

BÁO CÁO PHÂN TÍCH DỮ LIỆU

NGÂN HÀNG ACB

by Quang Vinh

CONTENT

- BỐI CẢNH - MỤC ĐÍCH
- DATA
- EDA
- MÔ HÌNH DỰ ĐOÁN
- KẾT LUẬN

I. BỐI CẢNH - MỤC ĐÍCH



- **Bối cảnh:**

Các ngân hàng muốn giữ chân khách hàng của mình để duy trì hoạt động kinh doanh

- **Mục đích:**

Phân tích đặc điểm và tìm hiểu nguyên nhân tại sao khách hàng lại có xu hướng churn, đưa ra đề xuất cải thiện tình trạng khách hàng churn

II. DATA

BANK CUSTOMER DATA DICTIONARY

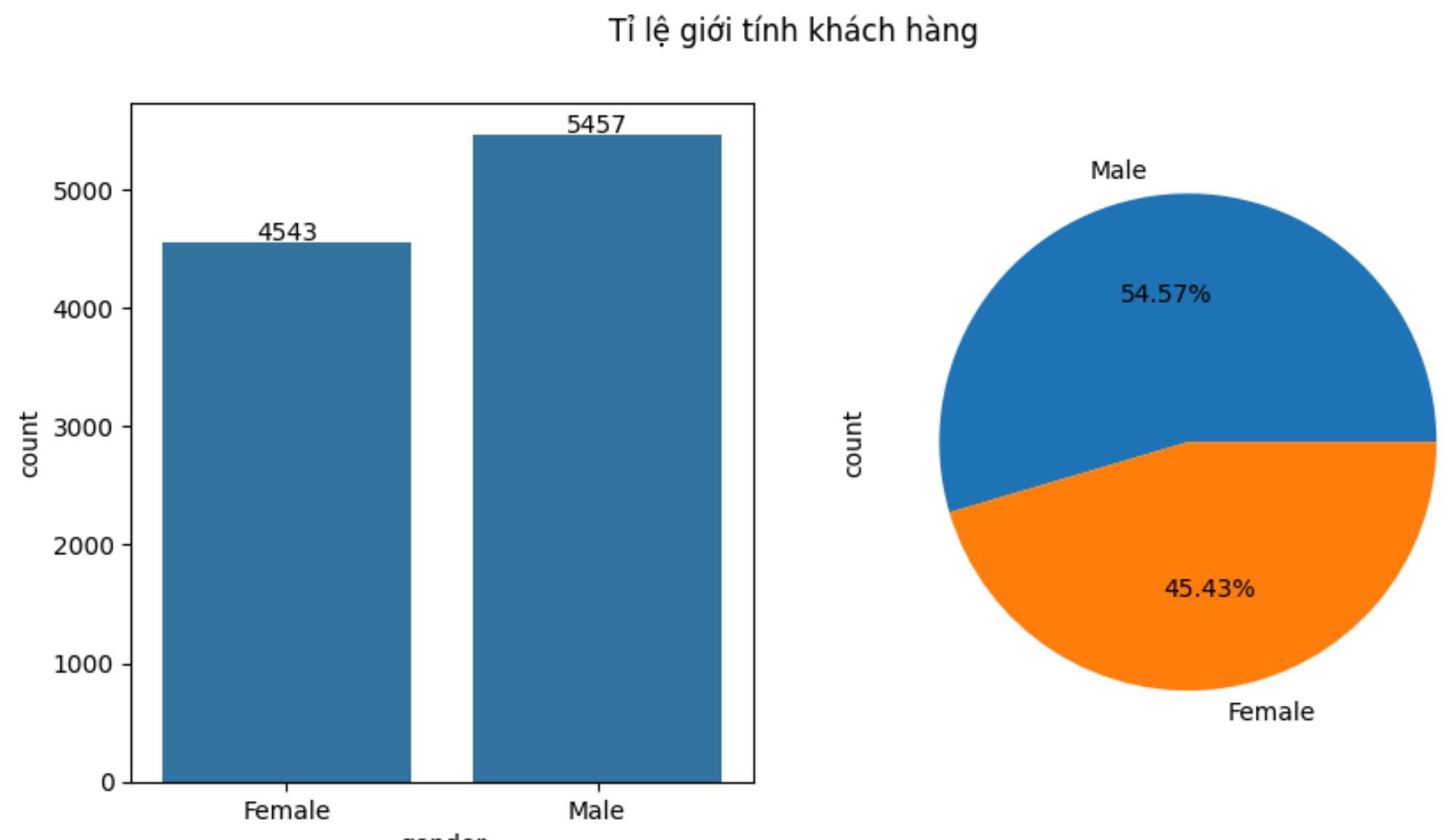
Column name	Description
customer_id	Account Number
credit_score	Credit Score
country	Country of Residence
gender	Sex
age	Age
tenure	From how many years he/she is having bank acc in ABC Bank
balance	Account Balance
products_number	Number of Product from bank
credit_card	Does this customer have a credit card ?
active_member	Is he/she an active Member of the bank ?
estimated_salary	Salary of Account holder
churn	Churn Status

```
Data columns (total 12 columns):
 #   Column           Non-Null Count Dtype  
 --- 
 0   customer_id      10000 non-null  int64  
 1   credit_score     10000 non-null  int64  
 2   country          10000 non-null  object  
 3   gender           10000 non-null  object  
 4   age              10000 non-null  int64  
 5   tenure           10000 non-null  int64  
 6   balance          10000 non-null  float64 
 7   products_number  10000 non-null  int64  
 8   credit_card      10000 non-null  int64  
 9   active_member    10000 non-null  int64  
 10  estimated_salary 10000 non-null  float64 
 11  churn            10000 non-null  int64  
 dtypes: float64(2), int64(8), object(2)
 memory usage: 937.6+ KB
```

