



Nhóm 1

BANK CUSTOMERS CHURN

**Tô Huỳnh Ngọc Ngân
Trương Thị Thanh Thùy
Nguyễn Thị Hồng Phúc
Trần Ngọc Minh Khanh**



Bank Customer Churn

Mục lục

Tổng quan về chủ đề

Xử lý dữ liệu

Phân tích dữ liệu

Kết luận và giải pháp

Tổng Quan Về Chủ Đề

Giới Thiệu Về Chủ Đề:

- Trong bối cảnh cạnh tranh khốc liệt, khách hàng dễ dàng rời bỏ ngân hàng nếu có trải nghiệm không tốt.
- Việc mất khách hàng sẽ làm giảm doanh thu và làm tăng mạnh chi phí tìm kiếm khách hàng mới.
- Các ngân hàng hiện đại phải cần có các giải pháp chủ động để phát triển và giữ chân khách hàng sớm nhất có thể.

Mục tiêu:

- Hiểu rõ các yếu tố và nguyên nhân thúc đẩy quyết định rời bỏ của khách hàng.
- Đề xuất giải pháp giữ chân khách hàng.



Chi tiết bộ dữ liệu

- 01 Nguồn: Kaggle.com
- 02 Dataset: Bank Customer Churn
- 03 Số lượng bảng: 1, số lượng cột: 18
- 04 Số lượng hàng: 10000

