

PHÂN TÍCH SỰ RỜI BỎ CỦA KHÁCH HÀNG TRONG NGÀNH NGÂN HÀNG

Presented by: Nhóm 03



01

Giới thiệu đề tài

02

Chân dung khách hàng

03

Phân tích các yếu tố ảnh hưởng đến sự rời bỏ

04

Dự đoán tỷ lệ rời bỏ

05

Đề xuất giải pháp

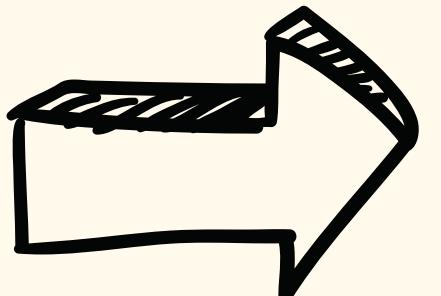
GIỚI THIỆU ĐỀ TÀI

1

Đặt vấn đề & mục tiêu đề tài

MỤC TIÊU ĐỀ TÀI

1. Minh họa chân dung khách hàng của ngân hàng từ tập dữ liệu trên
2. Phân tích những yếu tố có tác động đến sự rời bỏ của khách hàng
3. Xây dựng mô hình dự đoán để có thể dự đoán Tỉ lệ rời bỏ của khách hàng
từ những dữ liệu đầu vào
4. Nhận xét, đưa ra những đề xuất, giải pháp cho doanh nghiệp giúp giảm thiểu sự rời bỏ.



Các **công cụ** được dùng: **Tableau** (trực quan hóa dữ liệu),
Python (xây dựng mô hình dự đoán)

MÔ TẢ DỮ LIỆU

No	Column	Data type	Available values	Note
1	RowNumber	Integer		Số hàng trong dataset
2	CustomerId	Integer		
3	Surname	Text		Họ của khách hàng
4	CreditScore	Integer		
5	Geography	Text	France, Spain, Germany	
6	Gender	Text	Male, Female	
7	Age	Integer		
8	Tenure	Integer		
9	Balance	Float		
10	NumOfProducts	Integer		

11	HasCrCard	Integer	0, 1	1: khách hàng có thẻ tín dụng0: ngược lại
12	IsActiveMember	Integer	0, 1	1: khách hàng hoạt động0: ngược lại
13	EstimatedSalary	Float		
14	Exited	Integer	0, 1	1: khách hàng đã rời bỏ ngân hàng0: ngược lại
15	Complain	Integer	0, 1	1: khách hàng đã phàn nàn0: ngược lại
16	Satisfaction Score	Integer		
17	Card Type	Text	DIAMOND, GOLD, SILVER, PLATINUM	
18	Point Earned	Integer		

CUSTOMER PROFILE & BEHAVIOR

2

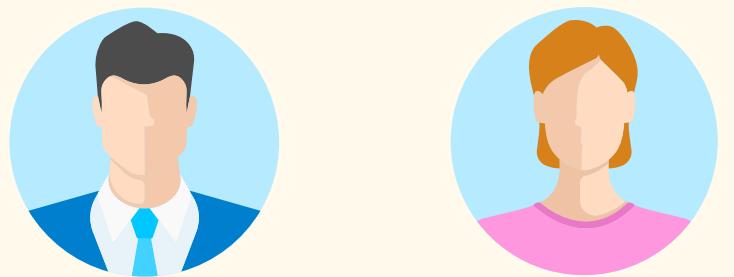
Chân dung khách hàng của ngân hàng

CUSTOMER DEMOGRAPHICS

10.000 CUSTOMERS

39 AVG AGE

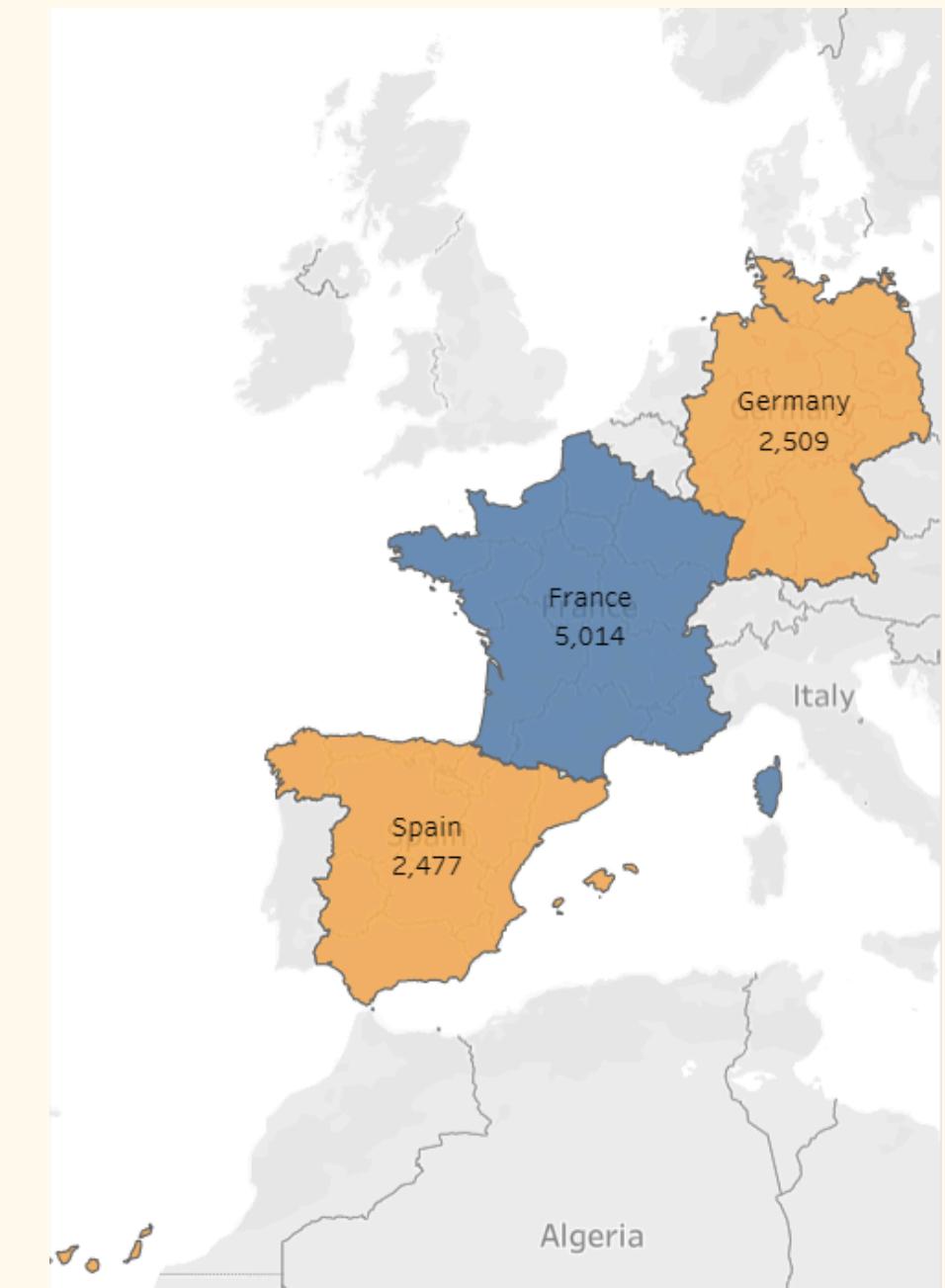
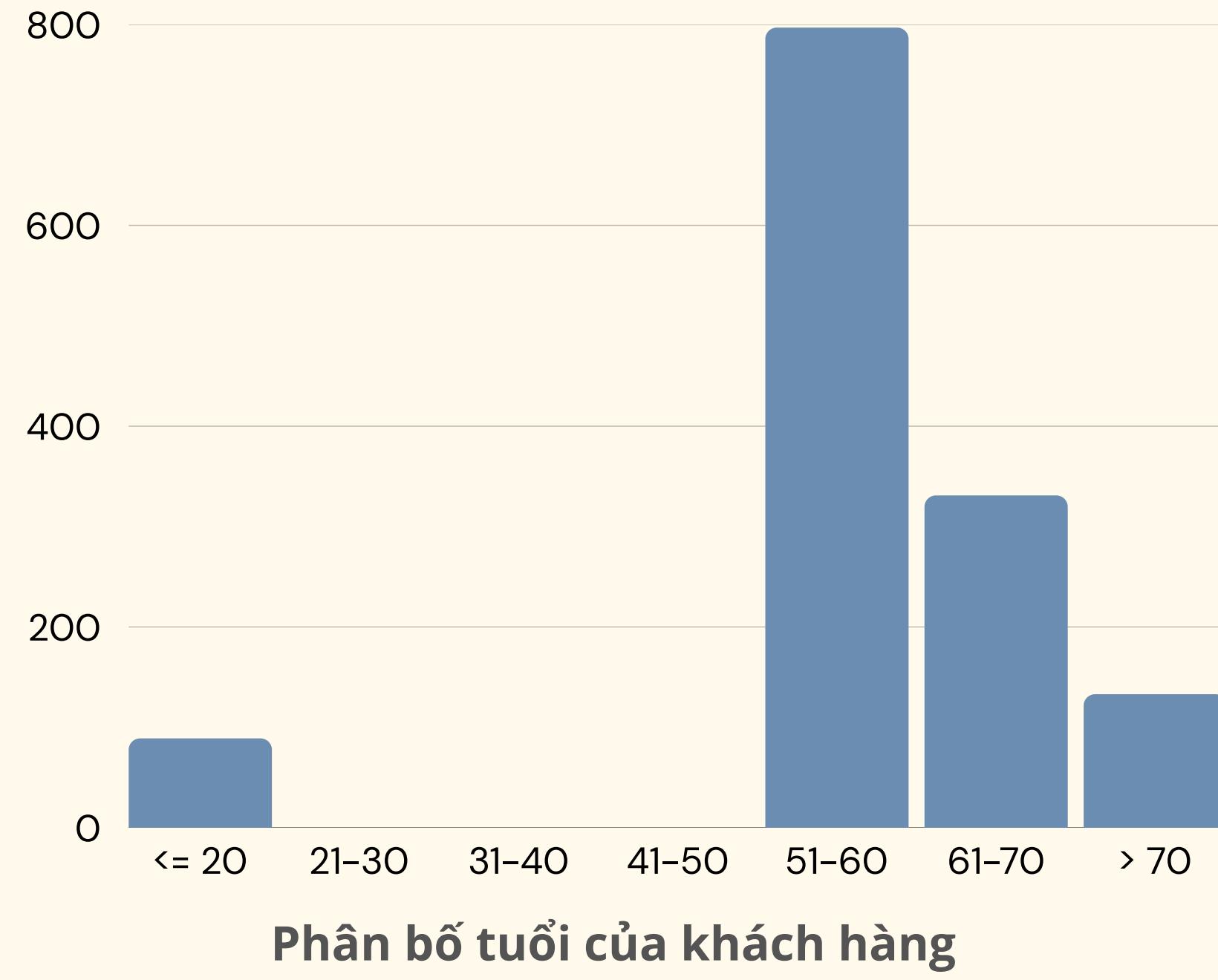
3 COUNTRIES



54,6% **45,4%**

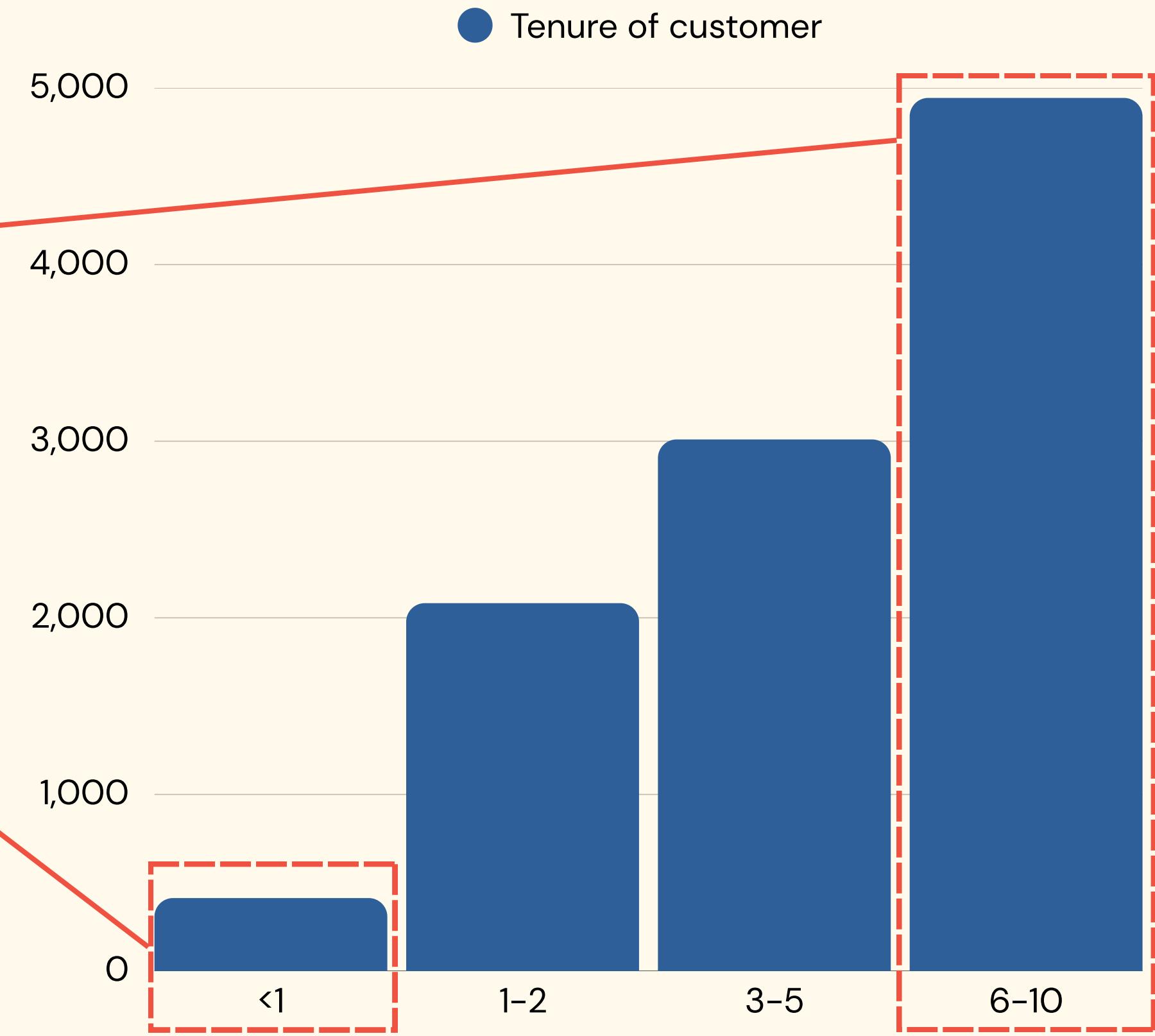
100,090 AVG SALARY

76,486 AVG BALANCE



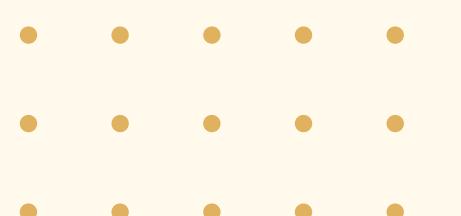
CUSTOMER BEHAVIOR

Thâm niên của khách hàng trong ngân hàng



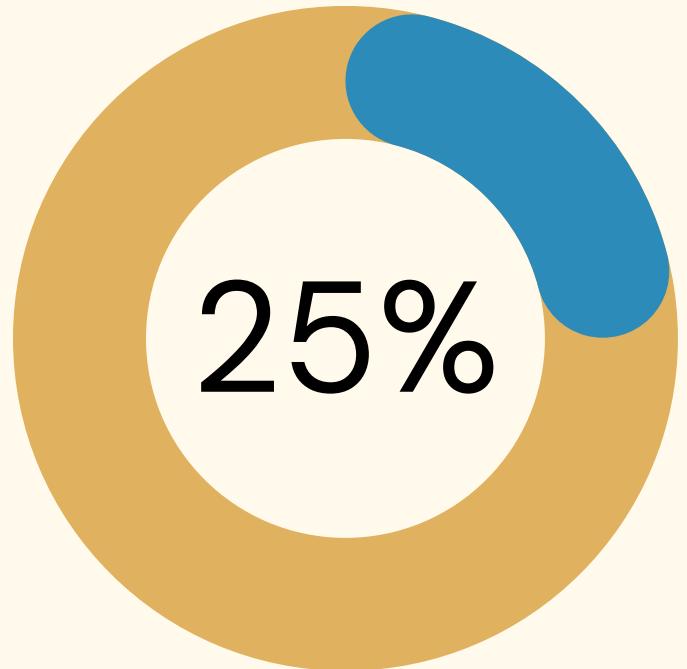
Số khách hàng có thâm niên từ **6 đến 10** năm chiếm số lượng cao nhất trong toàn bộ khách hàng với **4,494** (chiếm gần đến **45%**)

Ngược lại số năm gắn bó với ngân hàng **dưới 1 năm** là **thấp nhất** với **413** khách hàng (**0,04%**)

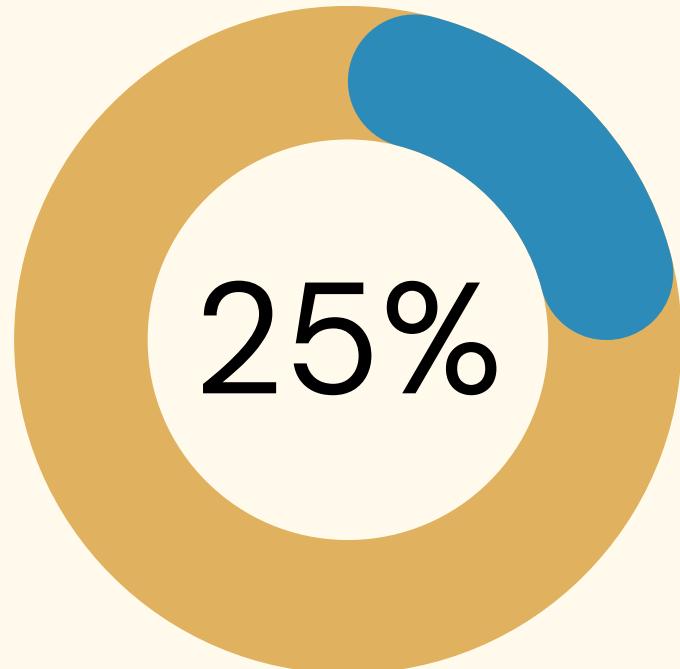


CUSTOMER BEHAVIOR

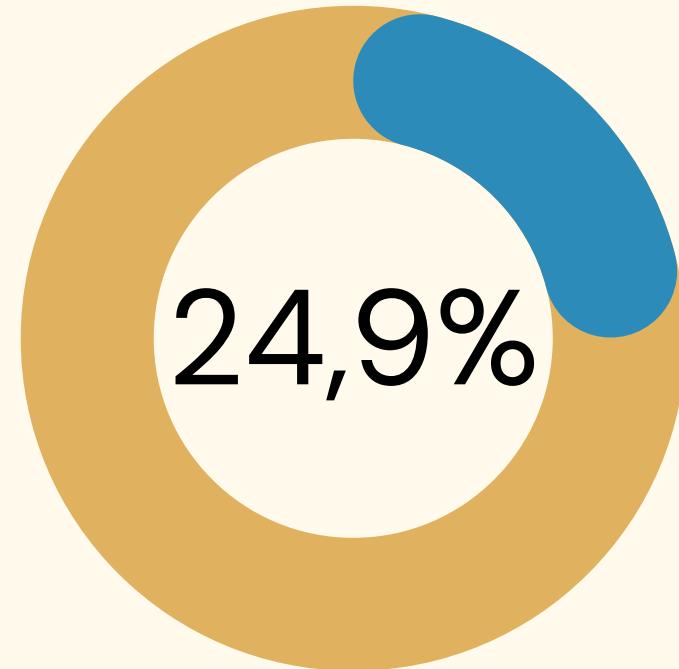
Có 4 loại dịch vụ cấp độ thẻ được ngân hàng áp dụng dựa trên nhu cầu và khả năng chi tiêu của khách hàng là : **Silver, Gold, Platinum** và **Diamond** (theo mức độ từ Silver đến Diamond)



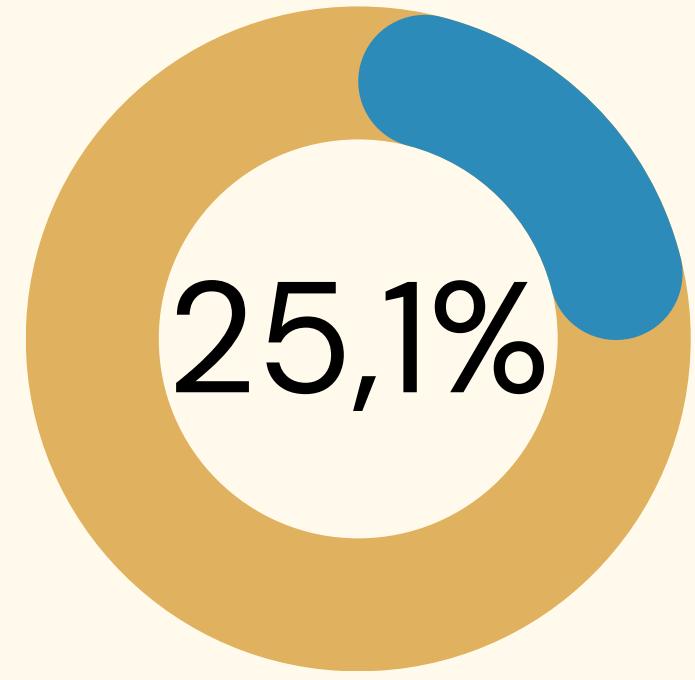
SILVER



GOLD



PLATINUM

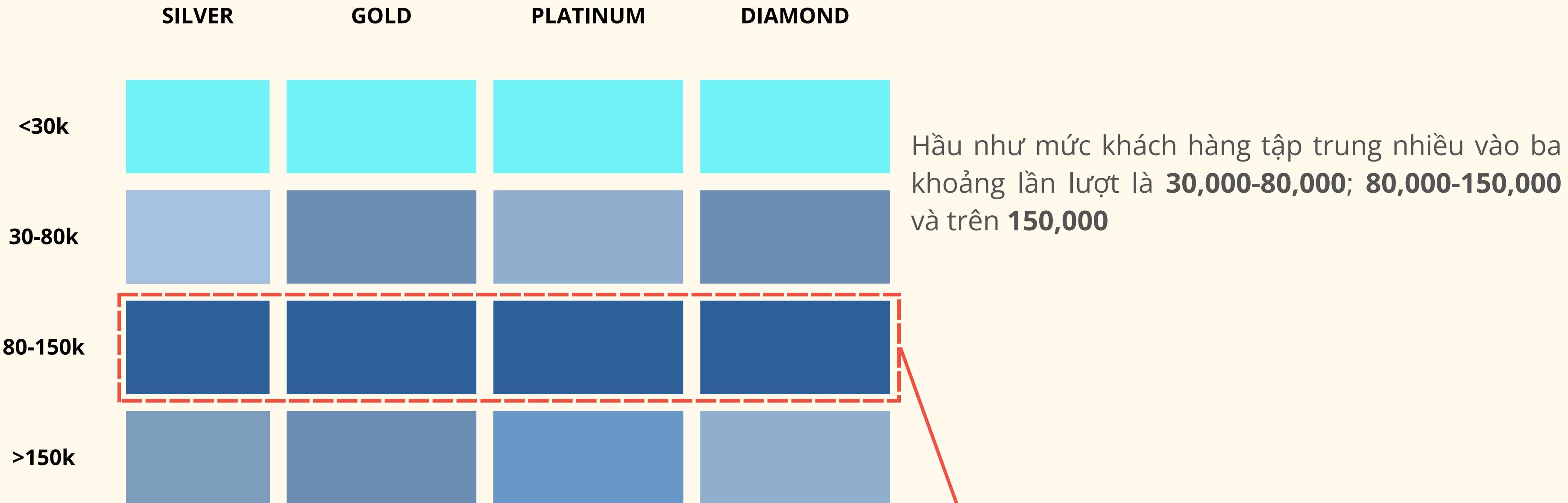


DIAMOND

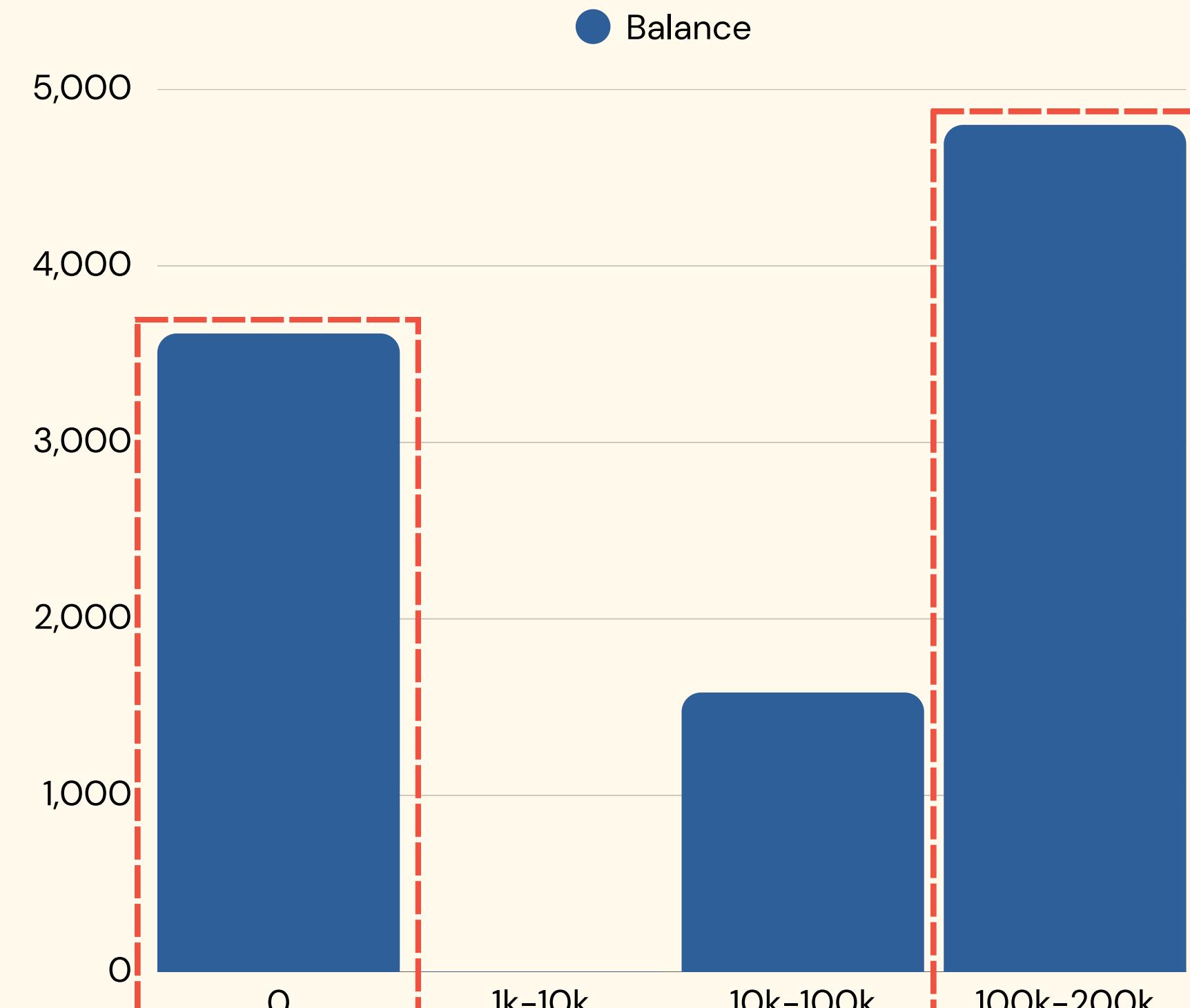
4 cấp thẻ đều có mức độ khách hàng sử dụng **đồng đều** (trải đều từ trên dưới 25% với mỗi cấp độ thẻ, hầu như không có dịch vụ thẻ nào quá nhiều hoặc quá ít khách hàng sử dụng đến)