Dữ liệu có 12 cột, 10 000 dòng
Dữ liệu không bị null, không bị duplicate

III. EDA

Tỷ lệ giới tính khách hàng

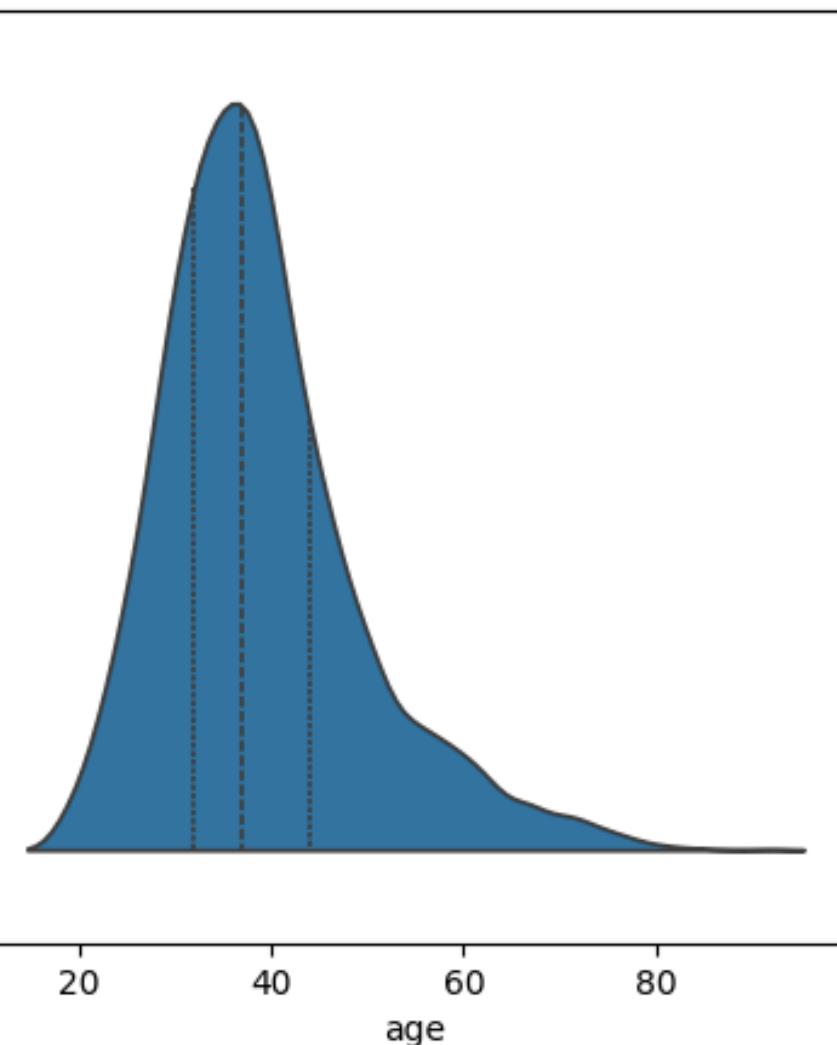
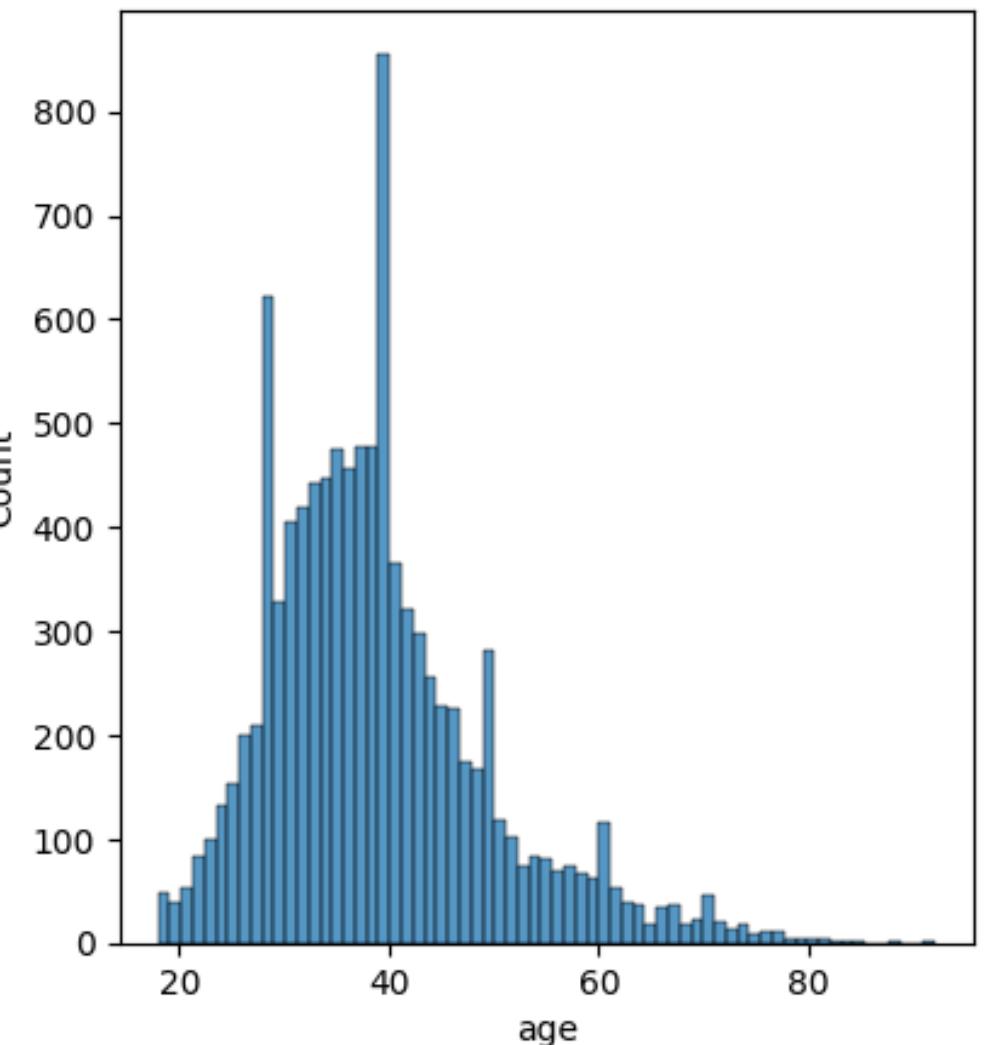


Hơn 1 nửa khách hàng sử dụng là NAM

Độ tuổi của khách hàng

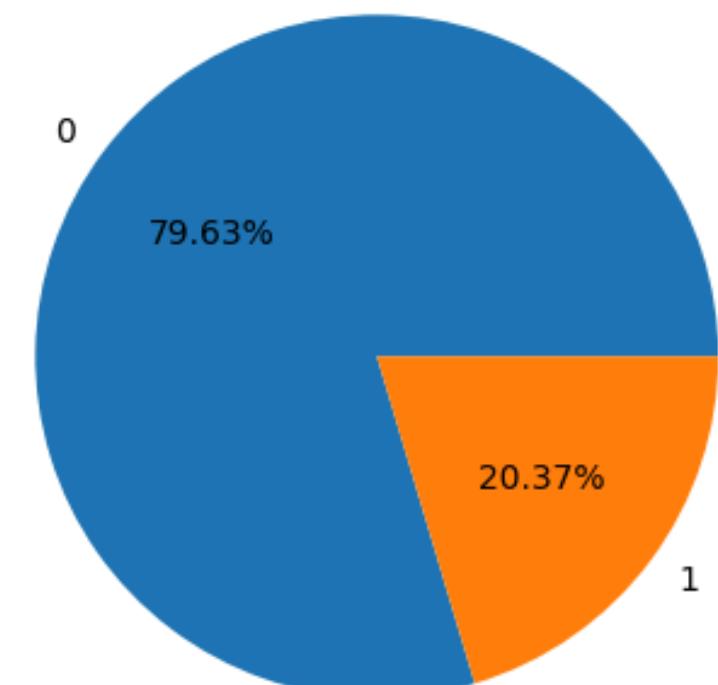
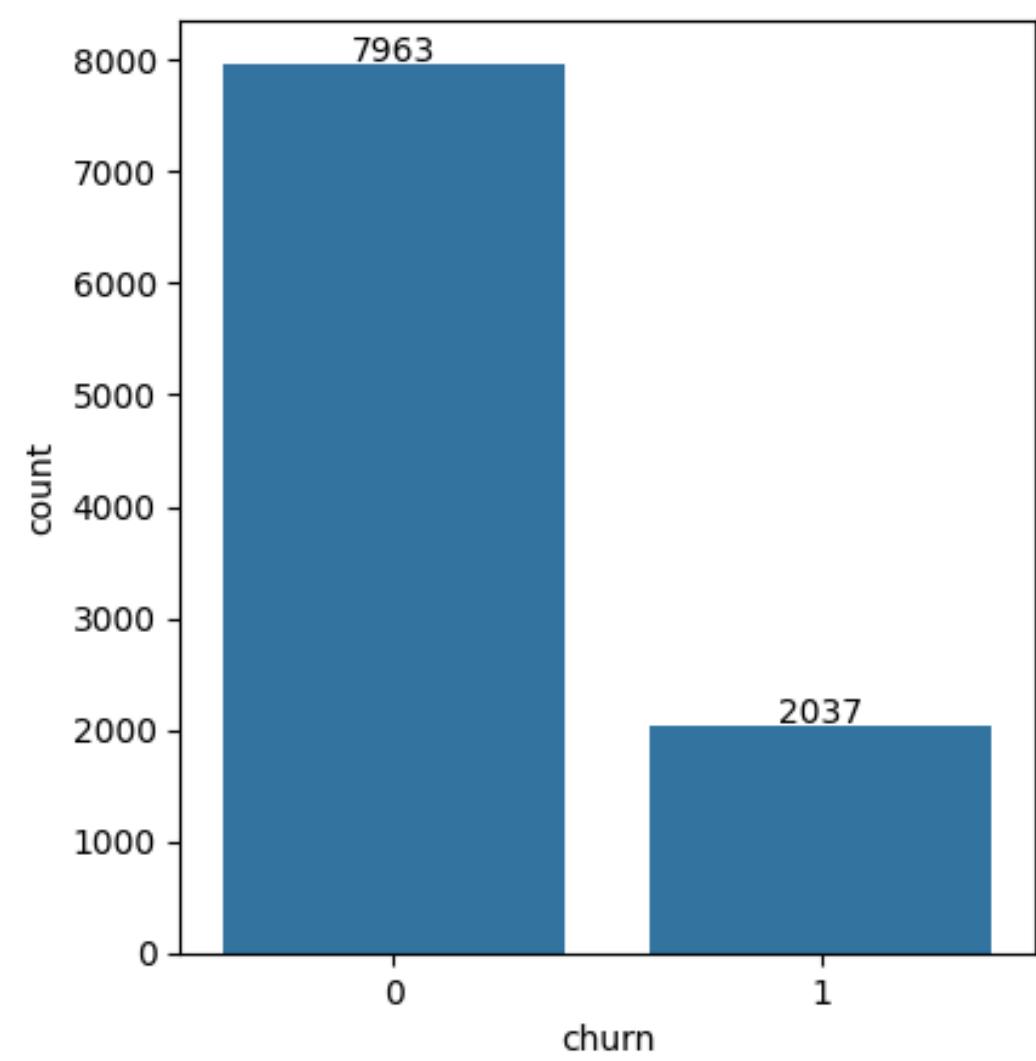
Tập trung từ 30 - gần 50 tuổi, thưa dần ở các người lớn hơn 60 tuổi

