



Customer Segmentation Report

Prepared by Bùi Ngọc Dũng



1.Xác định vấn đề

Trong bối cảnh cạnh tranh ngày càng tăng của thị trường bán lẻ trực tuyến, Công ty X, đang đối mặt với thách thức làm thế nào để tăng cường động cơ mua sắm và cải thiện sự trung thành của khách hàng. Dù có một lượng lớn dữ liệu khách hàng, công ty chưa tận dụng hiệu quả những thông tin này để cá nhân hóa trải nghiệm mua sắm.

2.Mục đích phân tích



Xác định các nhóm khách hàng dựa trên hành vi mua sắm, giá trị đóng góp để cá nhân hóa các chiến lược tiếp thị và sản phẩm.



Sử dụng phân khúc để tối ưu hóa các chiến dịch tiếp thị, tăng tỷ lệ chuyển đổi và cải thiện sự hài lòng của khách hàng.



Phát hiện những cơ hội mới thông qua việc hiểu rõ các nhu cầu và mong muốn chưa được đáp ứng của các phân khúc khách hàng.

1. Tổng quan dữ liệu

Báo cáo sẽ sử dụng dữ liệu từ hệ thống CRM, trang web và các nền tảng mạng xã hội của công ty, được tổng hợp lại thành 2 bảng là Customer_Registered (Danh sách khách hàng đăng kí) và Customer_Transactions (Danh sách giao dịch của khách hàng). Dữ liệu hiện có phục vụ lưu trữ các giao dịch của khách hàng, thuộc hệ thống OLTP (Online Transactional Processing).

1.1. Bảng Customer_Registered

Cột	Kiểu dữ liệu	Ý nghĩa
ID	Int	Mã khách hàng
Contract	Varchar	Mã hợp đồng
LocationID	Int	Mã địa chỉ
BranchCode	Int	Mã chi nhánh
Status	Int	Trạng thái
Created_date	Date	Ngày đăng kí
Stop_date	Date	Ngày hủy

	ID	Contract	LocationId	BranchCode	Status	CreatedDate	StopDate
1	1	SGDN00215	8	1	0	2011-11-25	2012-01-05
2	2	SGDN00214	8	1	0	2012-06-14	0000-00-00
3	3	SGD374348	8	1	0	2012-11-01	0000-00-00
4	4	SGD022064	8	1	2	2011-06-22	2013-05-29
5	5	SGD041015	8	5	2	2011-12-17	2014-11-11
6	6	SGDN00211	[NULL]	[NULL]	2	2015-06-09	2015-09-09
7	7	SGD374348	8	1	3	2012-11-26	2012-12-13
8	8	SGDN00013	0	0	1	2017-12-11	0000-00-00
9	9	BEAAA1809	75	0	1	2022-04-04	0000-00-00
10	10	BNAAA1298	241	2	1	2022-04-19	0000-00-00

1.2. Customer_Transaction

Cột	Kiểu dữ liệu	Ý nghĩa
Transaction_Id	Int	Mã giao dịch
CustomerId	Int	Mã khách hàng
PurchaseDate	Date	Ngày giao dịch
GMV	Int	Tổng giá trị giao dịch

	¹²³ Transaction_ID	¹²³ CustomerId	¹²³ PurchaseDate	¹²³ GMV
1	0	1,327,813	2022-06-01	95,000
2	1	1,157,830	2022-06-01	75,000
3	2	873,915	2022-07-01	95,000
4	3	3,505,071	2022-07-01	90,000
5	4	2,930,918	2022-07-01	109,091
6	5	899,882	2022-06-01	105,000
7	6	2,248,818	2022-06-01	75,000

Tuy nhiên, ta cần xử lý dữ liệu để có thể phân tích một cách hiệu quả (đưa dữ liệu về dạng OLAP). Ở đây, ta sử dụng phương pháp RFM Metrics để đưa dữ liệu về trạng thái sẵn sàng cho hoạt động phân tích và tìm ra insight. Ví dụ:

	¹²³ CustomerId	¹²³ Recency	¹²³ Frequency	¹²³ monetary	¹²³ R	¹²³ F	¹²³ M
1	855,918	31	0.25	0	4	4	1
2	1,040,767	47	0.33	0	3	4	1
3	1,036,806	45	0.33	0	3	4	1
4	1,044,142	31	0.33	0	3	4	1
5	214,484	31	0.33	0	4	4	1
6	235,758	31	0.33	0	4	4	1
7	267,349	31	0.33	0	4	4	1
8	282,119	31	0.33	0	4	4	1
9	306,300	31	0.33	0	4	4	1