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10000 entries, 0 to 9999
Data columns (total 18 columns):
#   Column                Non-Null Count  Dtype
---  -
0   RowNumber              10000 non-null  int64
1   CustomerId             10000 non-null  int64
2   Surname                10000 non-null  object
3   CreditScore             10000 non-null  int64
4   Geography              10000 non-null  object
5   Gender                 10000 non-null  object
6   Age                    10000 non-null  int64
7   Tenure                  10000 non-null  int64
8   Balance                 10000 non-null  float64
9   NumOfProducts          10000 non-null  int64
10  HasCrCard               10000 non-null  int64
11  IsActiveMember          10000 non-null  int64
12  EstimatedSalary         10000 non-null  float64
13  Exited                  10000 non-null  int64
14  Complain                10000 non-null  int64
15  Satisfaction Score     10000 non-null  int64
16  Card Type               10000 non-null  object
17  Point Earned            10000 non-null  int64
```

Xử lý dữ liệu

1. Tin giảm dữ liệu: Loại bỏ các cột không cần thiết/không liên quan, drop 2 cột RowNumber và cột Surname
2. Đổi kiểu dữ liệu của các cột: HasCard, IsActiveMember, Exited, Complain thành kiểu Text
3. Thêm cột phân nhóm theo độ tuổi, nhóm số dư (Balance) theo tứ phân vị, nhóm điểm tín dụng theo ngân hàng
4. Tách thêm 2 bảng con từ bảng chính dựa trên giá trị cột Exited (1: rời bỏ, 0: ở lại)

Bảng sau khi xử lý dữ liệu

CustomerId	CreditScore	Geography	Gender	Age	Tenure	Balance	NumOfProducts	HasCrCard	IsActiveMember	EstimatedSalary	Exited	Complain	Satisfaction Score	Card Type	Point Earned	AgeGroup	CreditScoreGroup	Balance_grp
15647311	608	Spain	Female	41	1	83807.86	1	0	1	112542.58	0	1	3	DIAMOND	456	Middle-Aged Adults	Fair	rat_it
15619304	502	France	Female	42	8	159660.8	3	1	0	113931.57	1	1	3	DIAMOND	377	Middle-Aged Adults	Poor	rat cao