Độ tuổi của khách hàng



Tỷ lệ khách hàng rời bỏ sử dụng dịch vụ

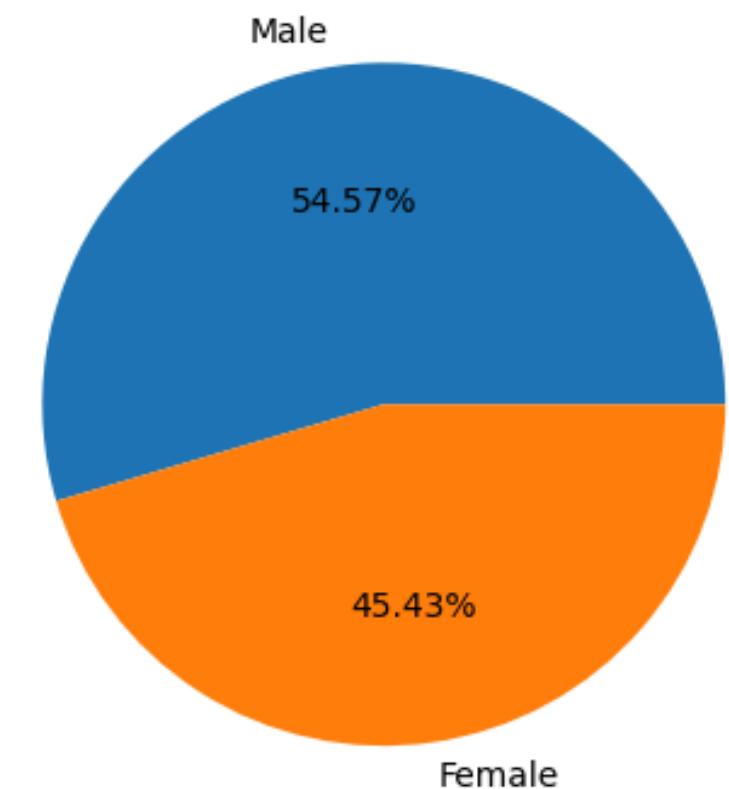
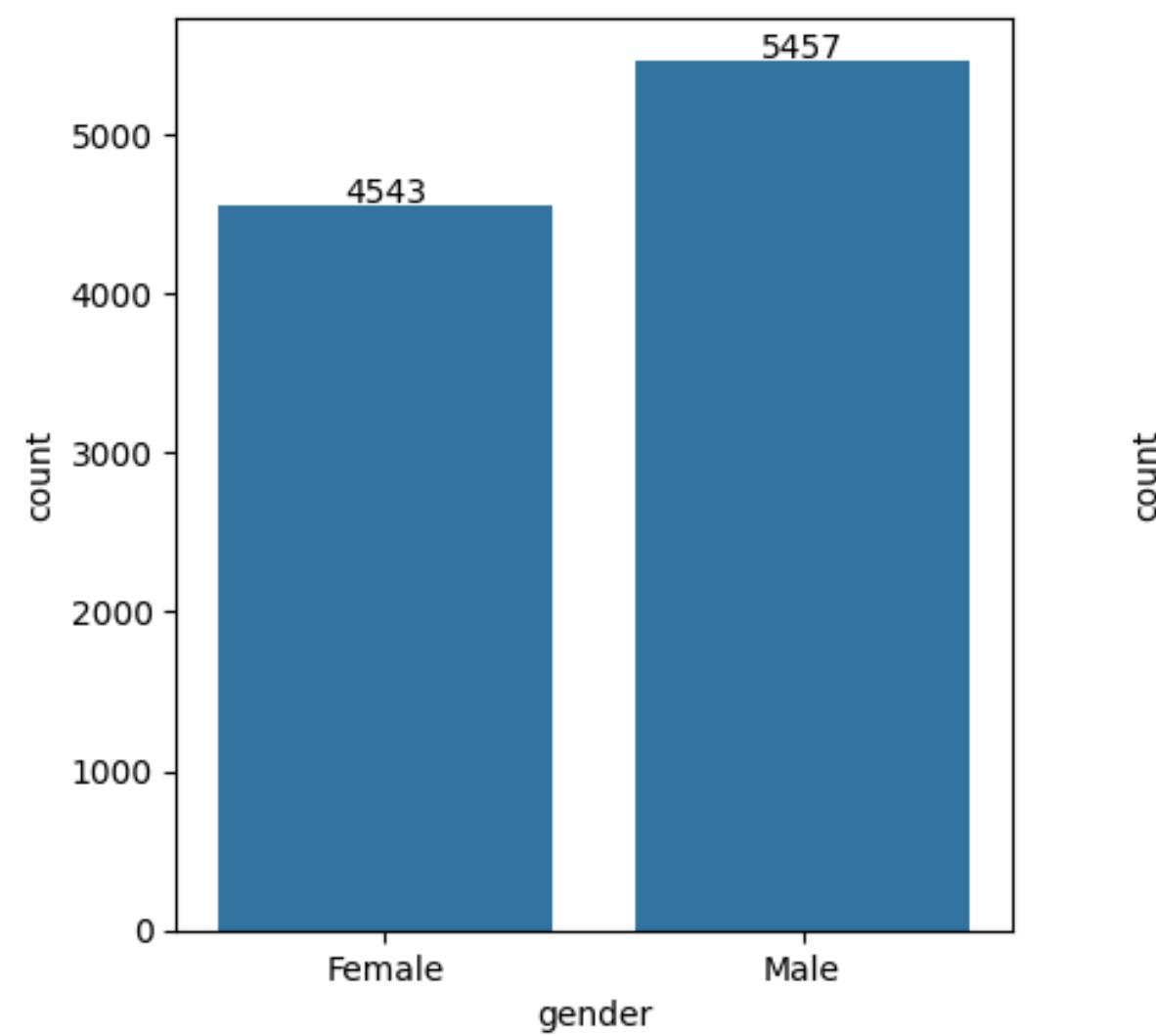
Tỷ lệ khách hàng rời bỏ



Churn rate của Bank:
Tỷ lệ khách hàng rời bỏ chiếm 20% trên
tổng số khách hàng của ngân hàng ACB