CUSTOMER BEHAVIOR

Xét đến cấp thẻ thành viên và mức lương ước tính



CUSTOMER BEHAVIOR



2 mức số dư đỗi lập là **0** và **100.000 - 200.000** chiếm tỉ lệ cao nhất trong toàn bộ khách hàng khi lên đến hơn **8000** khách hàng (**80%**)

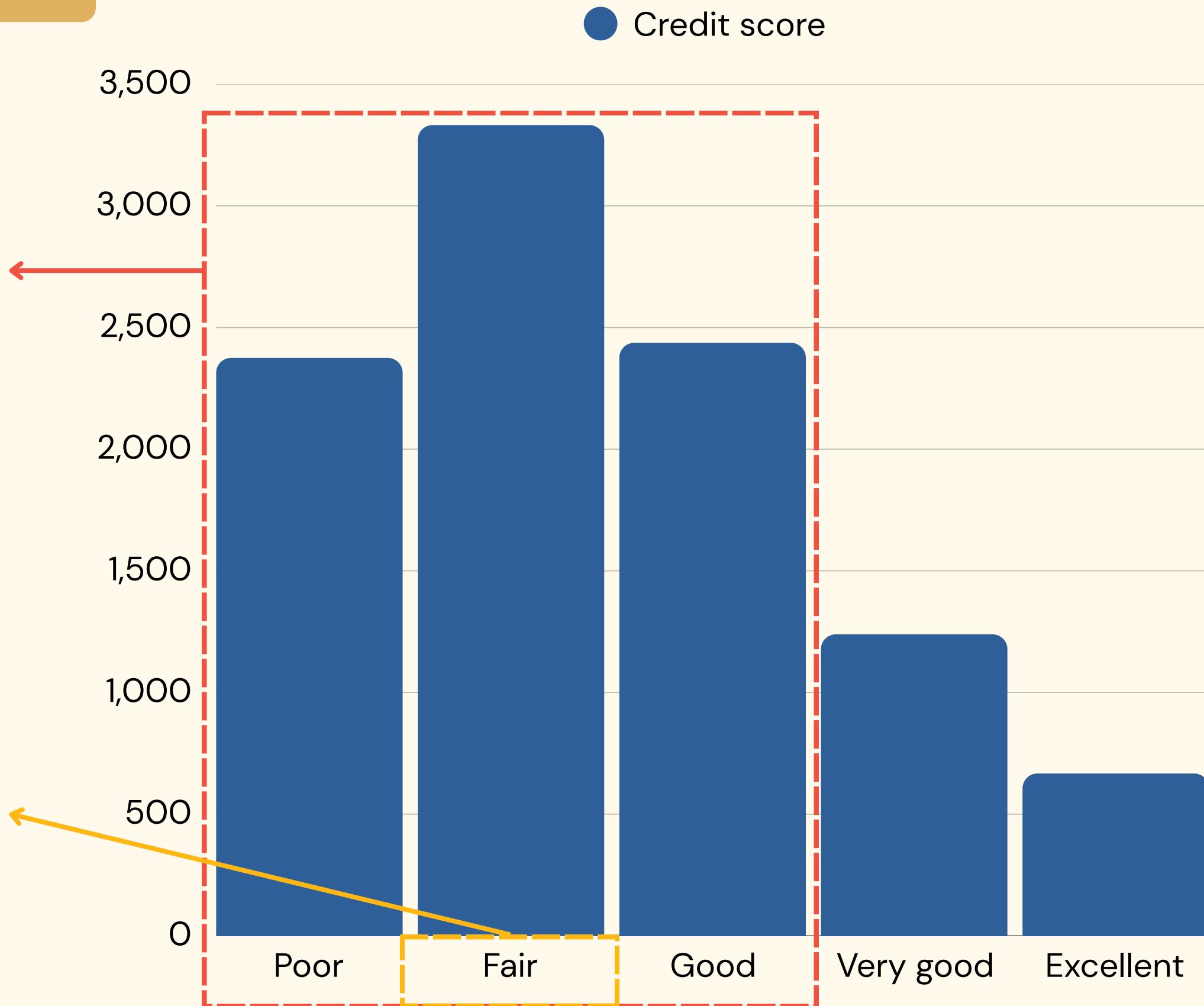
4,799(47%) với mức số dư **100.000 - 200.000**

⋮ ⋮ ⋮ ⋮ Trong đó mức số dư **0** có đến **3,617 (36%)**

CUSTOMER BEHAVIOR

3 mức cấp tín dụng: **Poor, Fair** và **Good** có số lượng tập trung khách hàng cao nhất

Trong đó hạng **Fair** đứng thứ nhất trong cả ba mức độ điểm tín dụng với **3,332** khách hàng với **(33,33%)** trên tổng số khách hàng



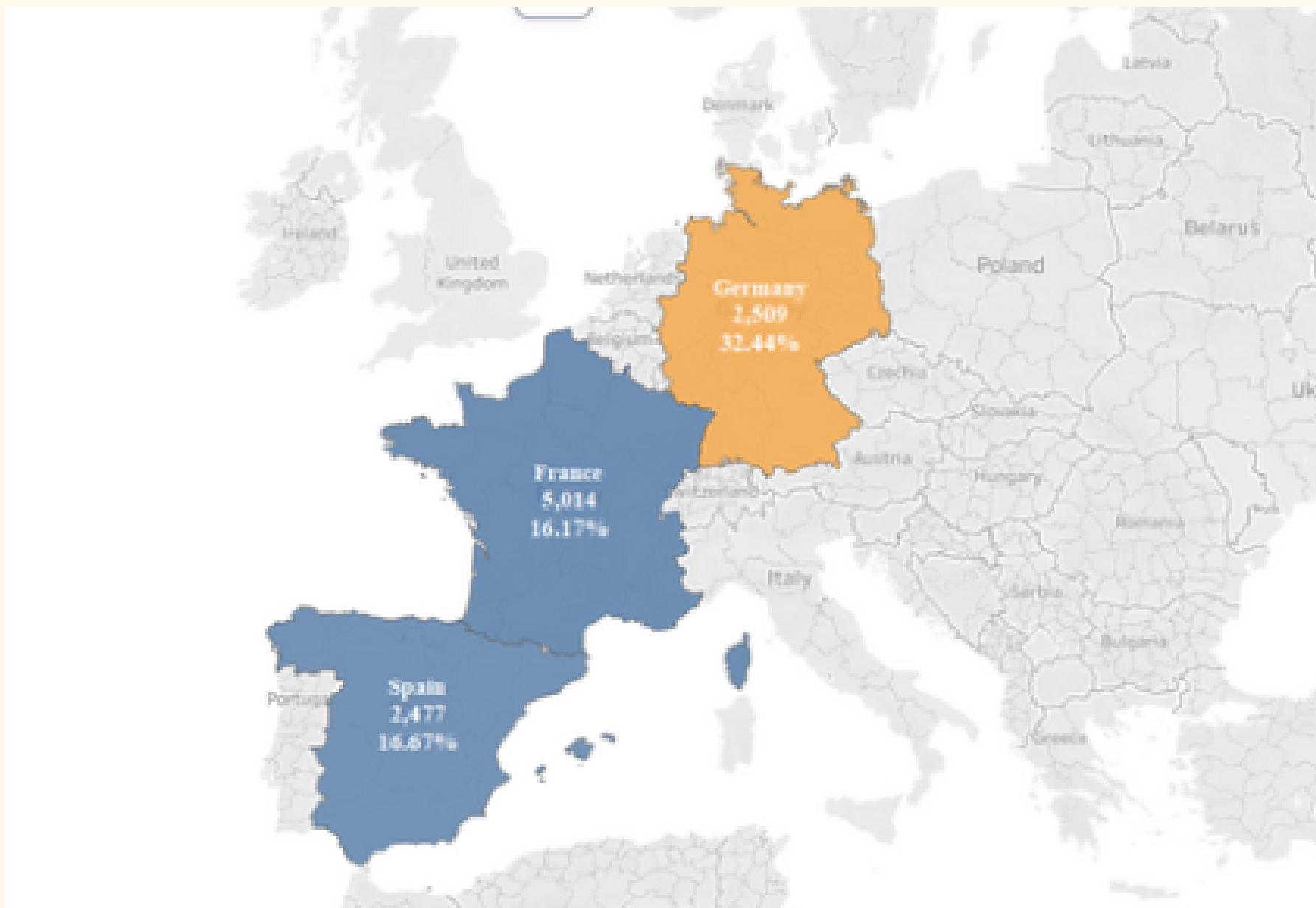
CUSTOMER CHURN PROFILE & BEHAVIOR



3

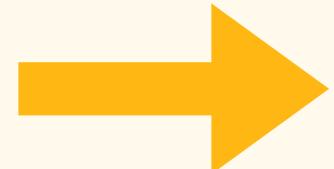
Chân dung khách hàng rời bỏ của ngân hàng

CUSTOMER CHURN PROFILE



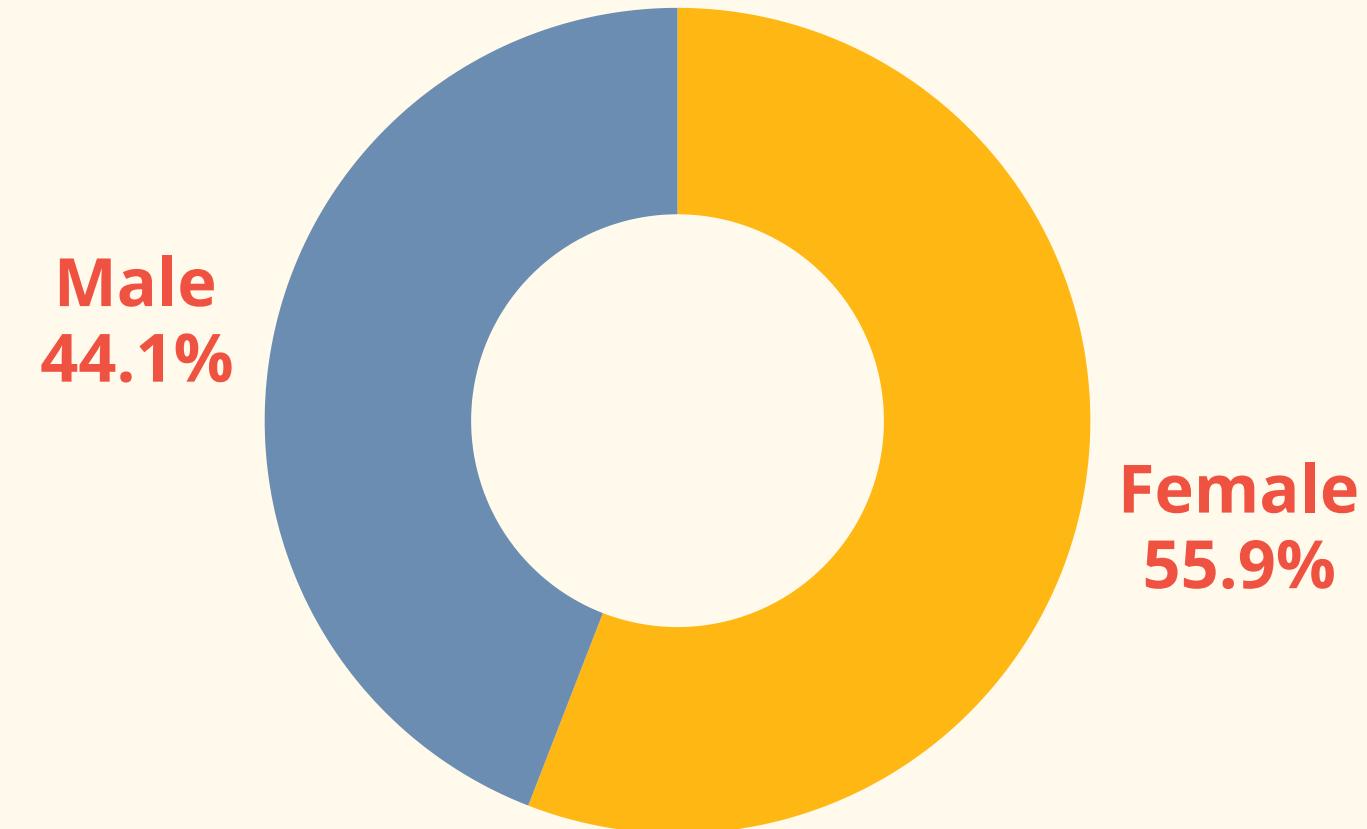
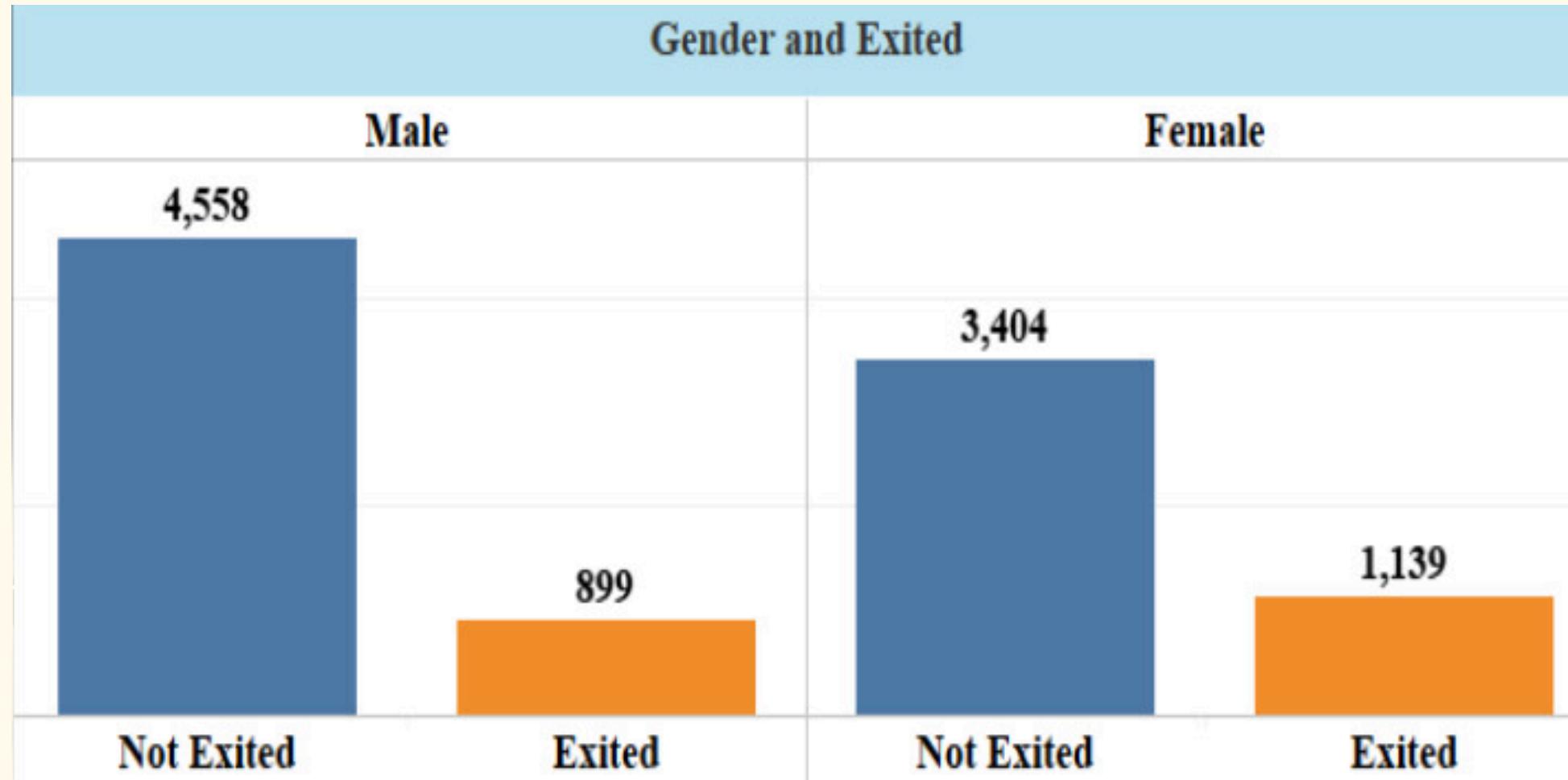
Tỷ lệ rời bỏ của khách hàng ở **Đức** là **cao nhất** với **32,44%** trong khi số khách hàng của họ (2,509) chỉ bằng **1/2** so với **Pháp** (5,014) và tỷ lệ rời bỏ của Pháp, **16,17%**

Tây Ban Nha với 2,477 khách hàng có tỷ lệ rời bỏ là **16,67%**, đứng vị trí thứ 2



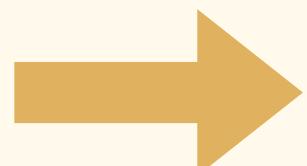
Tỉ lệ khách hàng rời bỏ từ đến từ nước **Đức** đang vượt qua mức trung bình là **19%**

CUSTOMER CHURN PROFILE



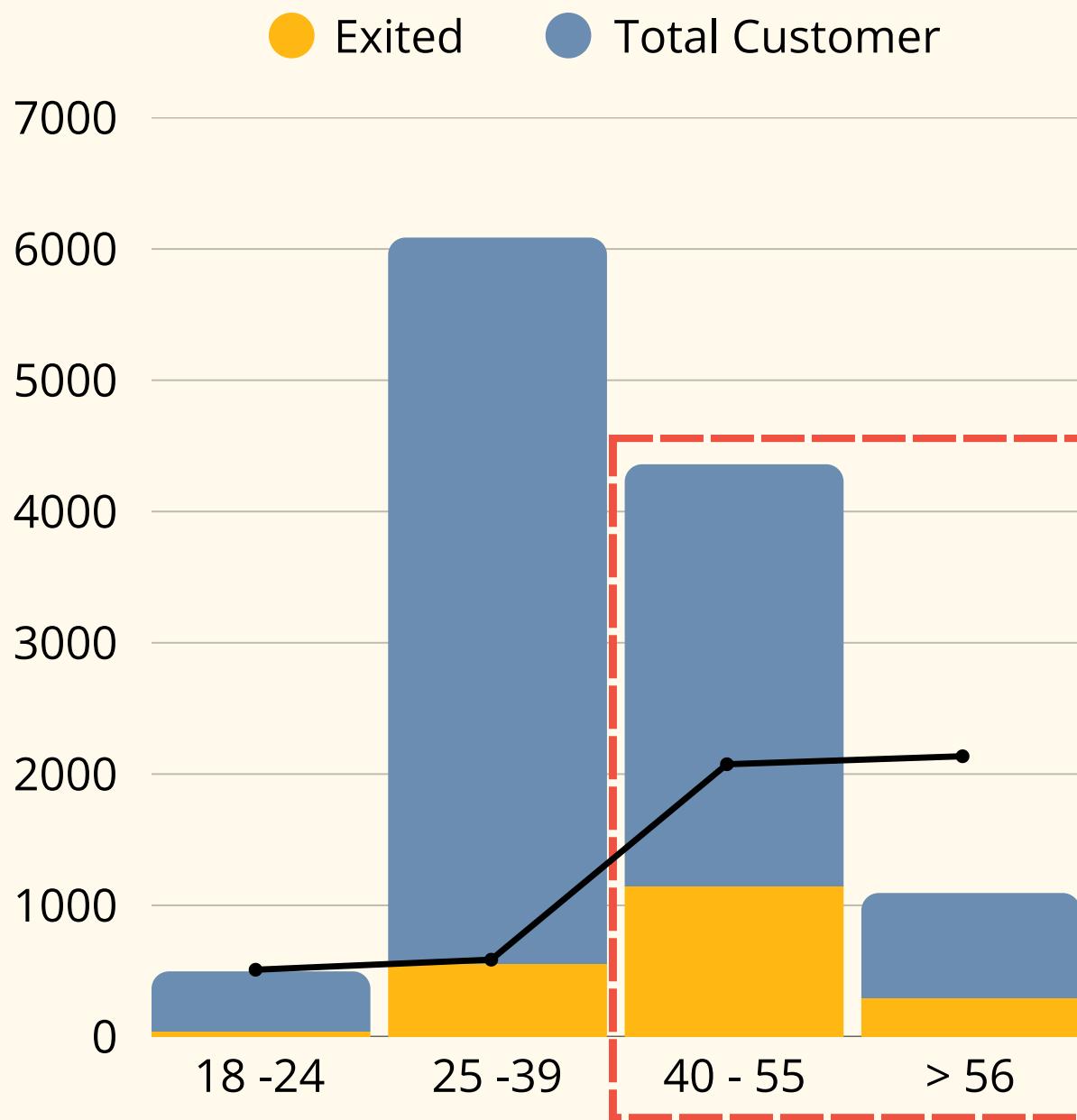
Tỉ lệ rời bỏ theo giới tính

Có đến **1,139** khách hàng **đã rời bỏ** trong **4,543** tổng số khách hàng nữ, chiếm tỉ lệ lên đến **55.89%** (**một con số rất cao**).
Tỉ lệ rời bỏ của nữ giới cao hơn **11%** so với nam (**44.11%**).