15737888	850	Spain	Female	43	2	125510.82	1	1	1	79084.1	0	0	5	GOLD	425	Middle-Aged Adults	Excellent	trung_binh
15574012	645	Spain	Male	44	8	113755.78	2	1	0	149756.71	1	1	5	DIAMOND	484	Middle-Aged Adults	Fair	trung_binh
15656148	376	Germany	Female	29	4	115046.74	4	1	0	119346.88	1	1	2	DIAMOND	282	Adults	Poor	trung_binh
15792365	501	France	Male	44	4	142051.07	2	0	1	74940.5	0	0	3	GOLD	251	Middle-Aged Adults	Poor	cao
15592389	684	France	Male	27	2	134603.88	1	1	1	71725.73	0	0	3	GOLD	342	Adults	Good	cao
15767821	528	France	Male	31	6	102016.72	2	0	0	80181.12	0	0	3	GOLD	264	Adults	Poor	it
15643966	616	Germany	Male	45	3	143129.41	2	0	1	64327.26	0	0	5	GOLD	308	Middle-Aged Adults	Fair	cao
15737452	653	Germany	Male	58	1	132602.88	1	1	0	5097.67	1	0	2	SILVER	163	Middle-Aged Adults	Fair	cao
15736816	756	Germany	Male	36	2	136815.64	1	1	1	170041.95	0	0	5	DIAMOND	236	Adults	Very Good	cao

Phân tích dữ liệu

Bank Customers Churn

1.Mối tương quan giữa các yếu tố với Churn

2.Phân tích các yếu tố nhân khẩu

3.Phân tích các yếu tố tài chính và sản phẩm

4.Phân tích trải nghiệm dịch vụ và khiếu nại

TỔNG QUAN VỀ DATA

10K

Total Customers

2038

Churned Customers

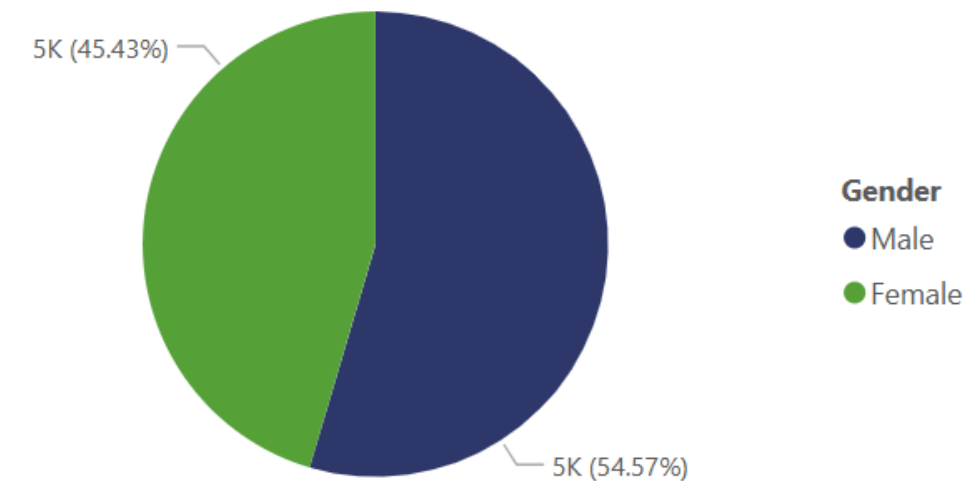
20.38%

Churn Rate

Total Customers by Age



Total Customers by Gender



Tổng số KH đang sử dụng : 10,000

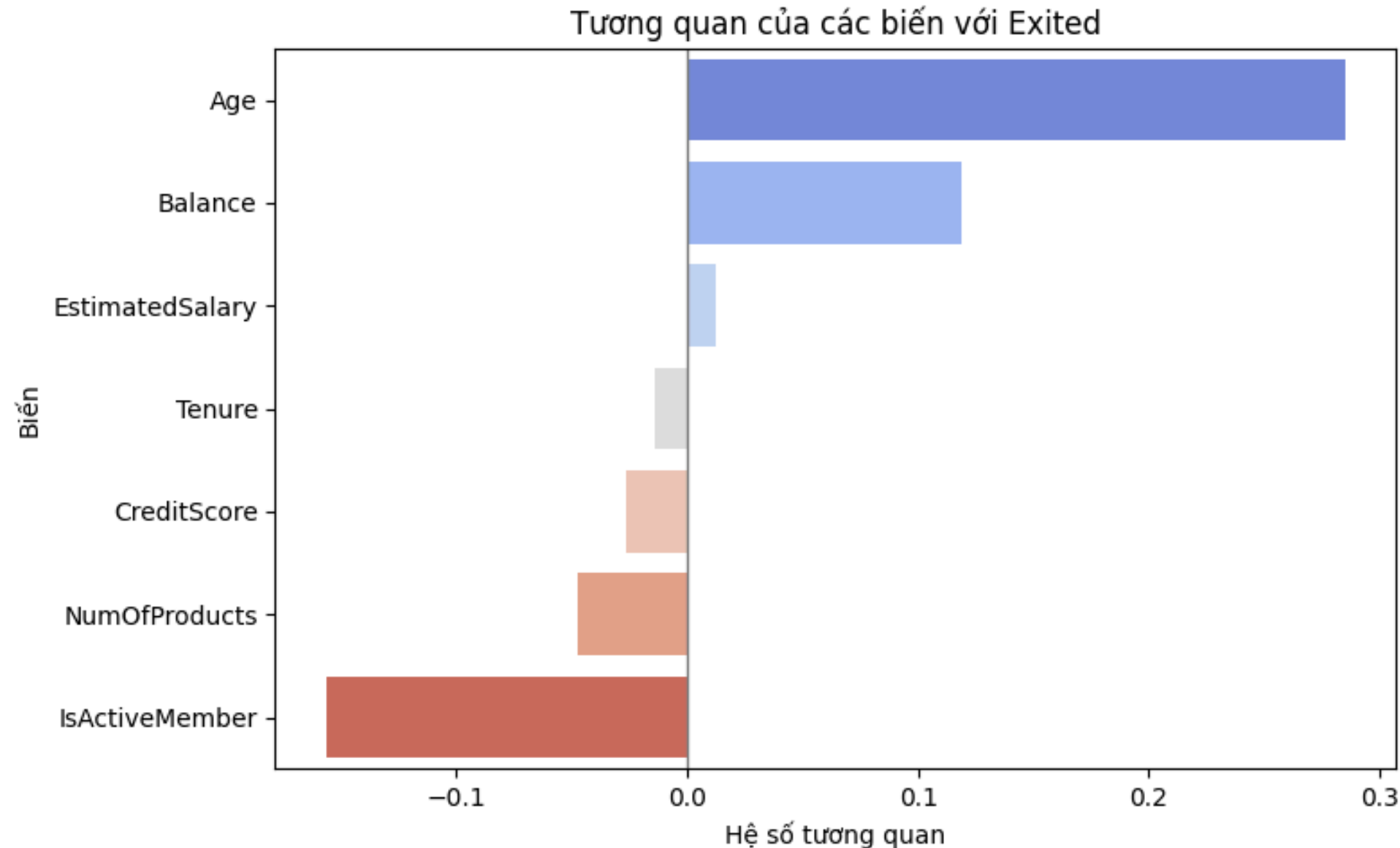
Tổng SL KH rời bỏ : 2038 người

Tỷ lệ rời bỏ : 20,38%

Độ tuổi sử dụng nhiều nhất : từ 32-45 tuổi

Tỷ lệ khách hàng Nam chiếm 54,57%

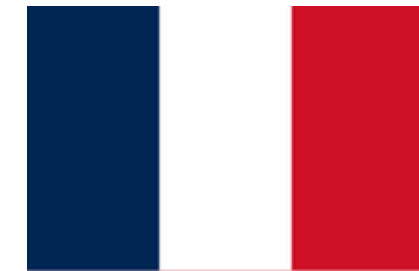
Mối tương quan giữa các yếu tố với Churn



- **Tuổi (Age)** là yếu tố có tương quan tuyến tính dương mạnh nhất với việc rời bỏ (khách hàng lớn tuổi dễ rời bỏ hơn).
- **Số dư (Balance)** cũng có tương quan dương đáng chú ý (số dư cao hơn dễ rời bỏ hơn).
- **Số lượng sản phẩm (NumOfProducts)** có tương quan âm nhẹ (dùng nhiều sản phẩm ít rời bỏ hơn).