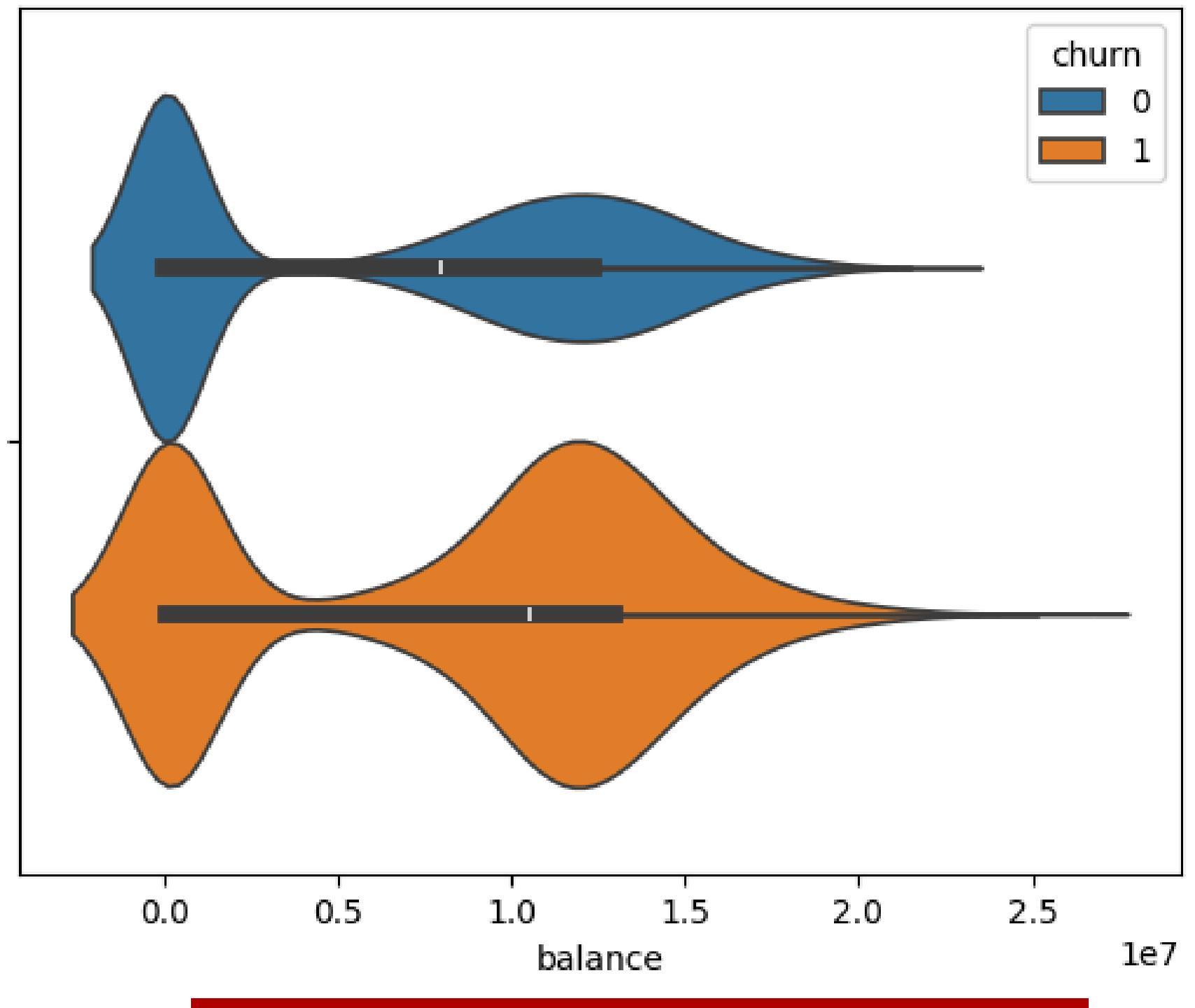
Tỷ lệ khách hàng phân bổ ở các quốc gia

Tỉ lệ giới tính khách hàng



Tập trung nhiều nhất ở France với con số
tỷ lệ tương ứng 5014 người, tiếp sau đó là
Spain và Germany => Khách hàng ở Pháp
chiếm hơn 50% trên tổng số

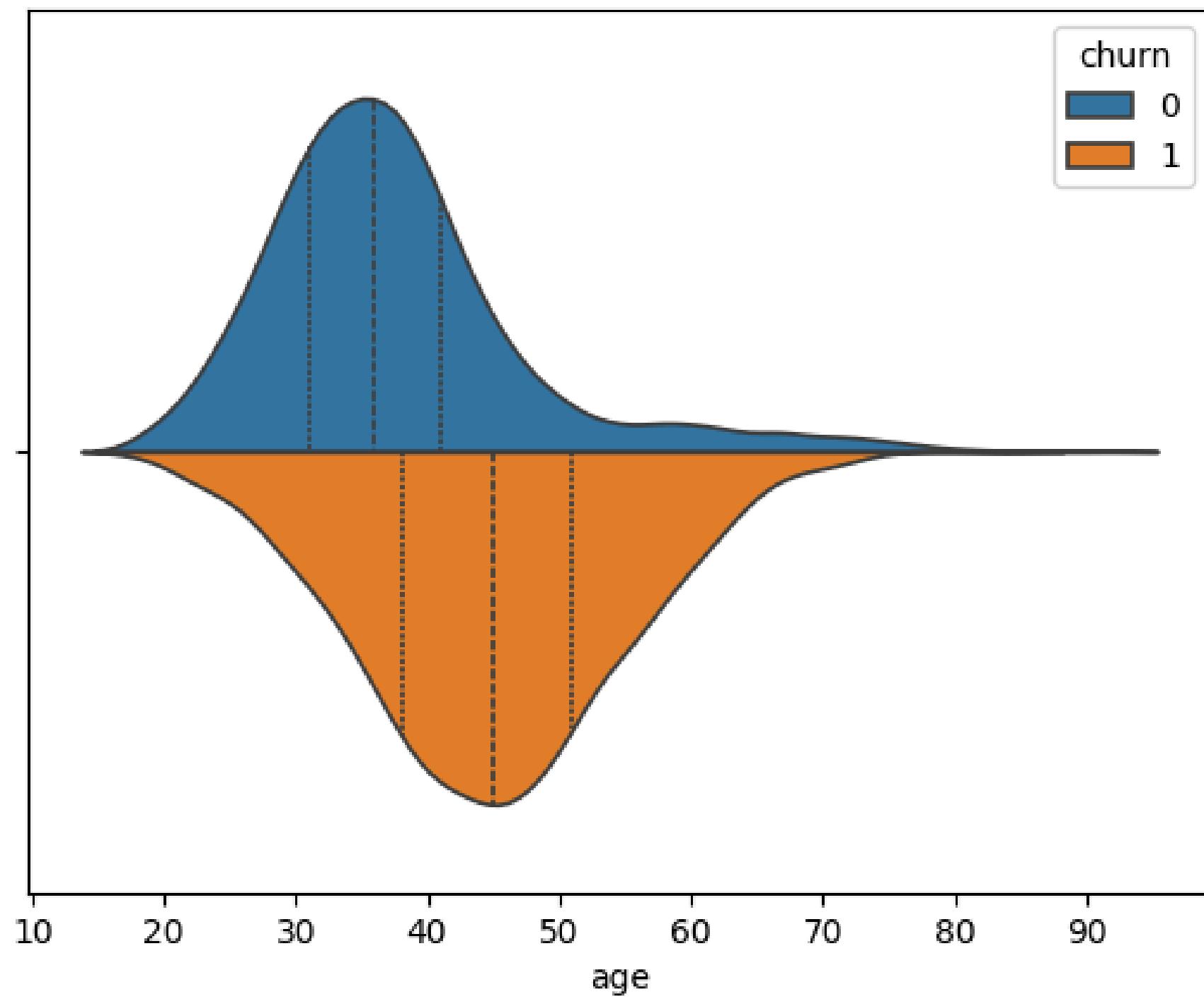
Số dư của khách hàng



Số dư của khách hàng

Số dư trong tài khoản ngân hàng của khách hàng tập trung nhiều nhất ở 0đ và từ 5 triệu đồng - 15 triệu đồng => khách hàng có số dư càng lớn, càng dễ có khả năng rời bỏ ngân hàng

