

22/03/2024

# CUSTOMER 360 REPORT

**Chuẩn bị bởi :**

Ngô Hoàng Anh

0966597718

[watchingbird02@gmail.com](mailto:watchingbird02@gmail.com)

# MỤC LỤC

01

## TỔNG QUAN VỀ MÔ HÌNH CUSTOMER 360

Mô hình Customer 360, ứng dụng của Customer 360

02

## TỔNG QUAN VỀ MÔ HÌNH RFM

Ma trận BCG, mô hình RFM, cách tính điểm R, F, M, phân khúc khách hàng

03

## QUY TRÌNH THỰC HIỆN

Làm sạch dữ liệu, visualize và phân tích, báo cáo

04

## SQL CODE QUERY

## I. TỔNG QUAN VỀ MÔ HÌNH CUSTOMER 360

- **Mô hình Customer 360** sử dụng dữ liệu về những gì khách hàng của bạn đã làm trước đây để dự đoán những gì họ có khả năng làm tiếp theo. Bằng cách tạo ra một bức tranh hoàn chỉnh về khách hàng, doanh nghiệp có thể hiểu rõ hơn về những gì khách hàng có thể cần trong tương lai. Hơn thế nữa, nó cung cấp cho các công ty cái nhìn sâu sắc về tương lai của chính họ bằng cách xác định vị trí hoặc cách thức họ có thể cần điều chỉnh các dịch vụ cung cấp đồng thời giảm thiểu tác động của những thay đổi trong tương lai bằng cách dự đoán cách khách hàng có thể phản ứng.



Hình 1: Mô hình Customer 360

- **Customer 360** được tích hợp từ nhiều nguồn để cho thấy được toàn diện về khách hàng. Trong đó có 4 nhóm dữ liệu chính:

Demographics Data	Transaction Data	Behavioral Data	Interaction Data
Thông tin khách hàng về nhân khẩu học như giới tính, tuổi tác, nghề nghiệp, tôn giáo,...	Thông tin về hành vi của khách hàng như thói quen mua hàng, tần suất mua hàng, sản phẩm hoặc dịch vụ ưa thích, thời gian và cách thức tương tác với doanh nghiệp.	Thông tin về các giao dịch mua hàng hoặc giao dịch tài chính của khách hàng như lịch sử mua hàng, số lượng mua hàng, giá trị giao dịch, phương thức thanh toán, v.v.	Thông tin về các tương tác giữa khách hàng và doanh nghiệp qua các kênh như điện thoại, email, trang web, mạng xã hội, v.v

Bảng 1: 4 yếu tố chính của Customer 360

Việc sử dụng mô hình **Customer 360** trong kinh doanh mang lại nhiều lợi ích cho doanh nghiệp để có thể tiếp cận, phân tích và dự đoán khách hàng một cách toàn diện và đầy đủ nhất, ngoài ra còn cải thiện và nâng cao trải nghiệm của khách hàng, và hoàn thiện liên kết khách hàng giữa những bộ phận.

## II. TỔNG QUAN VỀ MÔ HÌNH RFM:

- Ma trận BCG và ứng dụng:



Hình 2: Ma trận BCG

- Ma trận BCG là một công cụ được sử dụng để phân tích mô hình kinh doanh, đánh giá vị thế cạnh tranh và tiềm năng phát triển của danh mục sản phẩm hoặc đơn vị kinh doanh trong doanh nghiệp. Chính vì thế ma trận BCG còn được gọi là ma trận danh mục sản phẩm.

- Mô hình này được xây dựng dựa trên hai yếu tố là thị phần tương đối và tốc độ tăng trưởng của thị trường. Trên cơ sở hai yếu tố đó, BCG Matrix phân nhỏ danh mục kinh doanh của doanh nghiệp thành 4 nhóm bao gồm ngôi sao, dấu hỏi, bò sữa và con chó.