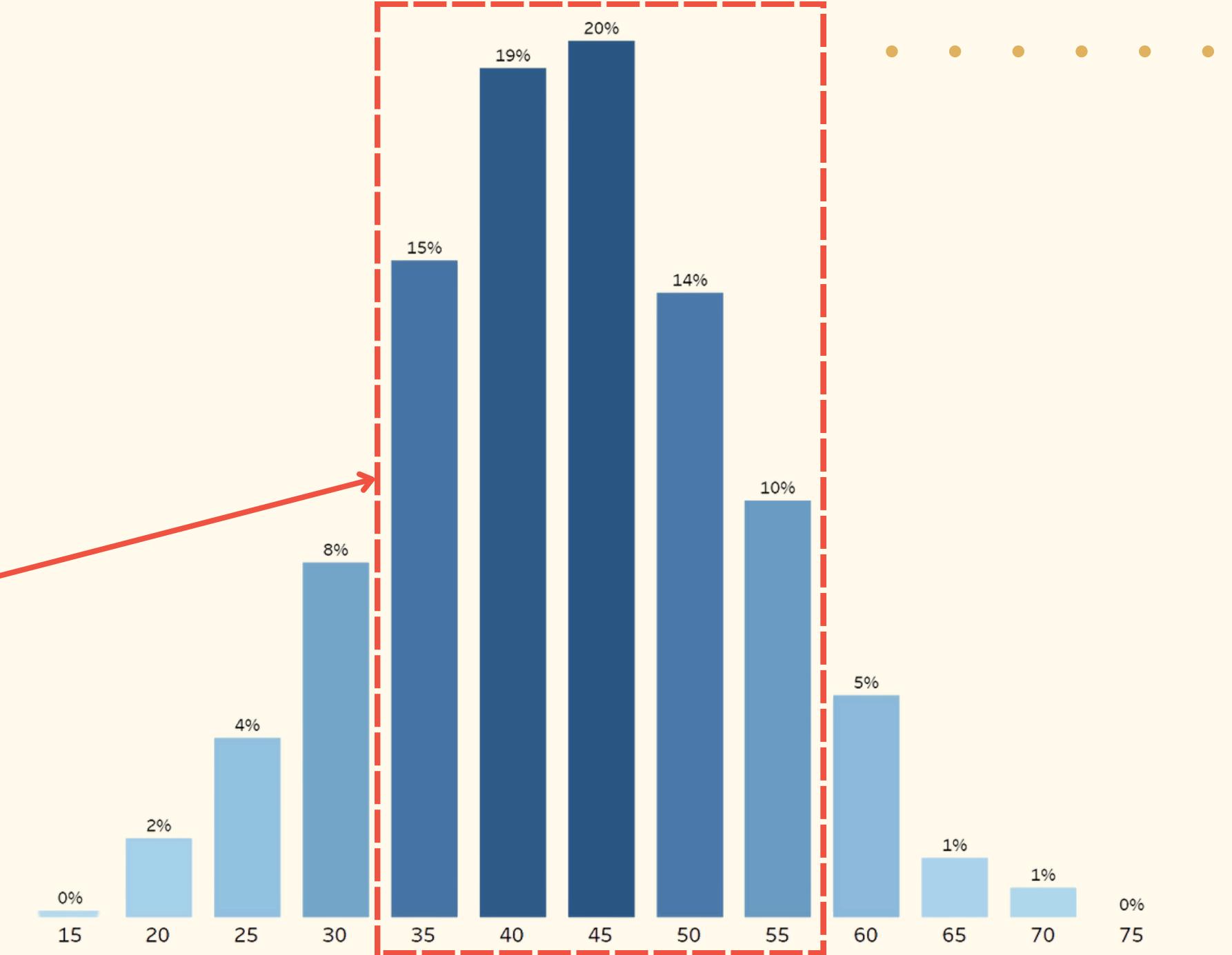
Điều này cho thấy khách hàng là nữ giới thường **dễ rời bỏ** hơn là nam giới.

CUSTOMER CHURN PROFILE



Tỷ lệ rời bỏ của nhóm tuổi trên 56 (**36,75%**) là **cao nhất** với xác suất rời bỏ **1/3**. Tương tự với nó là nhóm tuổi 40 - 55 (**35,7%**)

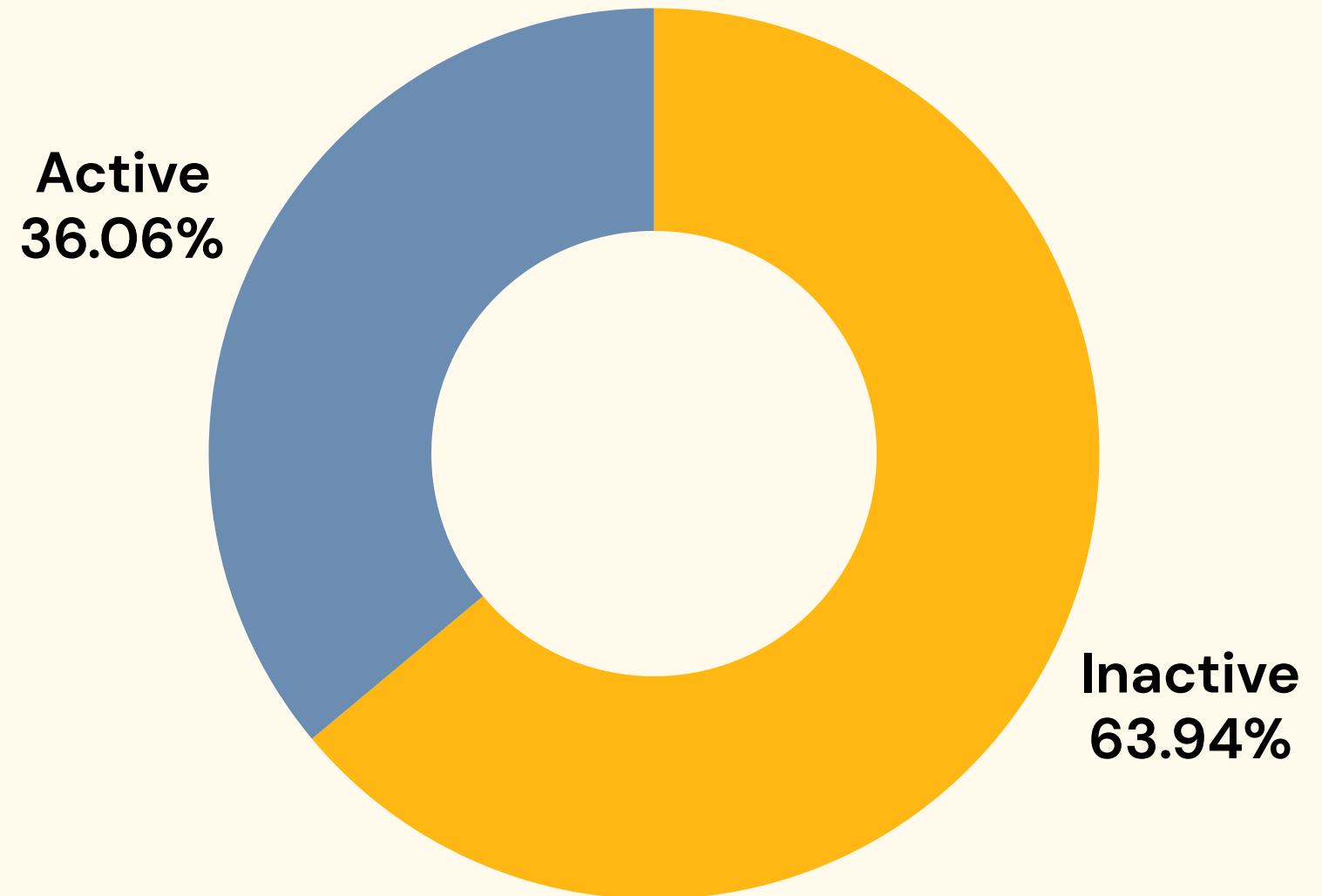
Mặc dù số khách hàng rời bỏ của nhóm tuổi 25 - 39 là **khá lớn** với **557** nhưng xác suất rời bỏ so với tổng khách hàng chỉ là **1/10** và tỷ lệ rời bỏ là **10,07%**



Phần lớn rơi vào độ tuổi **35, 40, 45 và 50** (tuổi trung niên)

CUSTOMER CHURN PROFILE

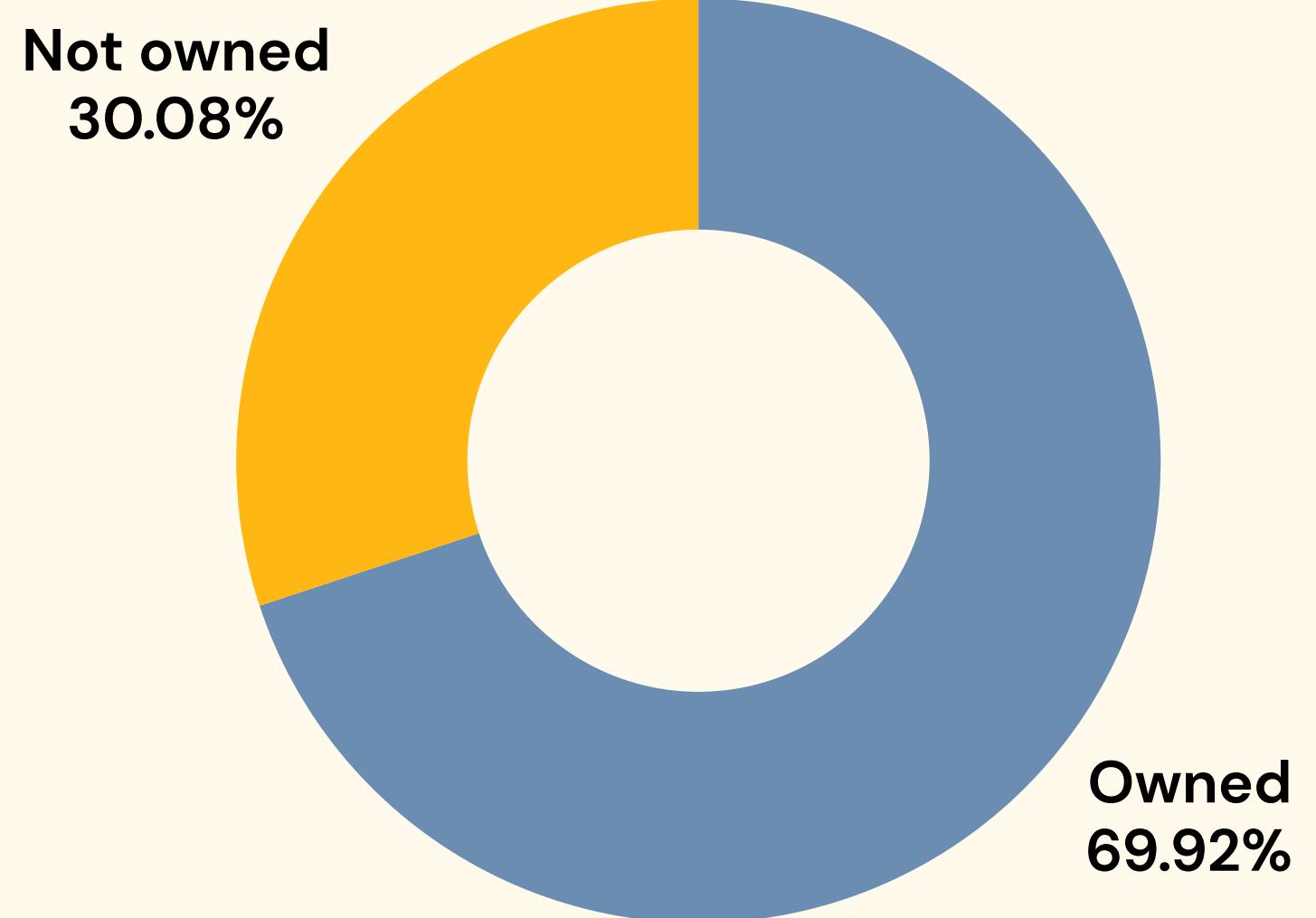
● Inactive ● Active



Trạng thái hoạt động

- Tỷ lệ tài khoản không hoạt động (63.94%)
- Tỷ lệ tài khoản hoạt động (36.06%)

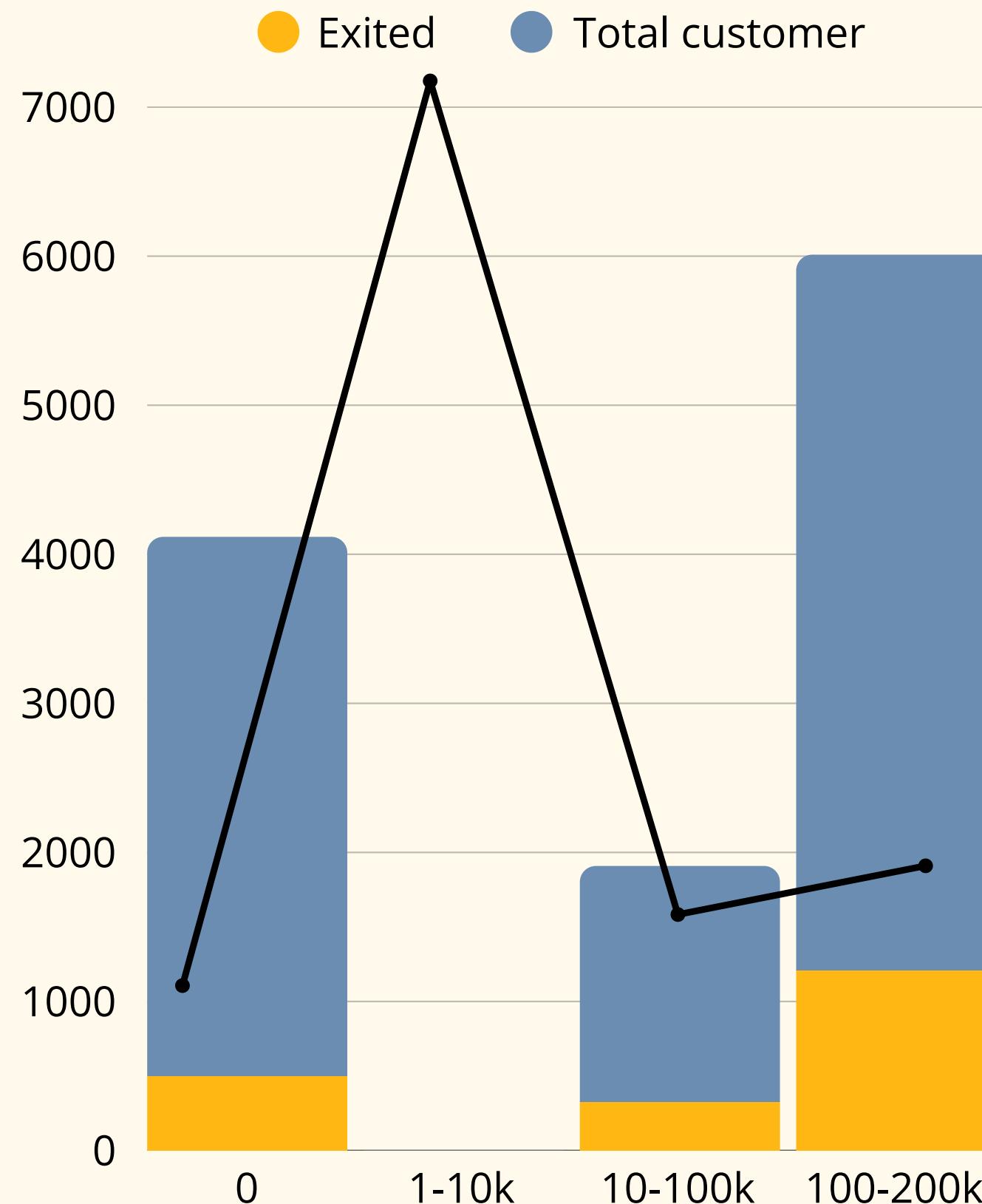
● Owned ● Not owned



Tình trạng sở hữu thẻ tín dụng

- Tỷ lệ khách hàng rời bỏ sở hữu thẻ là 69.92%
- Tỷ lệ khách hàng rời bỏ không sở hữu thẻ là 30.08%

CUSTOMER CHURN BEHAVIOR

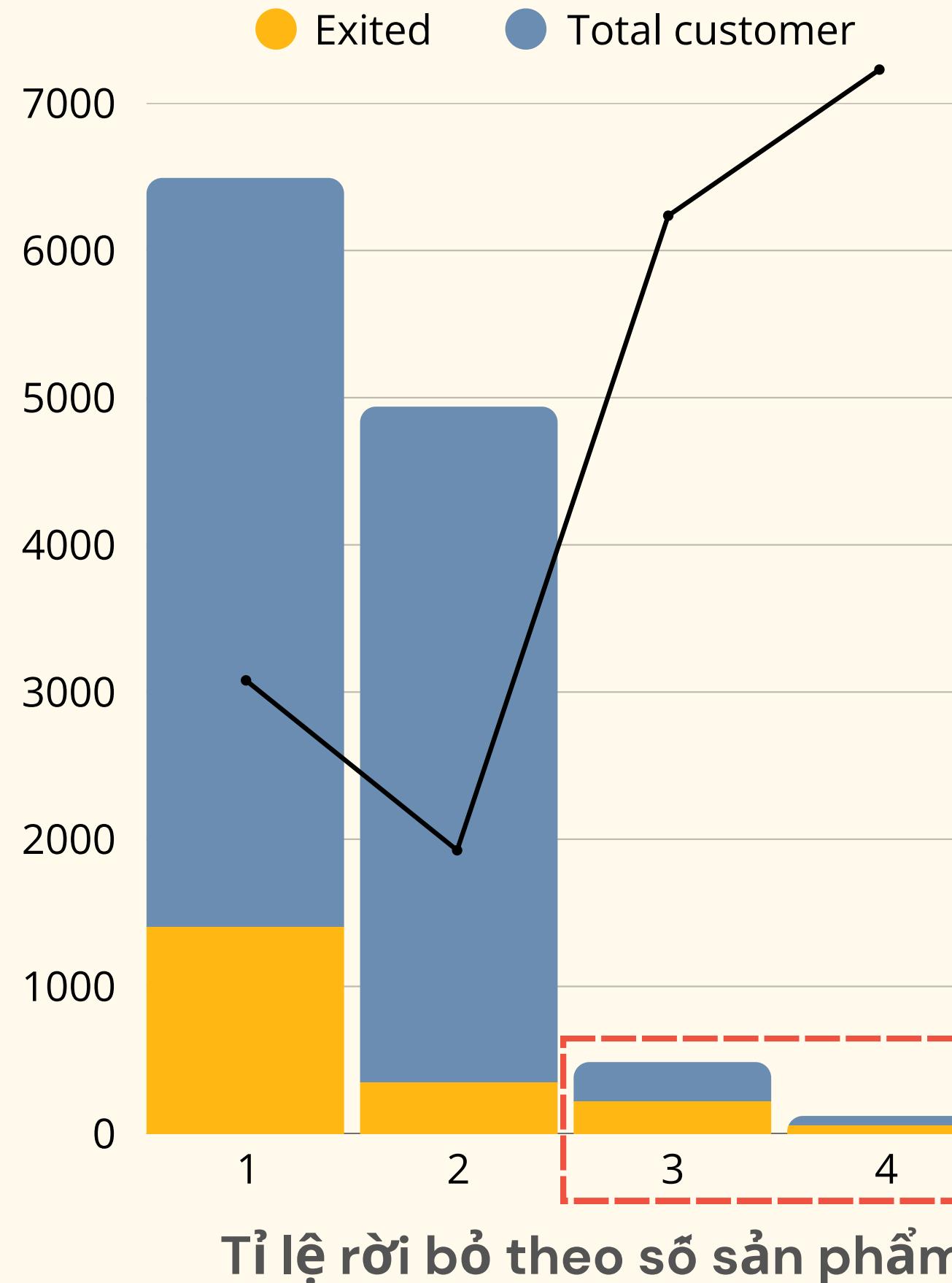


Số dư **có tác động** khá rõ đến tỉ lệ rời bỏ khách hàng. Một khách hàng nằm trong nhóm có số dư từ 1 - 10k đã rời bỏ, đây là một **giá trị ngoại lai**.

Đáng quan tâm nhất là nhóm khách hàng có số dư từ **100 - 200k**. Họ có xu hướng rời bỏ **cao nhất** với **25.23%**. **Giảm dần** nhóm khách hàng **10 - 100k** là **20.59%**, sau đó là khách hàng với số dư **0** chiếm **13.82%**.

TỈ LỆ KHÁCH HÀNG RỜI BỎ THEO SỐ DƯ TÀI KHOẢN

CUSTOMER CHURN BEHAVIOR



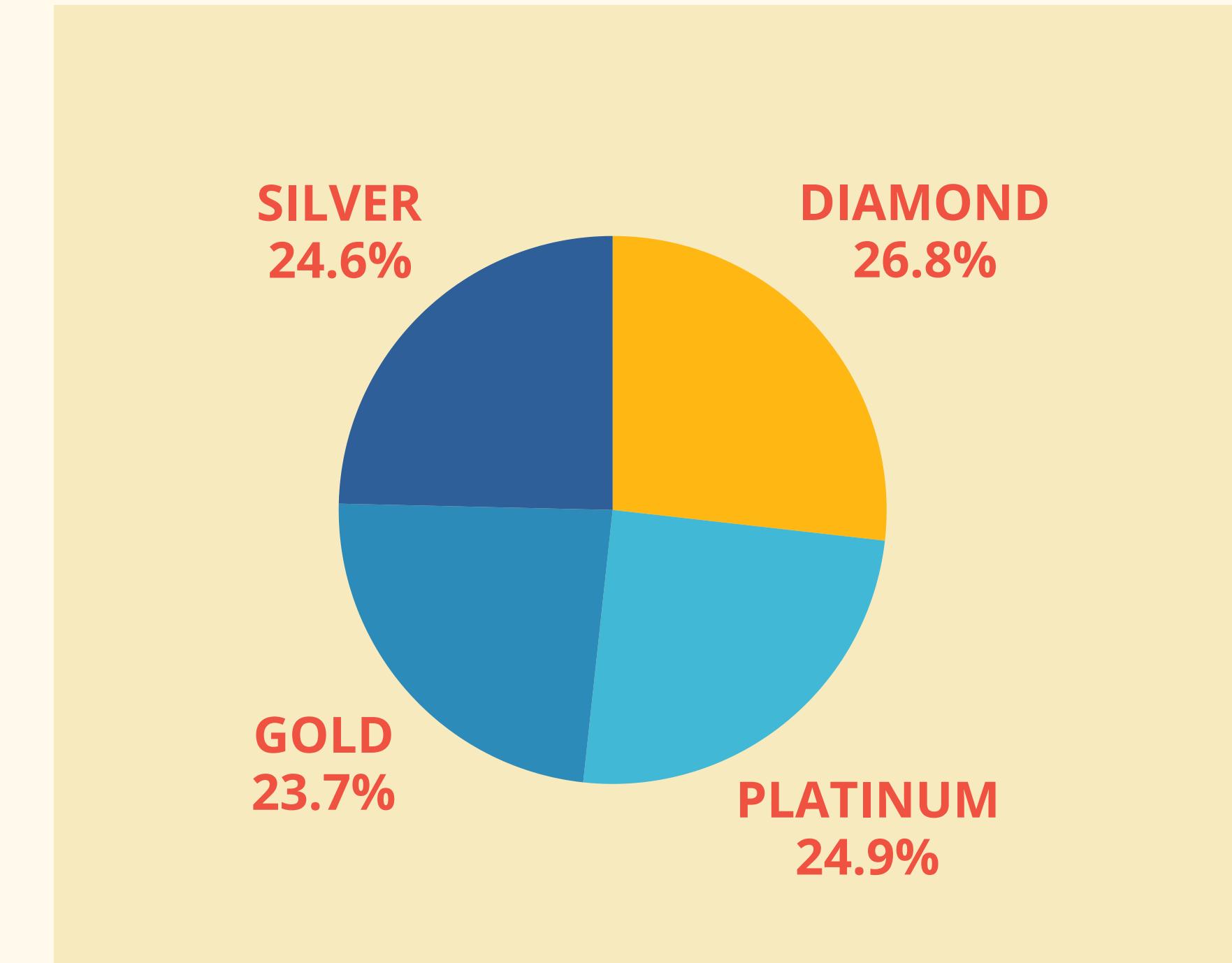
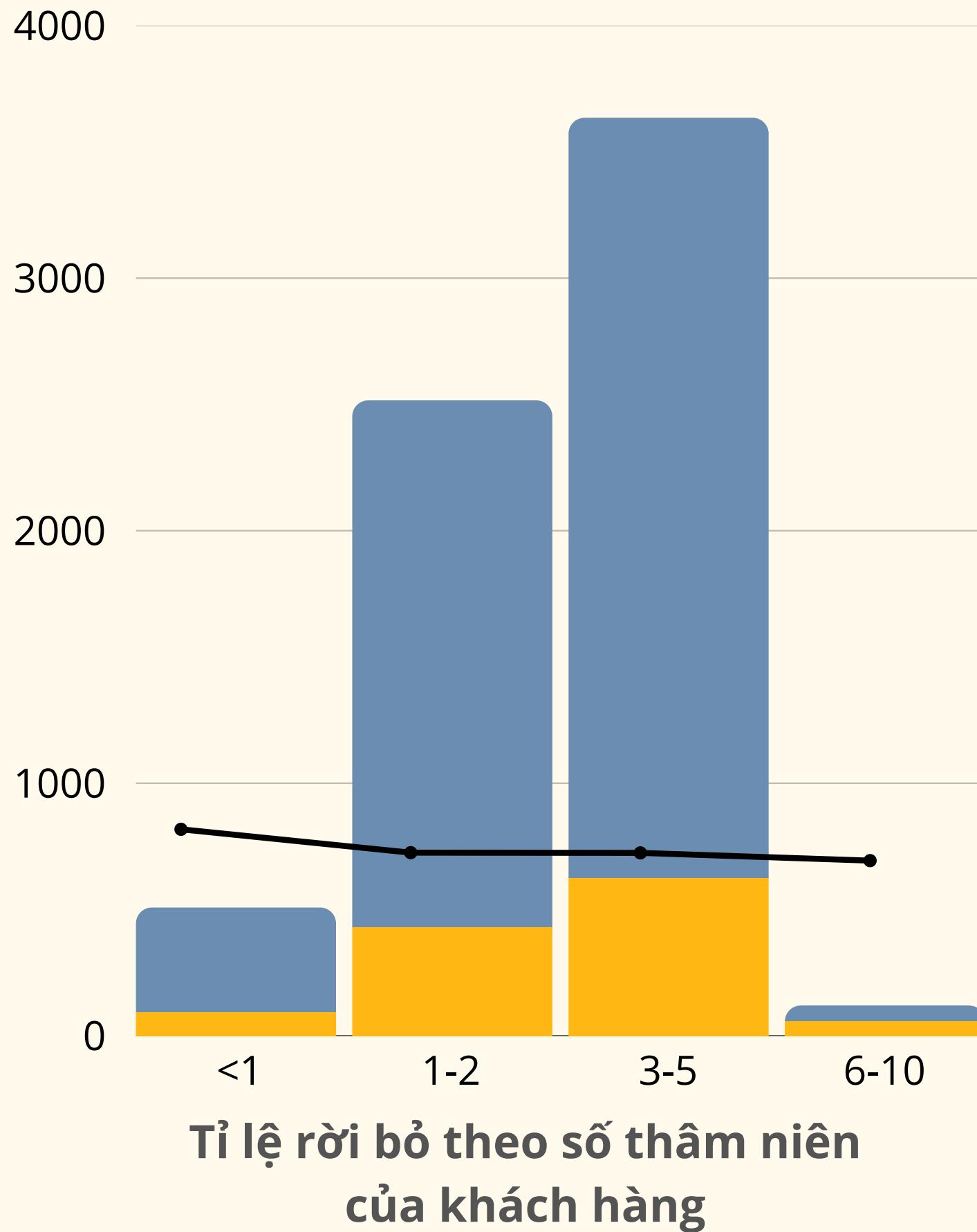
- Hầu hết khách hàng đều **chỉ sử dụng 1 hoặc 2 sản phẩm** trong ngân hàng
 - Số khách hàng chỉ sử dụng 1 sản phẩm chiếm tỷ lệ sử dụng **nhiều nhất**, với tỷ lệ rời bỏ là **27,71%**
 - Tỷ lệ rời bỏ của khách hàng sử dụng 2 sản phẩm dường như là **thấp nhất** với **7,6%**
- Tỷ lệ rời bỏ của nhóm **3 sản phẩm (87,5%)** và **4 sản phẩm (100%)** là cao nhất.