- **Thành viên tích cực (IsActiveMember)** có tương quan âm mạnh nhất (thành viên tích cực ít rời bỏ hơn).

Phân tích các yếu tố nhân khẩu

Phân tích theo yếu tố địa lý



Phân tích theo yếu tố giới tính



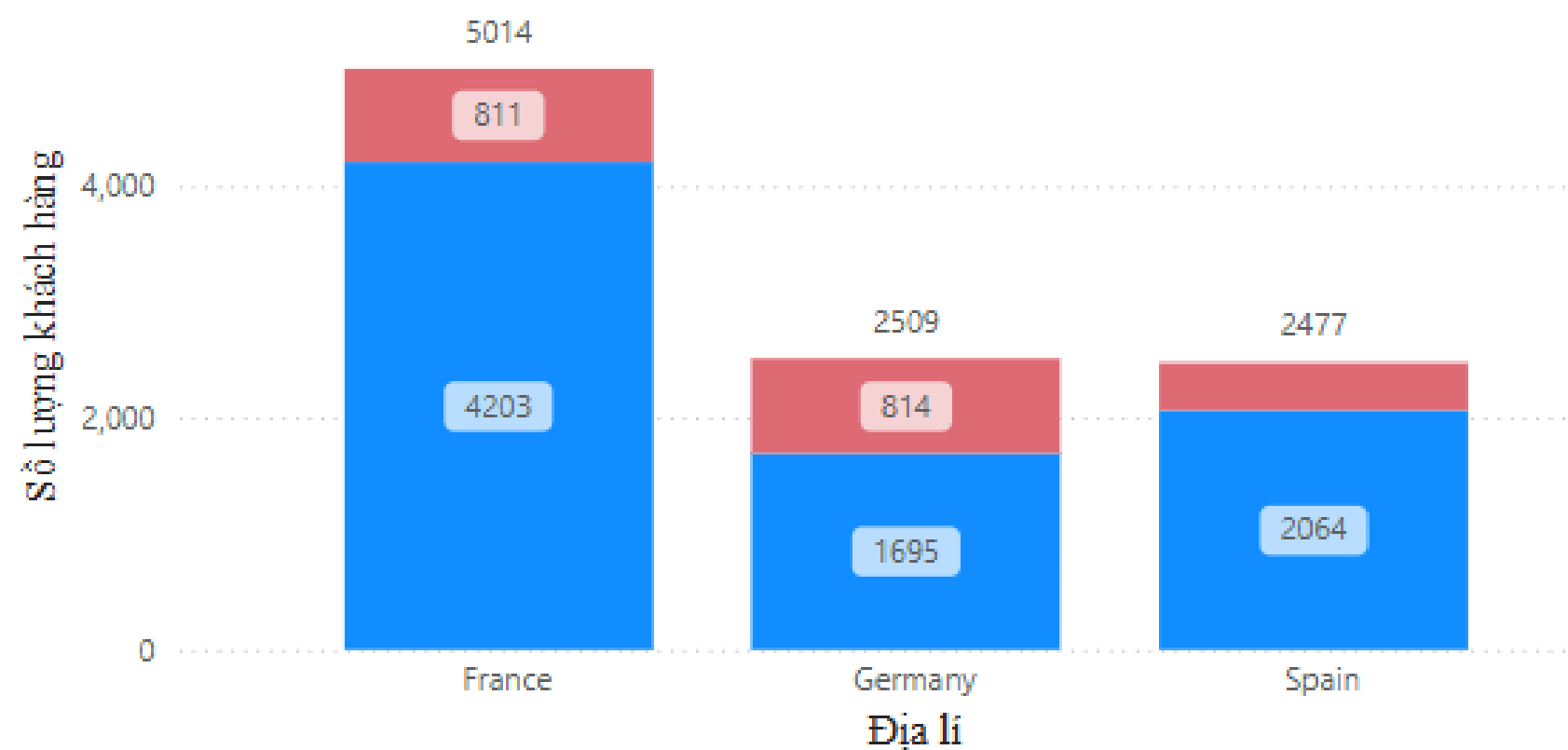
Phân tích theo yếu tố nhóm tuổi



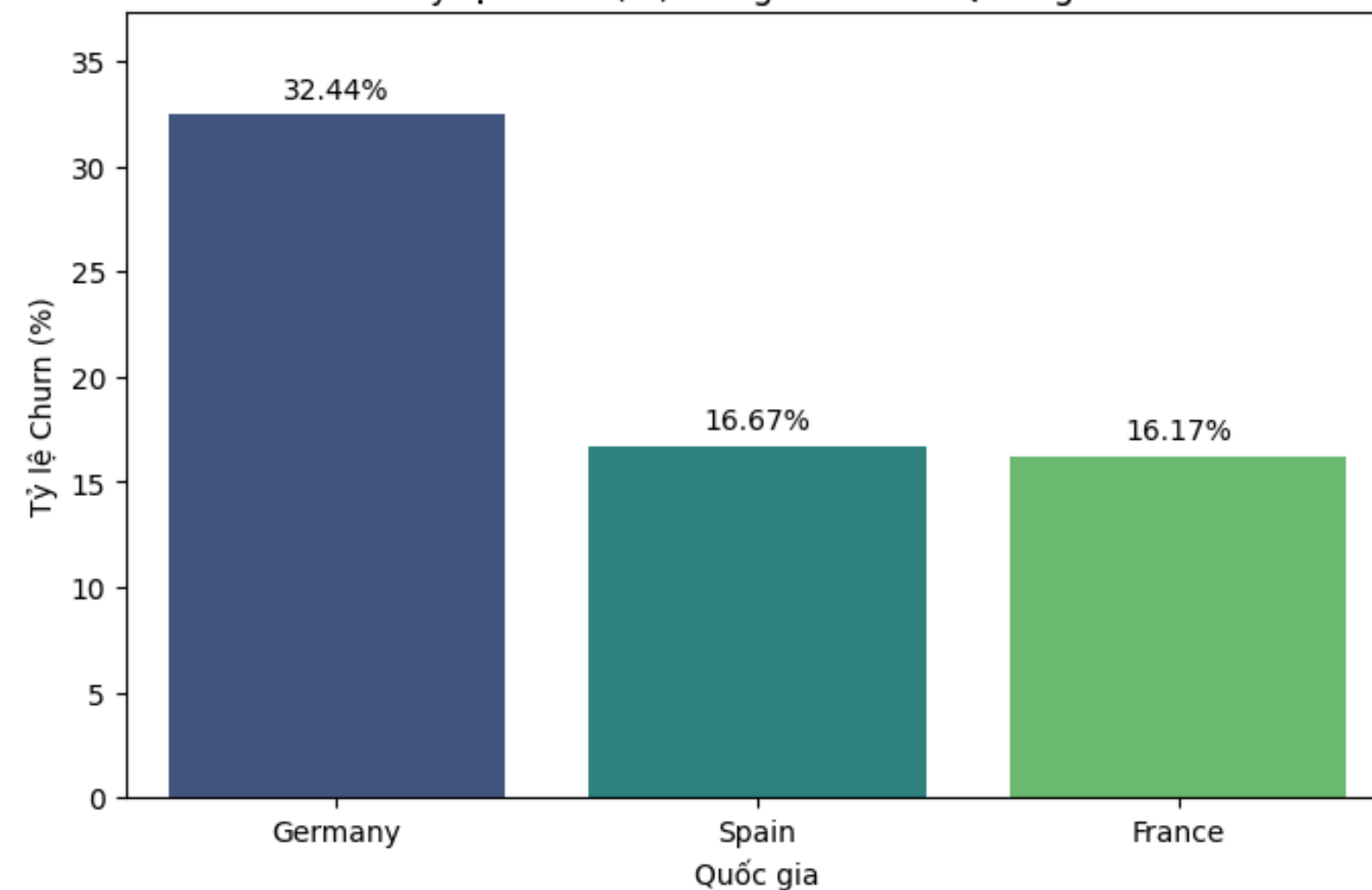
Yếu tố địa lý

Số lượng khách rời bỏ/ở lại theo địa lí

Exited ● 0 ● 1



Tỷ lệ Churn (%) trung bình theo Quốc gia

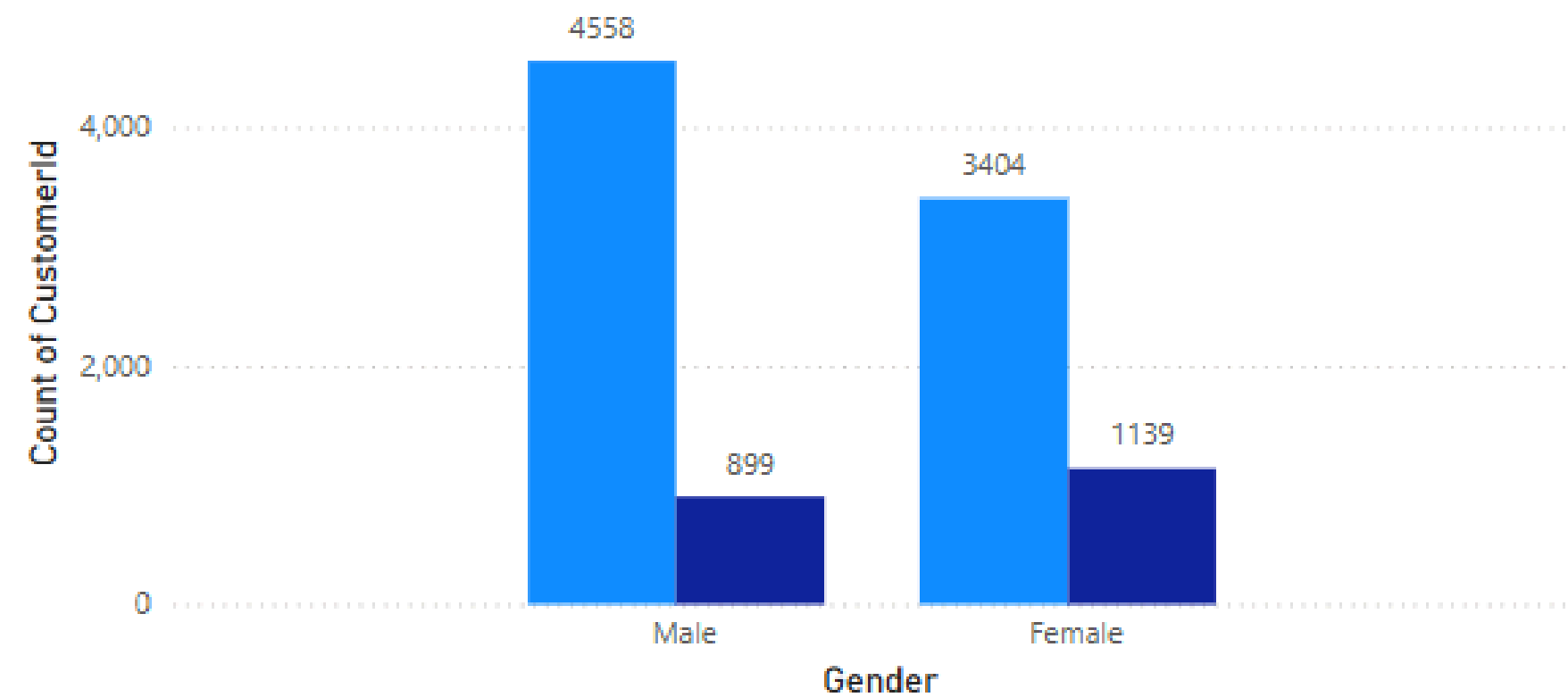


- Pháp là thị trường lớn nhất, chiếm 50,14% tổng số khách hàng
- Đức có tỷ lệ rời bỏ cao nhất 32,44%
- Tỷ lệ rời bỏ ở Pháp và Tây Ban Nha khá tương đồng

Yếu tố giới tính

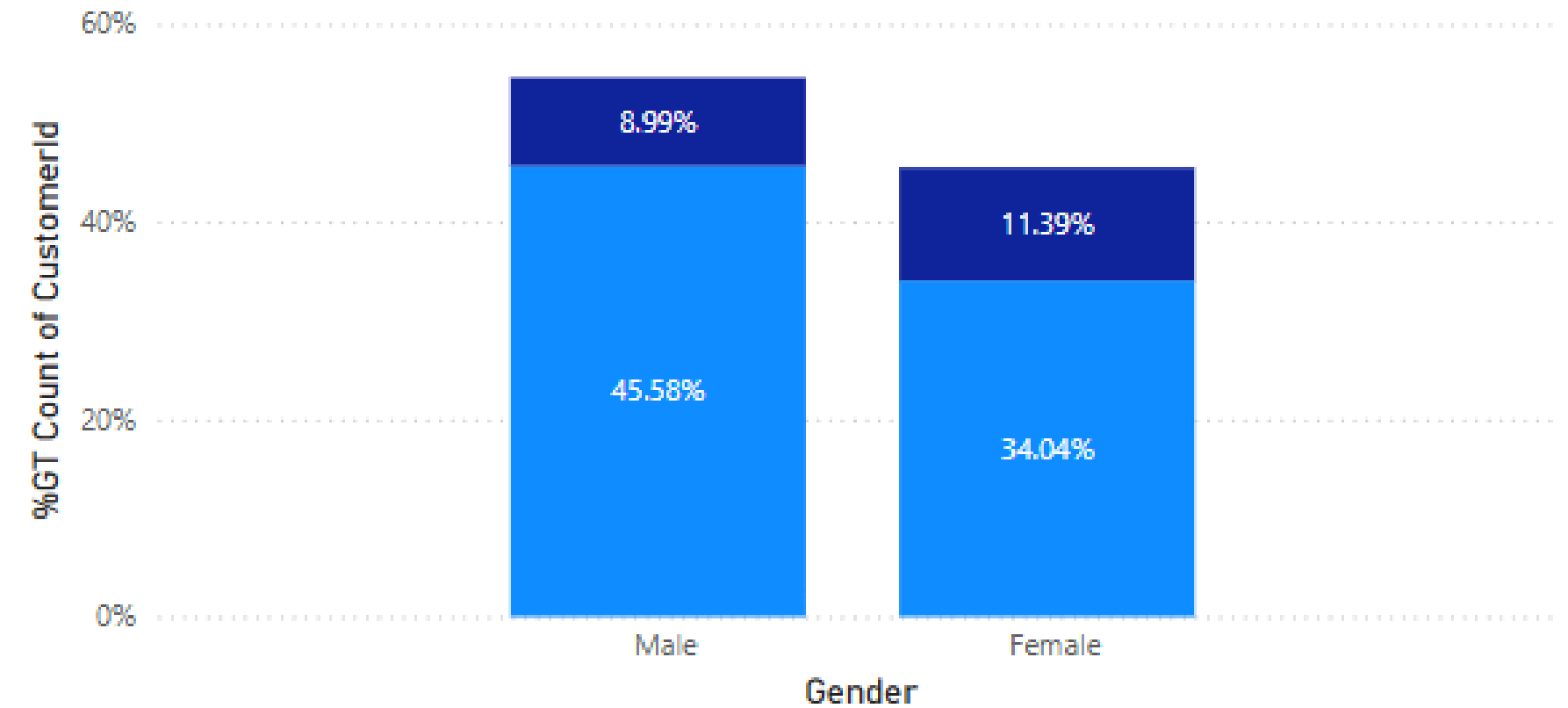
Số lượng rời bỏ/ở lại phân theo giới tính

Exited ● 0 ● 1



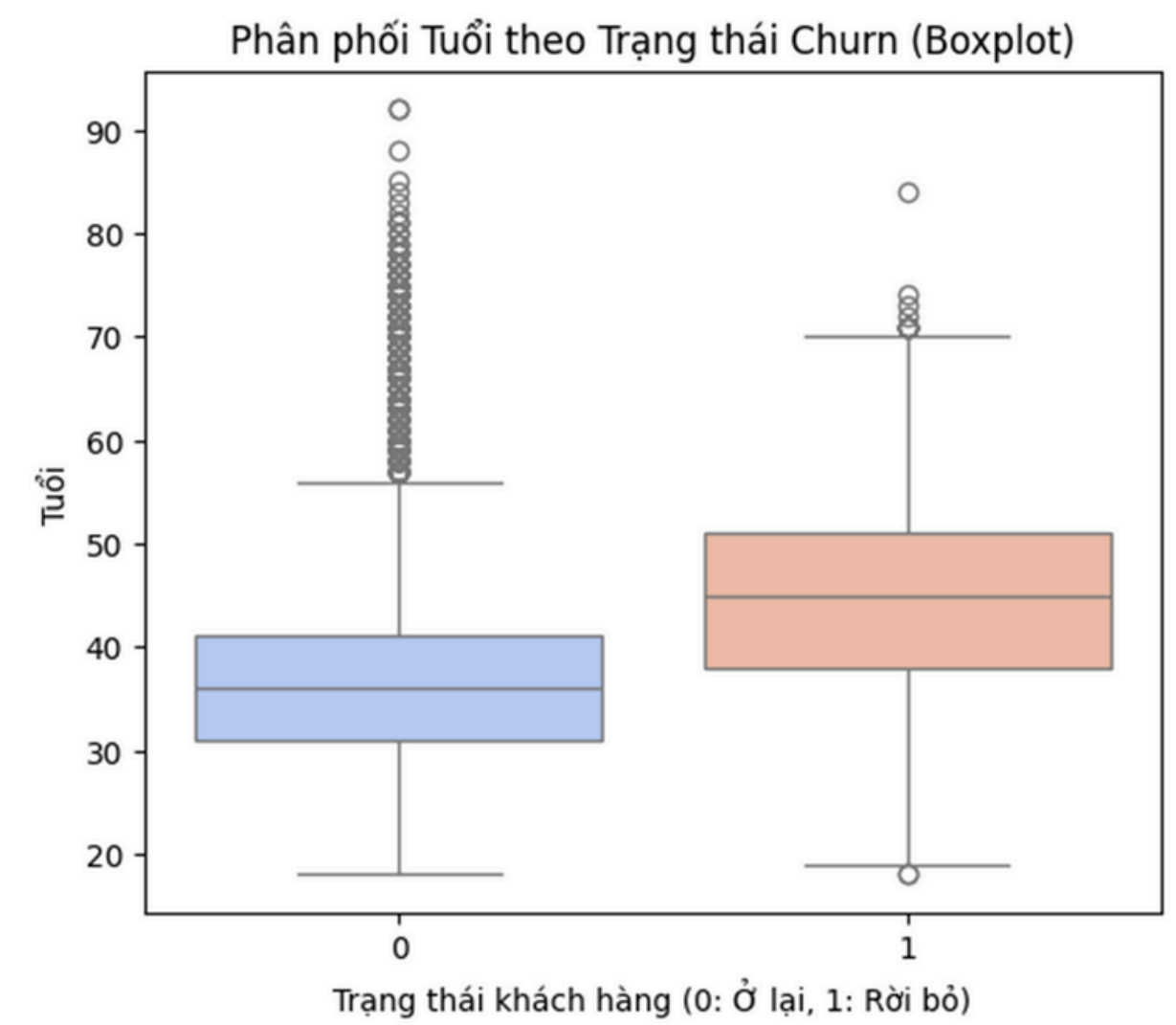
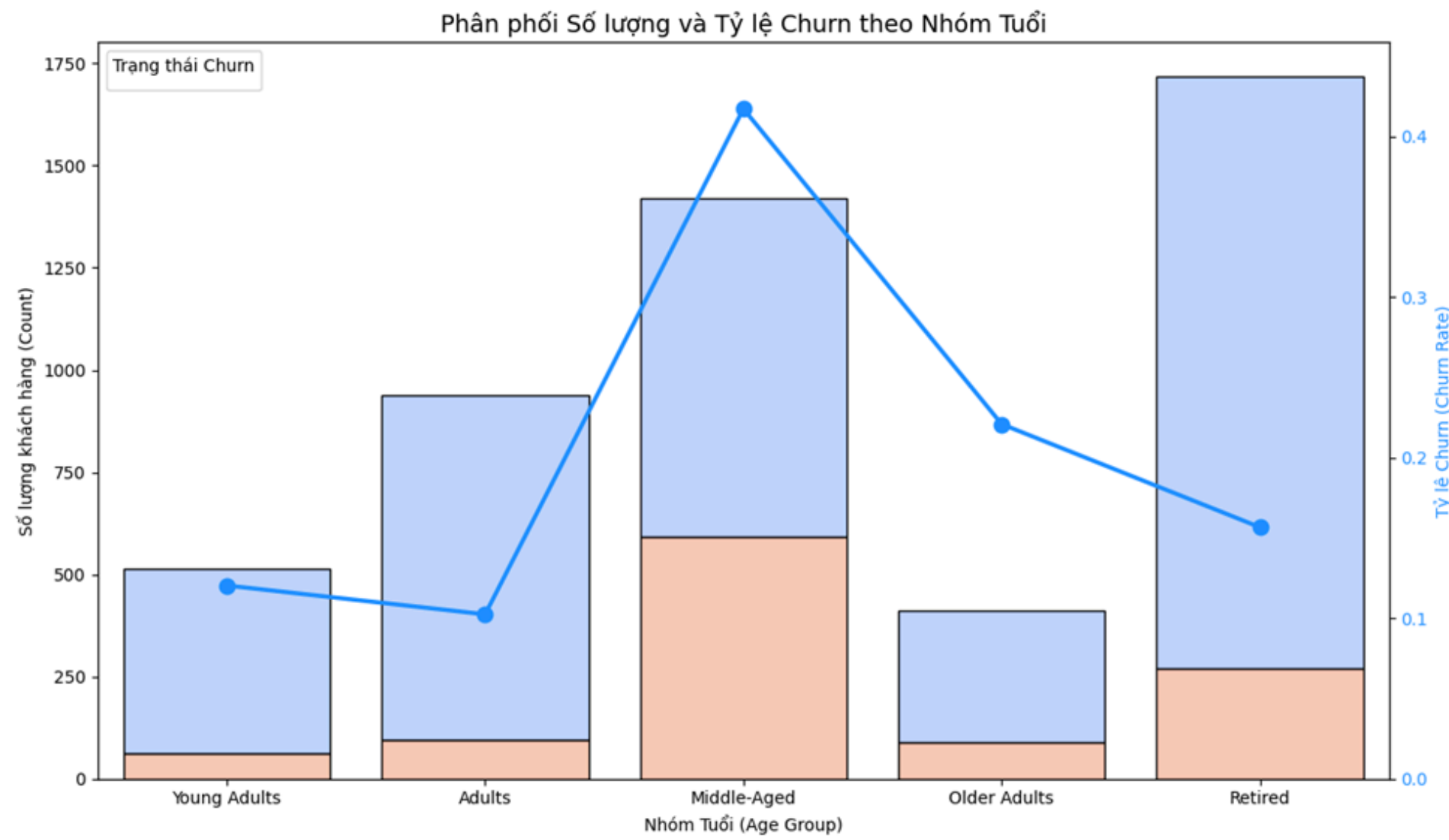
Tỷ lệ rời bỏ/ở lại phân theo giới tính

Exited ● 0 ● 1



- Tỷ lệ rời bỏ của nữ cao gấp đôi so với nam dù số lượng nữ thấp hơn
- Số lượng khách nữ rời bỏ lớn hơn nam (1,139 vs 899)
- Nữ có tỷ lệ giữ chân thấp hơn nam 11.54%

Yếu tố nhóm tuổi



- Khách hàng ở độ tuổi trung niên (đặc biệt là nhóm 40-59) có nguy cơ rời bỏ cao nhất cả về tỷ lệ lẫn số lượng
- Khách hàng lớn tuổi (Nghỉ hưu) tuy đông đảo nhưng tỷ lệ churn không quá cao
- Khách hàng trẻ và trưởng thành (dưới 39) có tỷ lệ churn thấp hơn : Đặc biệt nhóm Adults (25-38) là trung thành nhất.

Young Adults	Adults	Middle -Aged Adults	Older Adults	Retired
18–24	25–39	40–59	60–65	65+

Các yếu tố nhân khẩu

• Nhận xét:

Nhóm tuổi

- Tuổi trung niên (40–59) là nhóm rủi ro cao nhất cả về tỷ lệ và số lượng churn.
- Nhóm nghỉ hưu vẫn duy trì tốt (dù churn tuyệt đối cao nhưng tỷ lệ không lớn).
- Nhóm 25–38 tuổi ổn định nhất, có tỷ lệ churn thấp → nên được ưu tiên phát triển.

Giới tính:

- Nữ giới có churn rate cao hơn nam (25.06% vs 16.47%) → chênh lệch tới 11.5%