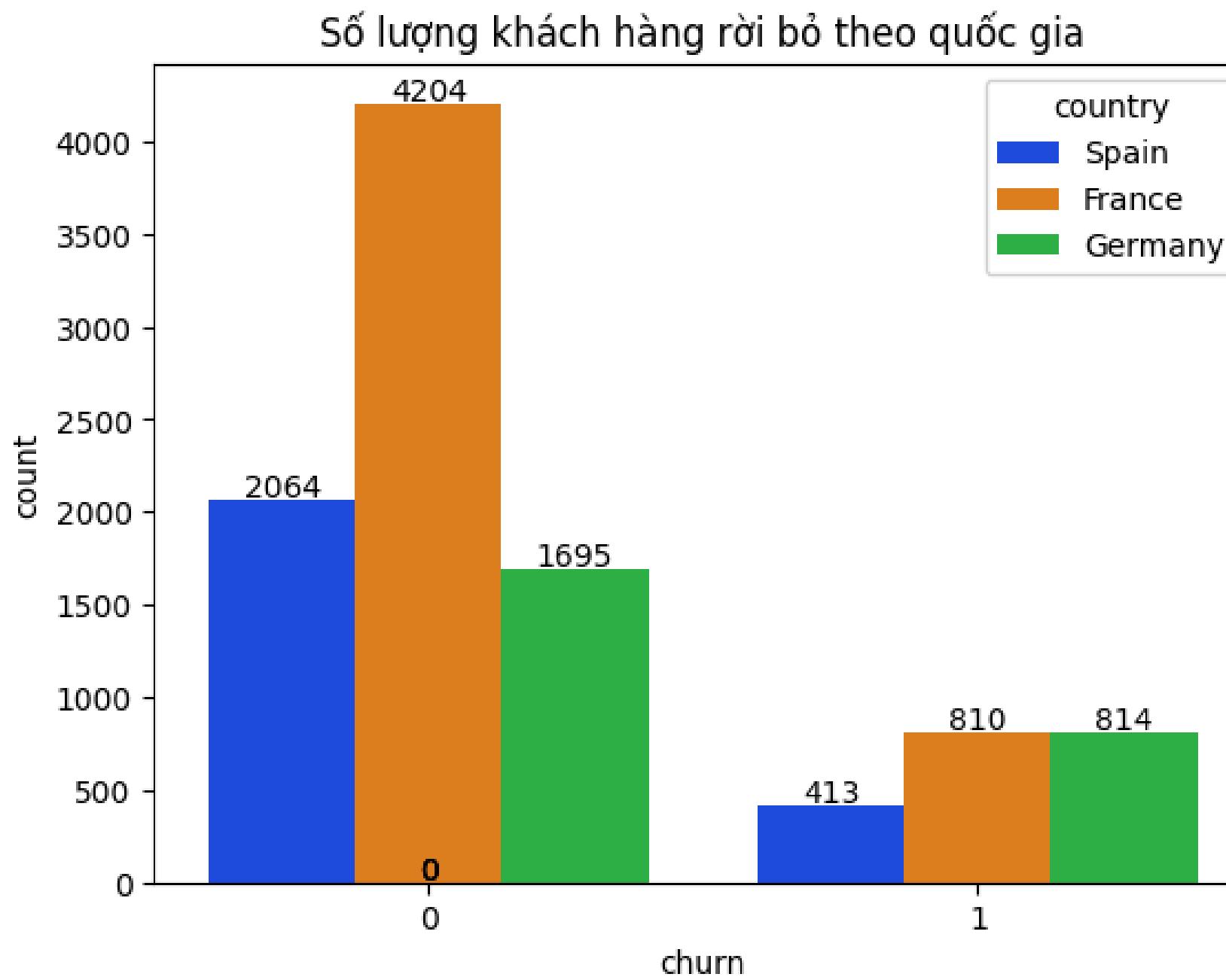
Độ tuổi khách hàng rời bỏ



Độ tuổi khách hàng rời bỏ

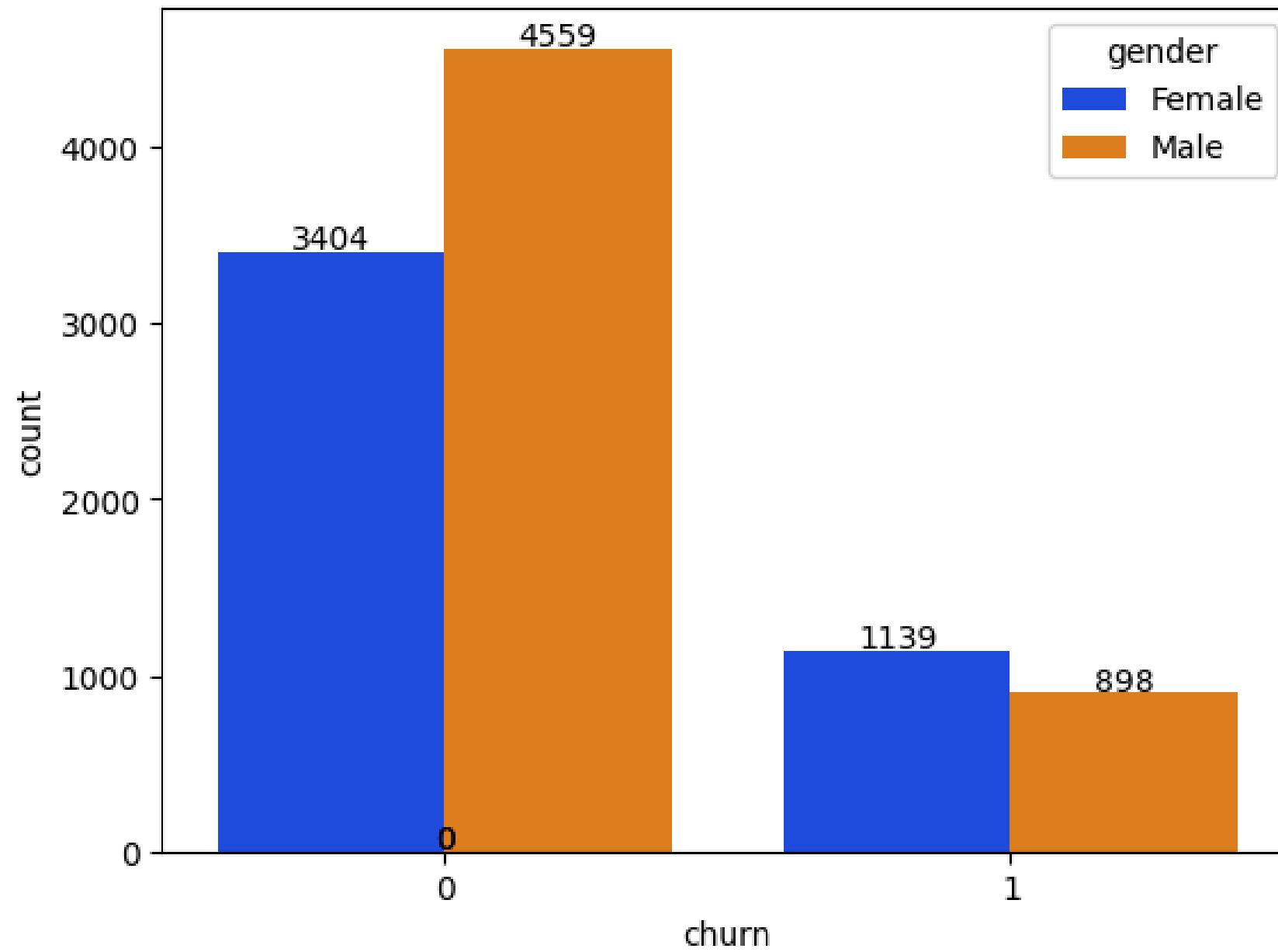
Độ tuổi từ 35-50 là khoảng thời gian được cho là khoảng tuổi mà người ta có xu hướng churn

