

Phân tích hành vi khách hàng và tối ưu hóa chiến lược bán lẻ



Tổng quát

1.

Giới thiệu data

2.

Làm sạch data

3.

Phân tích tổng
quan

4.

Phân tích chuyên
sâu

5.

Đưa ra các chiến
lược

6.

Tổng kết



1. Giới thiệu về Dataset

Đây là bộ dataset được lấy từ kaggle nói về giao dịch bán lẻ. Nó chứa thông tin về khách hàng, mua hàng, sản phẩm và chi tiết giao dịch của họ. Để biết thêm chi tiết có thể tìm [tại đây](#).



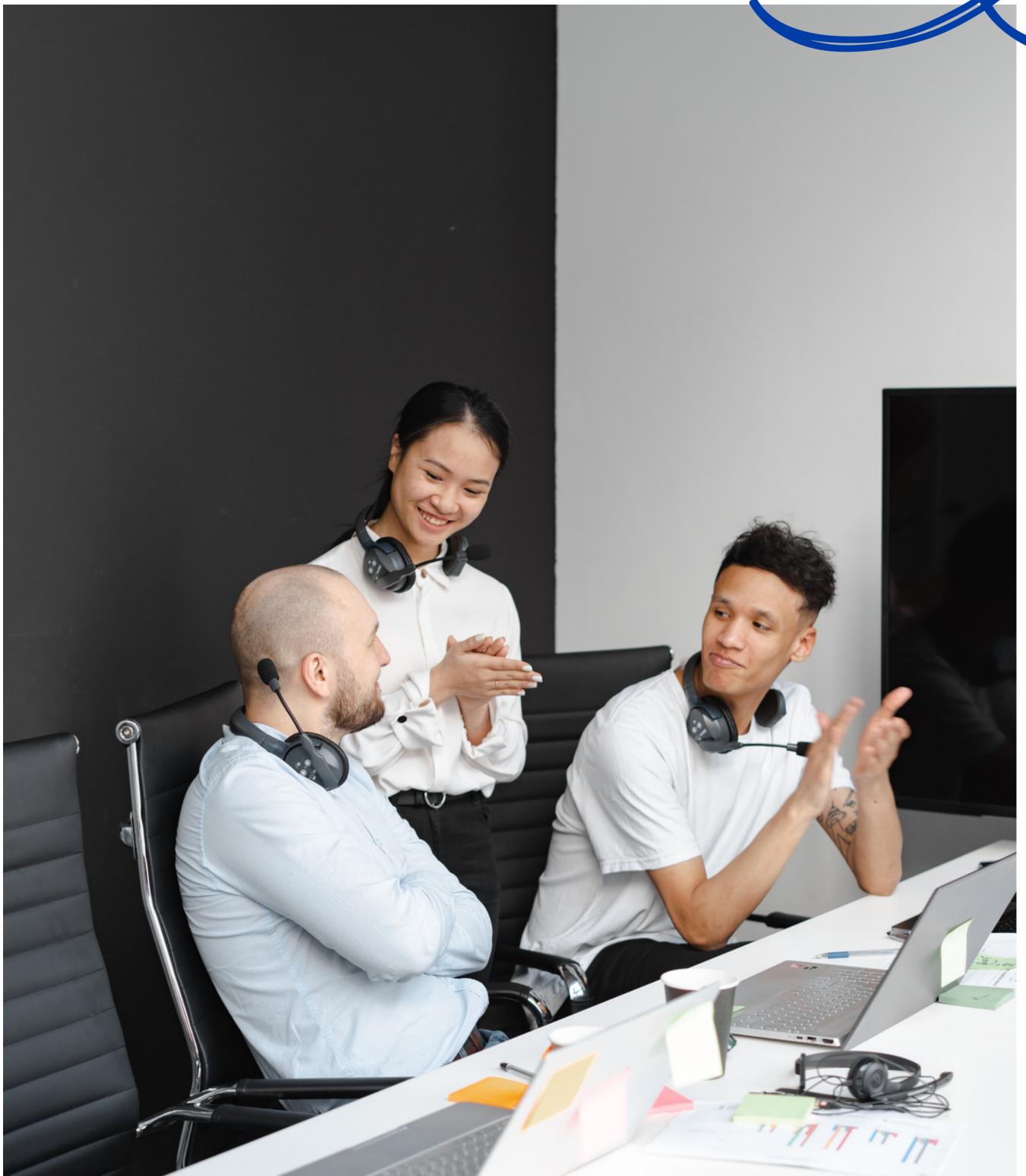
Giới thiệu về dataset

Dữ liệu bao gồm các 1 bảng và 30 cột chứa thông tin về khách hàng , trạng thái đơn hàng , sản phẩm . Đây là bộ dataset phù hợp dành cho ai muốn phân tích dữ liệu về tình hình mua sắm trên kênh bán lẻ

2.Xử lý dữ liệu

Các vấn đề của data

- Data khá là nặng
- Chứa giá trị null



Giải pháp

- Tách ra làm các bảng
- Drop các giá trị null
- Gom nhóm các khách hàng



Kết quả sau khi làm sạch data

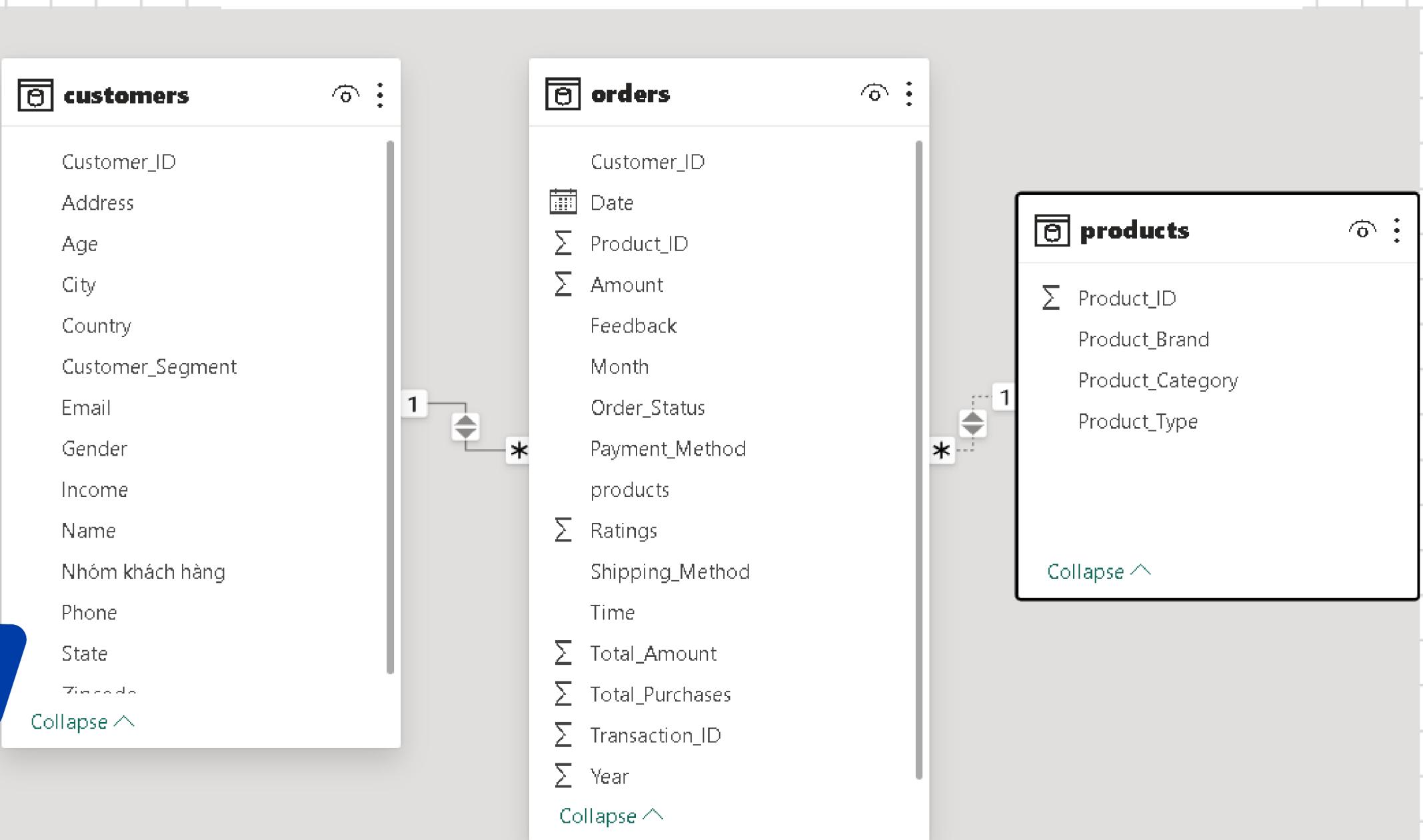
Số dòng , cột ban đầu

(302010, 30)

Sau khi xử lý

(293911, 30)

Tách 3 bảng bao gồm :
Customers, orders, products



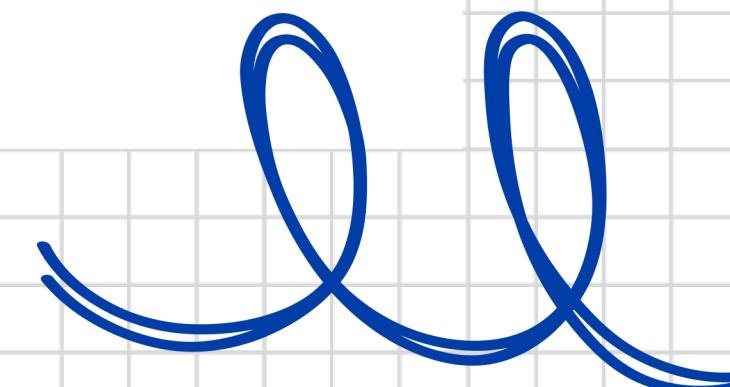
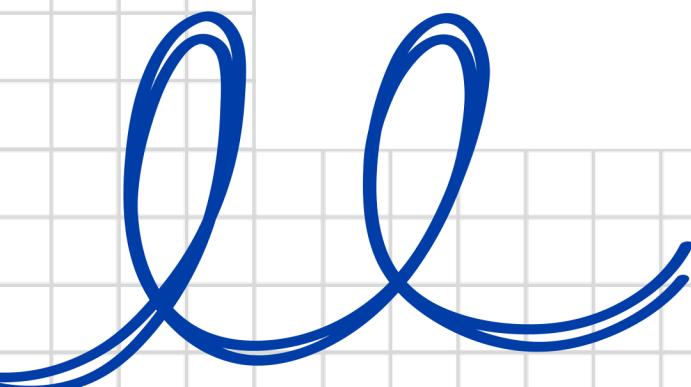
[Quay lại Trang Chương trình](#)



3.1 Phân tích tổng quan về khách hàng



Ở bảng customers chứa gồm 15 cột bao gồm:
**Customer_ID, Name, Gender, Age , Income, Address,
City, State, Address, Zipcode, Country, Phone, Email,
Customer_Segment, Nhóm khách hàng**