- Dù số lượng nữ thấp hơn, số khách nữ rời bỏ lại cao hơn nam → cần chiến lược cá nhân hoá theo giới tính.

Địa lý:

- Đức có tỷ lệ churn cao nhất (32.4%), dù quy mô nhỏ hơn → cần đặc biệt chú trọng giữ chân tại đây.
- Pháp và Tây Ban Nha ổn định hơn với churn ~16%, trong đó Pháp là thị trường lớn nhất (~50%).

Đề xuất chiến lược:

=> Tập trung giữ chân nữ giới, khách hàng tại Đức và nhóm trung niên, đồng thời đẩy mạnh phát triển nhóm trẻ ổn định như Adults (25–38 tuổi) để tăng trưởng bền vững.

Phân tích các yếu tố tài chính và sản phẩm

01



Balance
(Số dư tài khoản)

02



CreditScore
(Điểm tín dụng)

03



NumOfProducts
(Số sản phẩm
ngân hàng sử
dụng)

04



Tenure
(Số năm gắn bó
với ngân hàng)

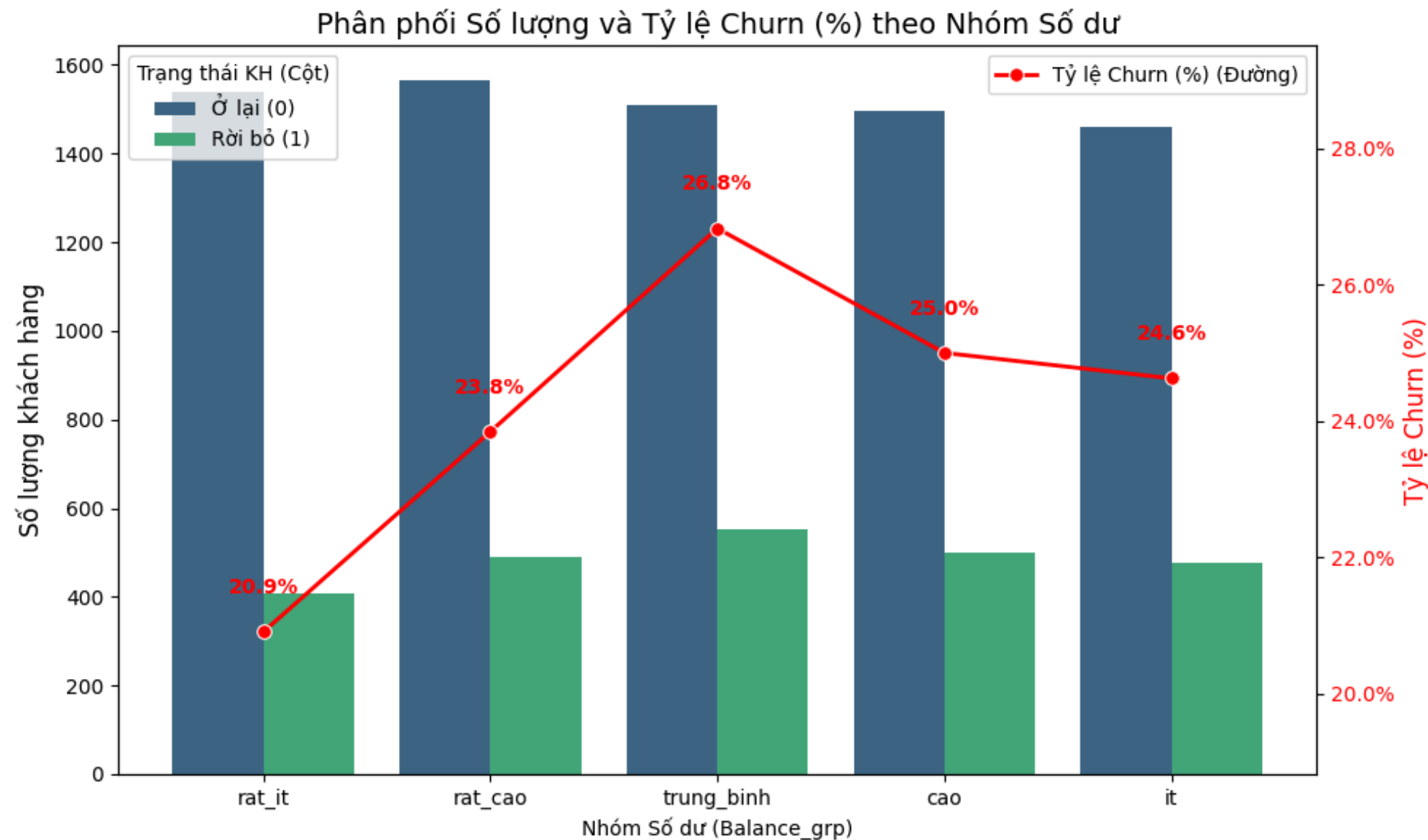
05



IsActiveMember
(Tài khoản được
kích hoạt)

	Tài khoản ngân hàng	Thẻ ATM
Khái niệm	Là một dạng tài sản tại ngân hàng, được thể hiện bằng dãy số do ngân hàng phát hành.	Công cụ để thực hiện các giao dịch thanh toán qua thẻ/ các giao dịch tại cây ATM theo nhu cầu của khách hàng
Chức năng	Khách hàng có thể gửi tiền vào tài khoản ngân hàng để thực hiện các giao dịch tại ngân hàng (chuyển tiền, thanh toán hoá đơn..) hoặc gửi tiết kiệm.	Công cụ để thực hiện các giao dịch thanh toán qua thẻ/ các giao dịch tại cây ATM theo nhu cầu của khách hàng

Balance (Số dư tài khoản)



Nhóm số dư thấp : Tỷ lệ churn thấp nhất 21.88%

Nhóm số dư trung bình thấp : Tỷ lệ churn 26.50%, Churn tăng mạnh so với nhóm trước, cho thấy đây là nhóm bắt đầu có những kỳ vọng nhất định, nhưng chưa được đáp ứng tốt.

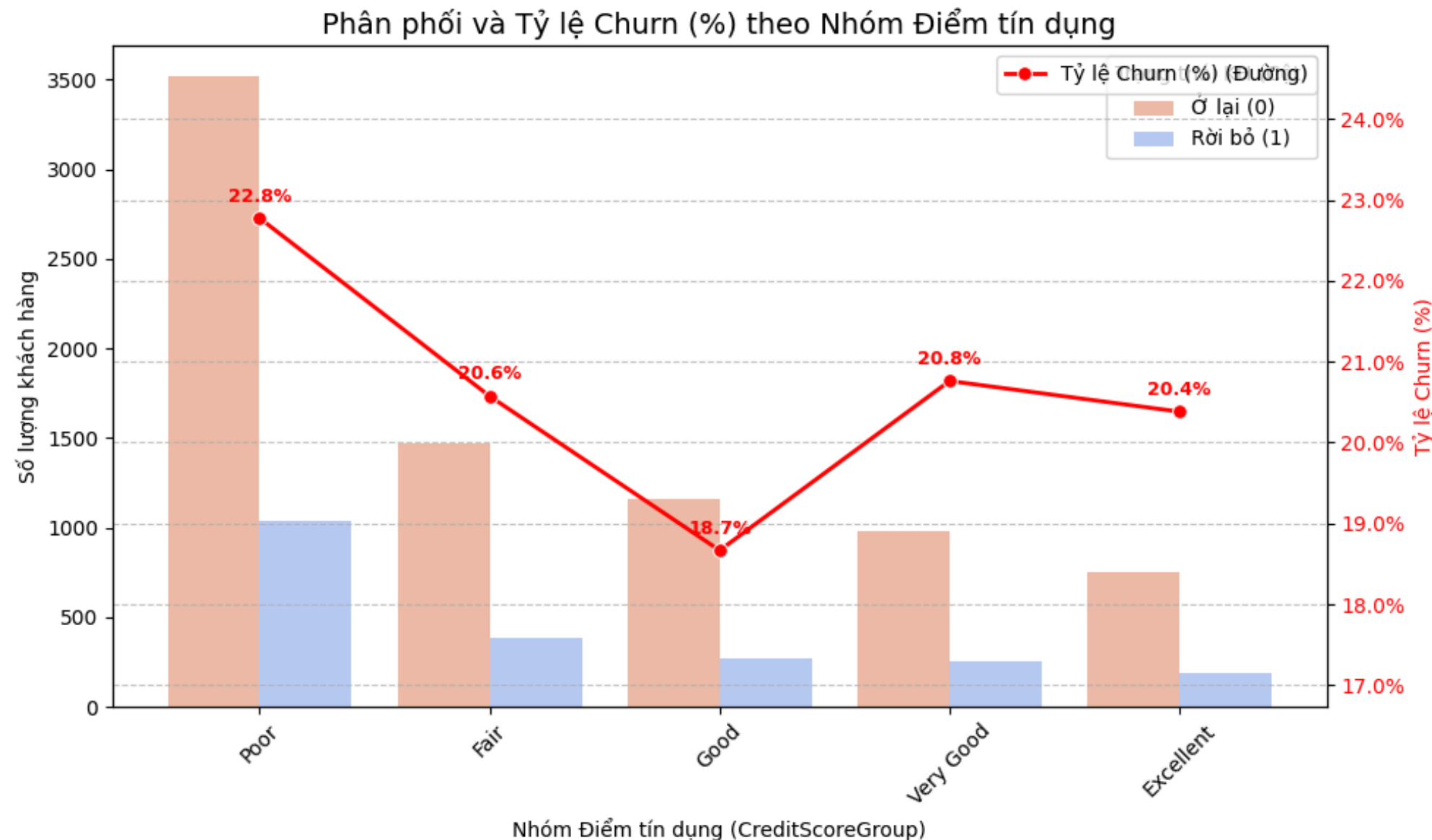
Nhóm số dư trung bình cao : Tỷ lệ churn 26.62% tiếp tục tăng nhẹ, đạt đỉnh cao nhất trong toàn bộ phân khúc.

Nhóm số dư cao : Tỷ lệ churn 25.75%

Tỷ lệ churn bắt đầu giảm nhẹ, tuy nhiên vẫn cao.

Nhóm số dư rất cao : Tỷ lệ Churn giảm đáng kể xuống còn 22.12% - đây là nhóm khách hàng tầm giá trị cao.