Số lượng khách hàng rời bỏ theo quốc gia



Ở Đức có lượng khách hàng rời bỏ cao nhất

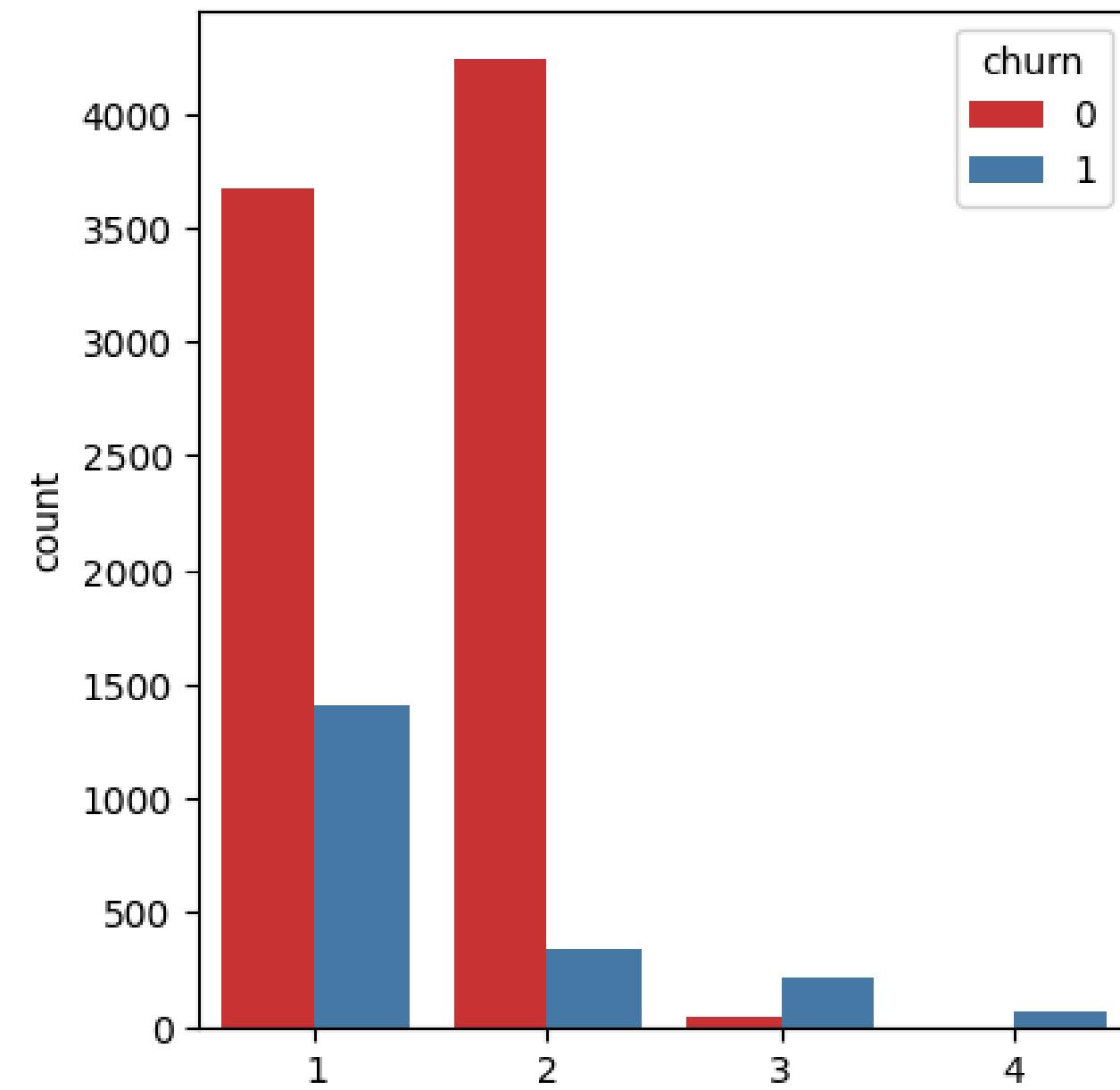
Số lượng khách hàng rời bỏ theo giới tính