- Trong doanh nghiệp, việc áp dụng ma trận BCG sẽ hỗ trợ nhà quản trị ra quyết định về cách phân phối tài nguyên, dự án đầu tư và chiến lược phát triển sản phẩm/dịch vụ, dựa trên sự hiểu biết về vị trí của mỗi sản phẩm hoặc đơn vị trong ma trận này.

Ma trận BCG là một mô hình kinh doanh hiệu quả được sử dụng rộng rãi bởi các doanh nghiệp trên thế giới. Mô hình này có một số ưu điểm sau:

- + **Dễ hiểu và sử dụng:** Ma trận BCG sử dụng hai yếu tố chính là thị phần tương đối và tốc độ tăng trưởng của thị trường để đánh giá vị thế cạnh tranh của sản phẩm/dịch vụ. Do đó, ma trận BCG dễ hiểu và sử dụng cho cả các doanh nghiệp lớn và nhỏ.

- + **Có thể được áp dụng cho nhiều ngành nghề:** Ma trận BCG có thể được áp dụng cho nhiều ngành nghề khác nhau, từ các ngành sản xuất đến các ngành dịch vụ.

- + **Có thể được sử dụng để hỗ trợ doanh nghiệp đưa ra các quyết định chiến lược hiệu quả:** Ma trận BCG có thể giúp doanh nghiệp phân tích và đánh giá vị thế cạnh tranh của các sản phẩm/dịch vụ trong danh mục kinh doanh. Từ đó, doanh nghiệp có thể đưa ra các quyết định chiến lược phù hợp để tối đa hóa lợi nhuận.

Ngược lại, ma trận BCG cũng có một số hạn chế sau:

- + Ma trận BCG chỉ dựa trên hai yếu tố là thị phần tương đối và tốc độ tăng trưởng của thị trường để đánh giá vị thế cạnh tranh của sản phẩm/dịch vụ. **Do đó, ma trận BCG có thể không phản ánh chính xác tiềm năng phát triển của sản phẩm/dịch vụ.**

- + **Ma trận BCG không tính đến các yếu tố khác như chi phí, lợi nhuận, khả năng cạnh tranh,...** Do đó, doanh nghiệp cần cân nhắc các yếu tố này khi đưa ra quyết định chiến lược dựa trên ma trận BCG.

- **RFM Customer 360** là một phương pháp phân tích khách hàng dựa trên mô hình RFM (Recency - Frequency - Monetary Value), kết hợp với các dữ liệu khách hàng khác để tạo ra một bức tranh tổng thể về khách hàng của một doanh nghiệp. Phương pháp này giúp doanh nghiệp có thể đánh giá giá trị của từng khách hàng dựa trên mức độ tương tác của họ với doanh nghiệp và đưa ra các chiến lược tiếp cận khác nhau để tối ưu hóa việc chăm sóc khách hàng.

Trong phân tích RFM Customer 360 có một số khái niệm như hình sau:

Thuật ngữ	Khái niệm
<b>Recency (R)</b>	Khoảng thời gian tính từ lần giao dịch cuối cùng của khách hàng với doanh nghiệp đến thời điểm hiện tại. Recency được sử dụng để đo lường mức độ liên quan của khách hàng đến doanh nghiệp trong thời gian gần đây.
<b>Frequency (F)</b>	Số lần giao dịch mà khách hàng đã có với doanh nghiệp trong một khoảng thời gian xác định (thường là một năm). Frequency được sử dụng để đo lường mức độ tương tác của khách hàng với doanh nghiệp.
<b>Monetary Value (M)</b>	Giá trị tiền khách hàng đã chi tiêu cho sản phẩm hoặc dịch vụ của doanh nghiệp trong một khoảng thời gian xác định (thường là một năm). Monetary Value được sử dụng để đo lường mức độ giá trị của khách hàng đối với doanh nghiệp.
<b>Segmentation</b>	Quá trình chia khách hàng thành các nhóm dựa trên các yếu tố RFM để phân tích và hiểu hơn về hành vi mua hàng và tương tác của khách hàng với doanh nghiệp.