2. Phương pháp phân tích

2.1. Phân tích customer 360

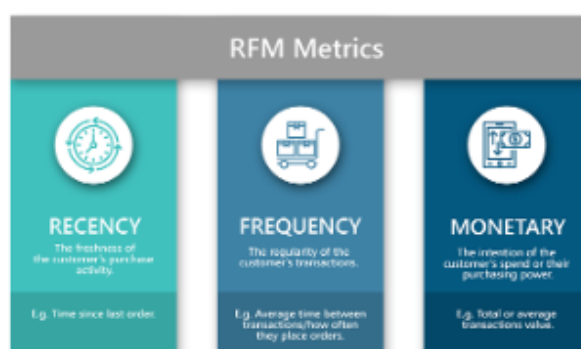


Phân tích Customer 360 là quá trình tổng hợp và phân tích dữ liệu khách hàng từ bốn nguồn chính để có cái nhìn toàn diện và sâu sắc về hành vi, sở thích, và nhu cầu của họ:

1. **Dữ liệu Nhân khẩu học:** Thông tin cơ bản như tuổi, giới tính, mức thu nhập, giúp xác định phân khúc khách hàng.
2. **Dữ liệu Giao Dịch:** Thông tin về mua hàng giúp hiểu thói quen mua hàng và độ trung thành của khách hàng.
3. **Dữ liệu Hành Vi:** Lịch sử duyệt web và tương tác trên mạng xã hội giúp hiểu sở thích và nhu cầu không rõ ràng qua giao dịch.
4. **Dữ liệu Tương Tác:** Thông tin về mọi tương tác giữa khách hàng và công ty qua các kênh, giúp cải thiện hành trình khách hàng.

Ở báo cáo này, thực hiện phân tích phân khúc khách hàng theo dữ liệu giao dịch phát sinh trong khoảng thời gian (01-06-2022 đến 31-08-2022).

2.2. Mô hình phân tích RFM (Recency-Frequency-Monetary)



RFM metrics là một phương pháp phân tích dựa trên ba tiêu chí: Recency (R) - thời gian kể từ lần giao dịch cuối cùng, Frequency (F) - tần suất giao dịch, và Monetary (M) - tổng giá trị giao dịch. Phương pháp này giúp doanh nghiệp phân khúc được từng nhóm khách hàng, từ đó đưa ra chiến lược riêng cho từng đối tượng cần nhắm đến.

Recency(R) - Thời gian mua hàng gần nhất

- **Vai trò:** Đo lường thời gian kể từ lần giao dịch cuối cùng của khách hàng. Recency giúp nhận diện khách hàng có khả năng quay lại mua sắm cao hơn vì họ đã tương tác gần đây.
- **Cách tính:** Lấy ngày hiện tại trừ đi ngày của giao dịch cuối cùng. Một giá trị thấp hơn (ngày gần đây) cho thấy khách hàng có khả năng cao để mua sắm lại.

Frequency (F) - Tần suất mua hàng

- **Vai trò:** Đo lường số lần khách hàng thực hiện giao dịch trong một khoảng thời gian nhất định. Frequency giúp xác định mức độ trung thành của khách hàng, với giả định rằng những khách hàng mua sắm thường xuyên hơn là có giá trị cao hơn.
- **Cách tính:** Tổng số giao dịch chia khoảng thời gian xác định.

Monetary (M) - Giá trị

- **Vai trò:** Đo lường tổng giá trị tiền mà khách hàng đã chi tiêu. Monetary giúp đánh giá giá trị tài chính mà khách hàng mang lại cho doanh nghiệp, với giả định rằng khách hàng chi tiêu nhiều hơn là quan trọng hơn.
- **Cách tính:** Tính tổng giá trị của tất cả giao dịch mà một khách hàng đã thực hiện.