TỔNG QUAN VỀ KHÁCH HÀNG

86.49K

Count of Customer_ID

5

Count of Country

130

Count of City

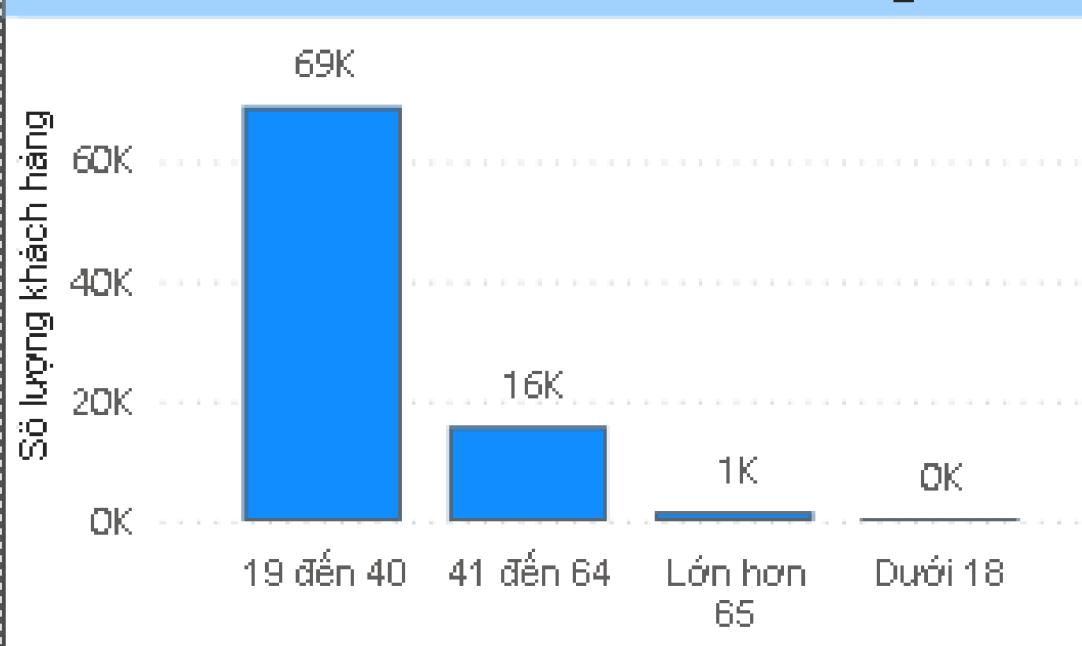
3

Count of Income

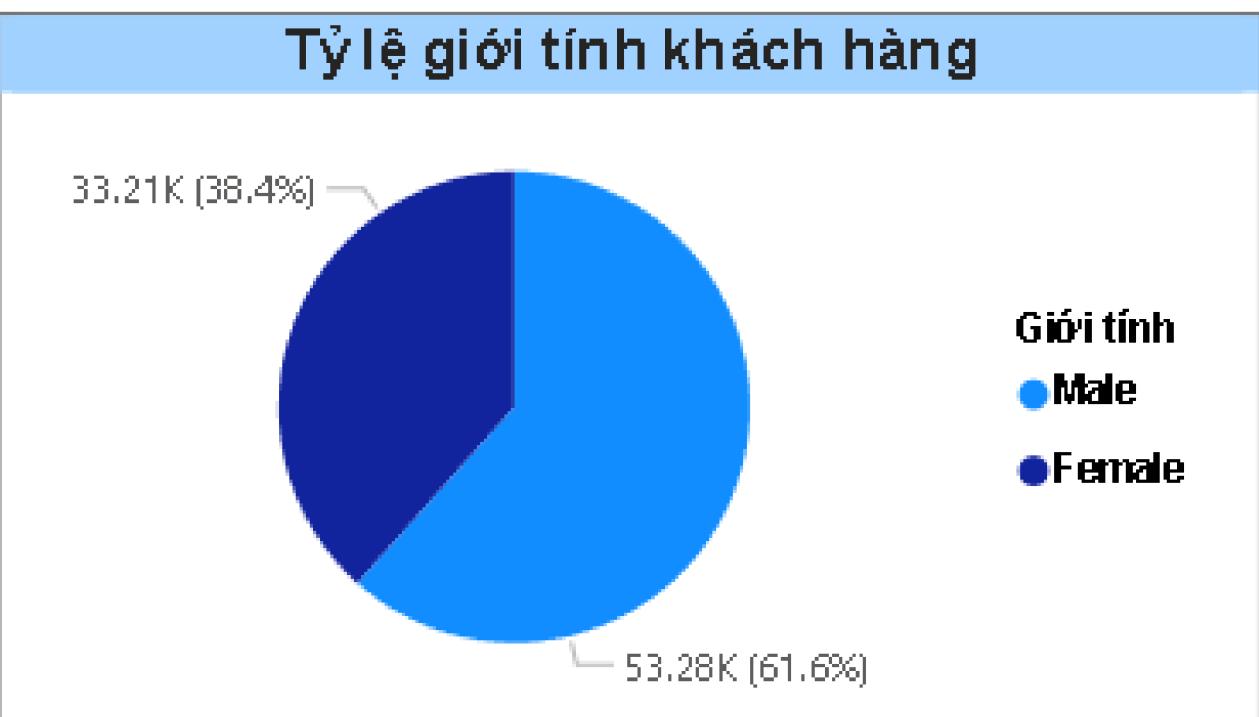
4

Count of Nhóm khách hàng

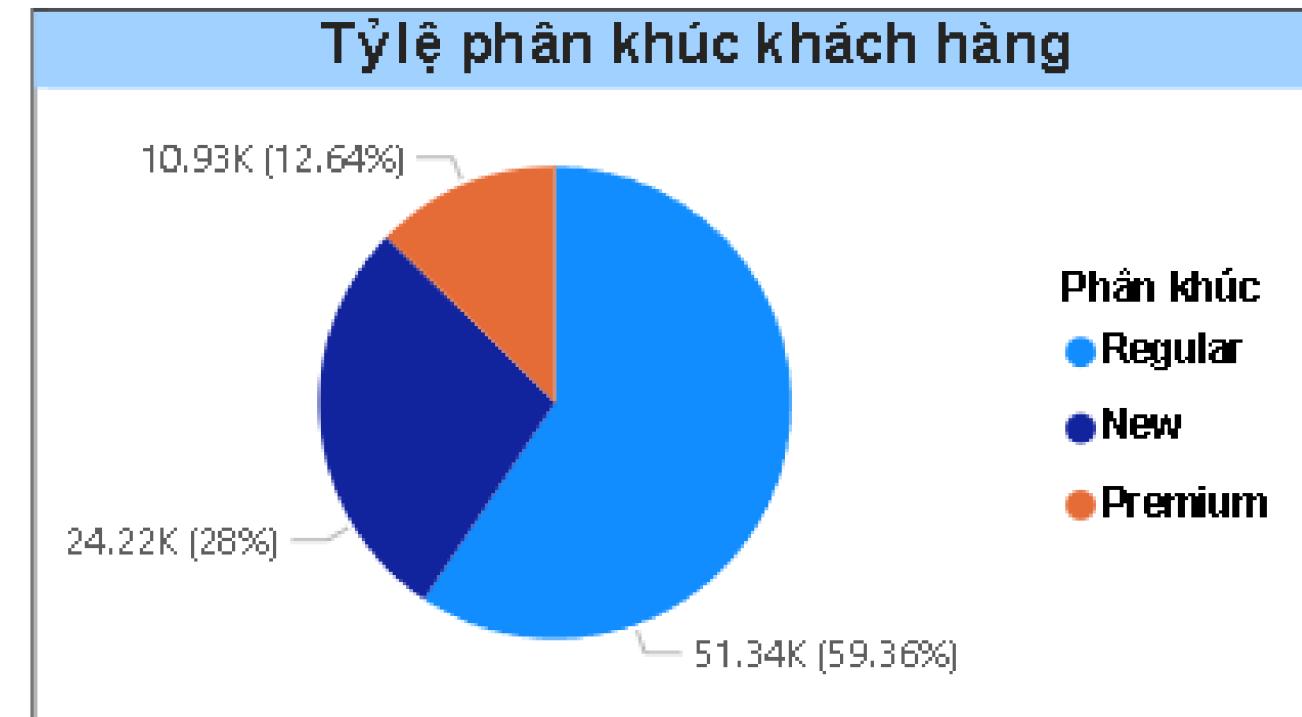
Phân bổ nhóm khách hàng



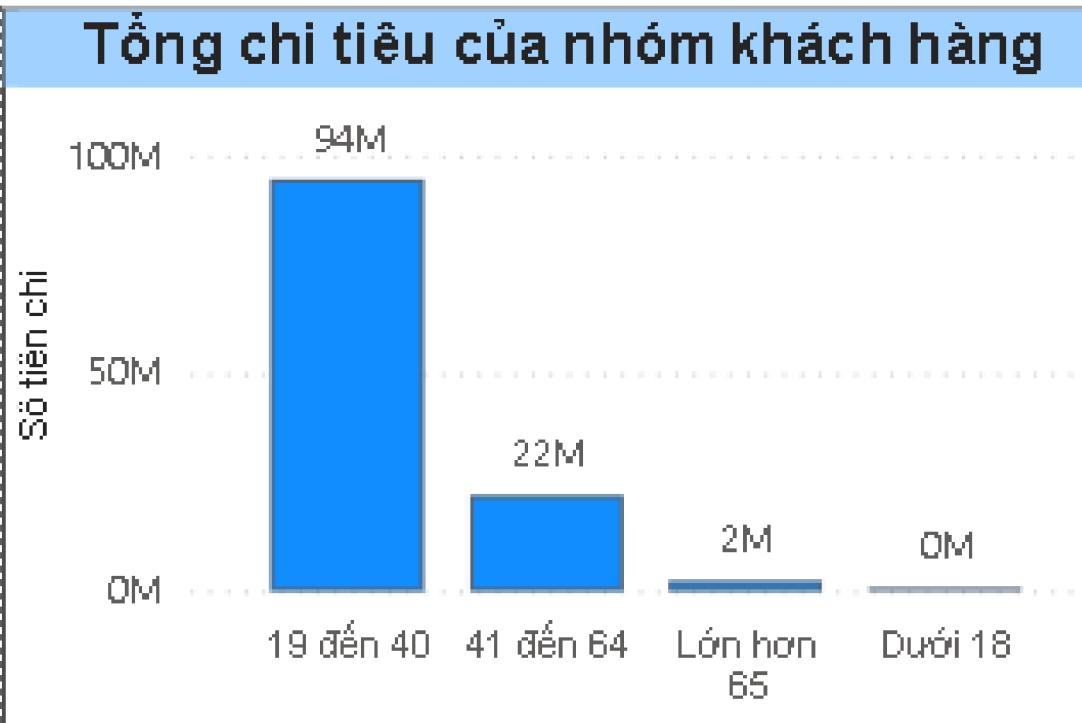
Tỷ lệ giới tính khách hàng



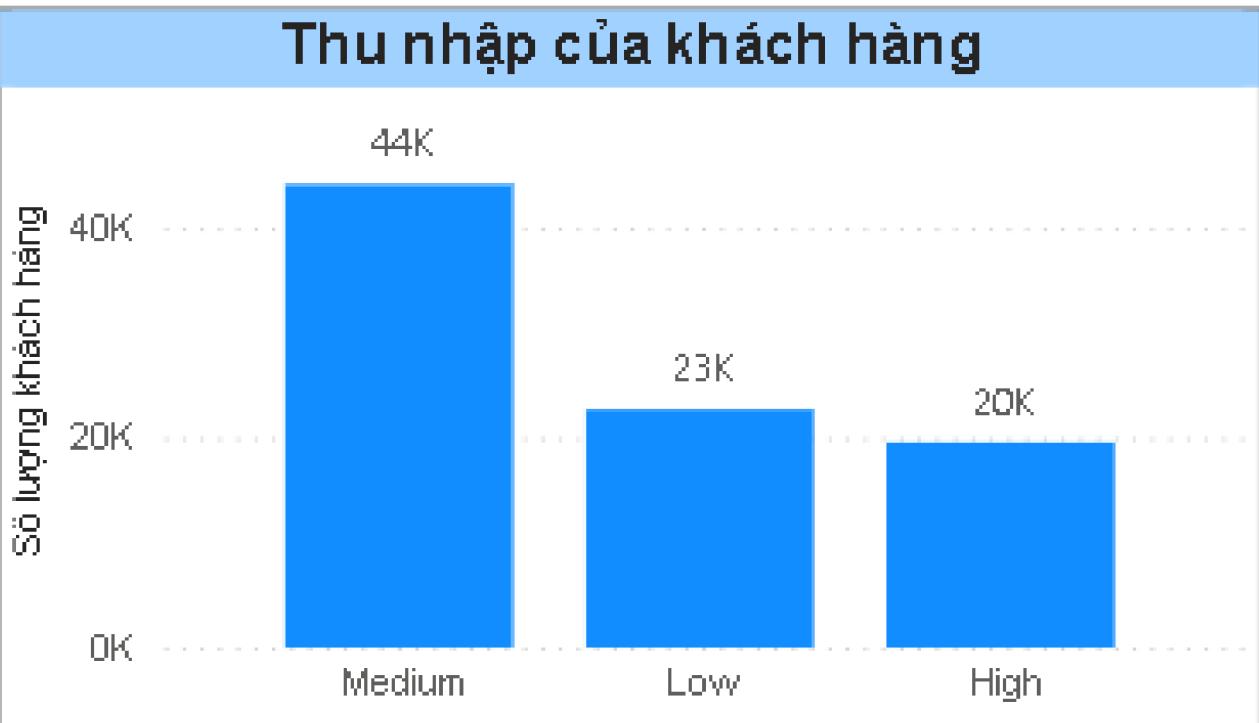
Tỷ lệ phân khúc khách hàng