CUSTOMER CHURN BEHAVIOR

		Not Complain	Complain
Churn	Not Churn	4	2,034
	Churn	0	0

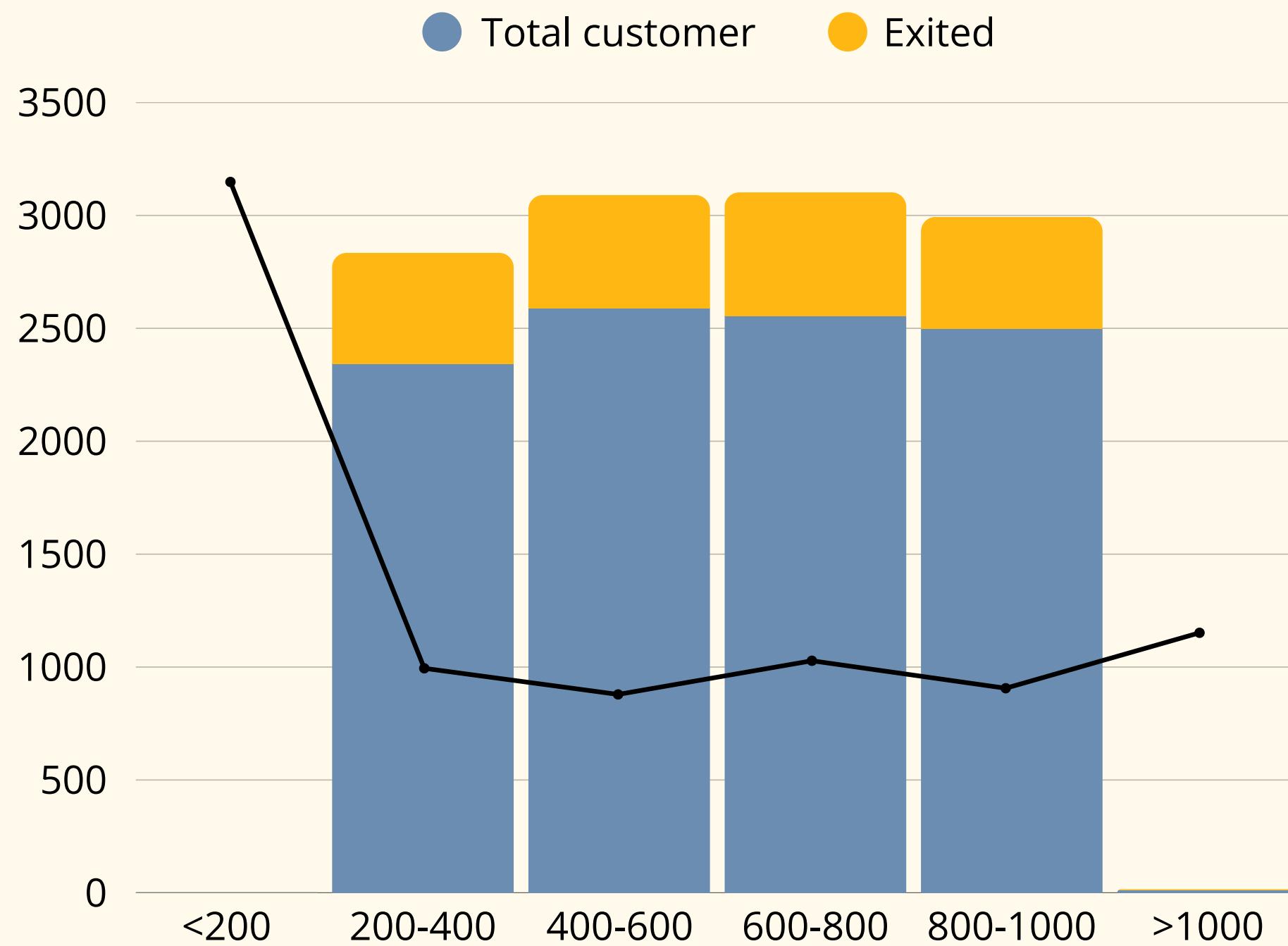
CUSTOMER CHURN BEHAVIOR

Exited Total customer

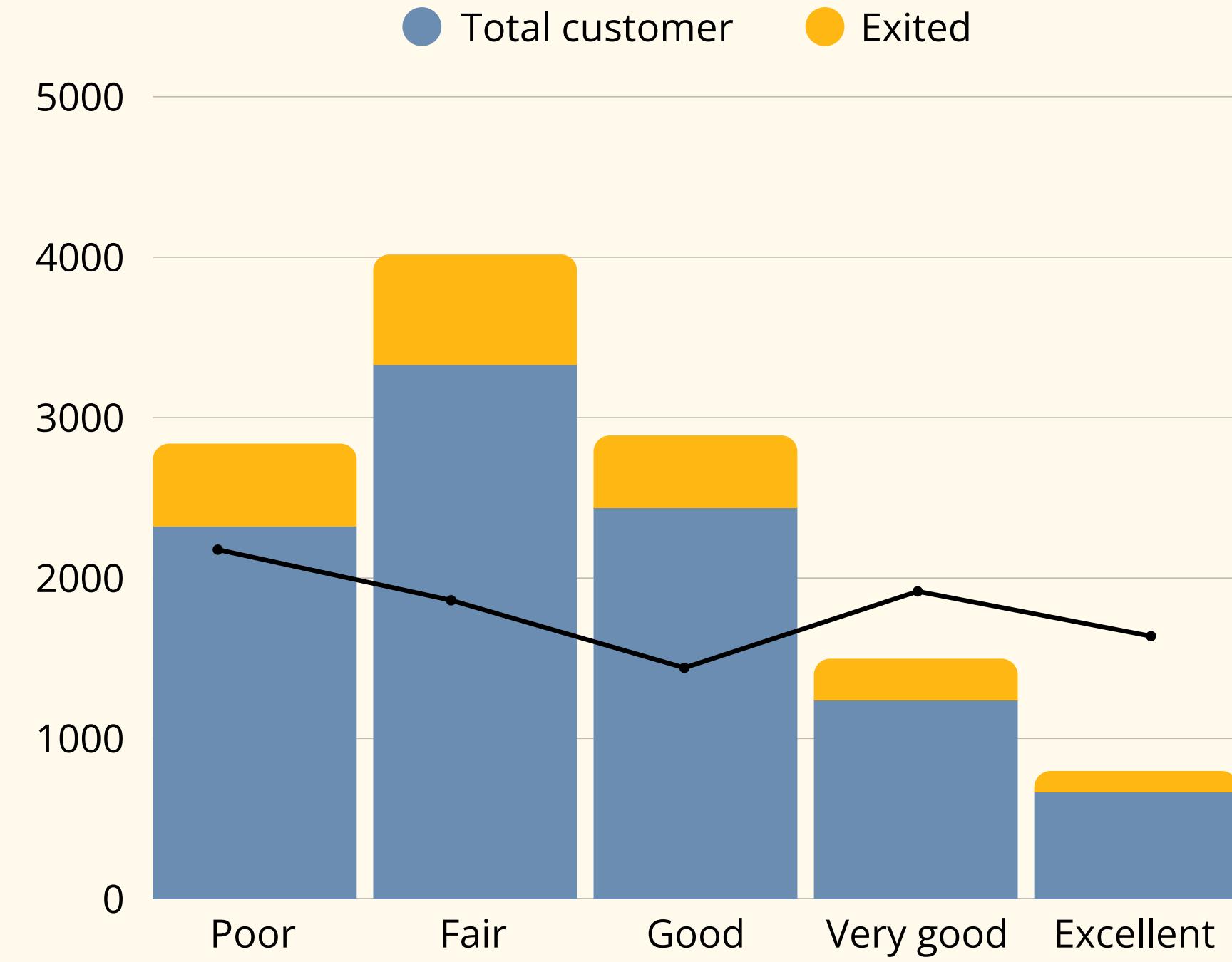


Tỉ lệ rời bỏ của khách hàng
theo cấp độ thẻ

CUSTOMER CHURN BEHAVIOR



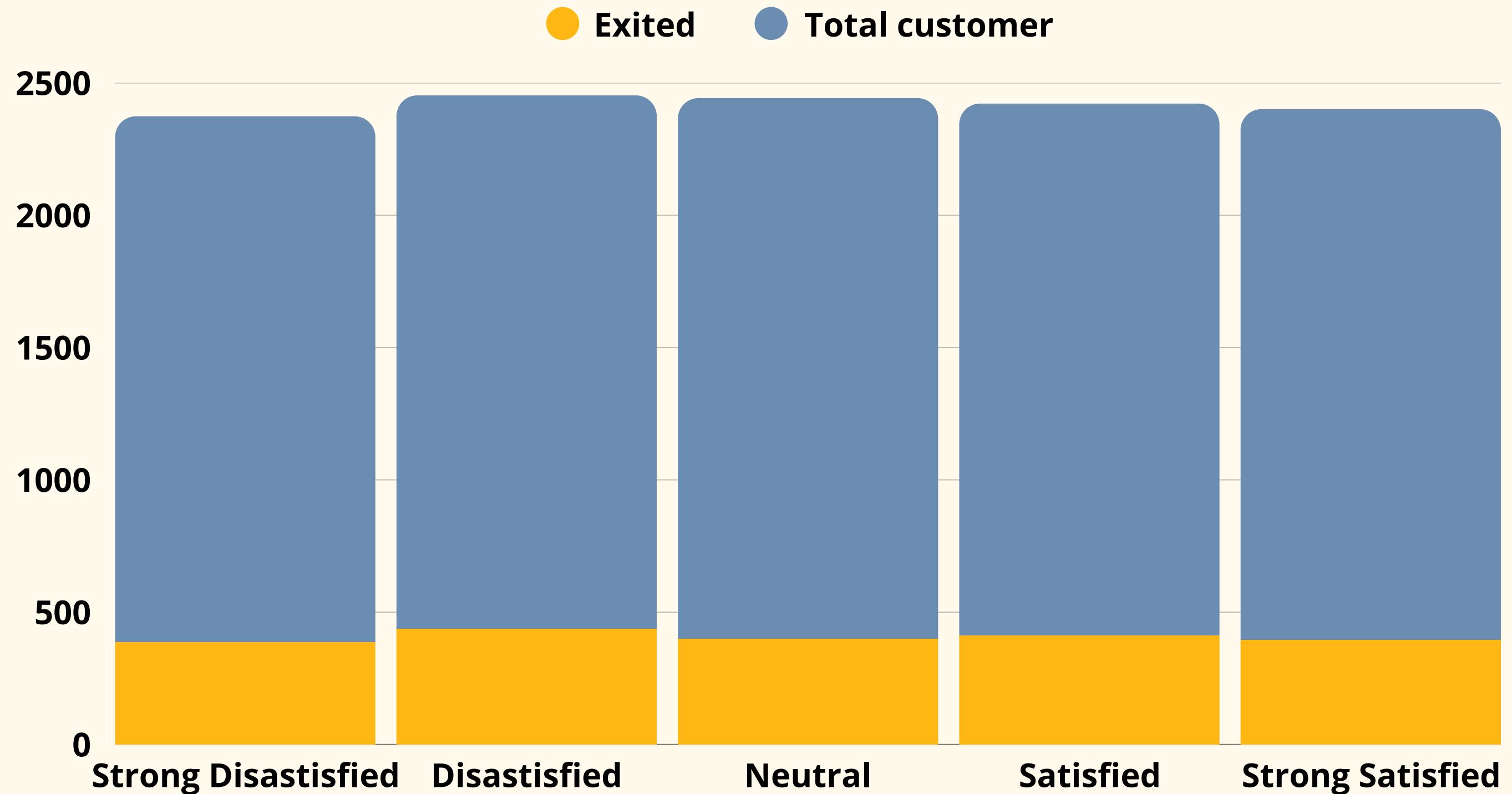
TỈ LỆ KHÁCH HÀNG RỜI BỎ THEO MỨC ĐIỂM NHẬN ĐƯỢC



TỈ LỆ KHÁCH HÀNG RỜI BỎ THEO MỨC ĐIỂM TÍN DỤNG

CUSTOMER CHURN BEHAVIOR

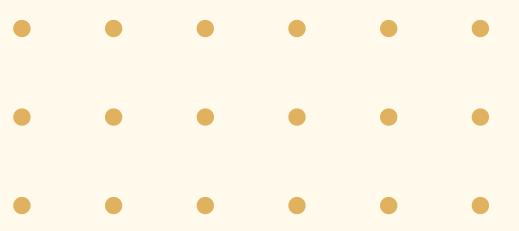
TỈ LỆ RỜI BỎ CỦA
KHÁCH HÀNG THEO
MỨC ĐỘ HÀI LÒNG



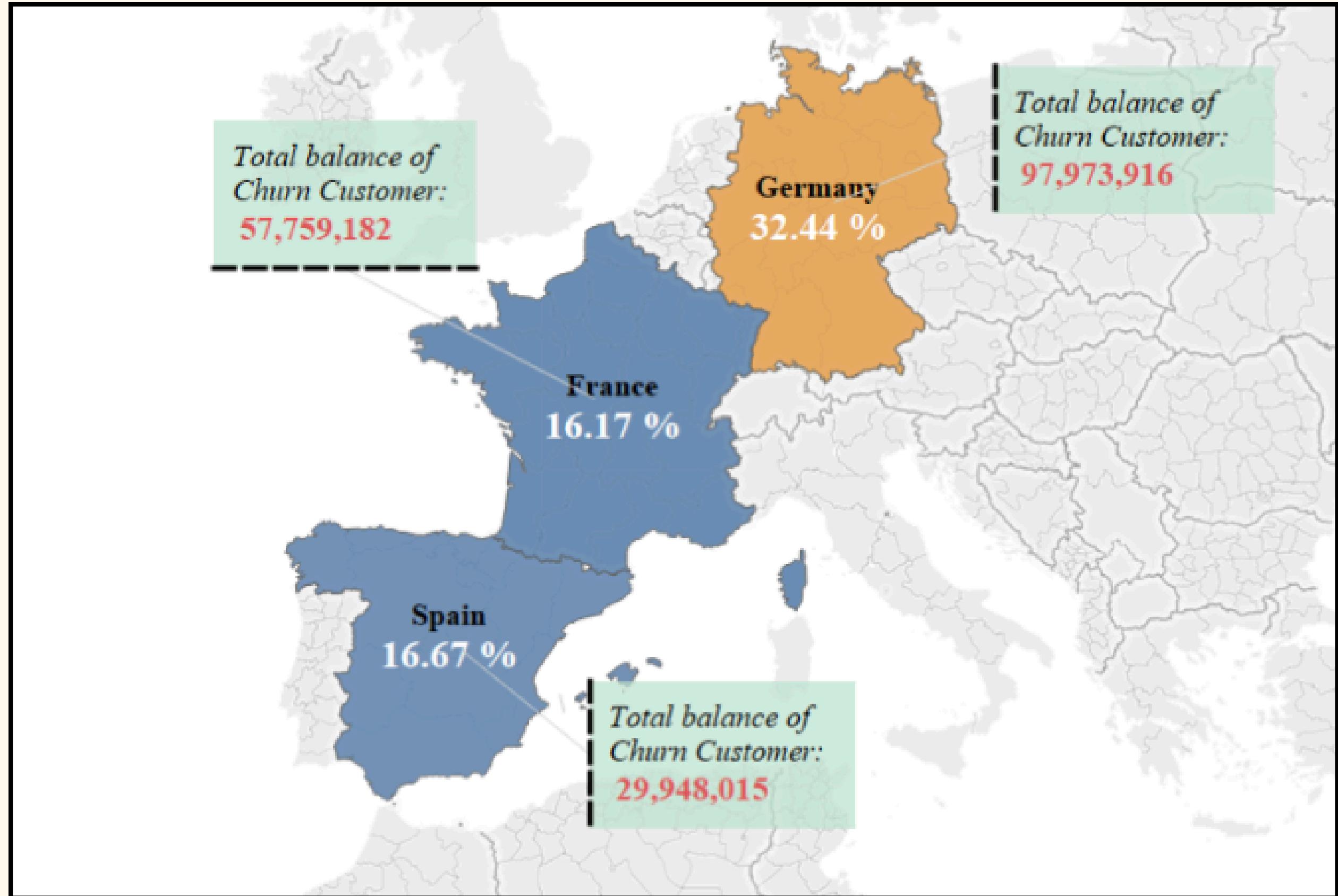
CUSTOMER CHURN BEHAVIOR

Khách hàng rời bỏ có những đặc tính sau đây:

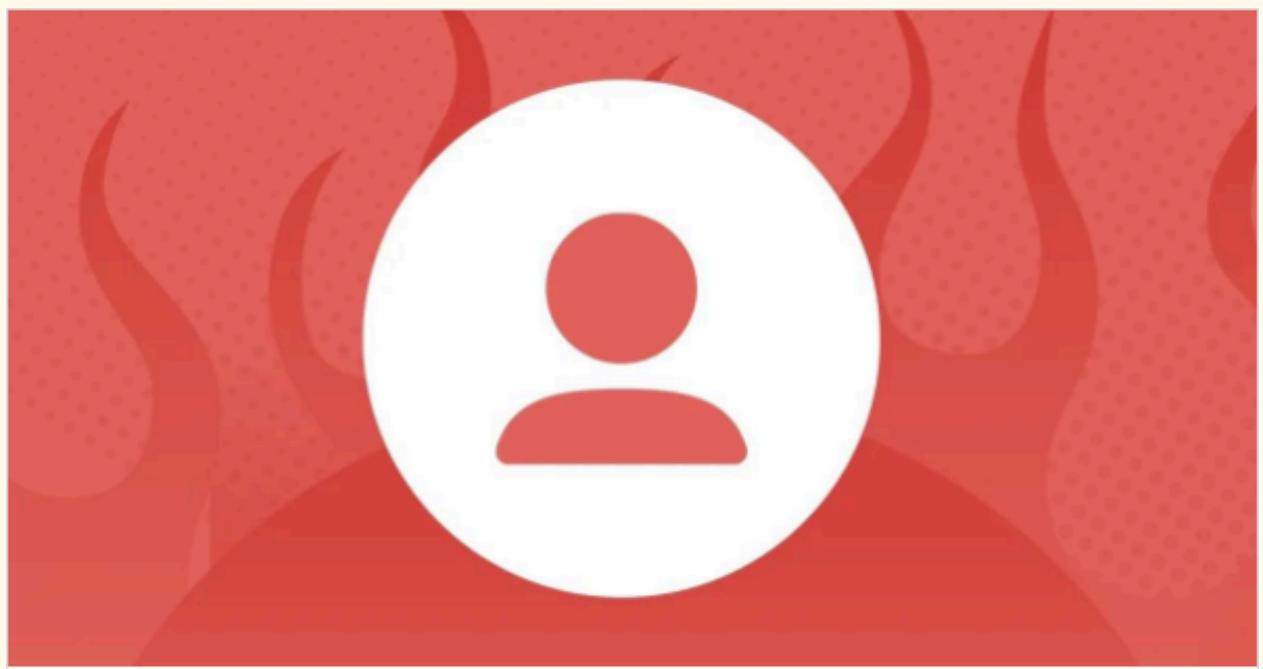
1. Khách hàng ở **Đức** có xu hướng rời bỏ cao hơn đáng kể ở **Pháp** và **Tây Ban Nha**
2. **Nữ giới** có xu hướng rời bỏ nhiều hơn nam giới
3. Độ tuổi của khách hàng rời bỏ từ **40** trở lên
4. Khách hàng đang **không hoạt động** và **có thẻ tín dụng**
5. Khách hàng với số lượng sản phẩm **từ 3 - 4**
6. Khách hàng với **thâm niên dưới 1 năm** dễ rời bỏ hơn các nhóm còn lại



GERMANY CUSTOMER CHURN BEHAVIOR



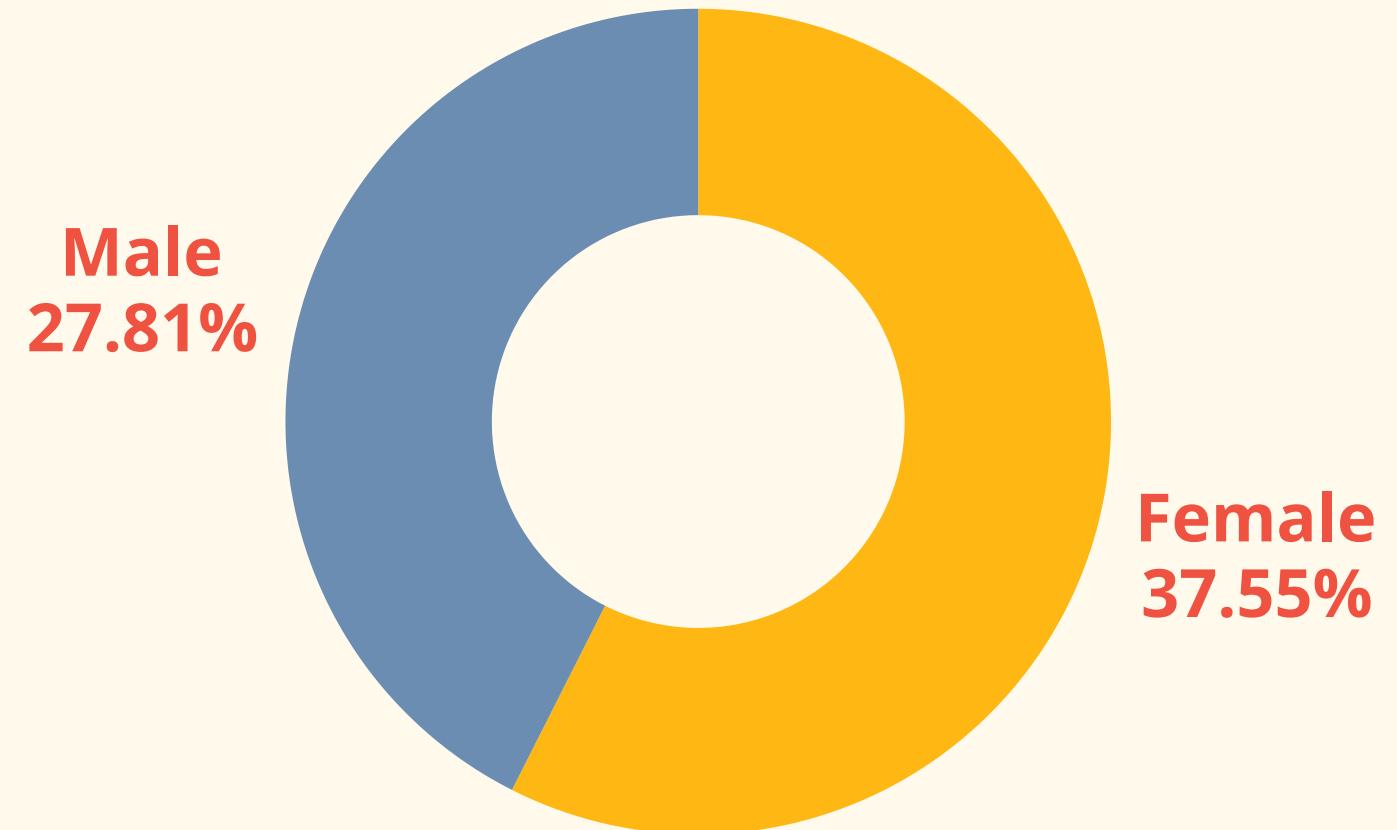
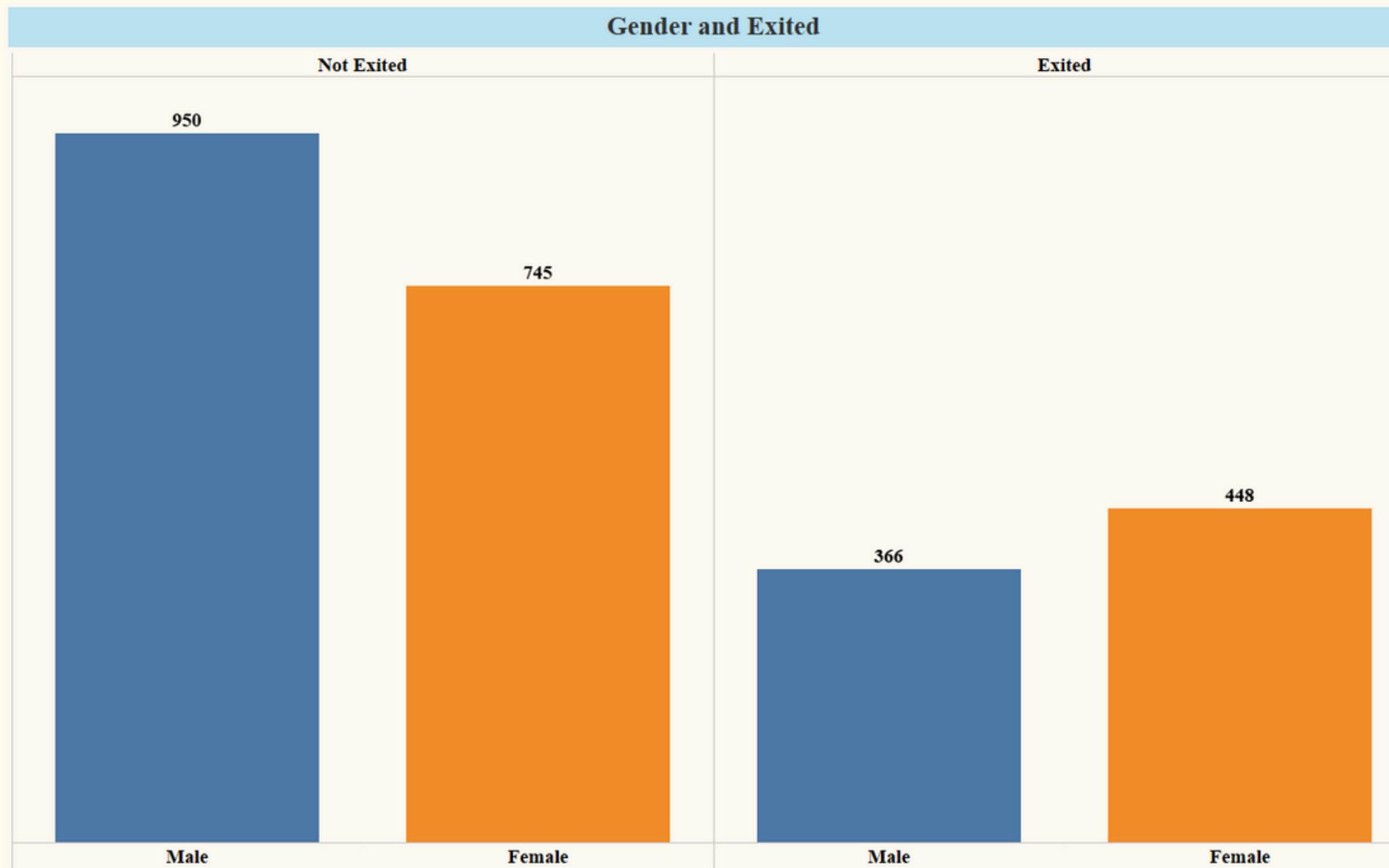
Năm 2022, theo báo cáo ngành
hàng có **tỷ lệ rời bỏ trung bình là 19%**