Số lượng khách hàng rời bỏ theo giới tính

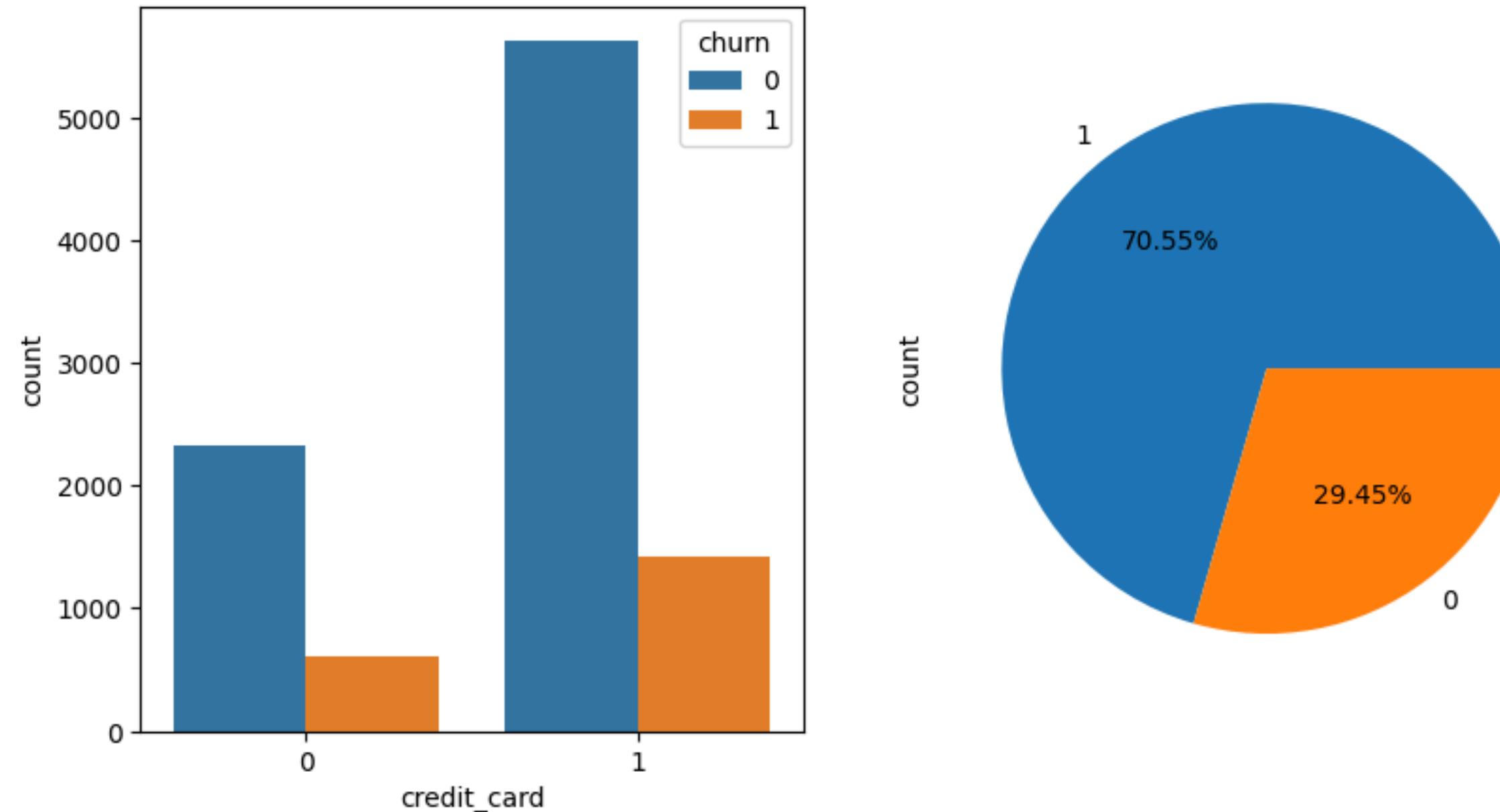
Nam có xu hướng churn nhiều hơn nữ

Số sản phẩm mà khách hàng sử dụng



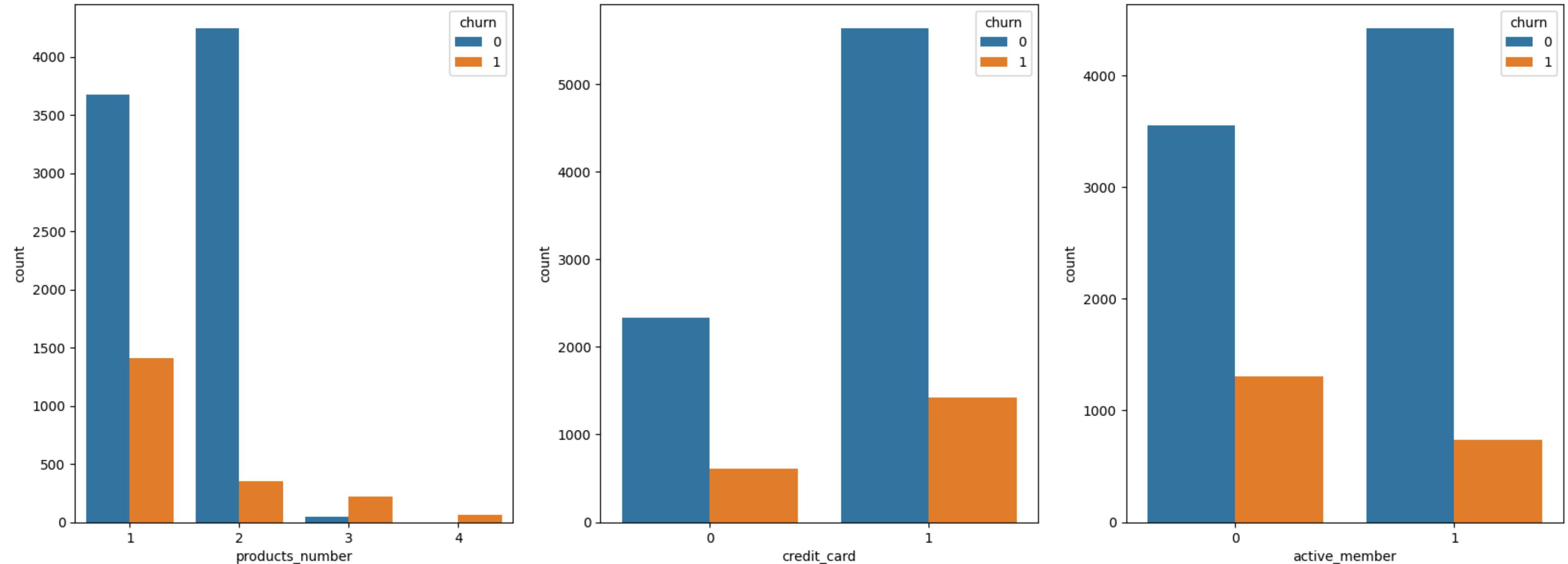
Tỷ lệ churn ở khách sử dụng 3-4 sản phẩm(dịch vụ) cao là 82,71% và 100% => Khách hàng càng sử dụng nhiều sản phẩm càng có khả năng churn thấp hơn

Số lượng khách hàng rời bỏ khi sử dụng thẻ tín dụng



Credit_card: Hơn 70% khách hàng sử dụng thẻ tín dụng
của ngân hàng và sau đó họ có khả năng rời bỏ thấp

Determining

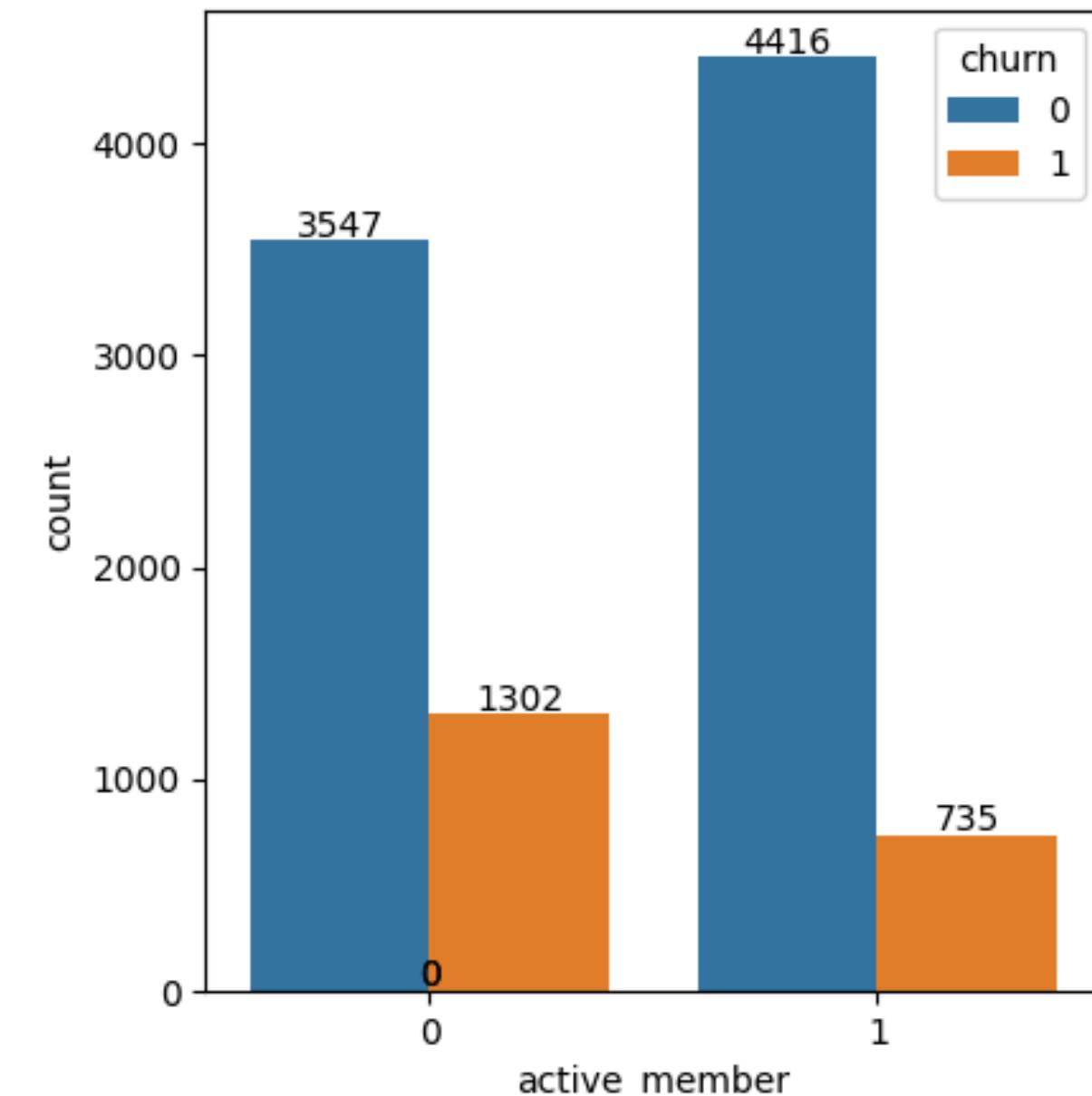
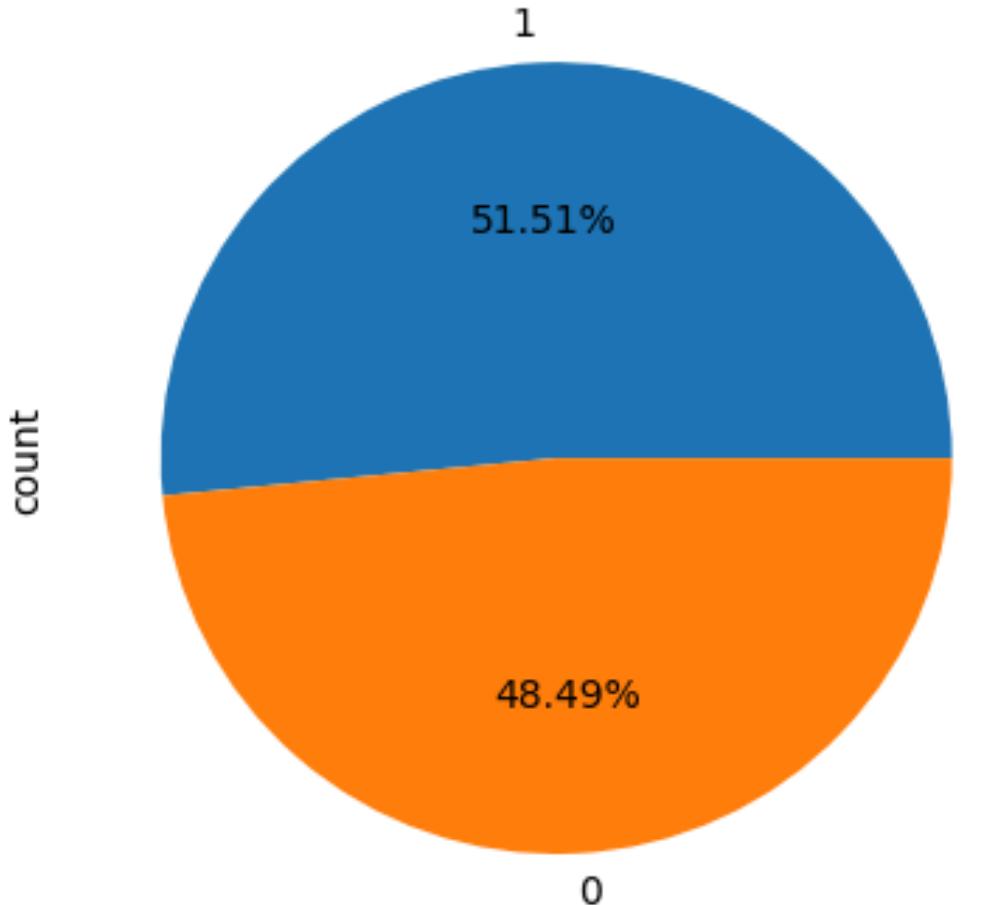