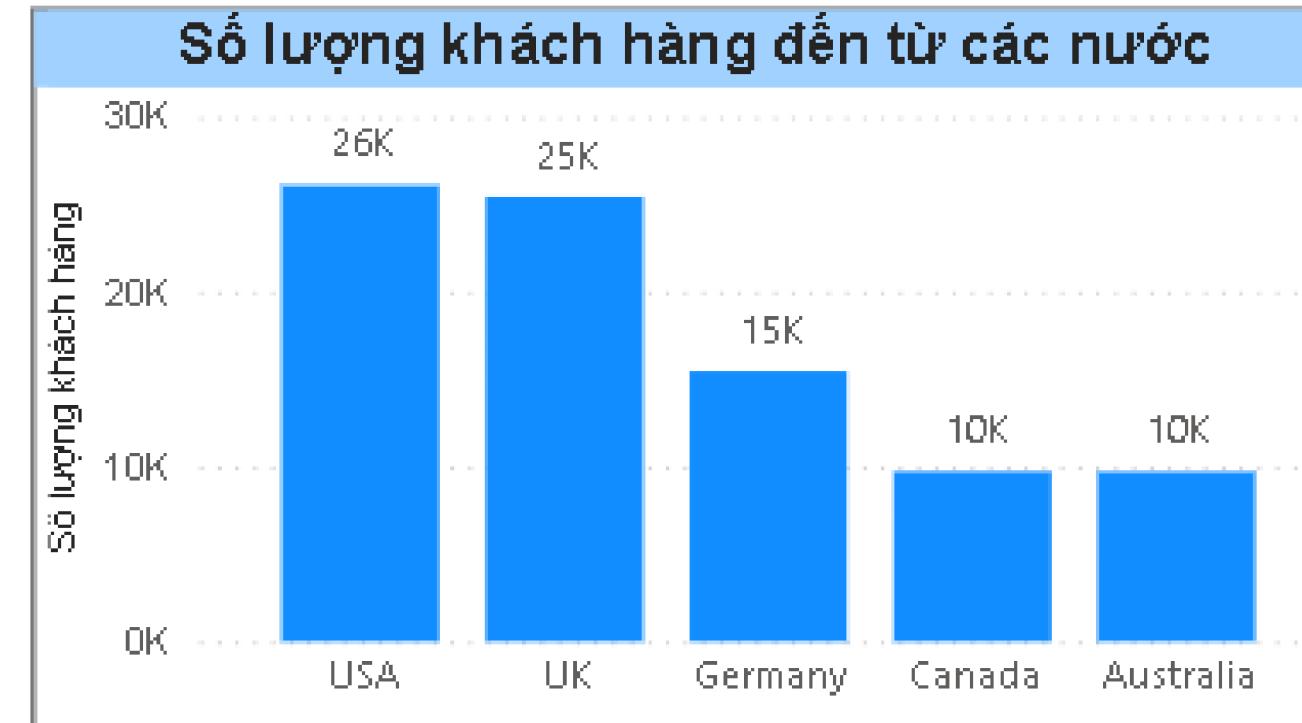
Tổng chi tiêu của nhóm khách hàng



Thu nhập của khách hàng

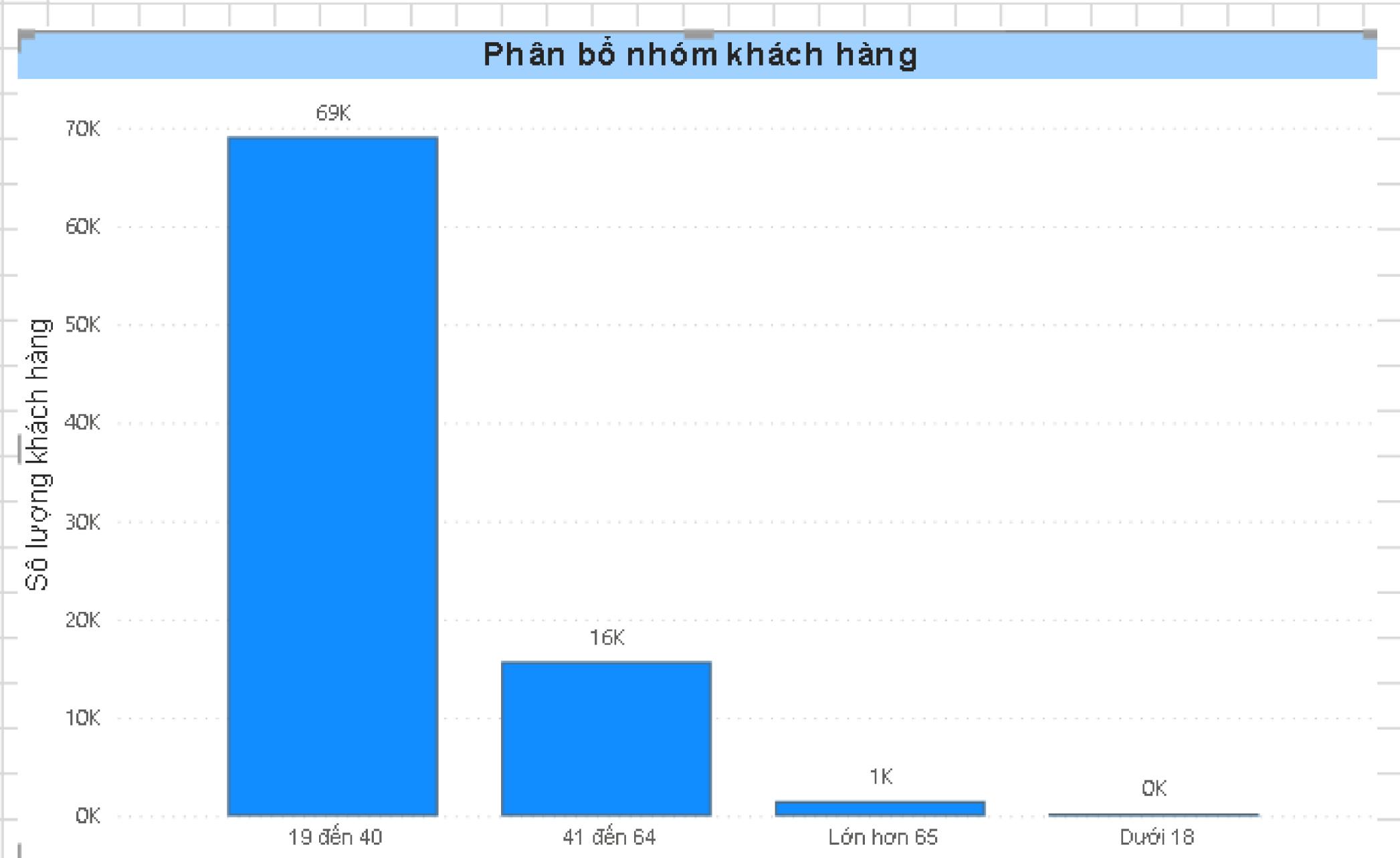


Số lượng khách hàng đến từ các nước





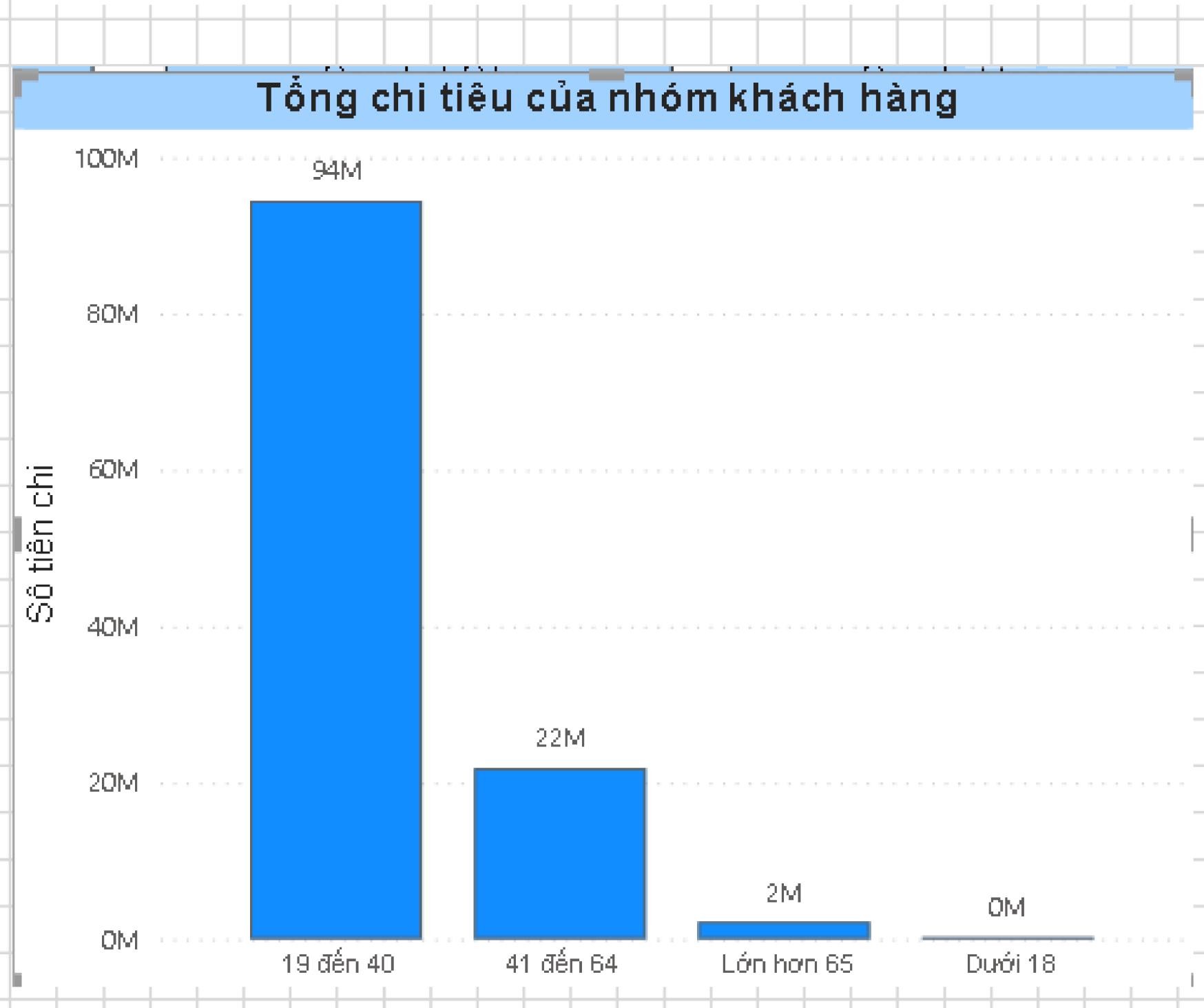
3.1 Phân tích tổng quan về khách hàng



Nhóm tuổi 19-40 chiếm tỷ lệ áp đảo, cho thấy đây là đối tượng khách hàng chính cần hướng đến trong các chiến lược tiếp thị và phát triển sản phẩm/dịch vụ.