CG CustomerGauge

Avg Balance per Churn Customer	
France	71,220
Germany	120,361
Spain	72,513

GERMANY CUSTOMER CHURN BEHAVIOR

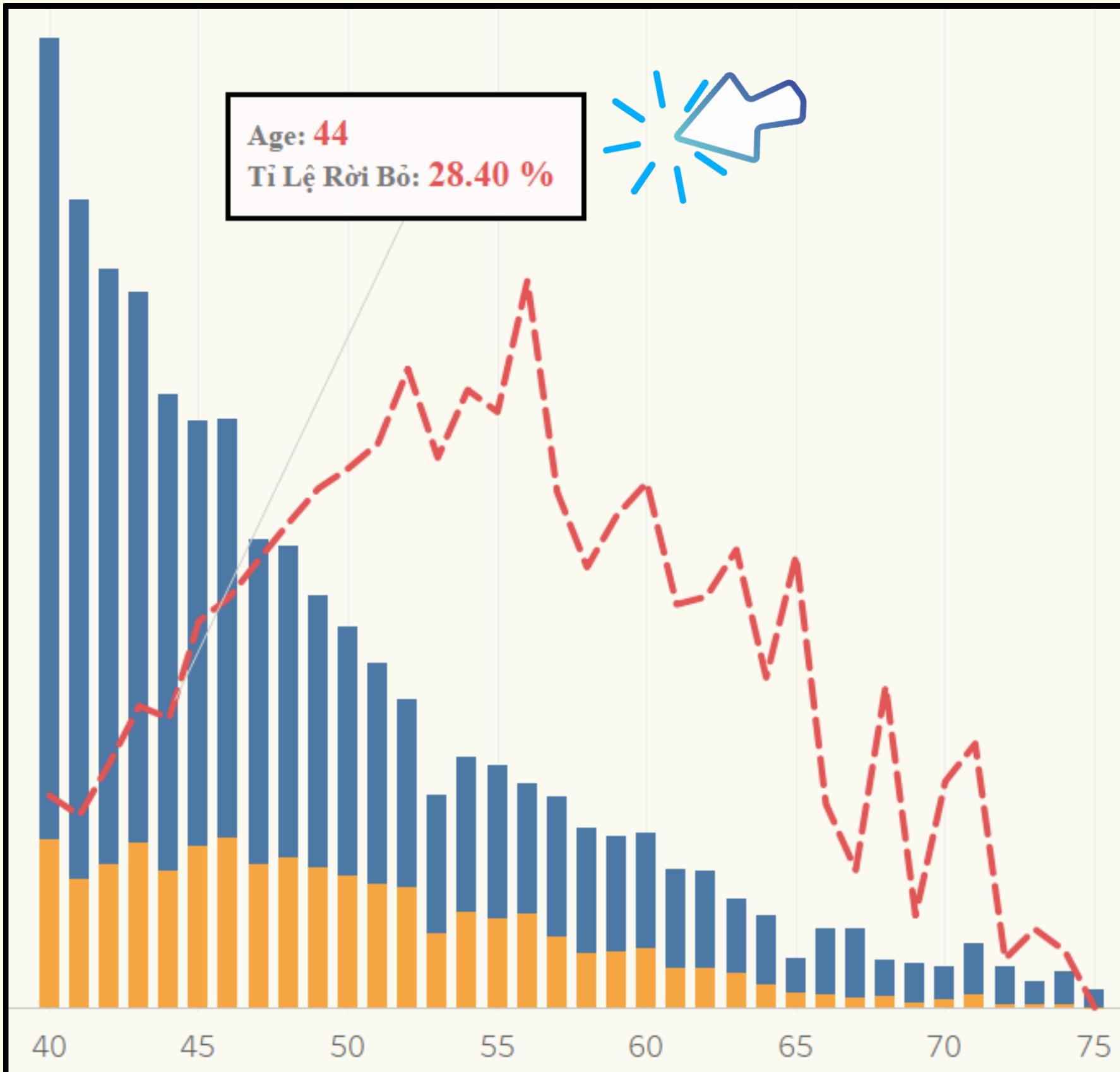


Có đến **448** khách hàng nữ đã rời bỏ trong **1,193** tổng số khách hàng nữ, chiếm tỉ lệ lên đến **37.55%** (**một con số rất cao**).
Tỉ lệ rời bỏ của nữ giới cao hơn **9.74%** so với nam (**37.55%**).

Tỉ lệ rời bỏ theo giới tính

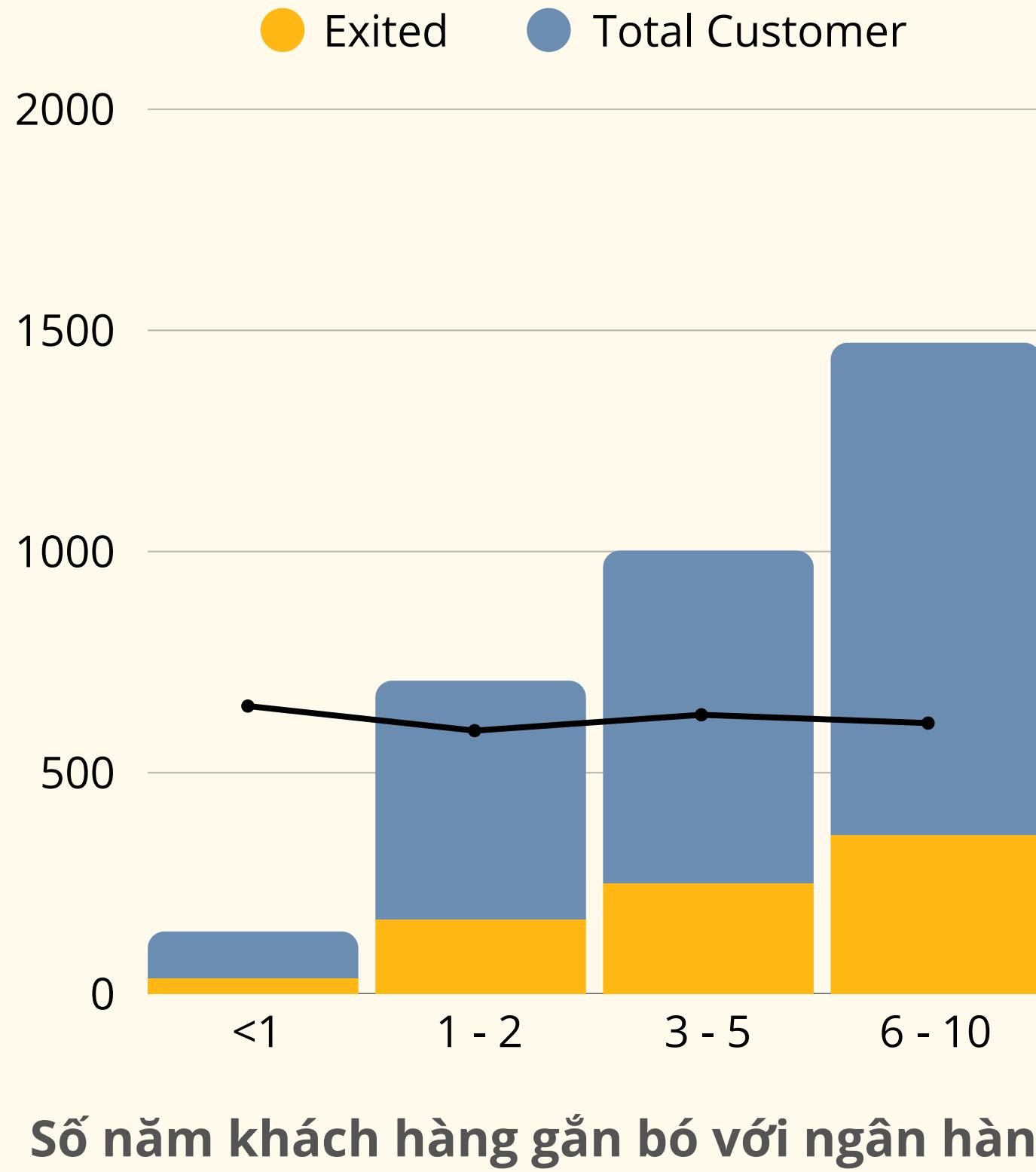
➡ Điều này cho thấy khách hàng là nữ giới thường **dễ rời bỏ** hơn là nam giới.

GERMANY CUSTOMER CHURN BEHAVIOR



- Phân tích sâu vào nhóm tuổi từ 40 trở lên, tỉ lệ rời bỏ khách hàng ở Đức bắt đầu cao ở **độ tuổi 44 (28,4%)**
- Sau đó có sự biến động về tỉ lệ rời bỏ ở các độ tuổi lớn hơn nhưng nhìn chung tỉ lệ rời bỏ ở nhóm tuổi từ 40 trở lên vẫn rất lớn

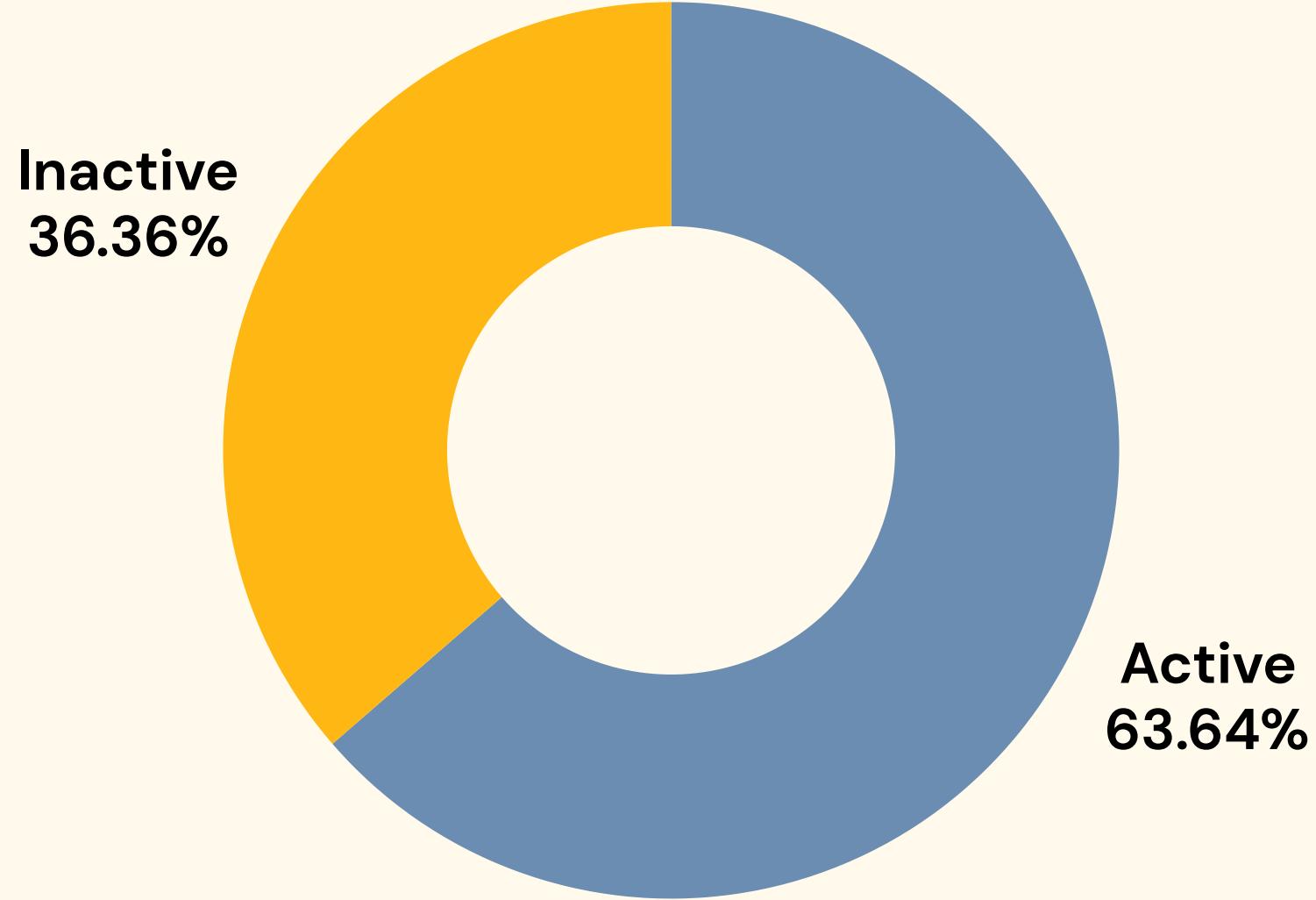
GERMANY CUSTOMER CHURN BEHAVIOR



- Tỷ lệ rời bỏ của khách hàng trong các năm **không có sự khác biệt** quá rõ rệt
- Tỷ lệ khách hàng rời bỏ:
 - Dưới 1 năm: 34.28%
 - 1-2 năm: 31.35%
 - 3-5 năm: 33.24%
 - 6-10 năm: 32.25%

GERMANY CUSTOMER CHURN BEHAVIOR

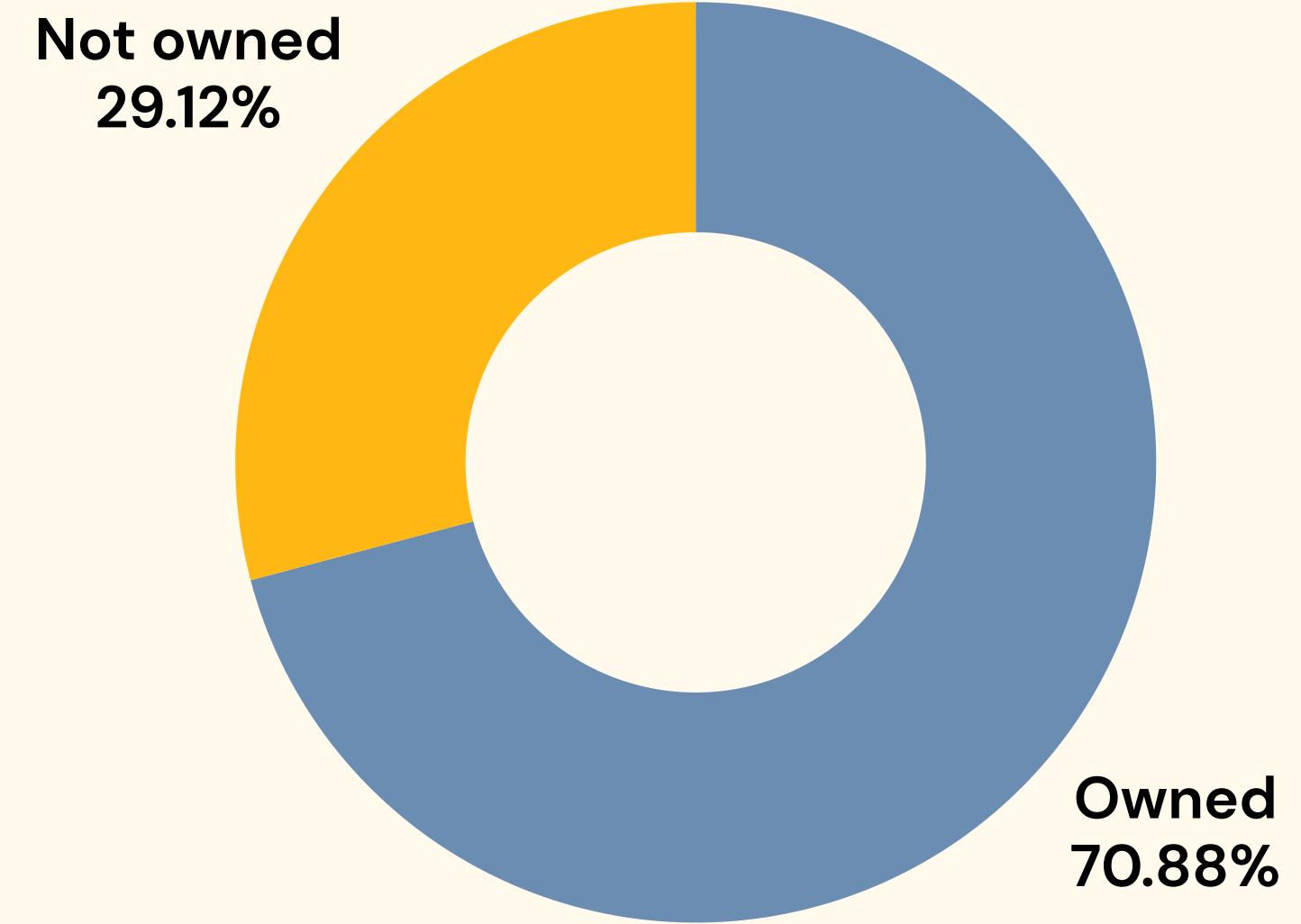
● Active ● Inactive



Trạng thái hoạt động

- Tỷ lệ tài khoản không hoạt động (36.36%)
- Tỷ lệ tài khoản hoạt động (63.64%)

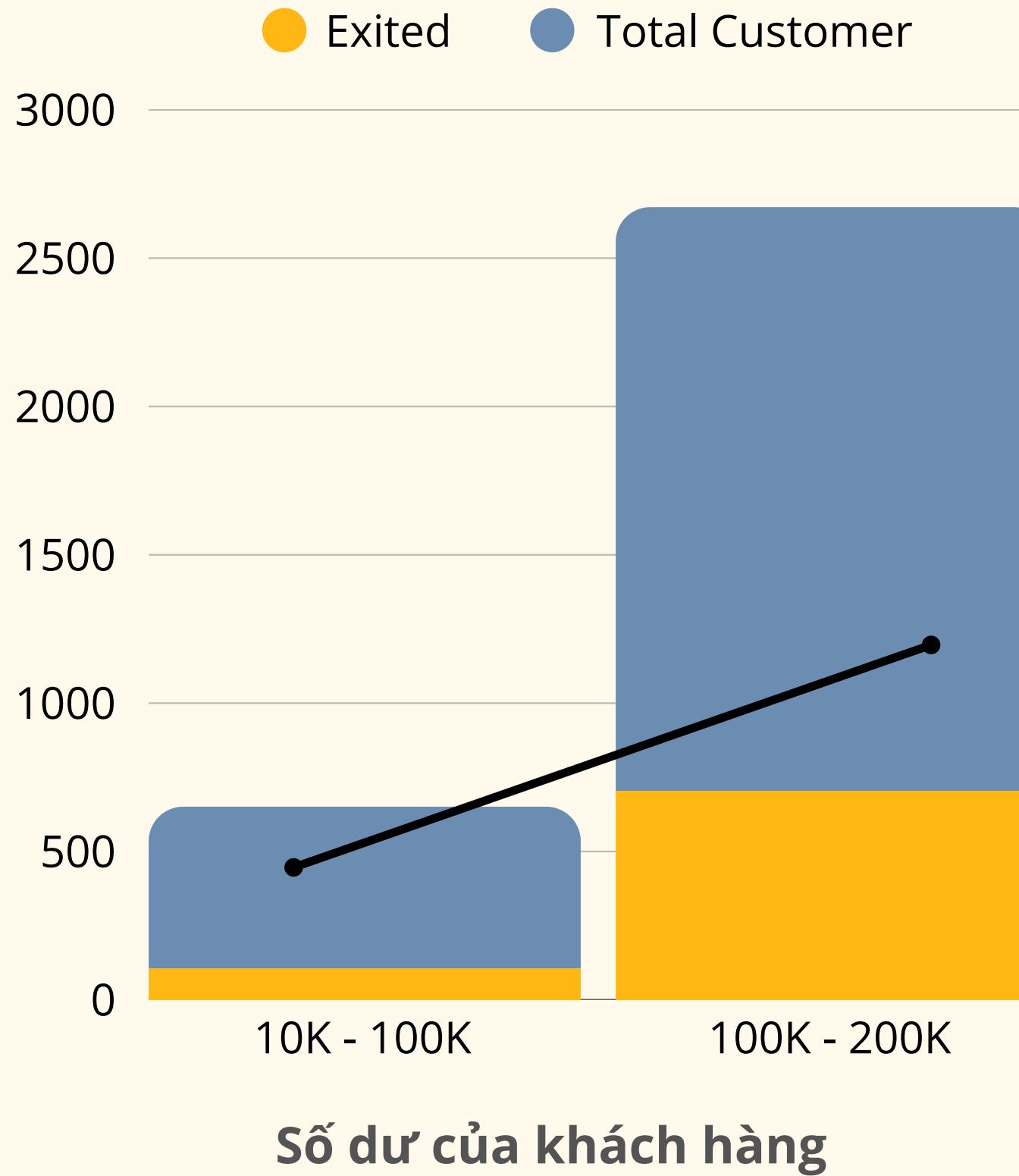
● Owned ● Not owned



Tình trạng sở hữu thẻ tín dụng

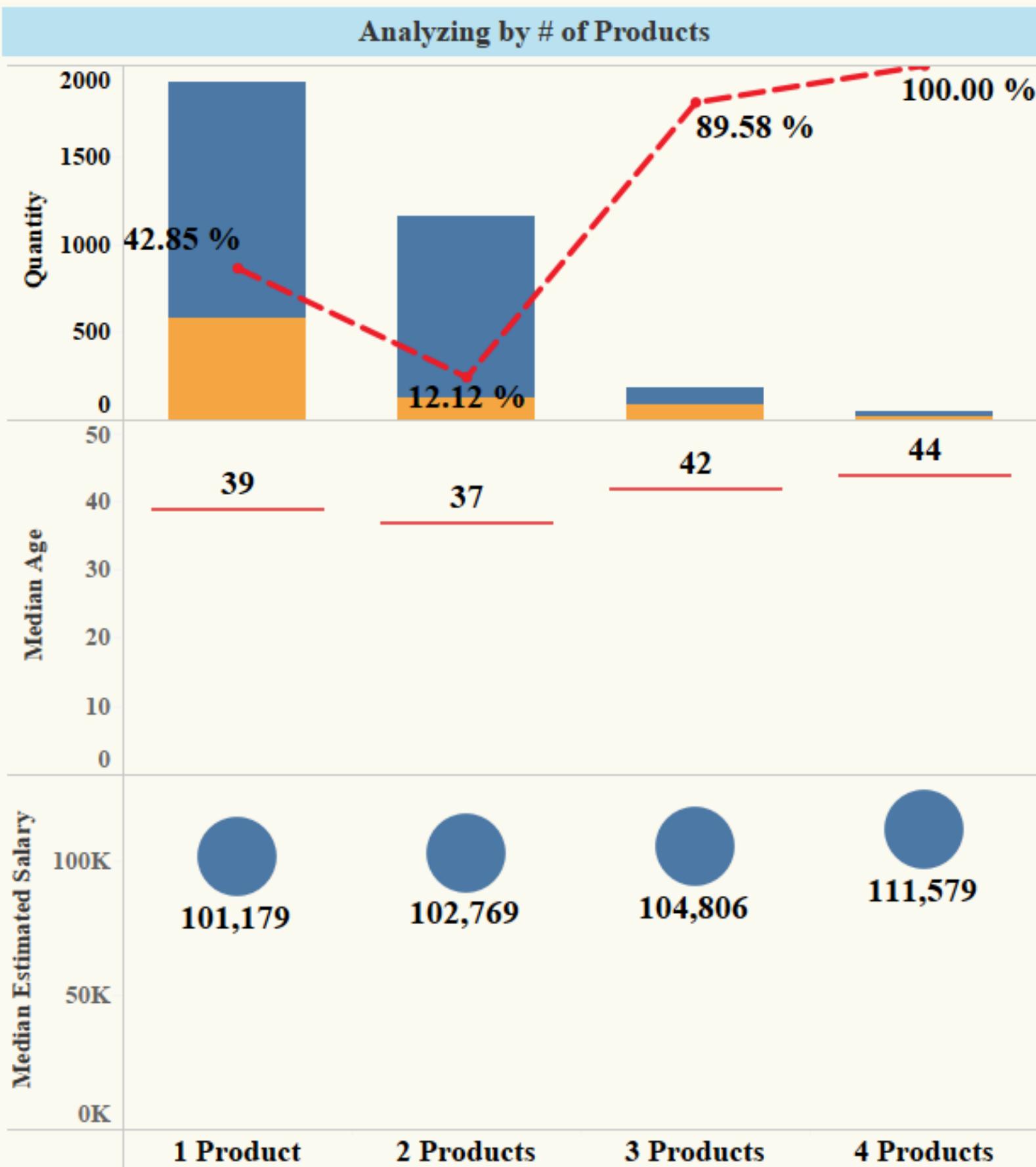
- Tỷ lệ khách hàng rời bỏ sở hữu thẻ là 70.88%
- Tỷ lệ khách hàng rời bỏ không sở hữu thẻ là 29.12%

GERMANY CUSTOMER CHURN BEHAVIOR



- Số dư chỉ còn lại 2 nhóm
- Số dư từ 100,000 đến 200,000 có tỷ lệ rời bỏ là **35.91%**
- Số dư từ 10,000 đến 100,000 tỷ lệ rời bỏ là **19.89%**

GERMANY CUSTOMER CHURN BEHAVIOR



- Những người có độ tuổi **từ 40 trở lên**
- Có trình độ giáo dục cao với **mức lương cao hơn trung bình**. Ngoài ra tỉ lệ rời bỏ cao nhất xảy ra trong suốt năm đầu gắn bó với ngân hàng.