Sử dụng IQR method để đánh giá phân bố dữ liệu cho từng chỉ số R,F,M

IQR	1	2	3	4	Đơn vị
Recency	≥ 92	63-91	32-62	1-31	Ngày
Frequency	≤ 0.0159	0.0159-0.0192	0.0192-0.0217	≥ 0.0217	Giao dịch/ Tuổi hợp đồng(tháng)
Monetary	≤ 75000	75000-85000	85000-105000	≥ 105000	Đvt

Sau khi dùng IQR method để phân bố các chỉ số của từng khách hàng, ta có được từng nhóm khách hàng có chung RFM. Áp dụng BCG matrix để thực hiện phân khúc khách hàng mục tiêu:



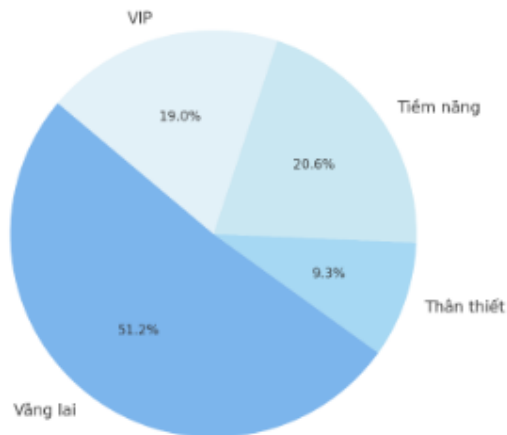
Dựa theo phương pháp trên có thể tạm chia khách hàng thành các nhóm

1. **Star - VIP:** Đây là những phân khúc có tỷ lệ tăng trưởng nhanh và mang lại lợi nhuận cao. Chúng là những phân khúc mà doanh nghiệp muốn đầu tư nhiều hơn để tận dụng tốc độ tăng trưởng và tăng cường thị phần. Các chiến lược có thể bao gồm tăng cường dịch vụ, tiếp thị hoặc phát triển sản phẩm đặc biệt cho những phân khúc này.
2. **Question Mark - Tiềm năng:** Đây là những phân khúc có tỷ lệ tăng trưởng cao, nhưng lợi nhuận chưa cao. Chúng có tiềm năng phát triển mạnh mẽ, nhưng cũng đòi hỏi nhiều đầu tư. Doanh nghiệp cần xem xét việc đầu tư thêm vào những phân khúc này để tăng cường thị phần và lợi nhuận trong tương lai.
3. **Cash Cow - Thân thiết:** Đây là những phân khúc có tỷ lệ tăng trưởng chậm, nhưng mang lại lợi nhuận cao. Chúng tạo ra lượng tiền mặt ổn định cho doanh nghiệp và thường không đòi hỏi đầu tư lớn. Doanh nghiệp cần duy trì và tối ưu hóa các phân khúc này để tận dụng lợi nhuận mà chúng mang lại.
4. **Dogs - Vãng lai:** Đây là những phân khúc có tỷ lệ tăng trưởng thấp và lợi nhuận thấp. Chúng thường không mang lại lợi nhuận đáng kể và có thể đòi hỏi nhiều chi phí duy trì. Doanh nghiệp có thể xem xét loại bỏ hoặc thu hồi vốn từ những phân khúc này.

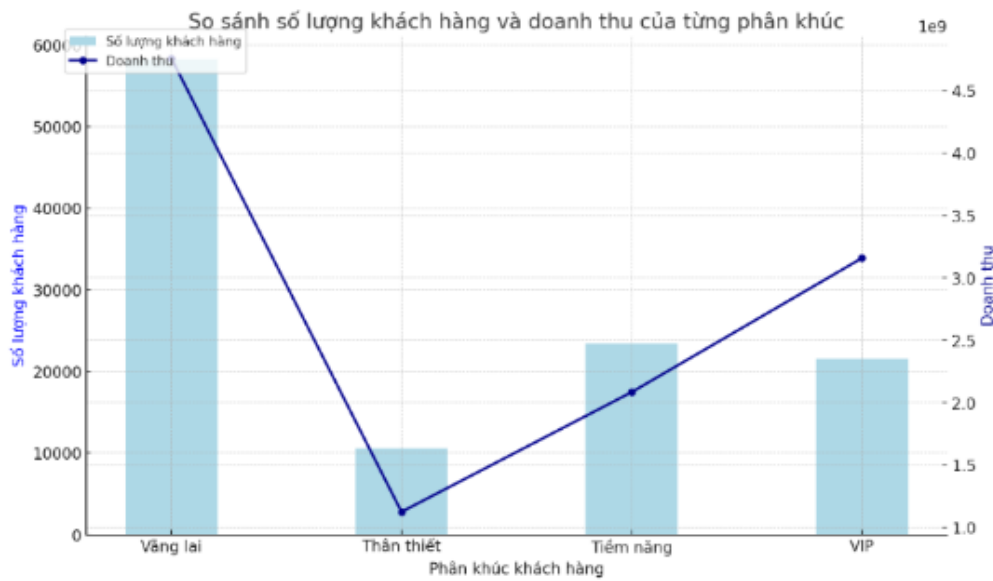
Phân loại khách hàng theo thang điểm RFM cụ thể theo 4 nhóm chính