- Product_number: tỉ lệ churn ở khách sử dụng 3 – 4 sản phẩm thấp
- Credit_card:
- Active_number: không phải là thành viên của bank thì có tỉ lệ churn cao hơn là 26.85%

Thành viên hoạt động tích cực

Thành viên hoạt động tích cực

Tỷ lệ khách hàng hoạt động tích cực



Khách hàng khi đã trở thành thành viên của ngân hàng năm/hoạt động tích cực có xu hướng churn ít hơn (chiếm hơn 50%)

IV. Mô hình dự đoán

```
▶ train_and_eval(SVC(C= 1.0, kernel= "rbf"))
```

	precision	recall	f1-score	support
0	0.93	0.73	0.81	1595
1	0.42	0.77	0.54	405
accuracy			0.74	2000
macro avg	0.67	0.75	0.68	2000
weighted avg	0.82	0.74	0.76	2000

```
▶ train_and_eval(  
DecisionTreeClassifier(max_depth=10))
```

	precision	recall	f1-score	support
0	0.92	0.76	0.83	1595
1	0.43	0.73	0.54	405
accuracy			0.75	2000
macro avg	0.67	0.74	0.69	2000
weighted avg	0.82	0.75	0.77	2000

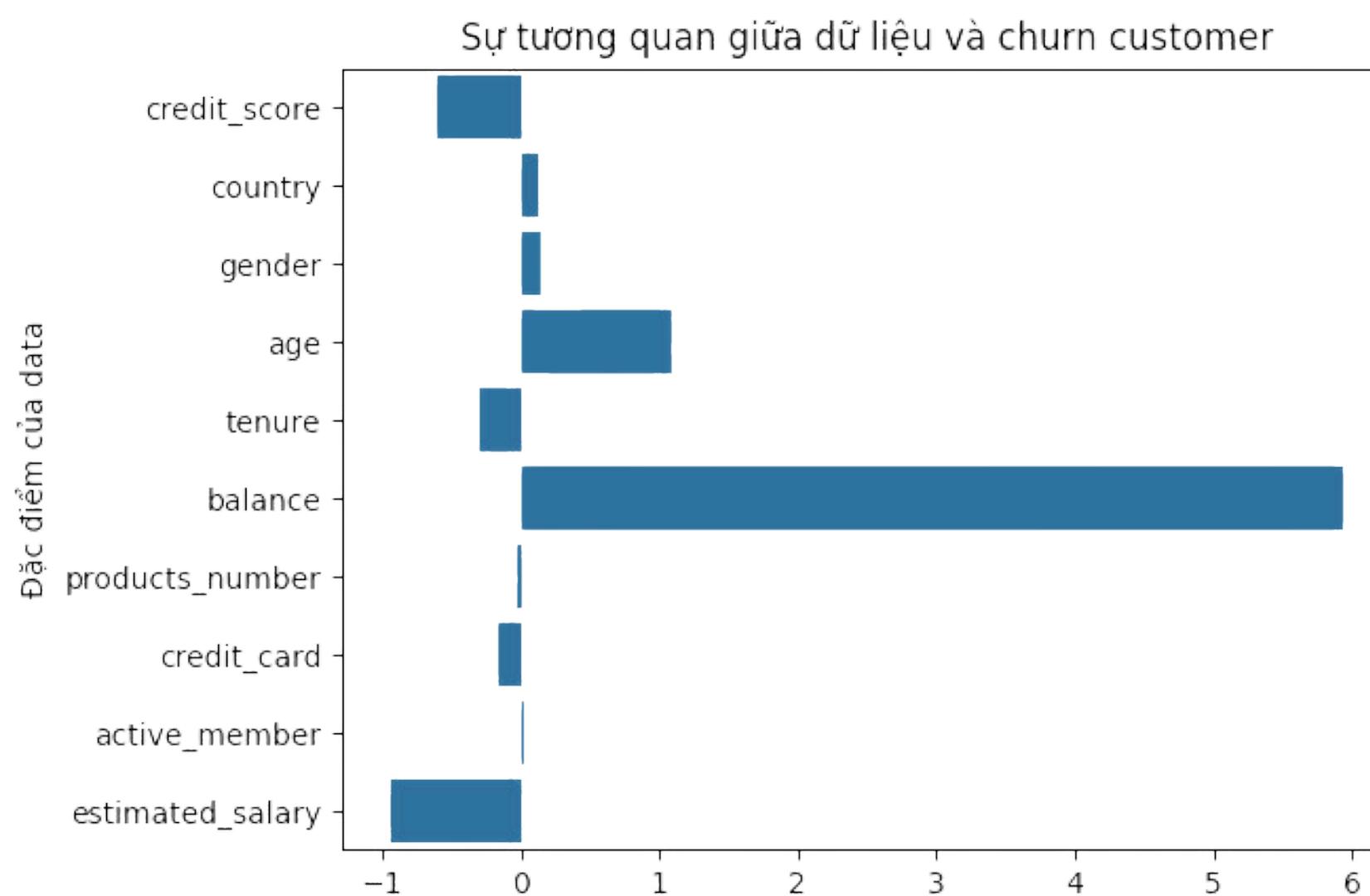
```
▶ train_and_eval(RandomForestClassifier  
(n_estimators=120, max_depth=10))
```

	precision	recall	f1-score	support
0	0.92	0.82	0.87	1595
1	0.51	0.73	0.60	405
accuracy			0.80	2000
macro avg	0.72	0.77	0.73	2000
weighted avg	0.84	0.80	0.81	2000

```
▶ lg_model = LogisticRegression()  
train_and_eval(lg_model)
```

	precision	recall	f1-score	support
0	0.90	0.71	0.79	1595
1	0.38	0.69	0.49	405
accuracy			0.70	2000
macro avg	0.64	0.70	0.64	2000
weighted avg	0.79	0.70	0.73	2000

Mối tương quan của các đặc điểm ảnh hưởng đến sự rời bỏ của khách hàng



Tương quan thuận: country, age, balance

Tương quan nghịch: product name, credit score, credit card, active member, gender

→ **Age** và **Balance** càng cao thì khả năng khách hàng rời bỏ càng lớn

→ **Product name** và **Active Member** càng thấp thì khả năng khách hàng rời bỏ càng cao

V. Đề xuất ý kiến

- Độ tuổi từ 35-50 là khoảng thời gian được cho là khoảng tuổi mà người ta có xu hướng churning vì thế nên có những chiến thuật và khuyến mãi khác biệt phù hợp cho từng độ tuổi để giữ chân khách hàng
- Cố gắng giữ chân những khách hàng cũ (ít hoặc miễn phí giao dịch, ck, quản lý,...)
- Lợi ích khi gia tăng lòng trung thành của khách hàng hiện tại bao gồm:
 - + Tiết kiệm chi phí thu hút khách hàng mới
 - + Giới thiệu truyền miệng (kênh mkt cực kỳ hiệu quả)
 - + Tăng số lượt mua trung bình/khách hàng

=> Trở thành thương hiệu ngân hàng số 1
- Tư vấn KH sử dụng đa dạng các sản phẩm của ngân hàng (training đội ngũ sale/bán hàng,...)





THANK YOU