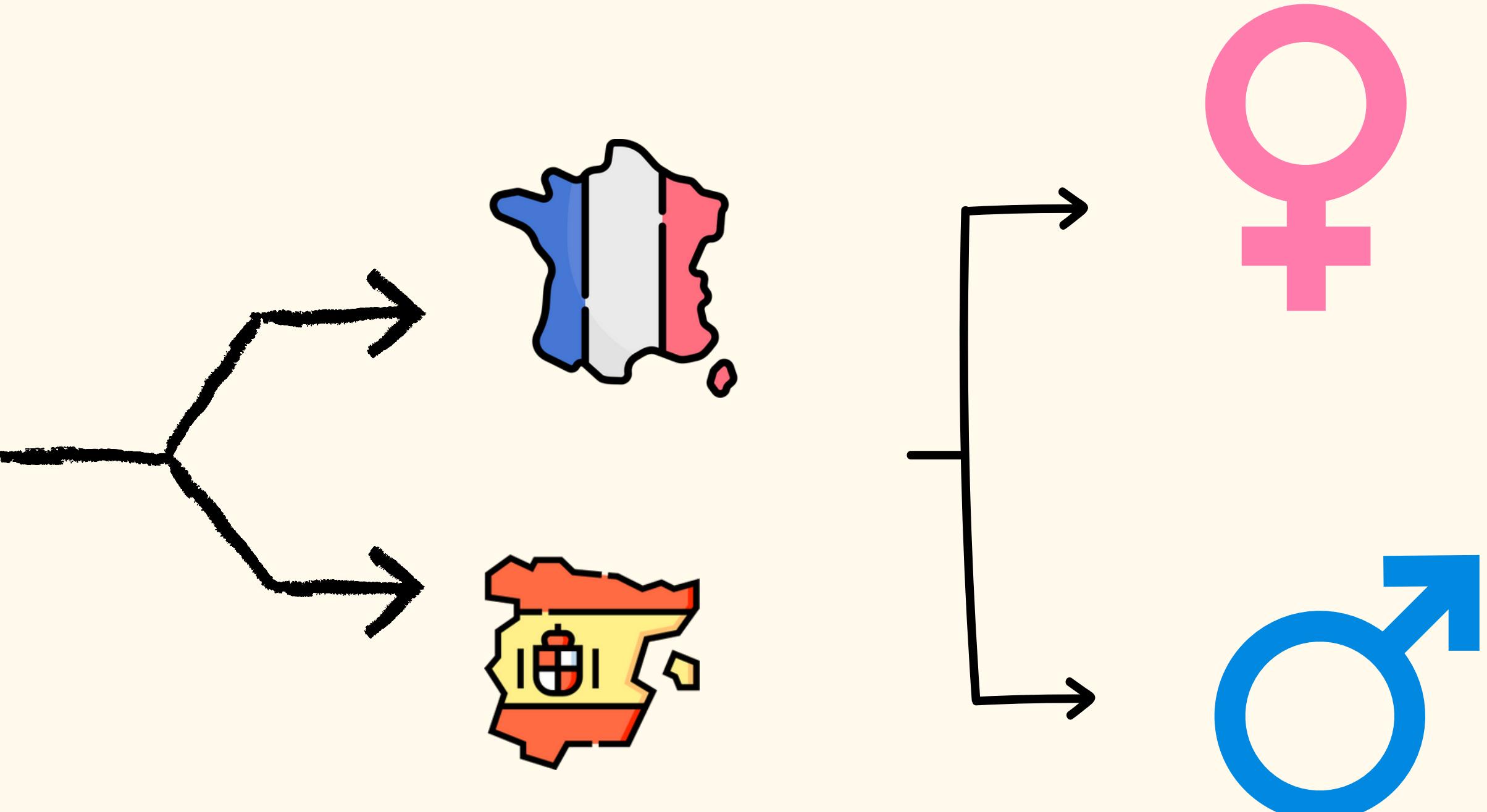
CUSTOMER CHURN BEHAVIOR

Khách hàng rời bỏ có những đặc tính sau đây:

1. Khách hàng **không hoạt động**
2. Khách hàng **sở hữu thẻ tín dụng**
3. Độ tuổi từ **44** trở lên
4. Khách hàng có số dư từ **10,000 đến 200,000**
5. Khách hàng với **số lượng** sản phẩm **từ 3 - 4**
6. Khách hàng với **thâm niên dưới 1 năm** dễ rời bỏ hơn các nhóm còn lại



Mức lương
80.000 - 150.000





METRICS



BANK CUSTOMER CHURN ANALYSIS



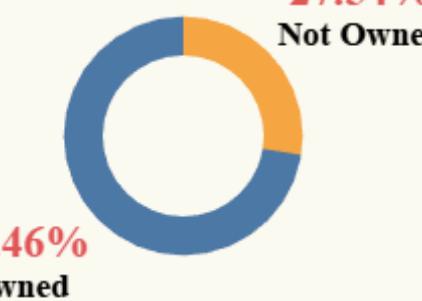
Estimated Salary

80000 150000

D

820
Customers167
Churn
Customers20.37 %
Churn Rate99,564
Avg Salary60,323
Avg Balance

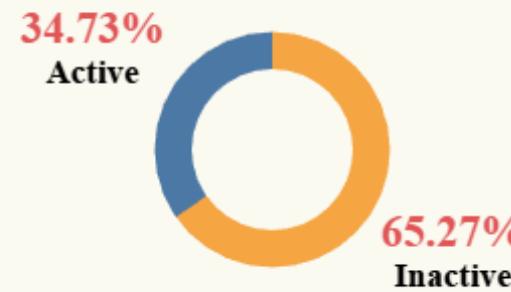
Churn By Credit Card Status



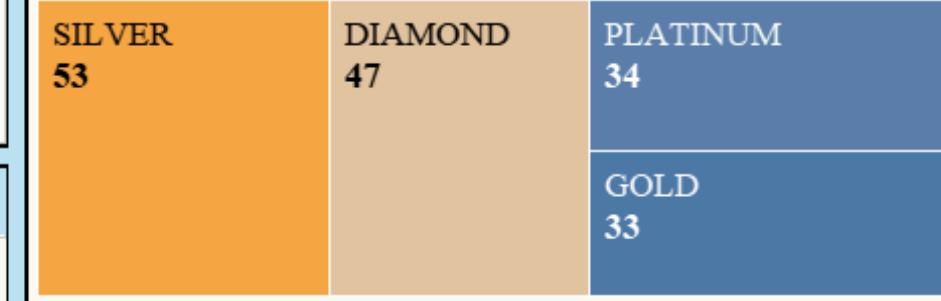
Churn By Complain



Churn By Active



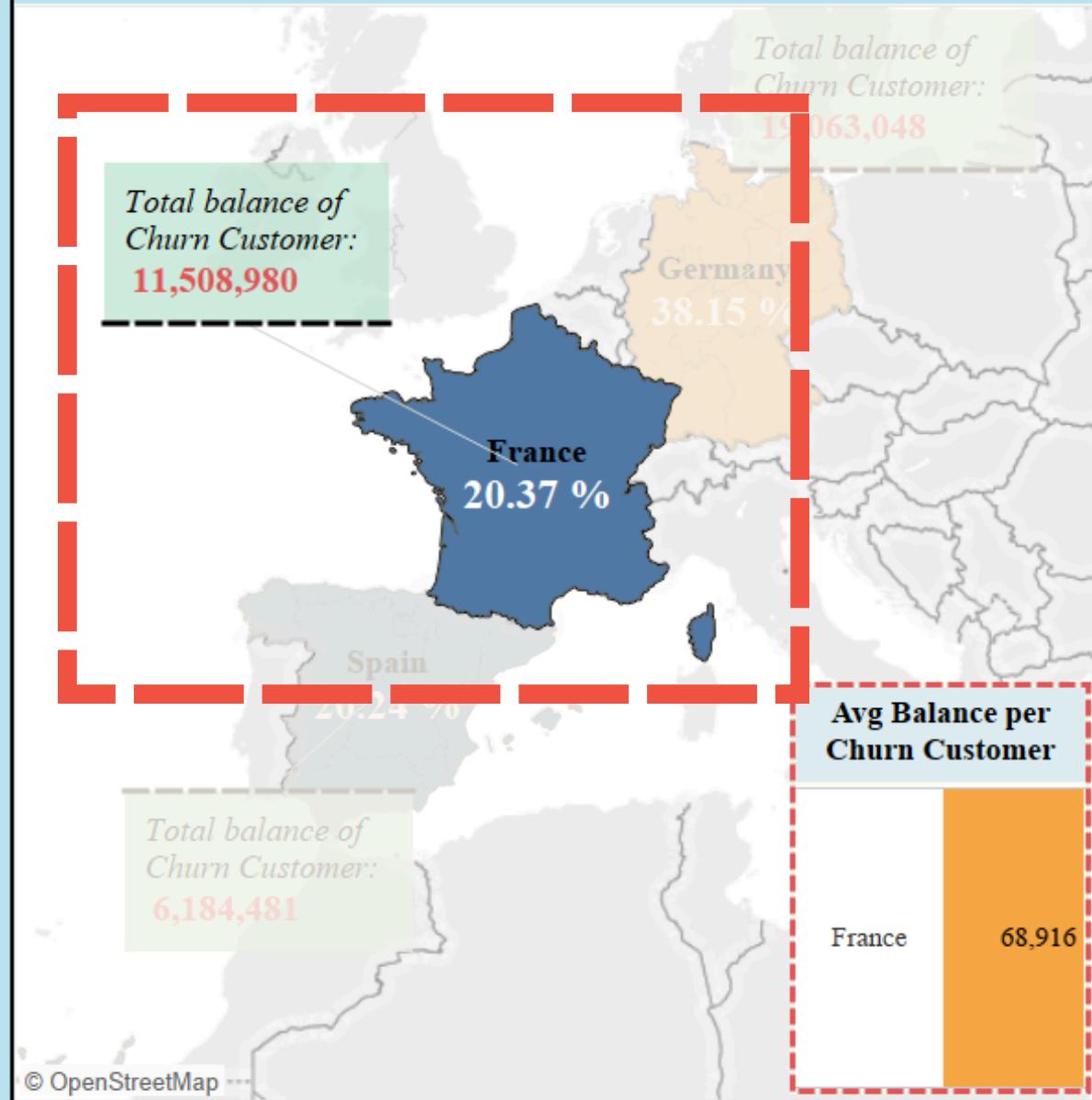
Churn By Card Type



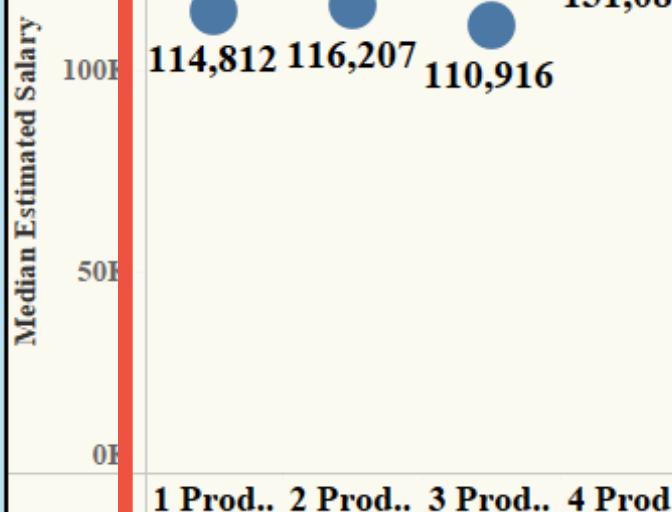
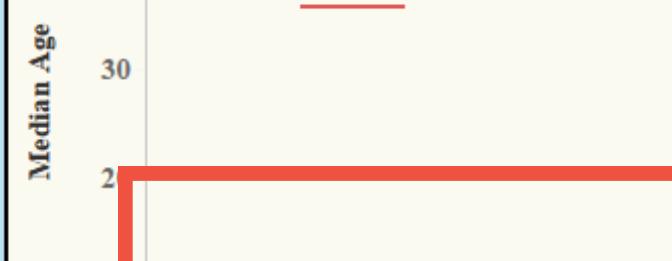
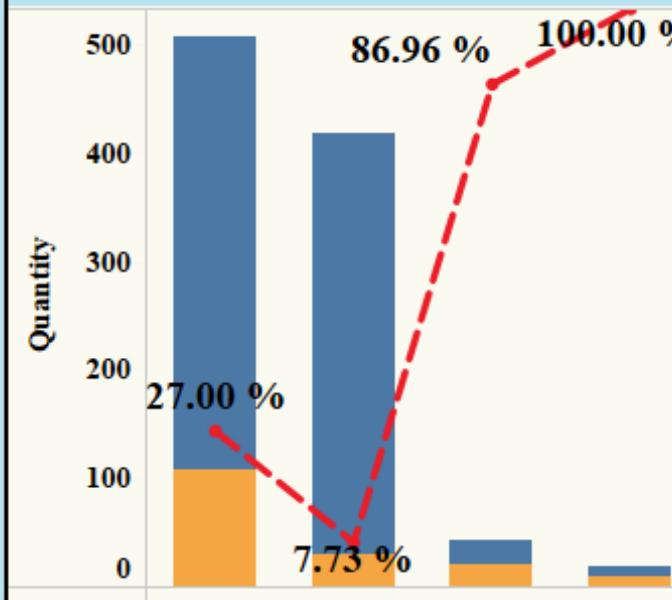
Churn By Satisfaction