[Quay lại Trang Chương trình](#)

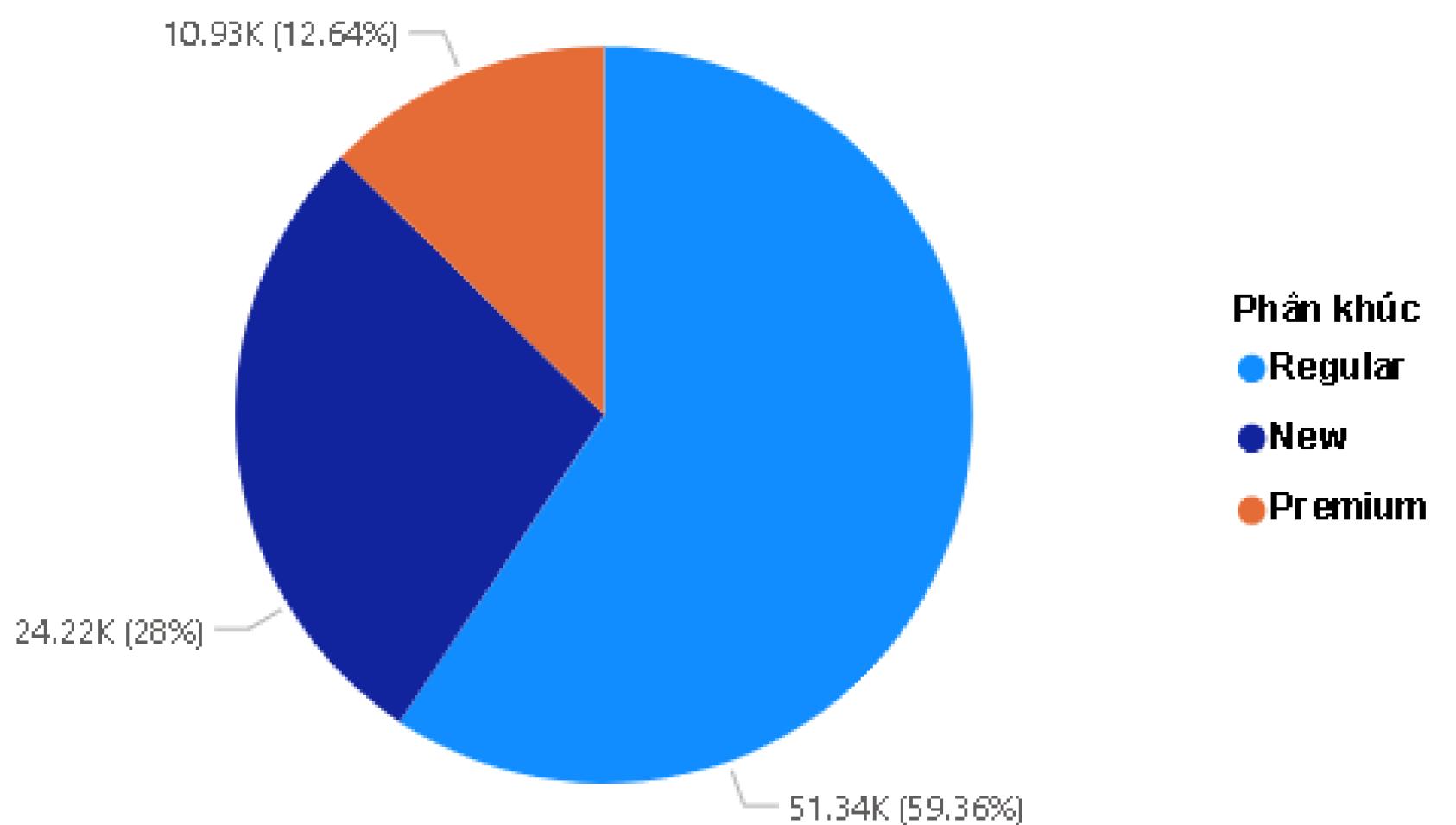
3.1 Phân tích tổng quan về khách hàng



Nhóm tuổi 19-40 cũng là nhóm chi tiêu nhiều nhất, khẳng định lại vị thế quan trọng của nhóm khách hàng này. Cần tập trung phân tích sâu hơn về nhu cầu, hành vi chi tiêu của nhóm này để tối ưu hóa lợi nhuận

3.1 Phân tích tổng quan về khách hàng

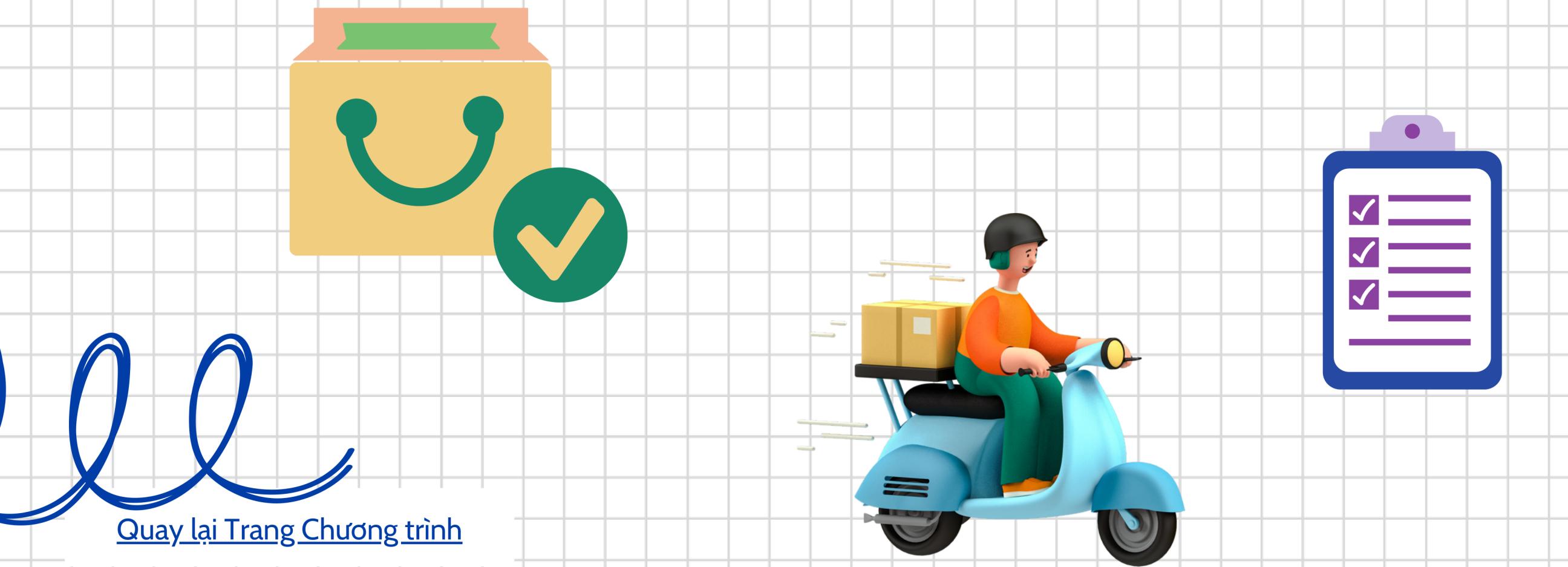
Tỷ lệ phân khúc khách hàng



- Regular chiếm tỷ lệ cao nhất: Cho thấy chiến lược duy trì khách hàng hiện tại đang hiệu quả.
- Premium chiếm tỷ lệ thấp: Cần xem xét các chương trình, ưu đãi đặc biệt để thu hút và giữ chân nhóm khách hàng tiềm năng này.

* 3.2 Phân tích tổng quan đơn hàng

Bảng đơn hàng chứa thông về ID của khách hàng , thông tin đơn hàng , vận chuyển , thanh toán và feedback của khách hàng



[Quay lại Trang Chương trình](#)

✓	orders	...
<input type="checkbox"/>	Σ Amount	
<input type="checkbox"/>	Customer_ID	...
>	<input type="checkbox"/> Date	
<input type="checkbox"/>	Feedback	
<input type="checkbox"/>	Month	
<input type="checkbox"/>	Order_Status	
<input type="checkbox"/>	Payment_Method	
<input type="checkbox"/>	Σ Product_ID	
<input type="checkbox"/>	products	
<input type="checkbox"/>	Σ Ratings	
<input type="checkbox"/>	Shipping_Method	
<input type="checkbox"/>	Time	
<input type="checkbox"/>	Σ Total_Amount	
<input type="checkbox"/>	Σ Total_Purchases	
<input type="checkbox"/>	Σ Transaction_ID	
<input type="checkbox"/>	Σ Year	