Nhóm	Mô tả	Tổ hợp RFM
VIP	Khách hàng mua hàng thường xuyên với giá trị cao và gần đây (R, F, và M cao). Đây là nhóm khách hàng quan trọng nhất cần được ưu tiên chăm sóc.	333,334,343,344,444
Thân thiết	Khách hàng có tần suất mua hàng cao (F cao) nhưng số tiền chi tiêu tương đối cao do chi tiêu lâu dài (M trung bình cao) và thời gian gần nhất mua hàng (R trung bình thấp). Đây có thể là nhóm khách hàng thân thiết mua hàng đều đặn với giá trị đơn hàng không quá cao.	133,134,143,144,233,234,243,244
Tiềm năng	Khách hàng ít mua hàng (F thấp) nhưng đã mua gần đây (R cao), mức chi tiêu ngẫu nhiên. Nhóm này có tiềm năng phát triển thành khách hàng thân thiết hoặc VIP nếu được khai thác đúng cách. Tuy nhiên hoàn toàn có thể thành Vãng lai nếu họ cảm nhận chất lượng sản phẩm tệ.	311,312,313,314,321,322,323,324,411,412,413,414,321,322,323,324,421,422,423,424
Vãng lai	Khách hàng có tần suất mua hàng thấp và thời gian gần nhất mua hàng xa (cả R và F thấp) nhưng chi tiêu ở mức trung bình (M trung bình). Có thể xem là nhóm khách hàng vãng lai mua ít nhưng khi mua thì chi tiêu	Còn lại

segmentation	total_number
Văng lai	58,223
Thân thiết	10,526
Tiềm năng	23,411
VIP	21,613



- Nhóm khách hàng Văng lai là lớn nhất:** Điều này cho thấy một phần lớn cơ sở khách hàng hiện tại của doanh nghiệp có thể chưa được kích thích mua sắm đủ mạnh hoặc chưa thấy đủ giá trị từ sản phẩm/dịch vụ của doanh nghiệp để trở nên thân thiết hơn.
- Nhóm Thân thiết và VIP có số lượng tương đối cân đối nhưng nhỏ hơn nhiều so với nhóm Văng lai:** Điều này cho thấy bạn đã có một cơ sở khách hàng trung thành nhất định, nhưng còn nhiều cơ hội để mở rộng và tăng cường mối quan hệ này.
- Nhóm Tiềm năng đáng chú ý với 23,411 khách hàng:** Đây có thể là nhóm mục tiêu quan trọng để chuyển đổi thành nhóm VIP thông qua các chiến dịch marketing cá nhân hóa, ưu đãi đặc biệt hoặc chương trình loyalty nhằm tăng cường sự tương tác và giá trị giao dịch.
- Ưu tiên phát triển và tối ưu hóa chiến lược cho từng nhóm:** Dựa trên số lượng và đặc điểm của mỗi nhóm, doanh nghiệp có thể ưu tiên phát triển các chiến lược cụ thể. Ví dụ, tập trung vào việc tăng cường giá trị và tần suất mua hàng cho nhóm VIP và Tiềm năng, trong khi tìm cách tăng cường mức độ tương tác với nhóm Văng lai và Thân thiết để chuyển đổi họ thành các nhóm có giá trị cao hơn.



segmentation	total_number	total_revenue	revenue_per_customer
Vãng lai	58,223	4,758,110,322	81,722.1772
Thân thiết	10,526	1,124,892,409	106,867.9849
Tiềm năng	23,411	2,084,050,027	89,020.1199
VIP	21,613	3,157,208,440	146,079.1394