CreditScore (Điểm tín dụng)



Nhóm "Poor"

♦ Tỷ lệ churn cao nhất : **23.7%**

→ Rủi ro cao, cần chiến lược giữ chân riêng.

Nhóm "Fair" & "Good"

♦ Tỷ lệ churn : 19.2% – 19.3%

→ Nhóm ổn định nhất, tiềm năng phát triển dài hạn.

Nhóm "Very Good"

♦ Tỷ lệ churn : 19.6% – tăng nhẹ, nhưng vẫn thấp.

Nhóm "Excellent"

♦ Tỷ lệ churn tăng trở lại : 20.6%

-> CreditScore càng cao tỷ lệ churn càng thấp

NumOfProducts (Số sản phẩm ngân hàng sử dụng)

Tương quan âm nhẹ (~ -0.05):

→ Khách dùng nhiều sản phẩm hơn có xu hướng gắn bó hơn

Nhóm ổn định nhất – 2 sản phẩm:

- Tỷ lệ churn thấp nhất: 8.2%
- Số khách ở lại cao nhất: 2,072 người

→ Nhóm trung thành, cần giữ chân và phát triển.

Nhóm 1 sản phẩm – Rủi ro cao:

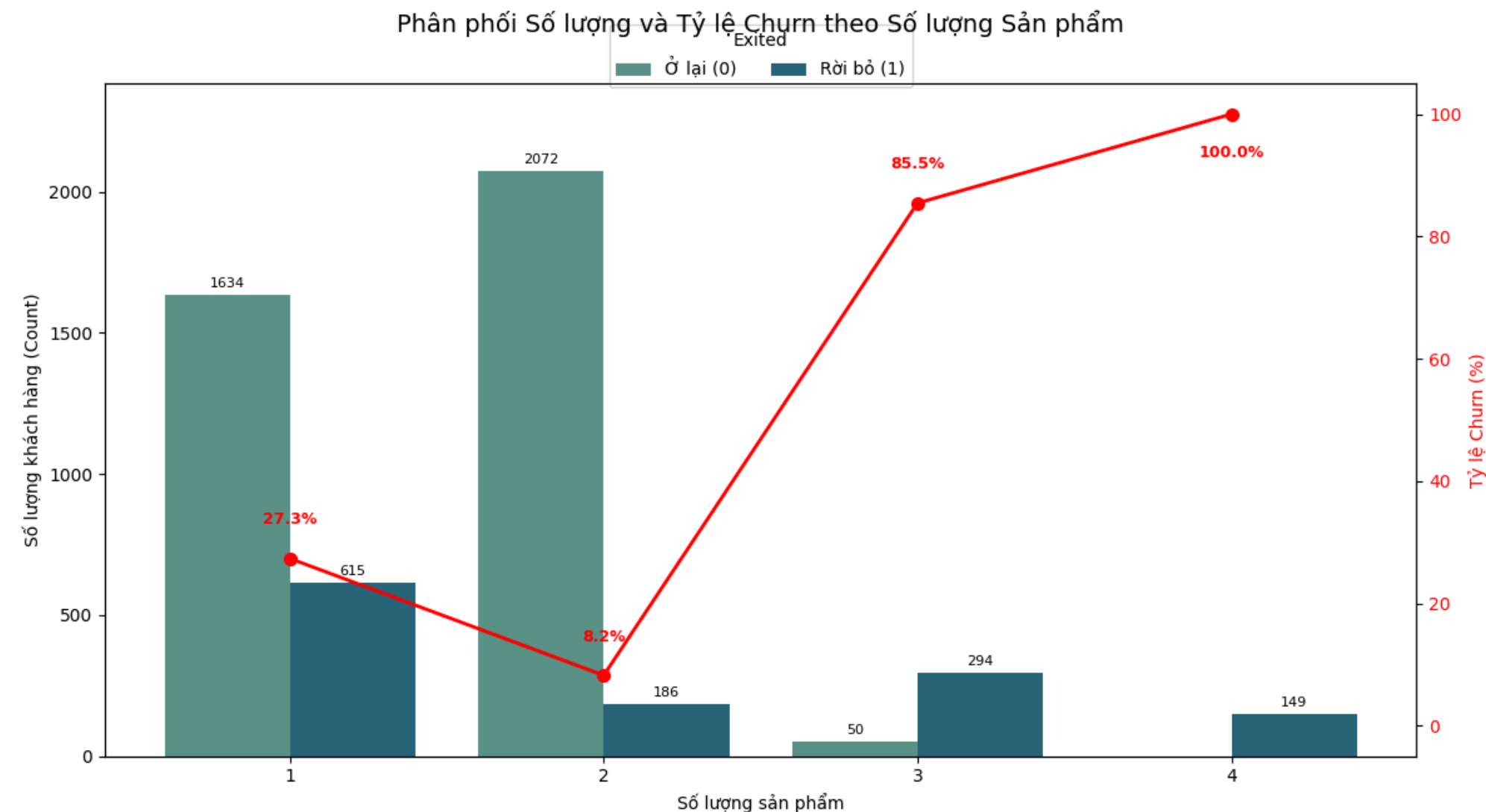
- Tỷ lệ churn: 27.3%
- Rời bỏ nhiều nhất: 615 người

→ Nên đẩy mạnh nâng lên 2 sản phẩm.

Báo động đỏ – 3 và 4 sản phẩm:

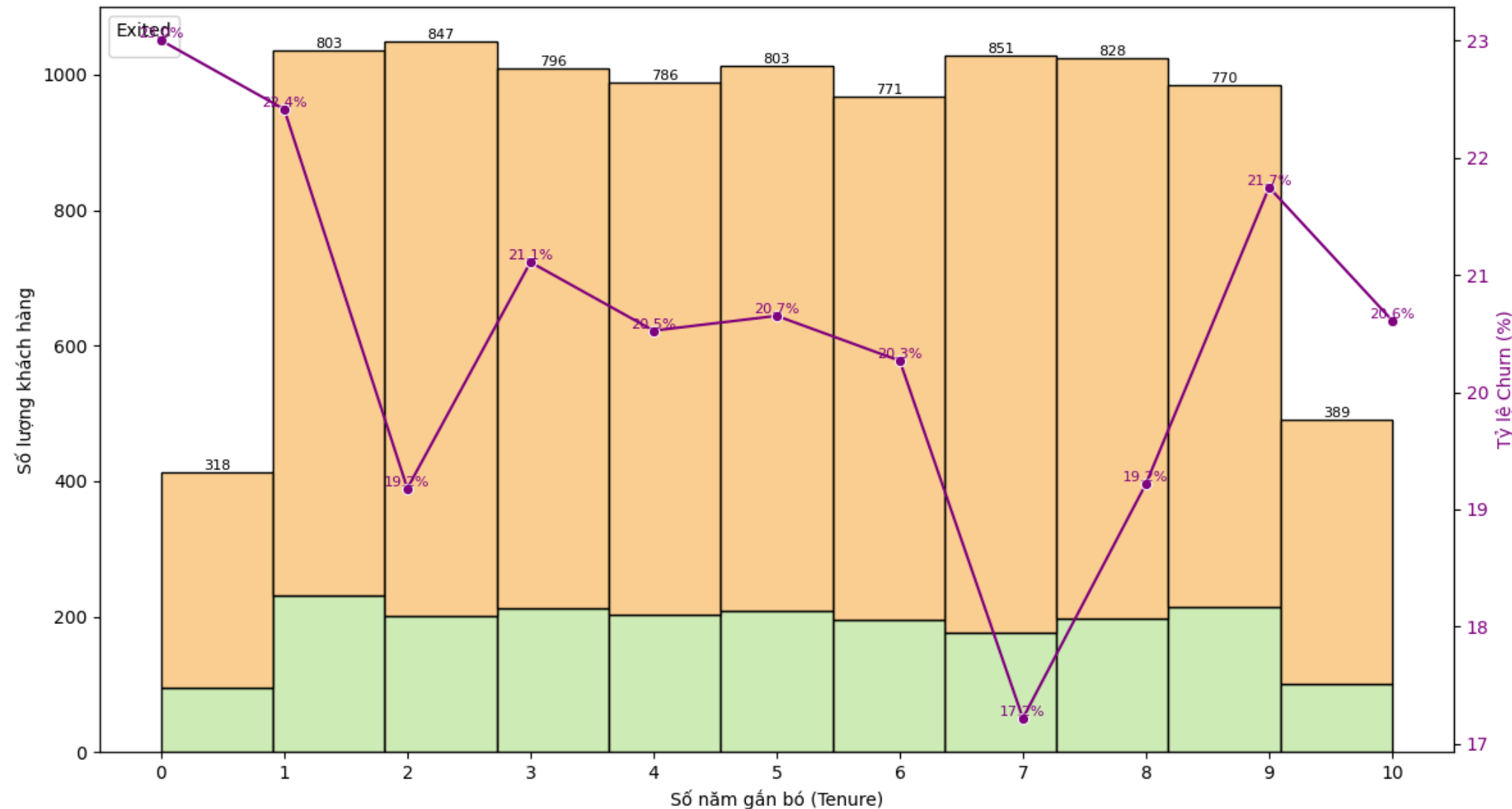
- 3 sản phẩm: Churn 85.5%
- 4 sản phẩm: Churn 100%

→ Phải điều tra gấp: Có thể do trải nghiệm kém hoặc dấu hiệu rút khỏi hệ thống.



Tenure (Số năm gắn bó với ngân hàng)

Phân phối và Tỷ lệ Churn theo Thời gian gắn bó (Tenure)



Nhóm khách hàng mới (gắn bó < 1 năm):

- Tỷ lệ rời bỏ cao nhất : 23%

→ là nhóm cần đặc biệt chú trọng trong việc giữ chân.

Nhóm khách hàng gắn bó từ 1-6 năm :

- Tỷ lệ rời bỏ trong dao động : 19%-22%

→ Có xu hướng giảm dần khi càng gắn bó lâu năm

Nhóm khách hàng gắn bó 7 năm :

- Tỷ lệ rời bỏ thấp nhất : 17,22%

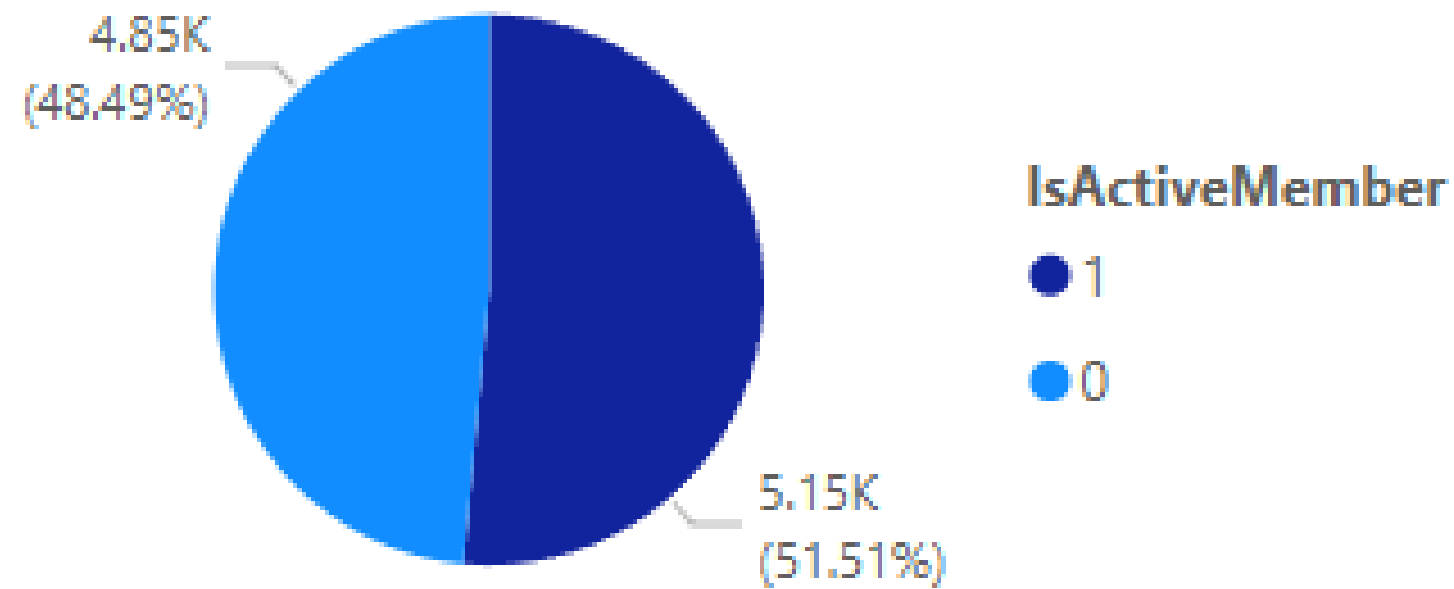
→ Đây là nhóm khách hàng trung thành cao

Nhóm khách hàng gắn bó sau 8 năm :

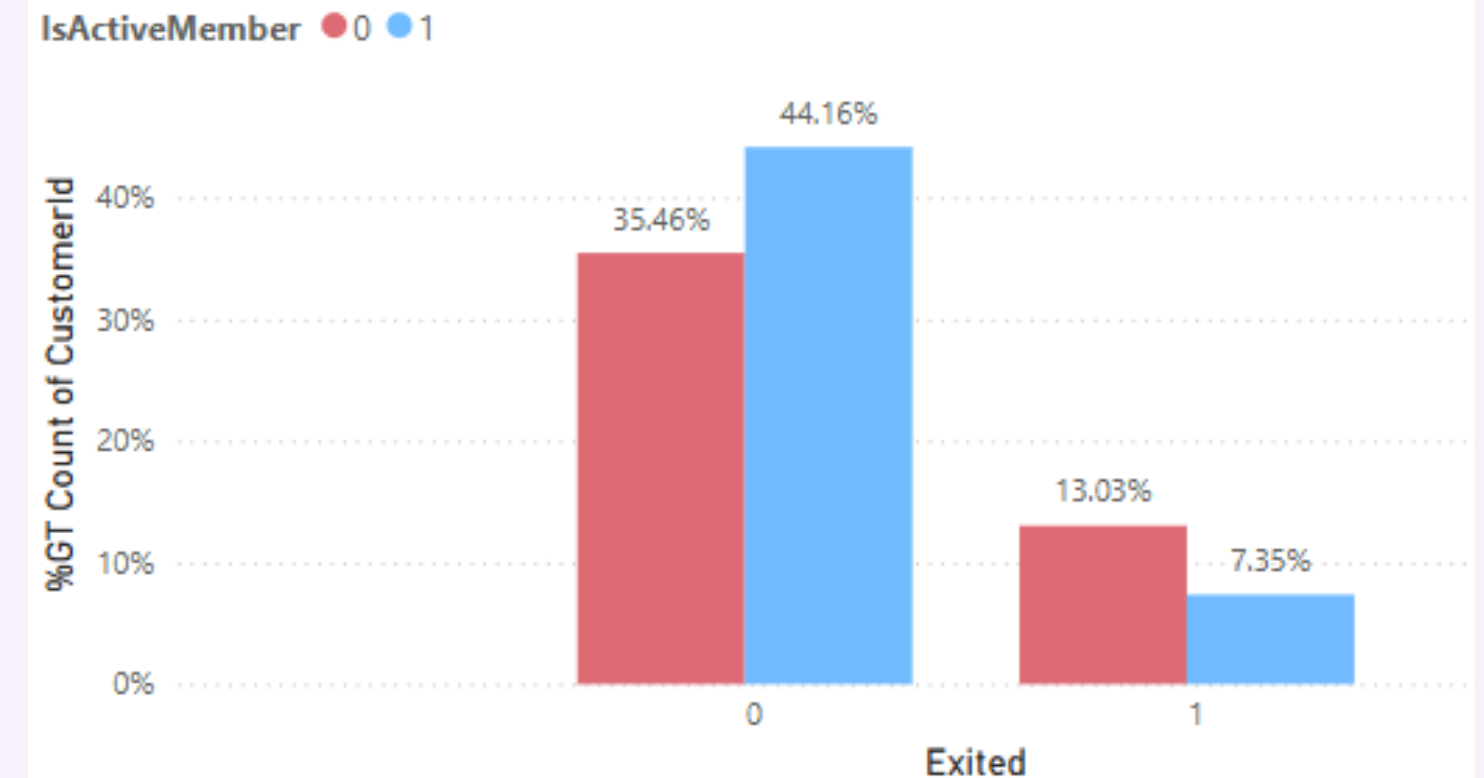
- Tỷ lệ rời bỏ có xu hướng tăng trở lại và năm thứ 9 có tỷ lệ tăng cao nhất

IsActiveMember (Tài khoản được kích hoạt)

Tỉ lệ khách hàng theo IsActiveMember



%GT Count of CustomerId by Exited and IsActiveMember



51,51% tổng số khách hàng vẫn còn sử dụng dịch vụ, bao gồm:

- Khách hàng hiện tại (chưa rời bỏ) có giao dịch đang hoạt động.
- Một phần khách hàng đã rời bỏ, nhưng vẫn duy trì tài khoản hoặc thực hiện giao dịch nhỏ lẻ.

Những khách hàng chưa rời bỏ :

- 35,46% không còn giao dịch tại ngân hàng , dù vẫn còn tài khoản
- 44,16% vẫn còn giao dịch bình thường, cho thấy mức độ hoạt động tốt.

Những khách hàng đã rời bỏ :

- 13,03% ngừng tất cả các giao dịch tại ngân hàng
- 7,35 % vẫn duy trì tài khoản mở nhưng không còn hoạt động chính (có thể chỉ để giữ tài khoản hoặc chờ rút hết tiền).

Các yếu tố tài chính và sản phẩm

- **Nhận xét các yếu tố ảnh hưởng đến churn:**

Balance:

- Churn cao nhất ở nhóm số dư trung bình (95,667k–144,57k).
- Nhóm có số dư cao nhất (>144,57k) churn thấp → khách VIP, cần duy trì.

CreditScore:

- Credit cao → churn thấp.
- Tập trung giữ chân nhóm “Good” và “Very Good”.
- Nhóm “Poor” rủi ro cao, cần hỗ trợ.

NumOfProducts:

- 2 sản phẩm churn thấp nhất (8.2%) → nhóm vàng.
- 1 sản phẩm churn cao (27.3%), từ 3 trở lên: rất rủi ro (85.5%–100%).

Tenure:

- Gắn bó lâu hơn → churn giảm.
- Cao nhất ở khách mới (<1 năm) và tăng trở lại sau 8 năm.

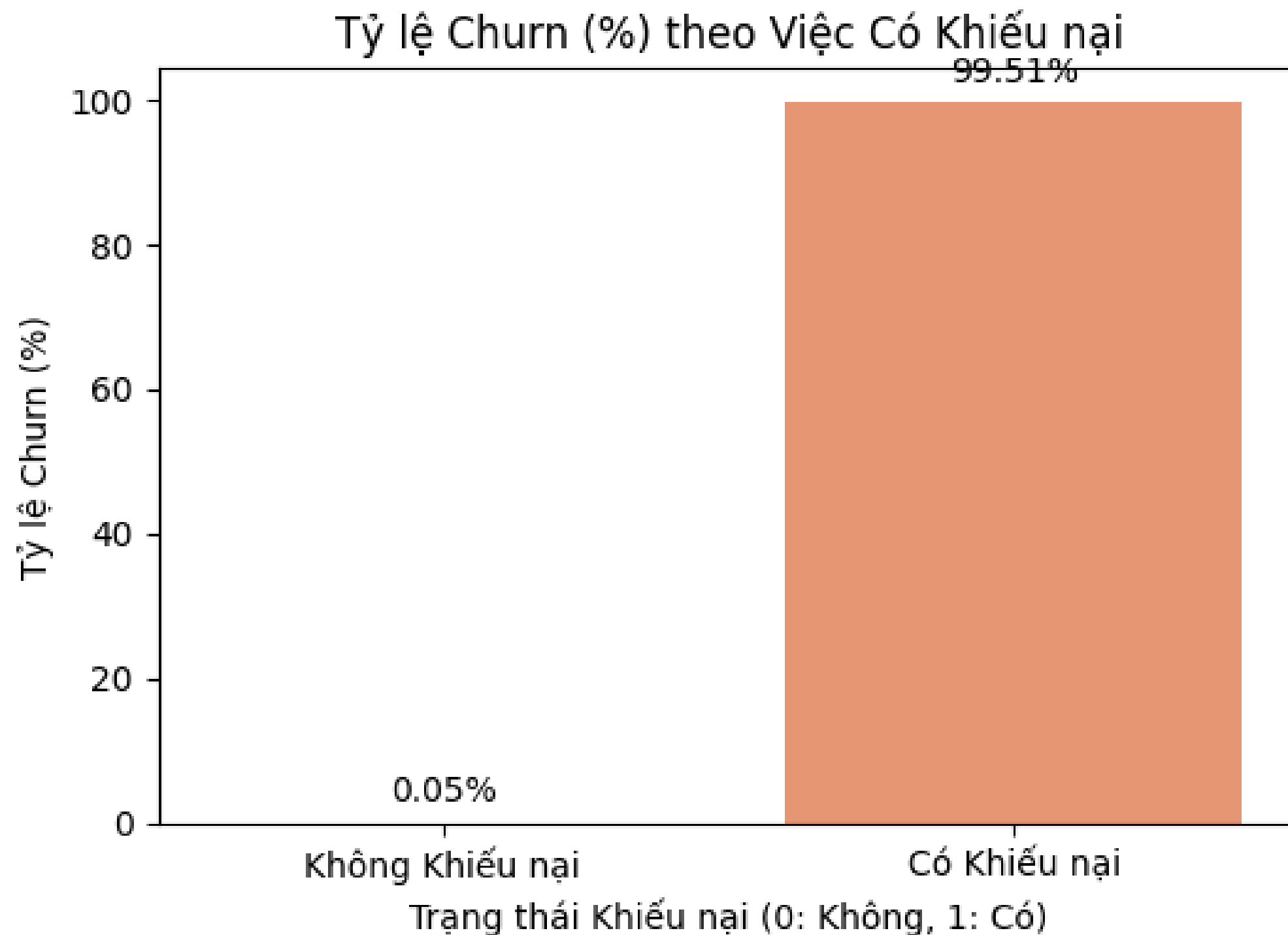
IsActiveMember:

- 51,51% còn hoạt động.
- Khách chưa rời: 44,16% còn giao dịch, 35,46% không hoạt động.
- Khách đã rời: 13,03% ngừng hẳn, 7,35% còn tài khoản nhưng không dùng.

Phân tích trải nghiệm dịch vụ và khiếu nại



Phân tích về tỷ lệ churn khi KH có khiếu nại



Có 2 nhóm:

Nhóm "Không Khiếu nại":

- Tỷ lệ churn chỉ ở mức **0.05%**
- Khách hàng không có khiếu nại có mức độ hài lòng hoặc trung thành rất cao.

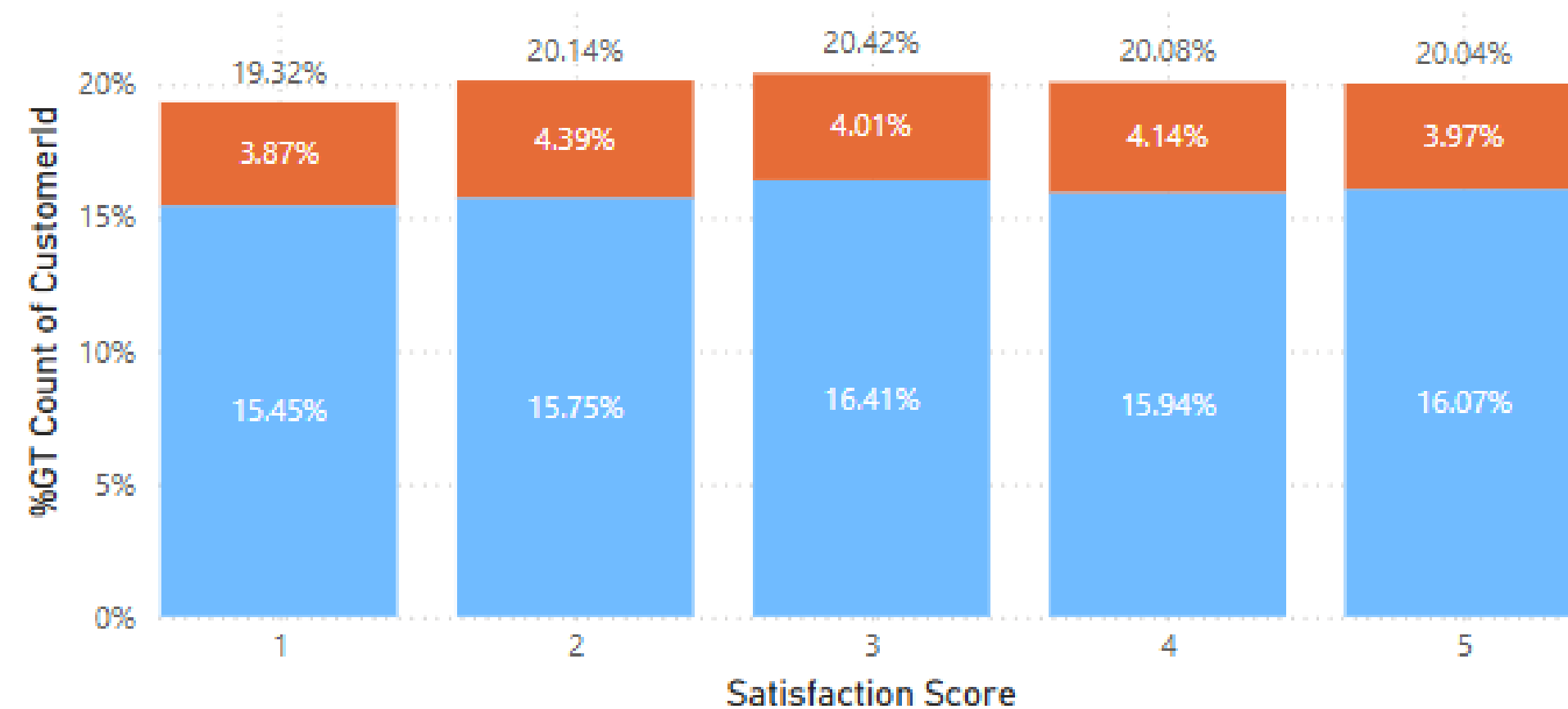
Nhóm "Có Khiếu nại":

- Tỷ lệ churn cao tới **99.51%**
- Hầu như 100% những khách hàng đã từng khiếu nại đều đã rời bỏ dịch vụ.

Phân tích về tỷ lệ churn khi KH có khiếu nại

Tỷ lệ rời bỏ/ở lại của khách hàng sau khi complain

Exited ● 0 ● 1