TỔNG QUAN VỀ ĐƠN HÀNG

48

Count of Product_ID

86.05K

Count of Transaction_ID

4

Count of Payment_Method

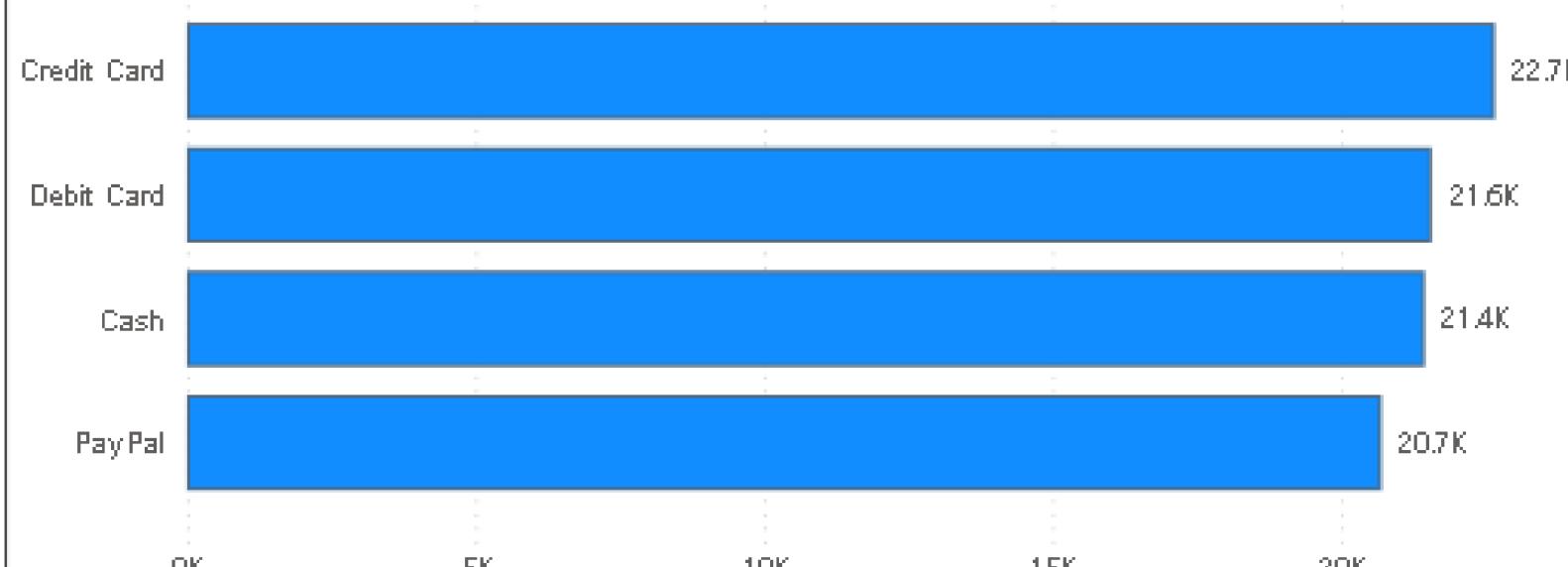
5

Count of Ratings

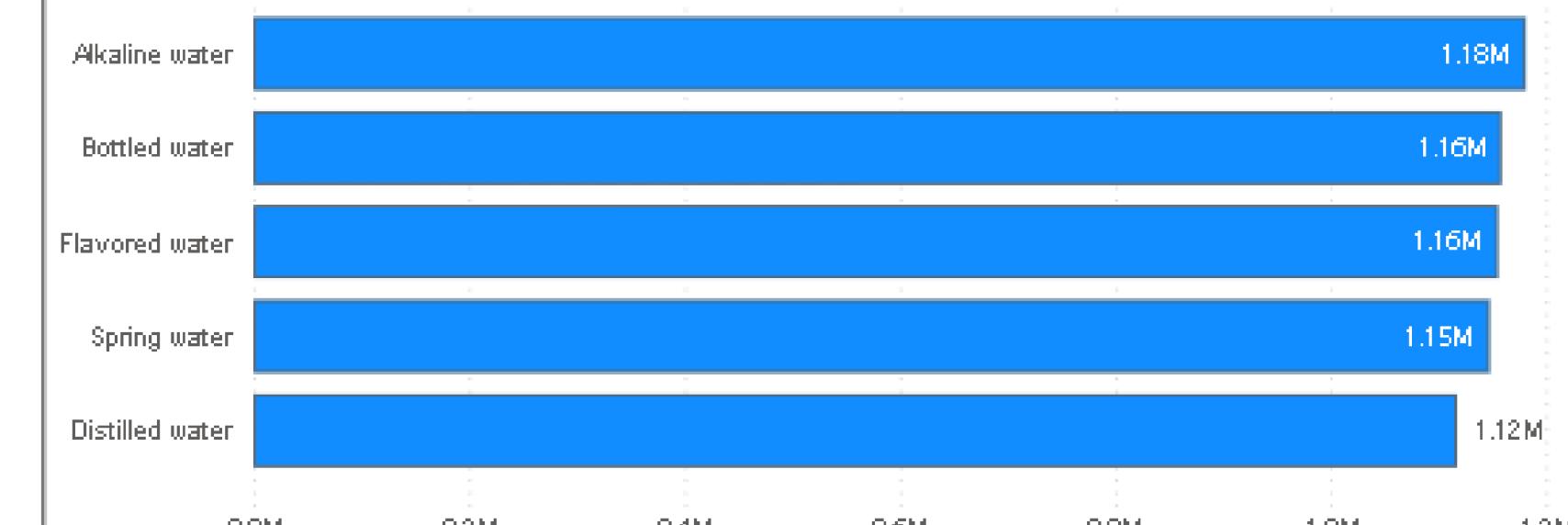
4

Count of Order_Status

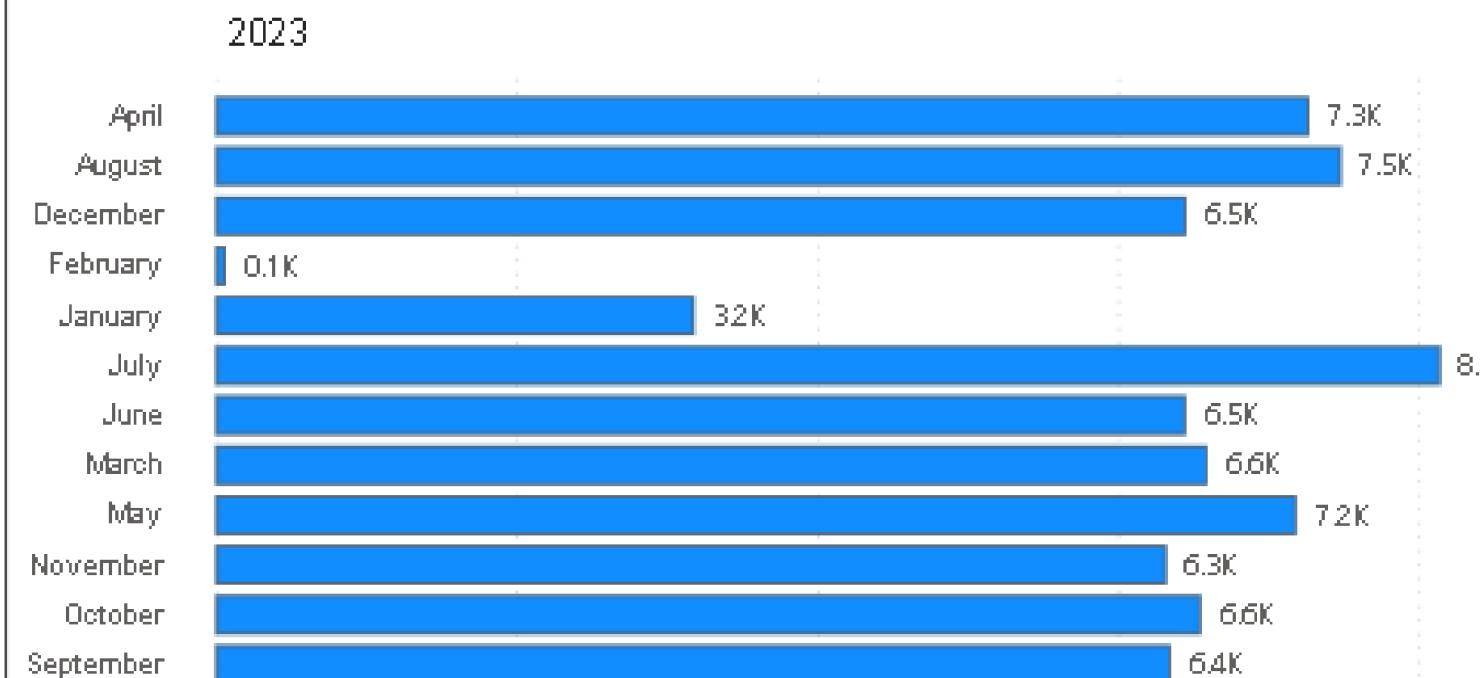
Lượng đơn hàng theo cách thức thanh toán



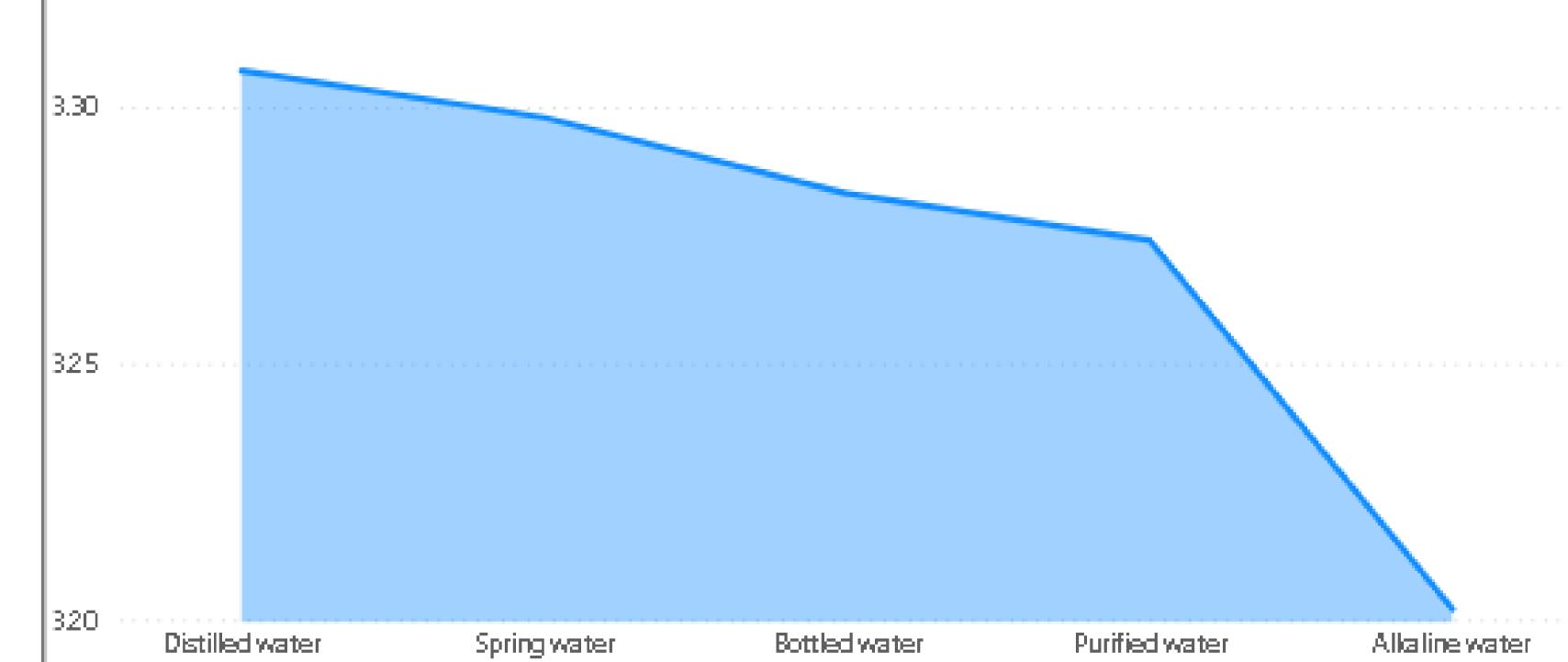
Top 5 doanh thu



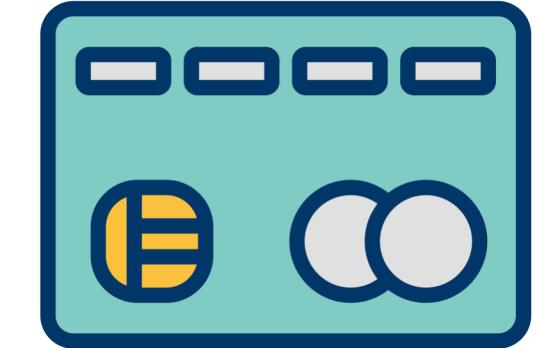
Lượng đơn hàng theo tháng của năm 2023



Top 5 sản phẩm có rating cao nhất

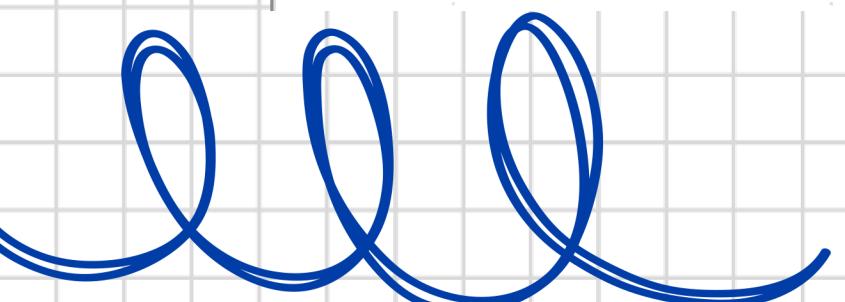
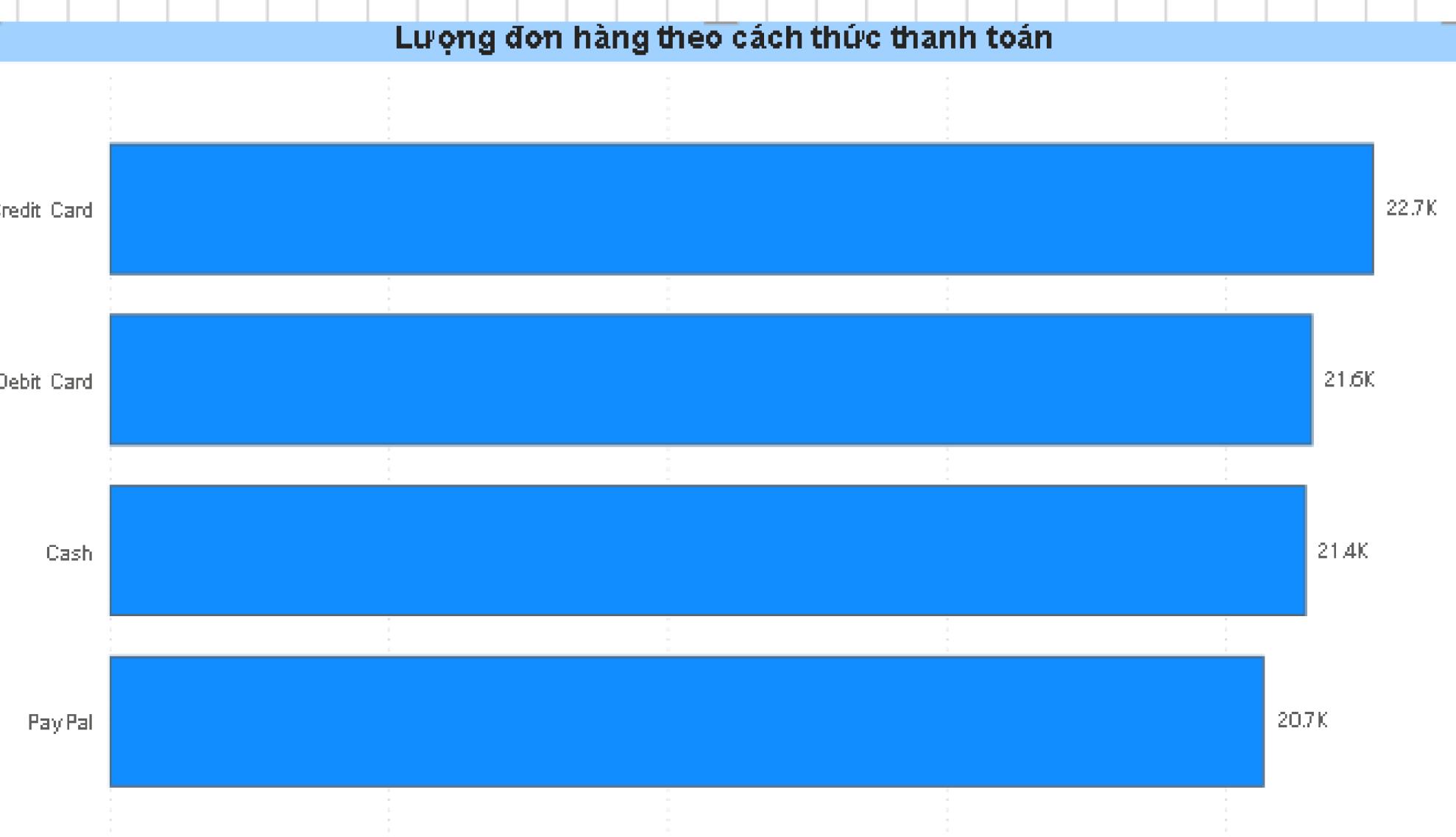


3.2 Phân tích tổng quan đơn hàng



*

- PayPal là phương thức thanh toán được ưa chuộng nhất.
- Credit Card và Debit Card có lượng sử dụng tương đương nhau và xếp thứ hai.
- Cash ít được sử dụng nhất, cho thấy xu hướng thanh toán không dùng tiền mặt đang phổ biến.