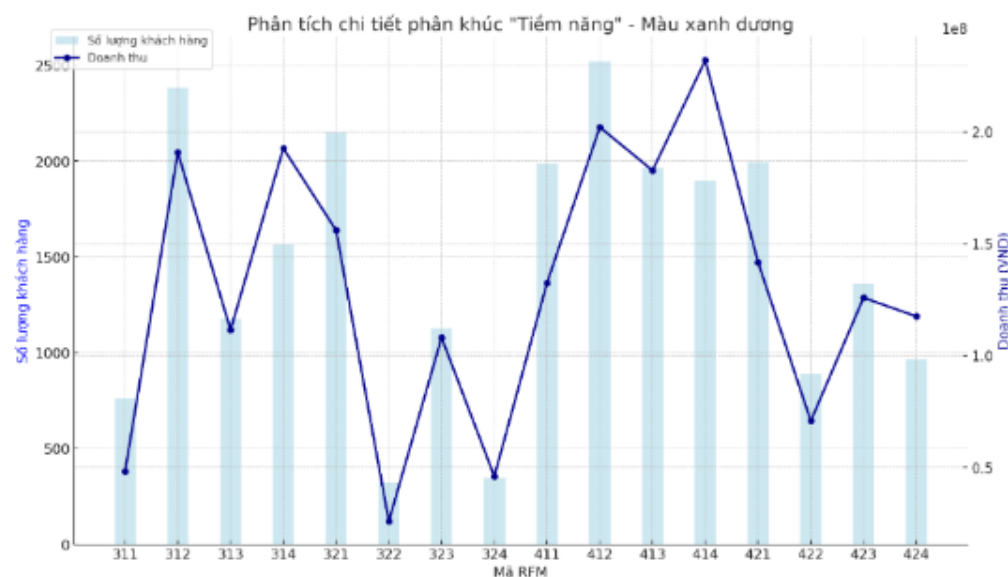
- **Vãng lai** có số lượng khách hàng lớn nhất nhưng cũng đóng góp một lượng doanh thu đáng kể, điều này cho thấy mặc dù đây là phân khúc với giá trị cá nhân thấp hơn, sự đóng góp tổng thể của họ vẫn quan trọng. Trung bình GMV của mỗi khách hàng **vãng lai** xấp xỉ bằng khách hàng **tiềm năng**. Từ đó có thể thấy, nếu có chiến lược tốt với nhóm này thì tiềm năng khai thác là hoàn toàn có.
- **Thân thiết** có ít khách hàng nhất nhưng vẫn đóng góp một lượng doanh thu đáng chú ý, điều này phản ánh giá trị cao từ mỗi khách hàng trong phân khúc này.
- **Tiềm năng** và **VIP** đều có số lượng khách hàng và mức độ đóng góp doanh thu tương xứng với kích thước của họ. Đặc biệt, phân khúc **VIP**, mặc dù không phải là lớn nhất về số lượng nhưng lại có doanh thu cao, cho thấy giá trị bình quân cao từ mỗi khách hàng.

Phân tích mức độ đóng góp doanh thu theo số lượng khách hàng



Qua biểu đồ trên ta có thể thấy, những nhóm RFM nhỏ thuộc phân khúc VIP có tỉ lệ số lượng khách hàng và doanh thu tương đối đồng đều nhau (Nhiều khách hàng đi kèm với doanh thu lớn). Biểu đồ cho thấy top 10 doanh thu xuất hiện tới 5 nhóm nhỏ của khách hàng Vãng lai. Vậy nên đây là nhóm khách hàng doanh nghiệp cũng chưa nên bỏ qua, vì hoàn toàn có thể nghiên cứu chiến lược để khai thác thêm từ họ.

Phân tích chi tiết nhóm khách hàng tiềm năng



Trong phân khúc **Tiềm năng**, có tổng cộng 16 mã RFM duy nhất với các thông số sau:

- **Số lượng khách hàng:** Có sự phân bố từ 321 đến 2521 khách hàng trong các nhóm RFM, với trung bình là khoảng 1463 khách hàng mỗi nhóm.
- **Doanh thu:** Doanh thu từ các nhóm này dao động từ khoảng 25,94 triệu đến 231,96 triệu, với giá trị trung bình là khoảng 130,25 triệu.

Có những nhóm RFM với số lượng khách hàng cao nhưng doanh thu không tương xứng, và ngược lại, một số nhóm có doanh thu cao với số lượng khách hàng tương đối ít. Điều này chỉ ra rằng giá trị trung bình từ mỗi khách hàng có sự chênh lệch giữa các nhóm.

Tổng thể, phân bố dữ liệu ở nhóm khách hàng tiềm năng còn đa dạng và chưa có 1 xu hướng cụ thể, cần thêm thời gian nghiên cứu về nhóm khách hàng này.