Churn By Country



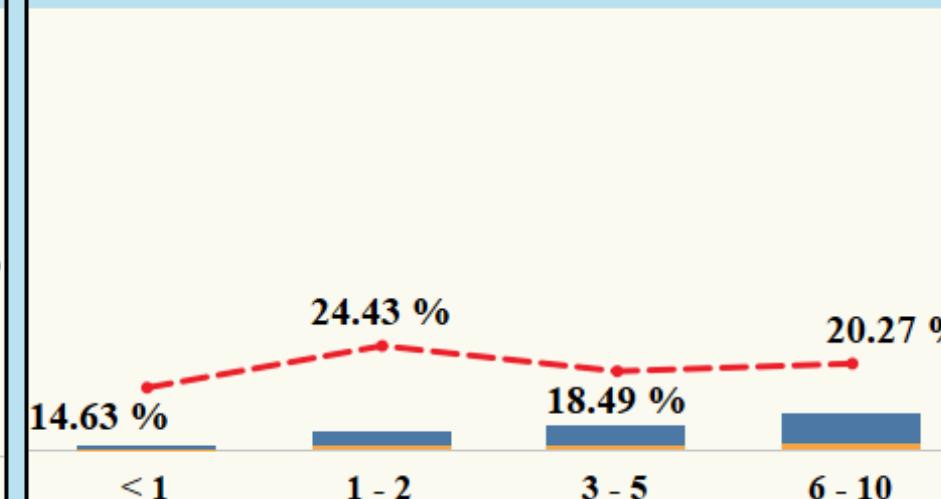
Analyzing by # of Products



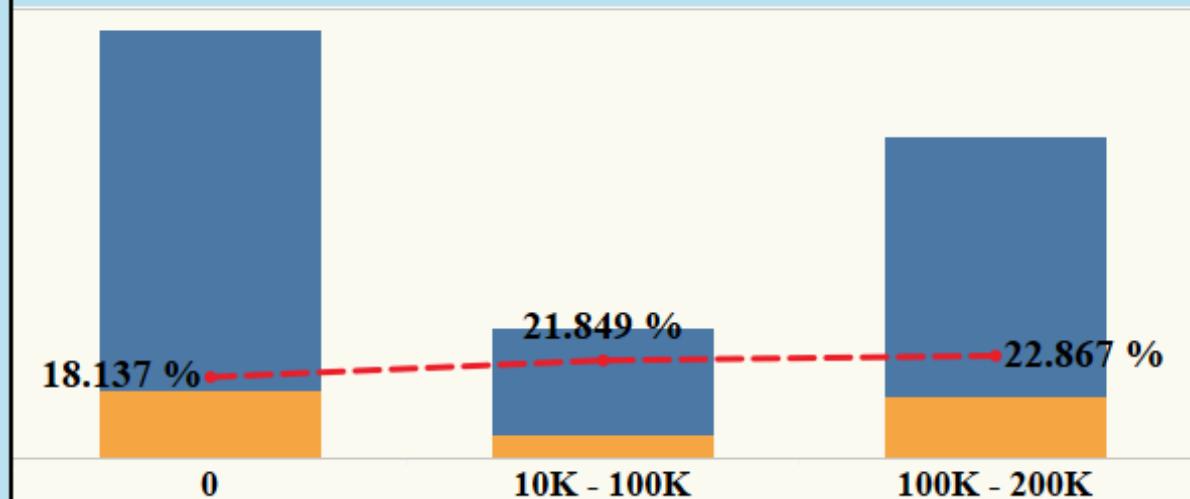
Churn By Age



Churn By Tenure



Churn By Balance



METRICS



BANK CUSTOMER CHURN ANALYSIS



Estimated Salary
80000 D 150000

820
Customers

167
Churn
Customers

20.37 %
Churn Rate

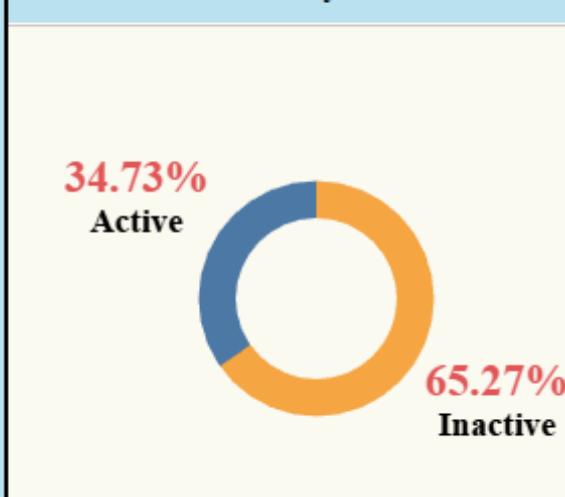
99,564
Avg Salary

60,323
Avg Balance

Churn By Credit Card Status



Churn By Active



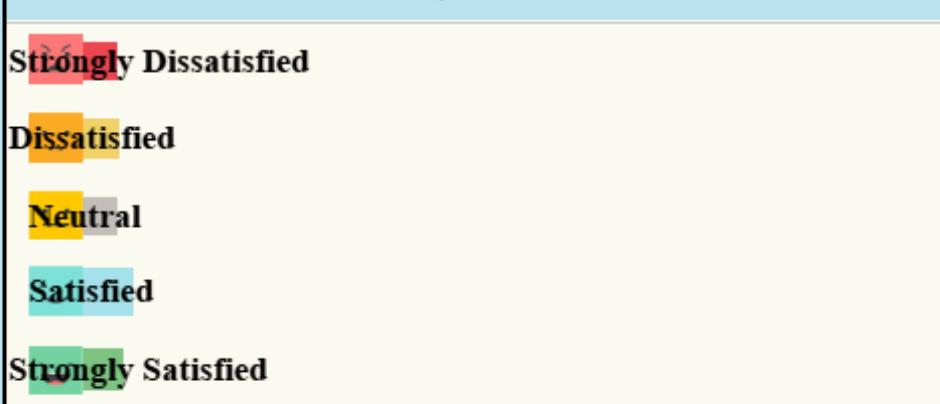
Churn By Complain

	Not Complain	Complain
Churned	1	166
Not churned	0	

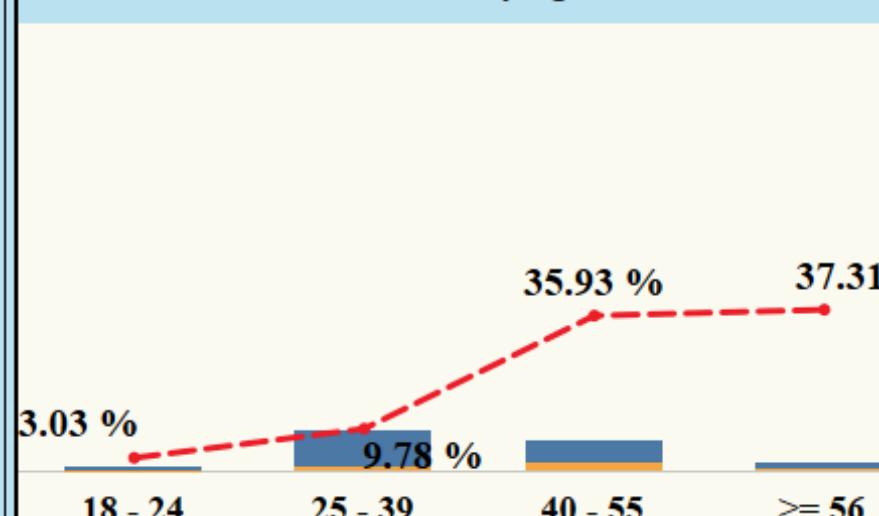
Churn By Card Type



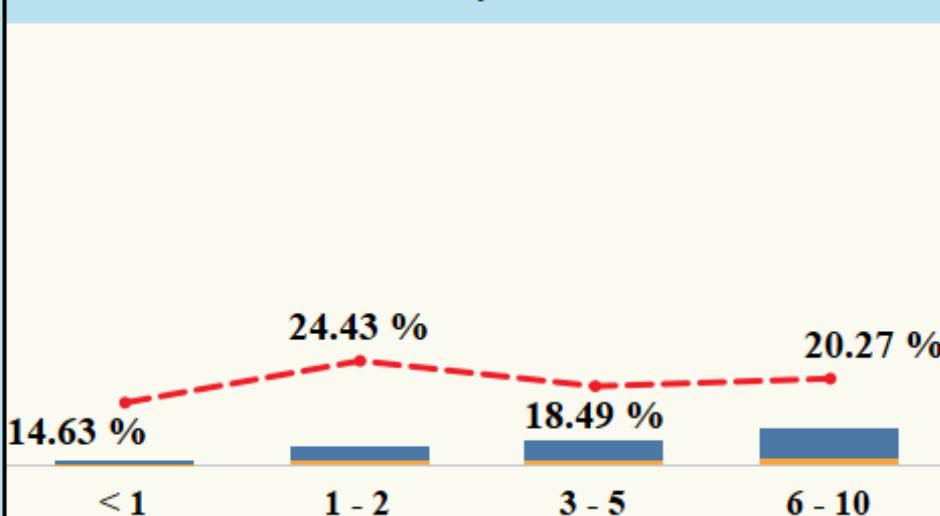
Churn By Satisfaction



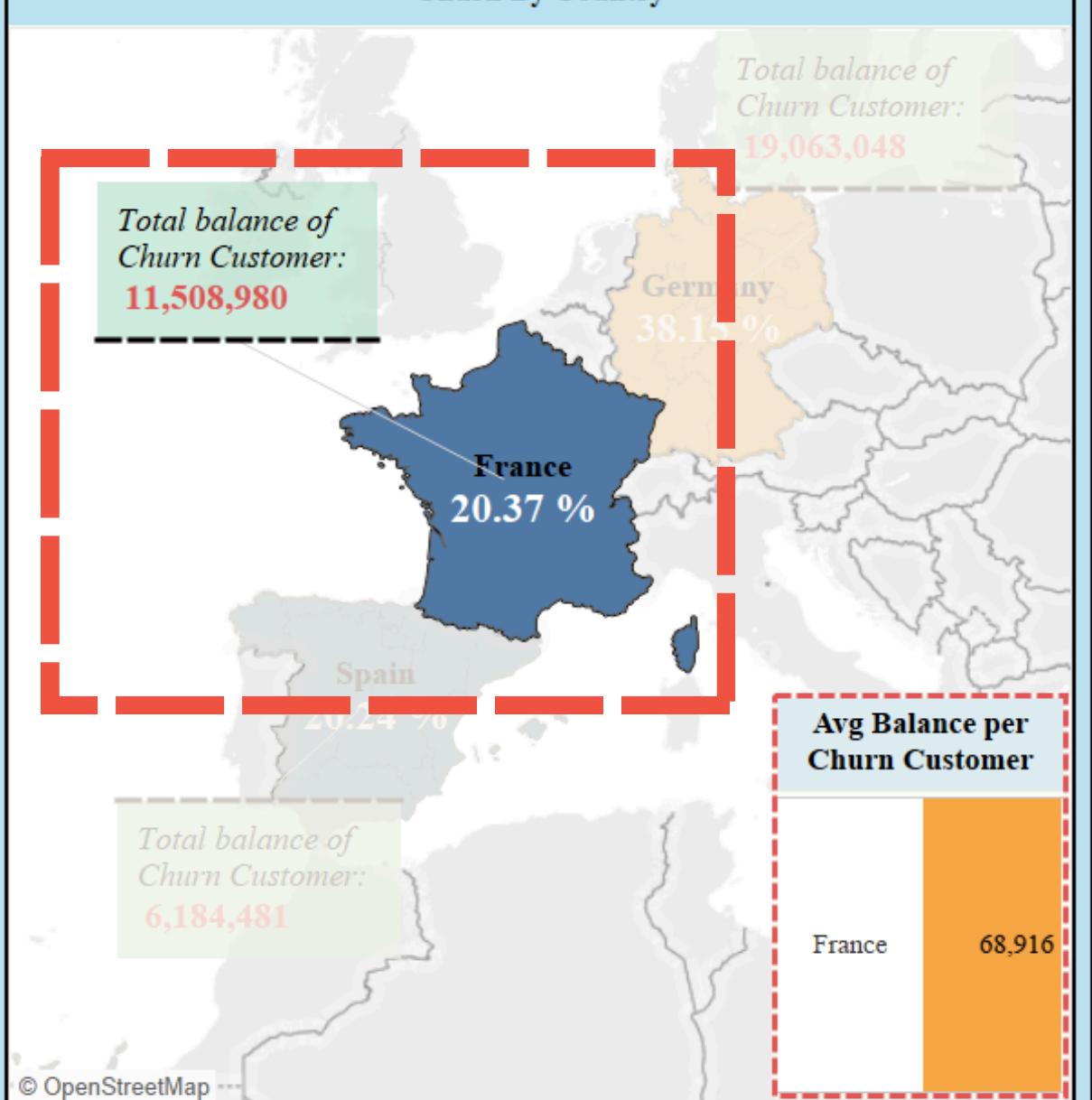
Churn By Age



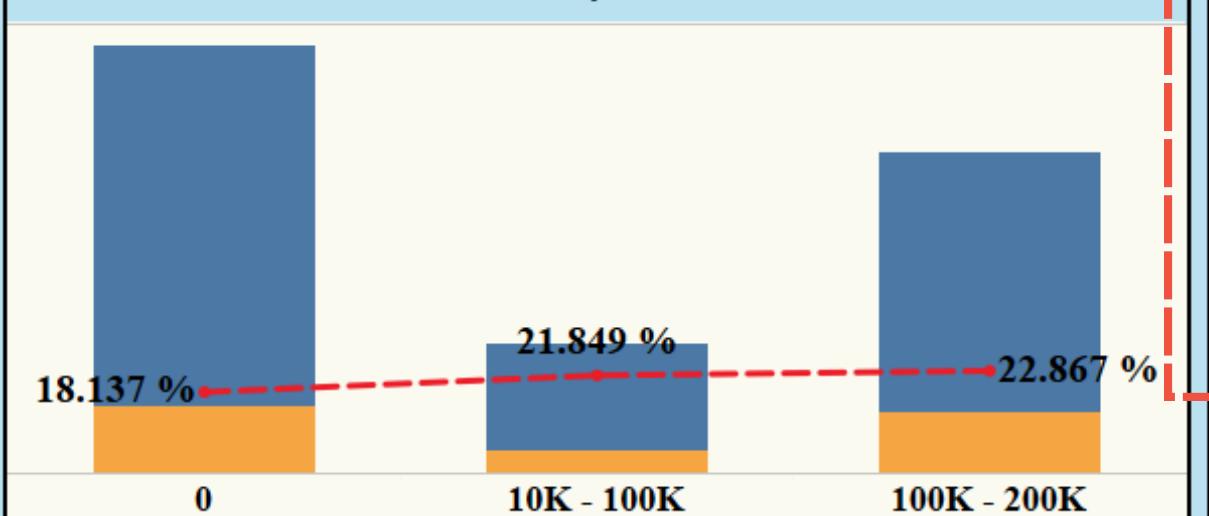
Churn By Tenure



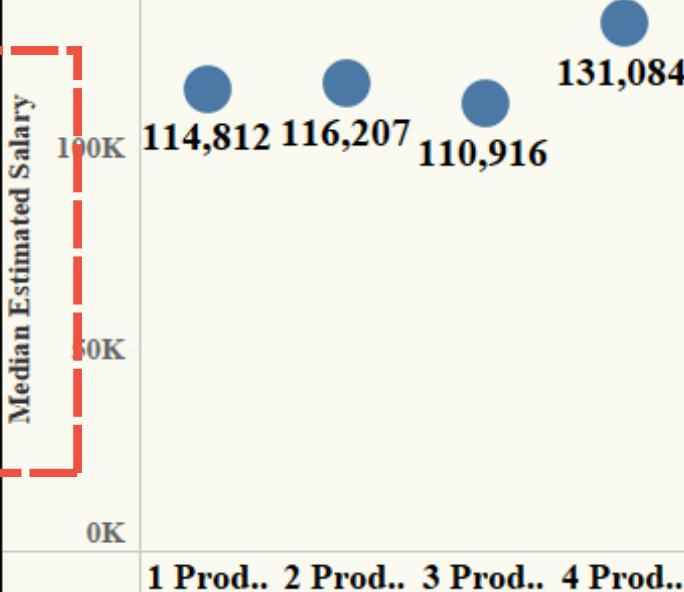
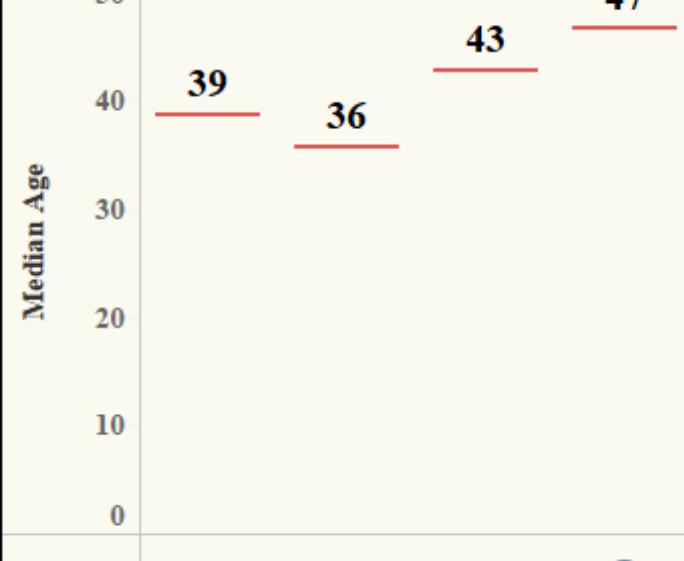
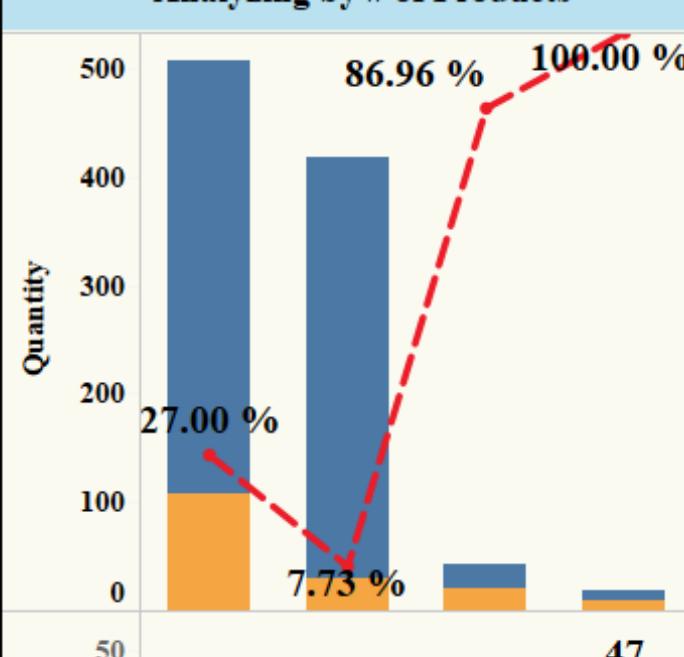
Churn By Country



Churn By Balance



Analyzing by # of Products





METRICS



BANK CUSTOMER CHURN ANALYSIS



Estimated Salary
80000 150000

420
Customers

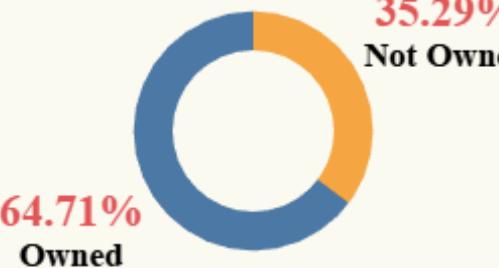
85
Churn
Customers

20.24 %
Churn Rate

100,734
Avg Salary

59,862
Avg Balance

Churn By Credit Card Status

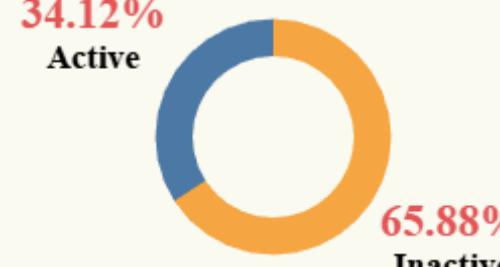


Churn By Complain

	Not Complain	Complain
Churned		85
Not churned	0	0

Card Type	Count
PLATINUM	28
SILVER	24
DIAMOND	19
GOLD	14

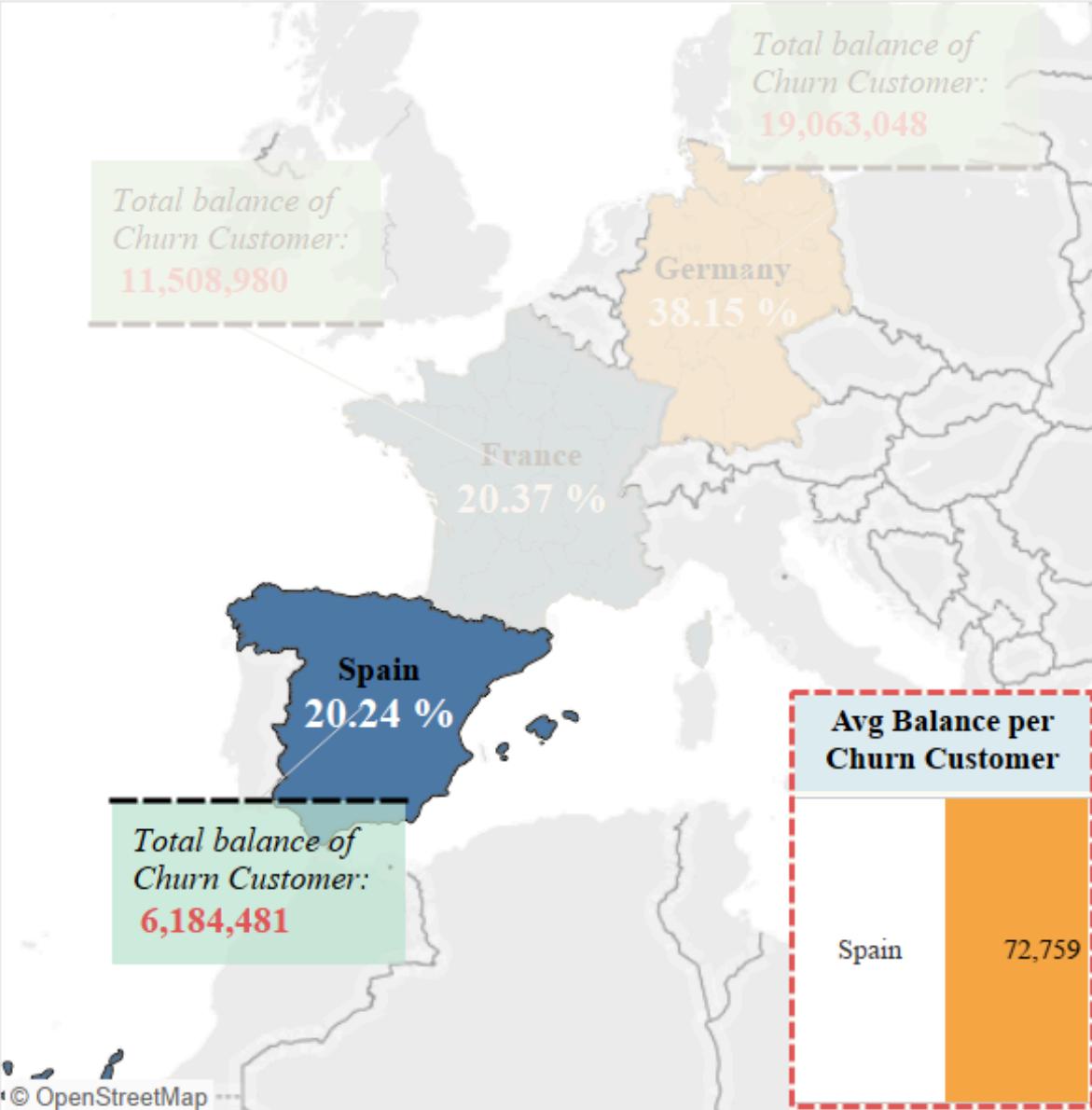
Churn By Active



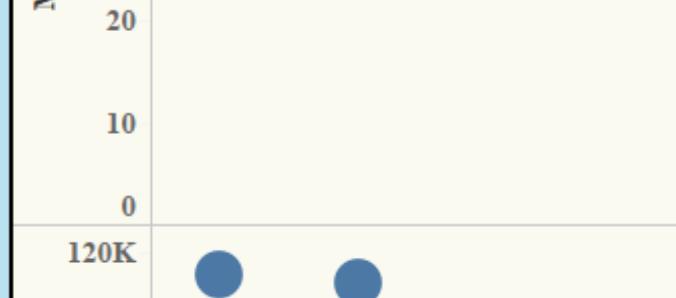
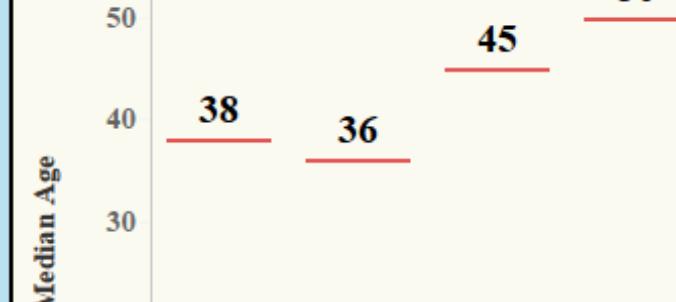
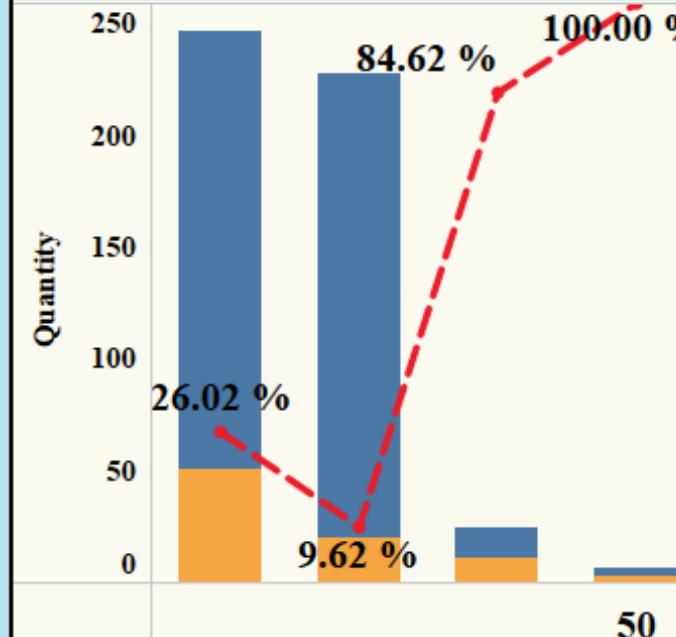
Churn By Satisfaction