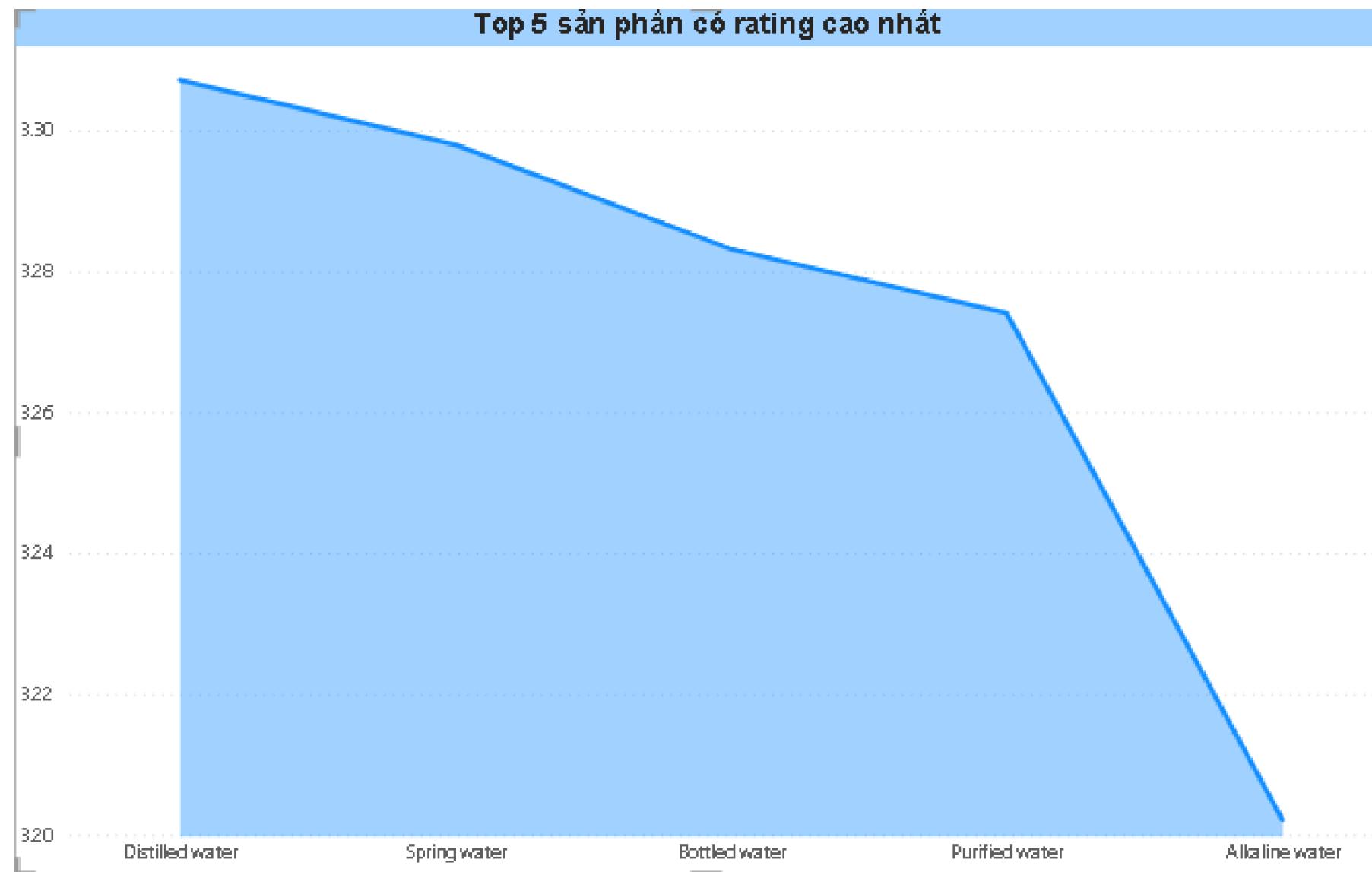
3.2 Phân tích tổng quan đơn hàng



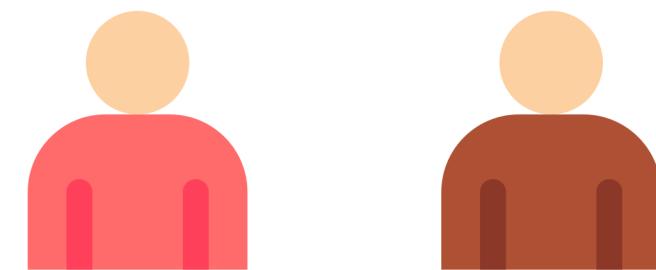
- Alkaline water là sản phẩm mang lại doanh thu cao nhất.
- Doanh thu của 4 sản phẩm còn lại (Bottled water, Flavored water, Spring water, Distilled water) khá đồng đều.

3.2 Phân tích tổng quan đơn hàng

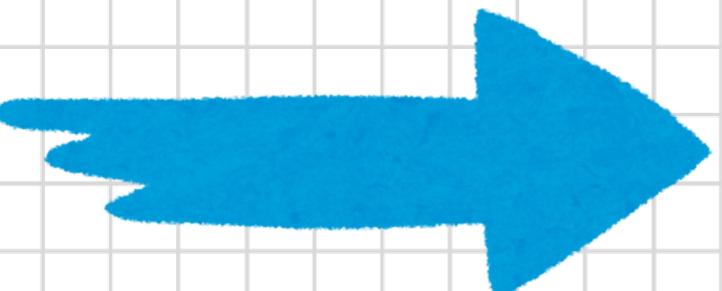
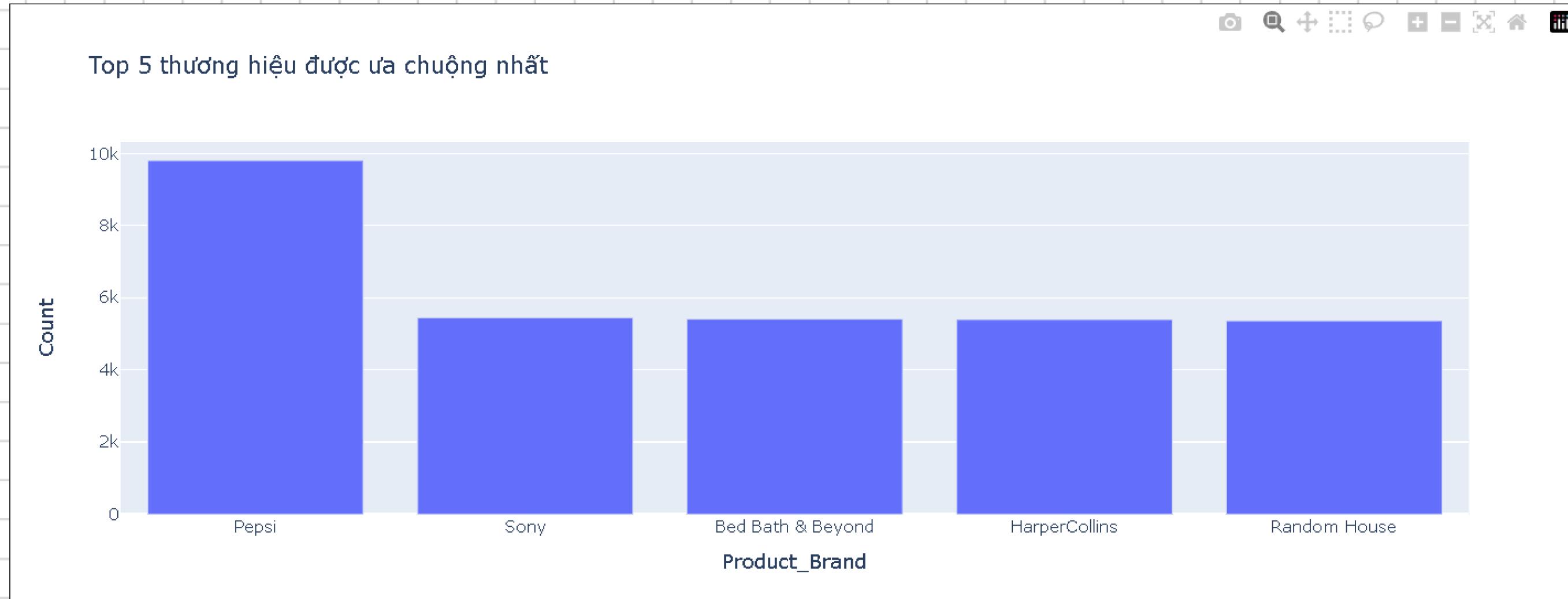
[Quay lại Trang Chương trình](#)



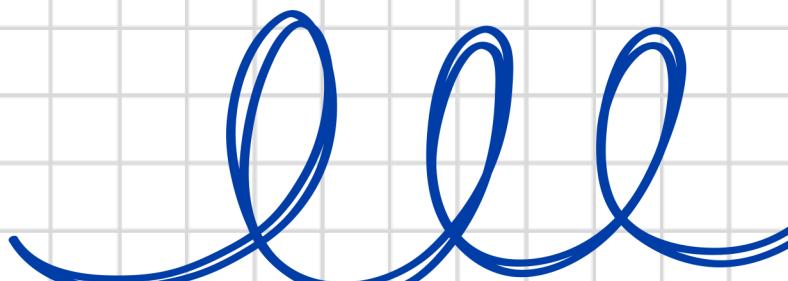
- Distilled water nhận được đánh giá cao nhất từ khách hàng.
- Alkaline water dù có doanh thu cao nhất nhưng lại có rating thấp nhất trong top 5, cho thấy có thể tồn tại một số vấn đề về chất lượng hoặc dịch vụ cần được cải thiện.



3.2 Phân tích tổng quan đơn hàng



Pepsi chiếm ưu thế vượt trội so với các thương hiệu còn lại. Điều này cho thấy Pepsi là thương hiệu rất phổ biến và được ưa chuộng bởi phần lớn khách hàng.



