** **Extreme salary groups** (both low and high) show higher churn rates, highlighting the need to address the **unique needs of both ends of the financial spectrum**.

This dashboard emphasizes the importance of **engaging inactive customers**, **offering tailored support to high-balance customers**, and **designing personalized retention strategies** for customers at both **low and high salary levels**.

7. ACT

Recommendations

Based on the insights gained, here are targeted strategies to reduce customer churn:

1. Enhanced Onboarding Process:

- Focus retention efforts on the **first three years**, especially during the initial months. Offer personalized onboarding experiences and proactive communication to engage new customers early.

2. Personalized Customer Engagement:

- Send targeted offers or discounts to at-risk customers (e.g., those with lower credit scores).
- Use CRM tools to send personalized messages for birthdays, milestones, or upcoming renewals.

3. Customer Feedback Programs:

- Set up feedback loops to capture customer pain points in real-time, allowing quick resolution.
- Implement NPS (Net Promoter Score) tracking to understand satisfaction trends.

4. Loyalty Programs:

- Introduce reward programs for long-term customers or those with high-value accounts.

5. Gender-Specific Strategies:

Develop targeted campaigns and engagement strategies for female customers, who show a higher churn rate. This could include personalized offers, enhanced customer service, and loyalty programs.

6. Region-Specific Retention Plans:

- Investigate the **German market** for factors driving churn and adapt services accordingly. This could involve competitive product offerings or improved customer service.

7. Product Bundling and Satisfaction Programs:

- Encourage customers to adopt multiple products through incentives and product bundles. Investigate reasons behind the high churn rate among customers with 4 products to address any product or service issues.

8. Engagement Programs for Mid-Tenure Customers:

- Introduce **mid-tenure loyalty programs** to keep customers engaged beyond the first few years. Offer exclusive benefits to customers with tenure between 4-6 years to prevent churn.

9. Financial Stability support:

- Provide financial counselling or credit score improvement programs to customers with **low credit scores** to reduce churn risks.

10. Proactive credit Support

- Provide financial advice or assistance to customers with lower credit scores to improve their standing.

11. Proactive Outreach for Non-Active Members:

- Implement re-engagement campaigns for **non-active members** through targeted communication and incentives, such as discounts on products or fees.

8. Conclusion

This analysis highlights several factors influencing bank customer churn, such as tenure, gender, geography, and product ownership. The first few years of the customer relationship are critical, requiring special focus to prevent churn. While multiple products improve retention, high churn among customers with many products points to potential service gaps. Germany's higher churn rate and female customers' propensity to leave also call for targeted strategies.

By implementing the recommended retention strategies, banks can not only reduce churn but also enhance long-term customer loyalty, leading to sustained growth and profitability. Continuous monitoring and analysis of customer behaviour will further refine these strategies over time.