- Điểm 3 chiếm tỷ lệ cao nhất **20.42%**, điểm 1 thấp nhất **19.32%**
 - Tỷ lệ khách hàng ở lại khá ổn định dao động 15.45% - 16.41%.
 - Tỷ lệ khách hàng rời bỏ **không biến động nhiều** dao động trong 3.87%- 4.39%
 - Tỷ lệ Churn theo nhóm khá tương đồng nhau, đều dao động quanh mức 20-22%.
- Không có xu hướng rõ ràng. Điểm hài lòng càng cao thì tỷ lệ churn càng thấp.

Trải nghiệm dịch vụ và khiếu nại

Nhận xét

- Khách hàng khiếu nại là tín hiệu cực kỳ nguy hiểm, khả năng rời bỏ gần như chắc chắn **99.51%**
- Mức độ hài lòng không phải là chỉ số đáng tin cậy để dự đoán khả năng giữ chân họ.
- Tỷ lệ rời bỏ ~20% ở nhóm hài lòng nhất → vấn đề nằm ở trải nghiệm tổng thể hoặc lòng tin đã mất

Kết luận chung và giải pháp

Tỷ lệ khách hàng rời bỏ dao động quanh mức ~20%

Ảnh hưởng trực tiếp đến doanh thu, chi phí thay thế khách hàng

Tỷ lệ khách hàng rời bỏ mang tính hệ thống, không đơn thuần là yếu tố cá nhân

Bao gồm đặc điểm cá nhân, tình hình tài chính, hành vi sử dụng sản phẩm, mức độ gắn kết và trải nghiệm của họ với dịch vụ ngân hàng.

Tỷ lệ khách hàng rời bỏ sau khi khiếu nại (99,51%)

Xử lý khiếu nại kém hiệu quả

Giải pháp



Xây dựng quy trình phản hồi về khiếu nại trong 24 giờ đầu tiên.



Nâng cao chất lượng dịch vụ về trải nghiệm khách hàng



Triển khai hệ thống cảnh báo khách hàng có nguy cơ rời bỏ



Xây dựng mô hình “Cá nhân hóa” cho từng khách hàng

The image features a solid blue background. In the top right corner, there is a cluster of three overlapping hexagonal shapes: a light blue one at the bottom, a purple one in the middle, and a white one at the top. In the bottom left corner, there is another cluster of three overlapping hexagonal shapes: a white one at the top, a purple one in the middle, and a light blue one at the bottom. Centered in the middle of the image is the text "Thanks for watching" in a white, sans-serif font.

Thanks for watching