4. Phân tích chuyên sâu



4. Phân tích chuyên sâu

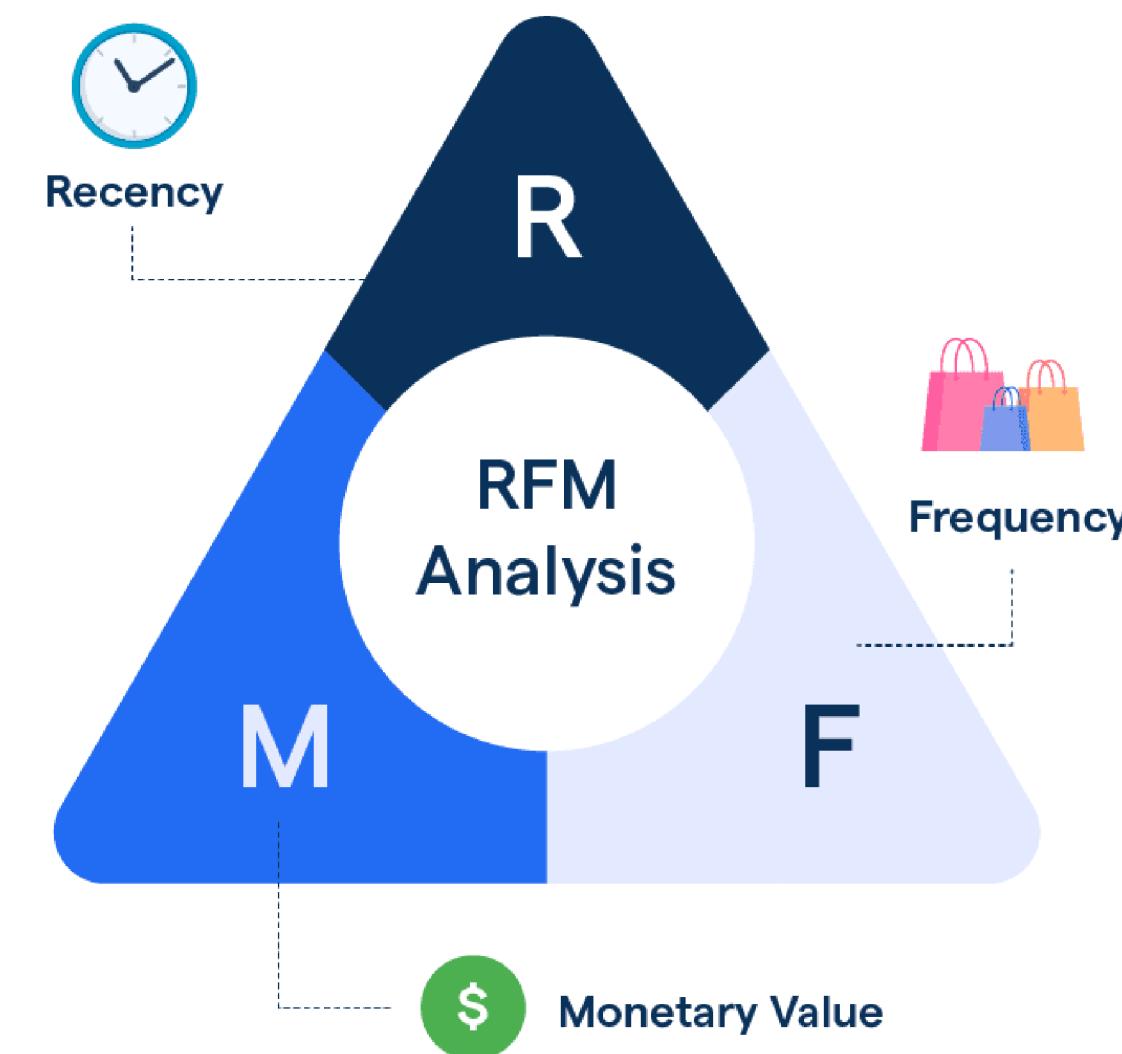
Mục đích để tìm hiểu sâu hơn hành vi của nhóm khách hàng , dự đoán phát hiện được các yếu tố giúp phát triển lợi nhuận cho doanh nghiệp.



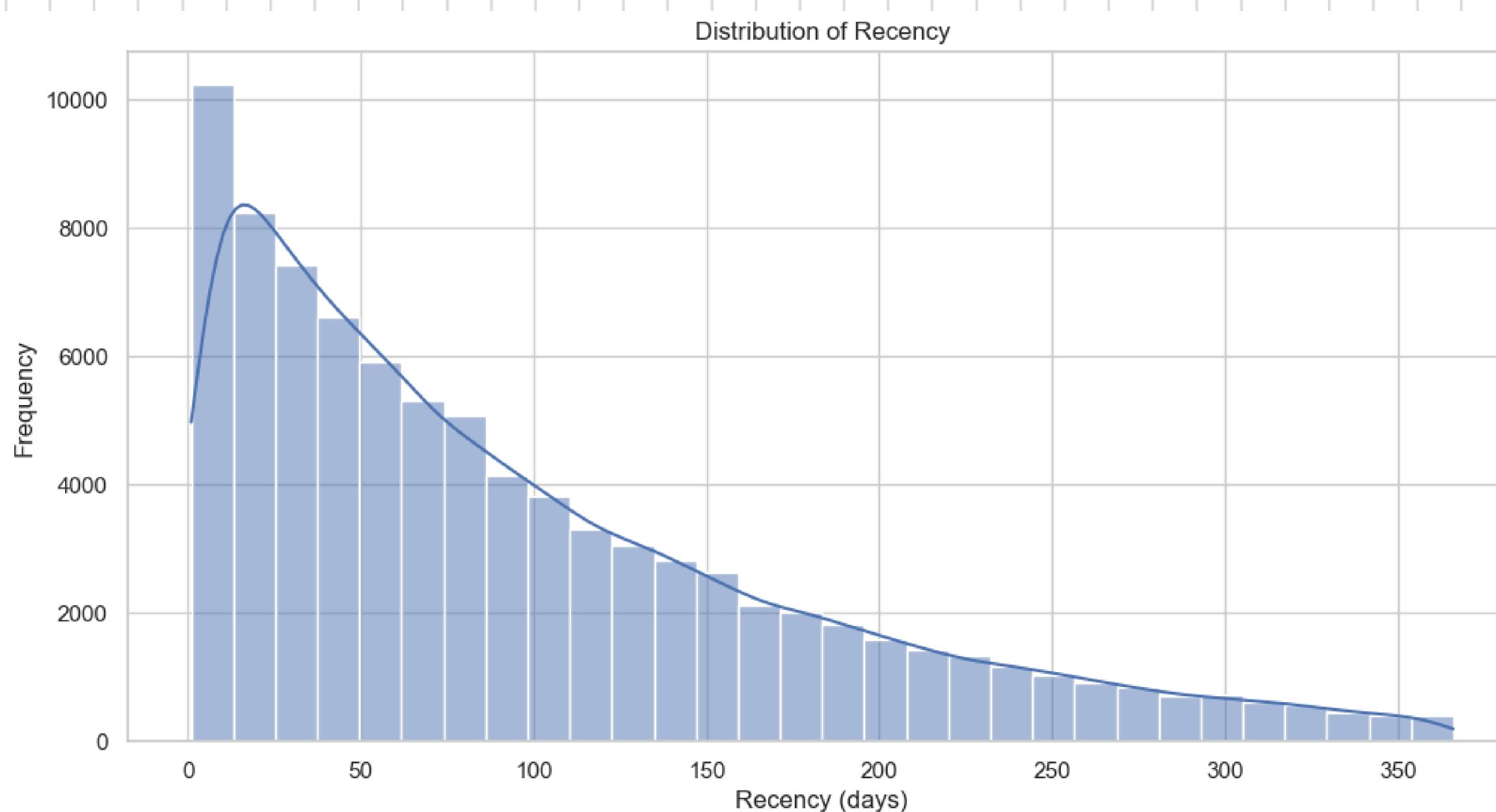
4.1 Phương pháp RFM



RFM là phương pháp phân tích hành vi khách hàng đơn giản nhưng hiệu quả, giúp doanh nghiệp đưa ra quyết định kinh doanh dựa trên dữ liệu và tối ưu hóa hoạt động marketing