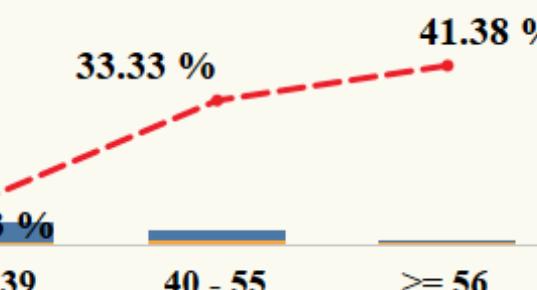
Churn By Country



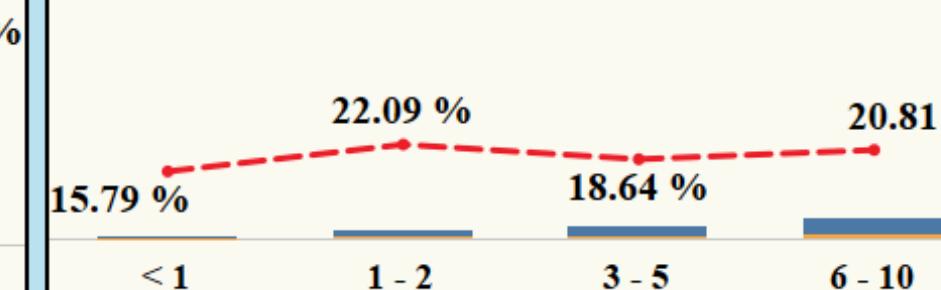
Analyzing by # of Products



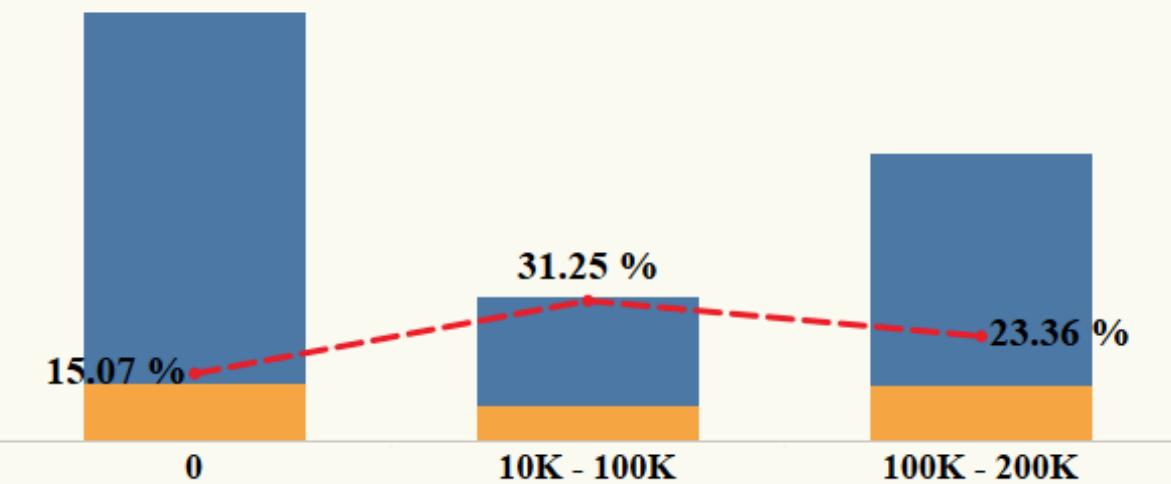
Churn By Age



Churn By Tenure



Churn By Balance



METRICS



BANK CUSTOMER CHURN ANALYSIS



Estimated Salary
80000 150000

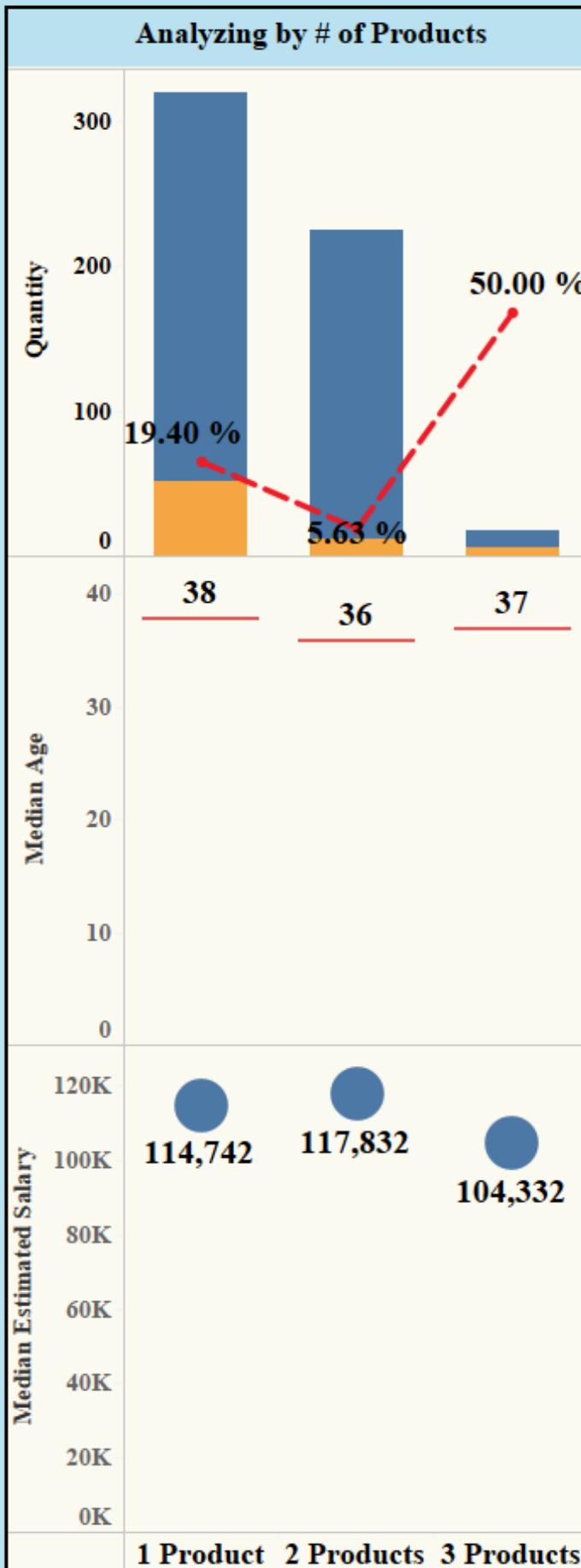
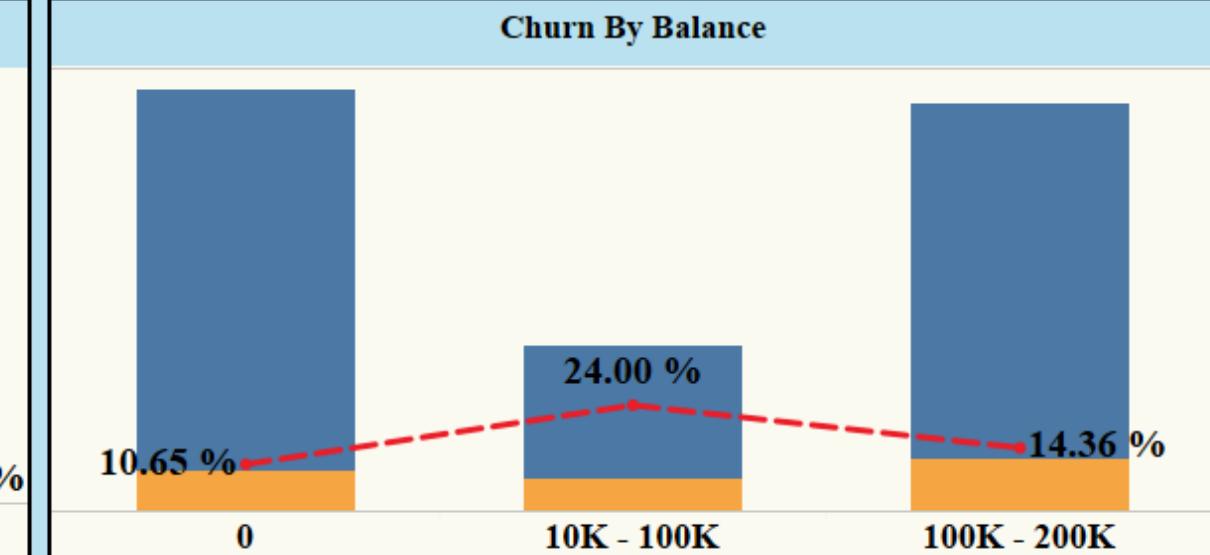
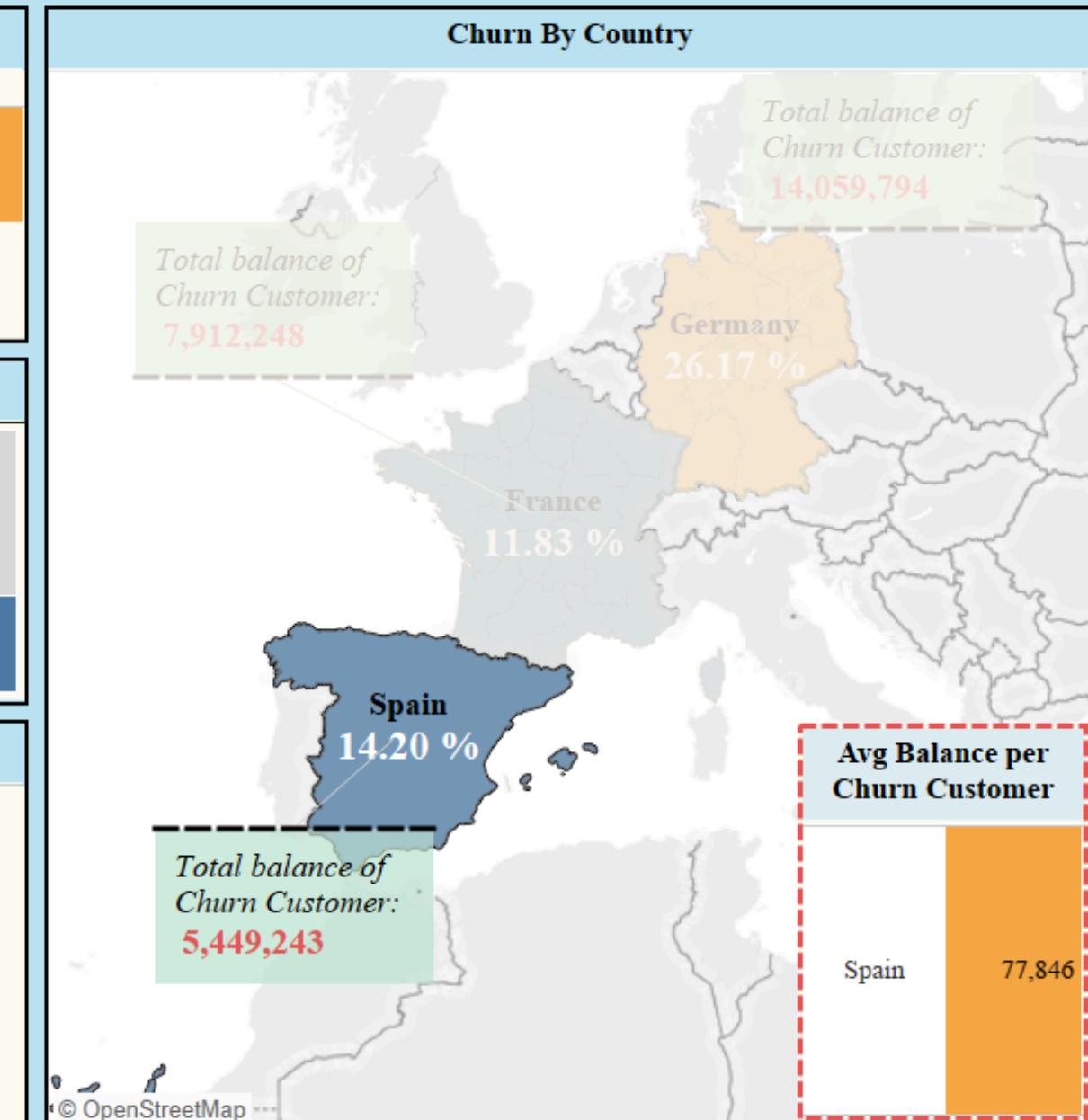
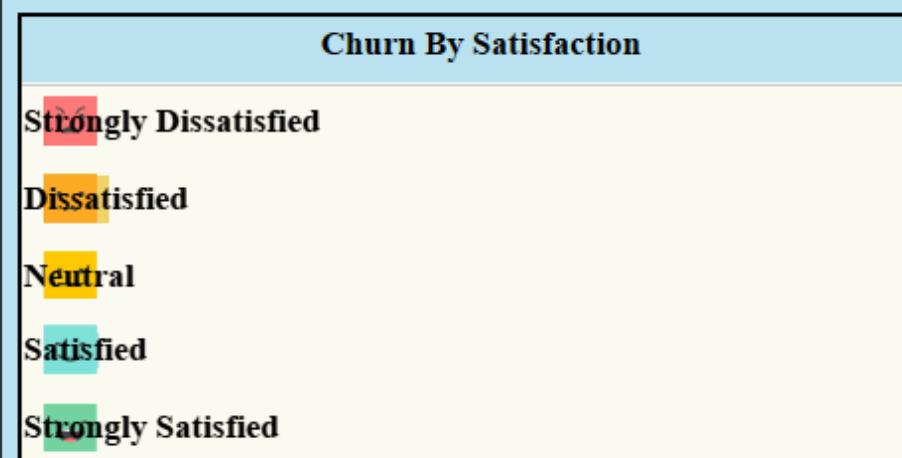
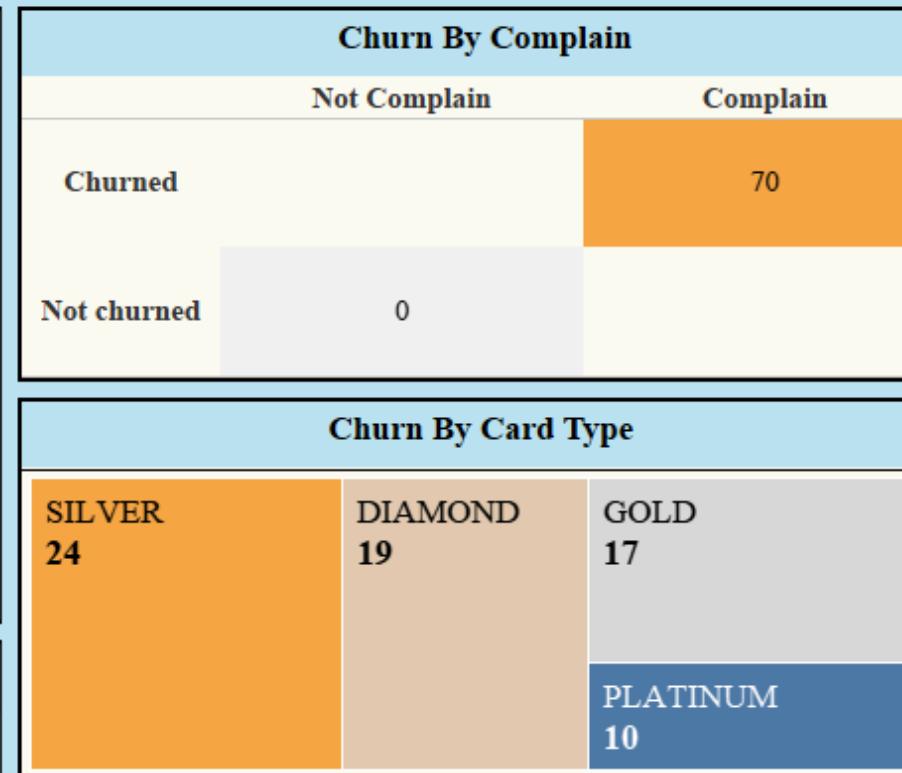
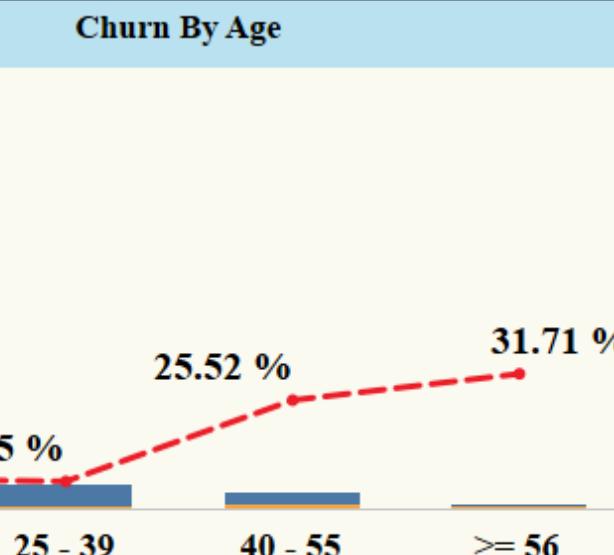
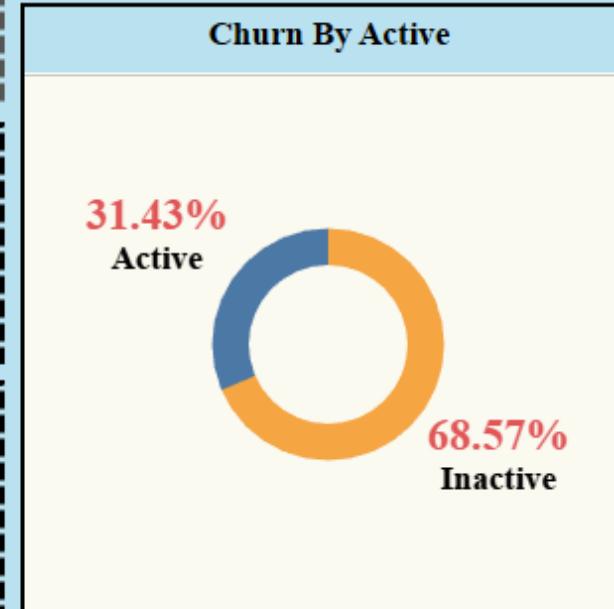
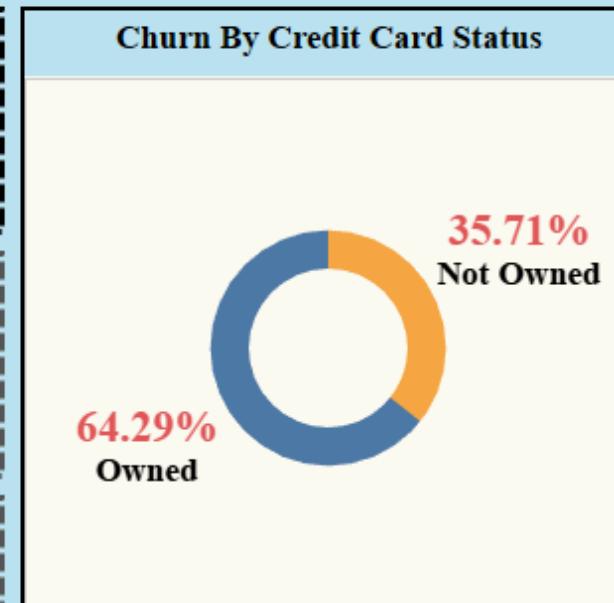
493
Customers

70
Churn
Customers

14.20 %
Churn Rate

98,426
Avg Salary

63,353
Avg Balance



BANK CUSTOMER CHURN PREDICTION



4

Dự đoán tỷ lệ khách hàng rời bỏ của ngân hàng

Dự đoán tỷ lệ rời bỏ khách hàng

Sử dụng mô hình **Random Forest** để dự đoán tỷ lệ rời bỏ của khách hàng

Classification Report:					
		precision	recall	f1-score	support
	0	0.89	0.94	0.91	1576
	1	0.94	0.88	0.91	1609
accuracy			0.91	0.91	3185
macro avg		0.91	0.91	0.91	3185
weighted avg		0.91	0.91	0.91	3185

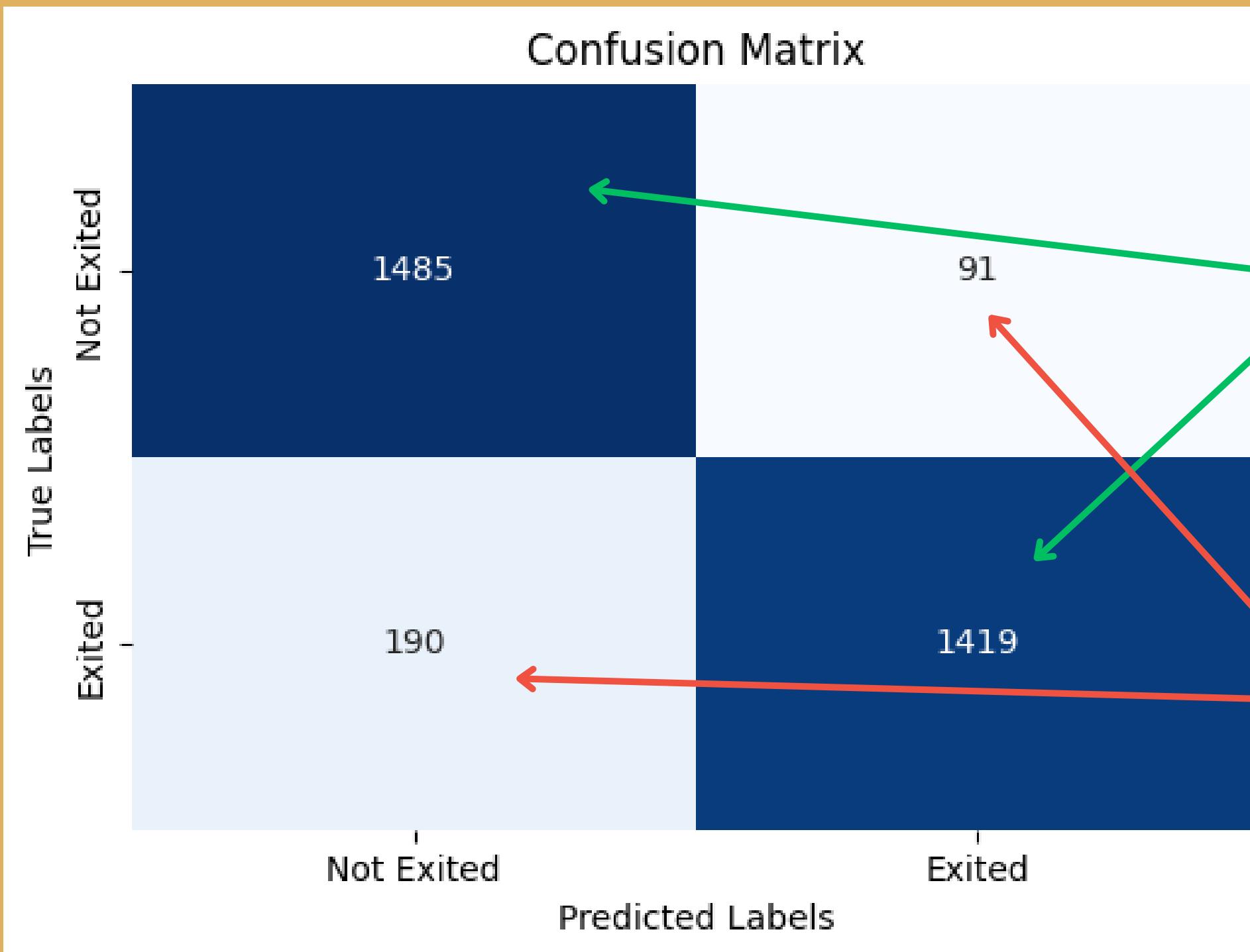
Độ chính xác: **91.18%**

Dự đoán chi tiết:

- **89%** khách hàng không rời bỏ được dự đoán chính xác (precision = **0.89**)
- **94%** khách hàng rời bỏ được dự đoán chính xác (precision = **0.89**)
- Mô hình có độ chính xác cao cho cả hai lớp 0 và 1, với giá trị **F1-score** là **0.91** cho cả hai lớp

Dự đoán tỷ lệ rời bỏ khách hàng

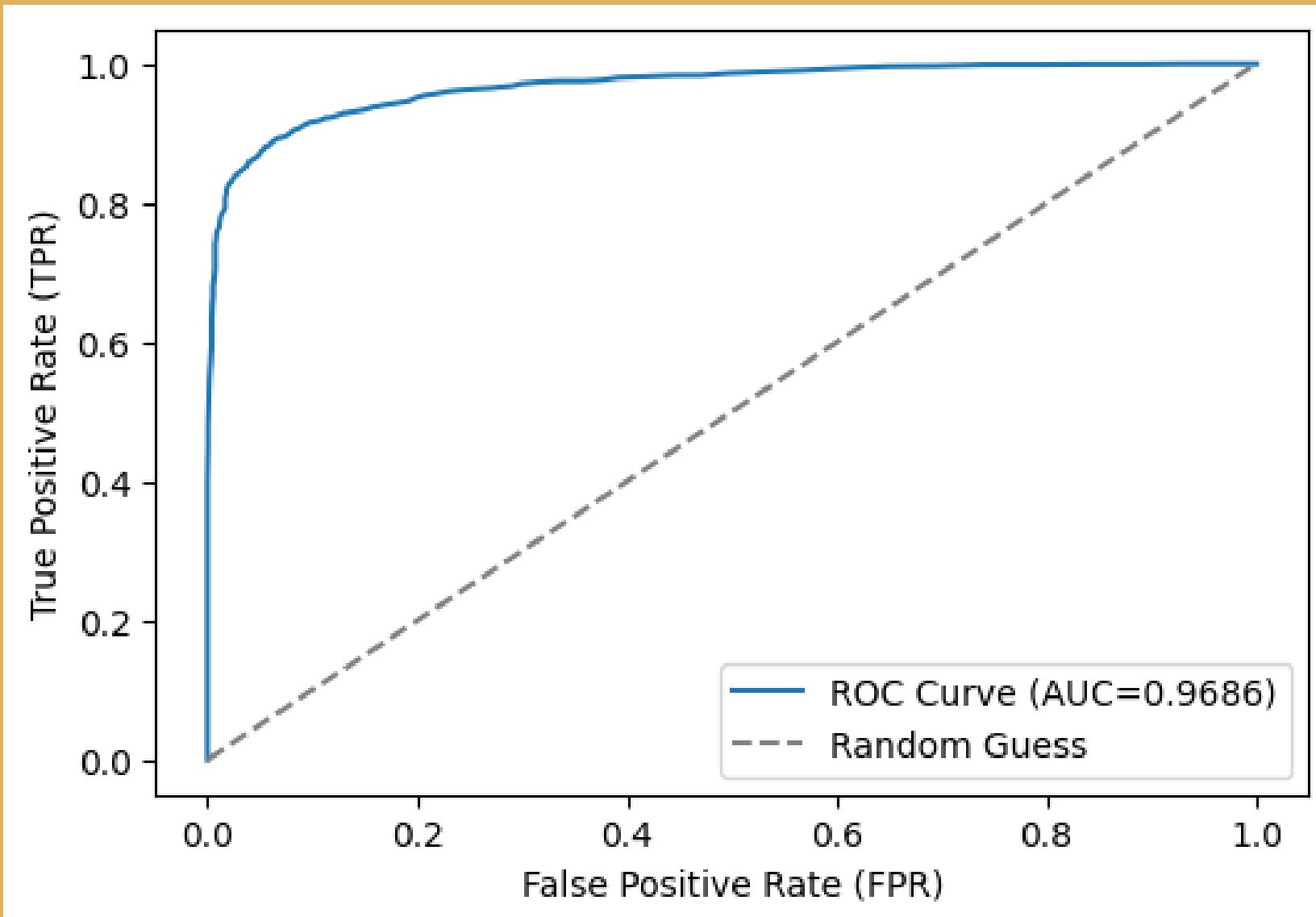
DỰA VÀO BẢNG CONFUSION MATRIX



Dự đoán **chính xác** (đường chéo chính):
Lớp Not Exited: 1485 mẫu được dự đoán chính xác (**True Positive**).
Lớp Exited: 1419 mẫu được dự đoán chính xác (**True Negative**).

Dự đoán **sai** (ngoài đường chéo chính):
Lớp Not Exited: 91 mẫu được dự đoán sai (dự đoán là lớp Exited nhưng thực sự là lớp not Exited) (**False Positive**).
Lớp Exited: 190 mẫu được dự đoán sai là Not Exited (nhưng thực sự là lớp Exited) (**False Negative**)

Dự đoán tỷ lệ rời bỏ khách hàng



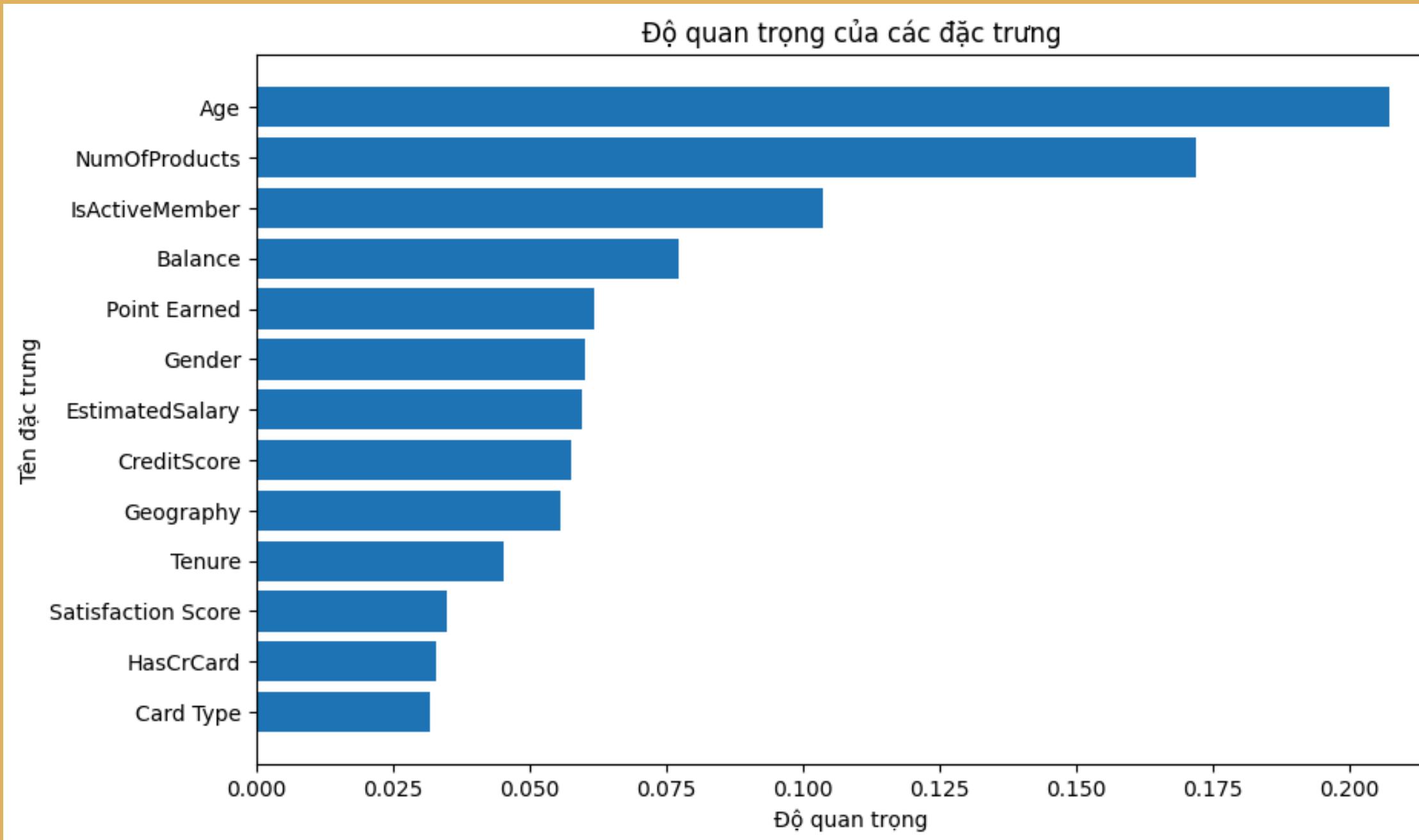
Mức độ ROC-AUC 0.9686 được đánh giá là **xuất sắc**.

Đường cong ROC:

- AUC = **0.9686**, cao hơn đường chéo (AUC = **0.5**) cho thấy mô hình dự đoán tốt hơn so với dự đoán ngẫu nhiên.
- Diện tích dưới đường cong càng lớn, khả năng phân biệt giữa hai nhóm khách hàng càng cao.

→ Mô hình dự đoán: có **độ hiệu quả cực kỳ cao** với **tỷ lệ sai rất thấp**.

Feature importance ranking



- **Age:** Yếu tố quan trọng nhất (0.2)
- **NumOfProducts:** Ảnh hưởng lớn (0.17)
- **IsActiveMember:** Quan trọng (0.1)
- **Balance:** Tương đối quan trọng (0.078)

→ Có độ ảnh hưởng lớn nhất đến khả năng dự đoán chính xác của mô hình.

RECOMMENDATION

5

Đề xuất giải pháp

1. **Chăm sóc** khách hàng tốt hơn ở nhóm **nữ** giới từ **40 tuổi trở lên**
2. **Chạy các chương trình khuyến mãi, kích thích** cho những khách hàng đang ở trạng thái không hoạt động như tặng các voucher
3. Cần **nâng cấp quyền lợi, lợi ích** và tính khác biệt rõ ràng giữa các nhóm khách hàng theo **thâm niên, các loại thẻ và điểm thưởng**
4. Cải thiện **trải nghiệm** mua hàng lần đầu và **nâng cấp dịch vụ, chăm sóc khách hàng** cho các sản phẩm thứ 3 trở lên
5. Nghiêm túc **xử lý** các **khiếu nại, phàn nàn** từ khách hàng

Thank's For Watching