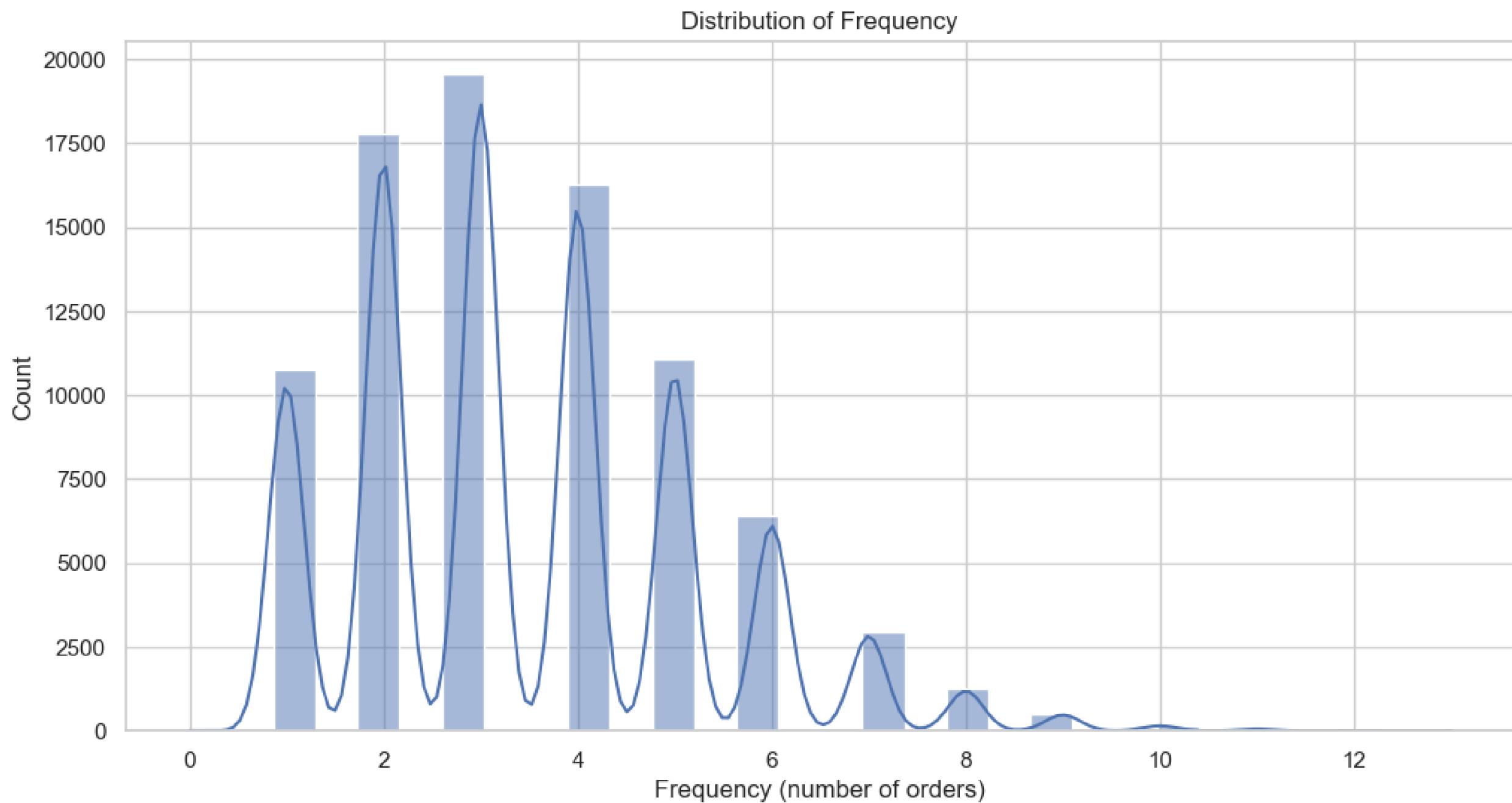


4.1 Phương pháp RFM



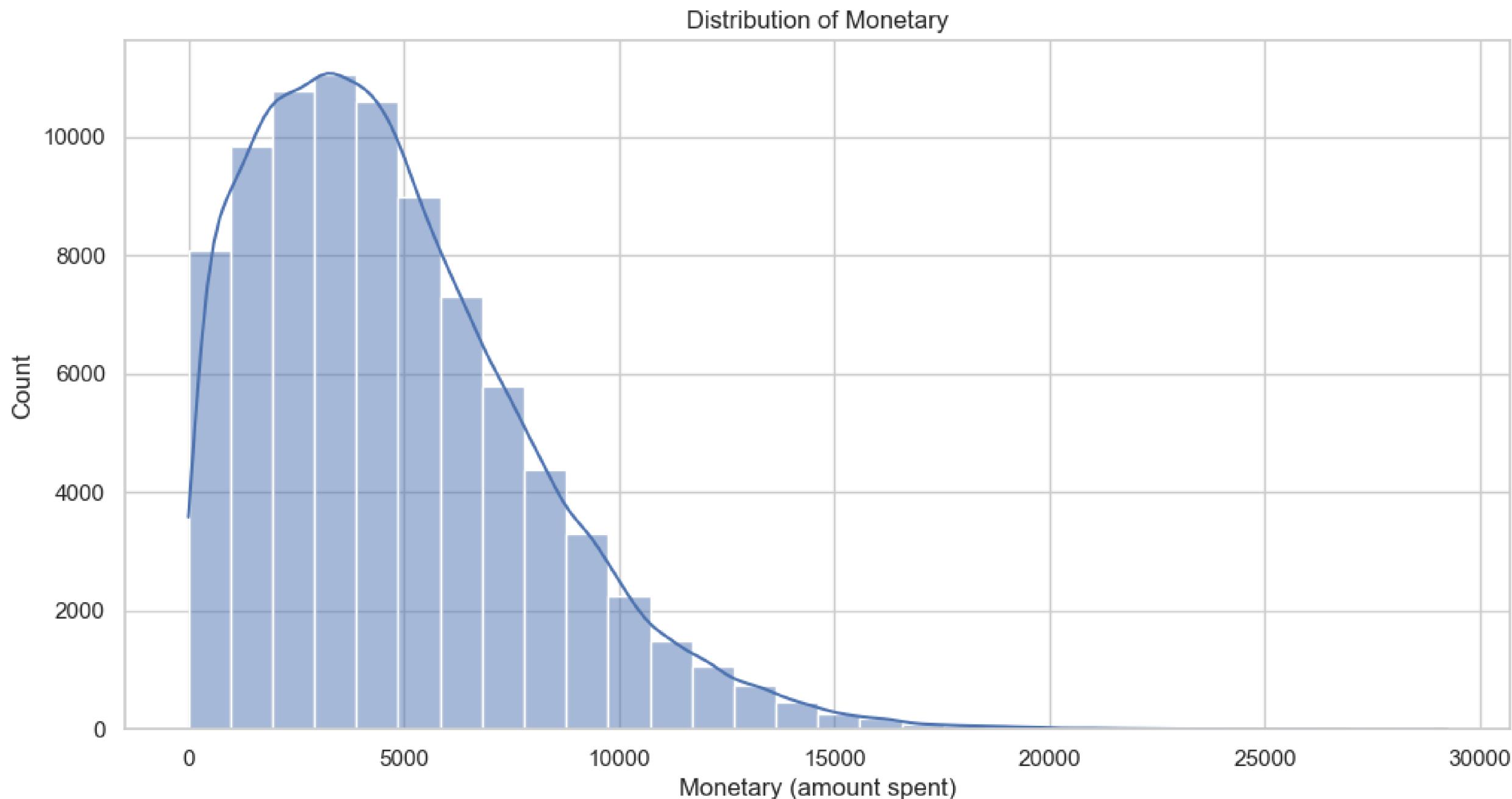
- Số lượng lớn khách hàng có Recency thấp (0-50 ngày) là dấu hiệu tốt.
 - Cho thấy cơ sở khách hàng khá năng động và tích cực
- Khách hàng trung bình (30-150 ngày): Nhóm tiềm năng để kích hoạt lại và tăng tần suất mua hàng.

4.1 Phương pháp RFM



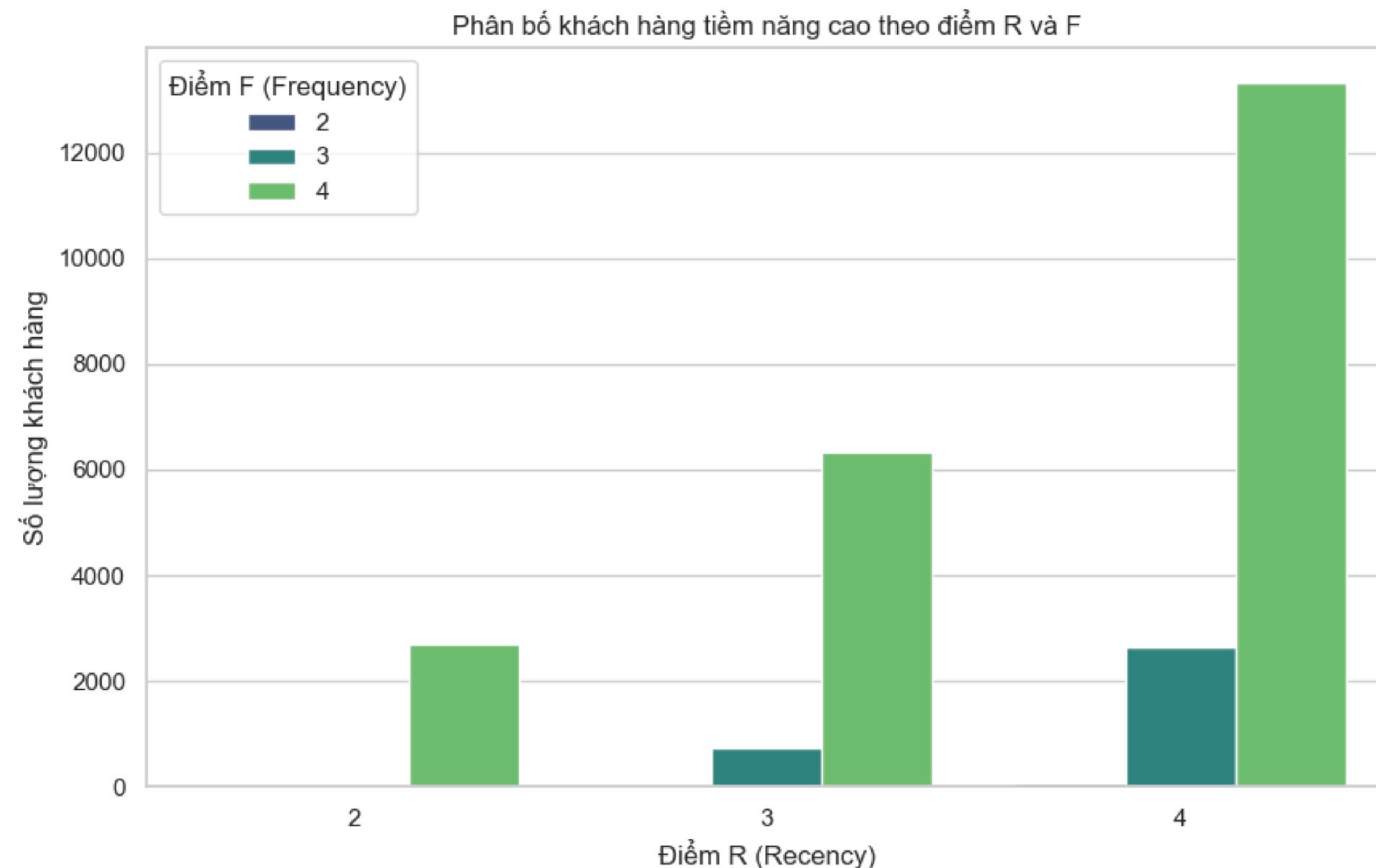
- Khách hàng một lần là nhóm lớn, cần chiến lược chuyển đổi thành khách hàng thường xuyên.
- Khách hàng thường xuyên Đây là nhóm chủ chốt, cần tập trung giữ chân và phát triển.
- Khách hàng trung thành là nhóm nhỏ nhưng có giá trị cao.

4.1 Phương pháp RFM



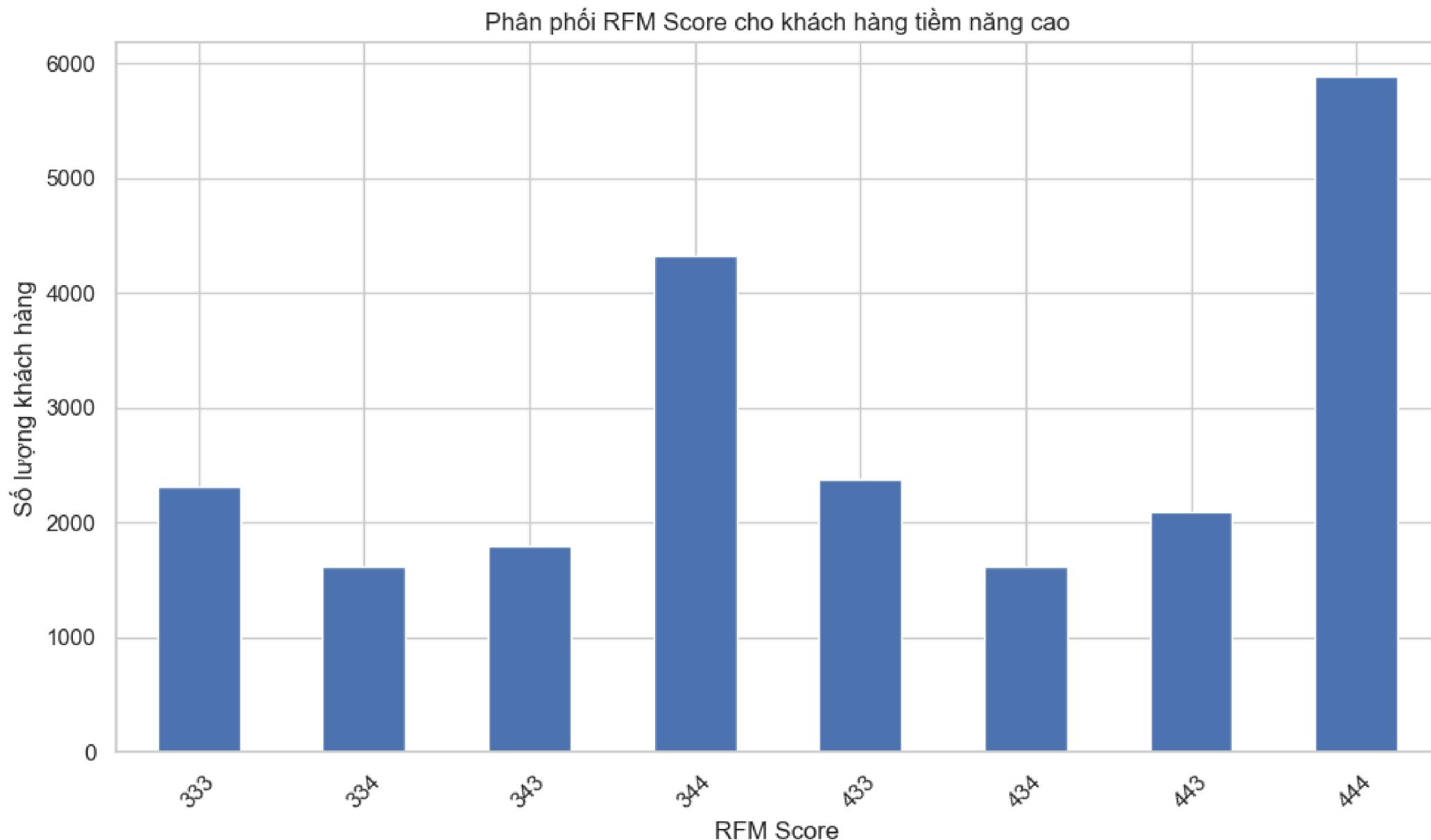
- Đỉnh của phân phối nằm trong khoảng 2000-4000 đơn vị tiền tệ.
- Số lượng khách hàng giảm nhanh khi giá trị chi tiêu tăng.
- Có một số ít khách hàng với giá trị chi tiêu rất cao (trên 20,000).

4.1 Phương pháp RFM



- Chọn ra các nhóm khách hàng có tiềm năng cao (có chỉ số RF cao) để phân tích
- Nhìn chung, nhóm khách hàng có điểm R cao (3 và 4) chiếm tỷ lệ lớn nhất, cho thấy phần lớn khách hàng đã mua hàng gần đây.

4.1 Phương pháp RFM



- Biểu đồ cho thấy rõ ràng nhóm khách hàng có điểm RFM cao chiếm tỷ lệ lớn trong tập khách hàng tiềm năng.
- Doanh nghiệp nên tập trung vào việc duy trì và phát triển nhóm khách hàng này. Bên cạnh đó, cần chú ý đến những nhóm có điểm R thấp nhưng F và M cao để có chiến lược thu hút phù hợp.

4.2 Chiến lược CRM



CRM là viết tắt của Customer Relationship Management

Mục tiêu của CRM:

- Nâng cao sự hài lòng của khách hàng
- Tăng tỷ lệ giữ chân khách hàng
- Thu hút khách hàng mới
- Tăng doanh thu và lợi nhuận