*Bảng 2: Các thuật ngữ trong RFM Model*

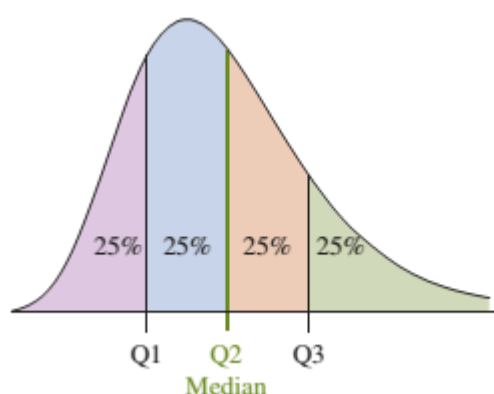
- Cách tính điểm R, F, M trong RFM Model:

Bằng cách sử dụng tứ phân vị trong xác suất thống kê, ta có thể chia R, F, M thành 4 phần, mỗi phần tương ứng với hành vi của khách hàng.

- Tứ phân vị trong xác suất thống kê là cách chia dữ liệu thành bốn phần bằng nhau dựa trên thứ tự của chúng. Cụ thể, tứ phân vị chia dữ liệu thành bốn phần có cùng số lượng quan sát, tức là mỗi phần chứa 25% tổng số quan sát. Tứ phân vị có 3 giá trị, đó là tứ phân vị thứ nhất (Q1), thứ hai (Q2) và thứ ba (Q3). Ba giá trị này chia một tập hợp dữ liệu (đã sắp xếp dữ liệu theo trật từ từ bé đến lớn) thành 4 phần có số lượng quan sát đều nhau. Giả sử tập dữ liệu có n quan sát.

+ Tứ phân vị thứ nhất được tính bằng công thức  $Q1 = 25 * (n+1) / 100$ .

- + Tứ phân vị thứ hai là giá trị trung vị  $Q2 = (n+1)/2$
- + Tứ phân vị thứ ba được tính bằng công thức  $Q3 = 75 * (n+1) / 100$ .
- + Tứ phân vị thứ 4 chính là phần còn lại



Hình 3: Tứ phân vị

Với mỗi biến (R, F, M), chúng ta chia thành các nhóm từ 1 đến 4, với nhóm 4 đại diện cho các giá trcao nhất và nhóm 1 đại diện cho các giá trị thấp nhất. Riêng với R, điểm 4 có nghĩa là khách hàng mới mua hàng gần đây, và 1 là khách hàng đã rất lâu chưa có phát sinh giao dịch với doanh nghiệp.

- Bảng phân khúc khách hàng:

Nhóm khách hàng	Mô tả nhóm khách hàng	RFM
Champions	Đây là nhóm khách hàng tốt nhất, họ đã mua hàng gần đây nhất, mua hàng thường xuyên nhất và chi tiêu nhiều tiền nhất.	444, 443, 434, 344
Loyal Customers	Khách hàng trung thành với thương hiệu; có thể họ không chi tiêu nhiều bằng các khách hàng VIP, nhưng họ thường xuyên sử dụng dịch vụ và sử dụng dịch vụ trong thời gian gần đây	442, 441, 432, 431, 433, 343, 342, 341
Potential	Khách hàng có tiềm năng trở thành các khách hàng trung thành; họ sử dụng dịch vụ ít thường xuyên hơn, tuy nhiên có mức chi tiêu cao	424, 423, 324, 323, 413, 414, 343, 334
Promising	Khách hàng chi tiêu ít hơn khách hàng trung thành, nhưng sử dụng dịch vụ khá thường	333, 332, 331, 313



	xuyên và khá gần đây	
Need Attention	Khách hàng đã chi tiêu nhiều và sử dụng dịch vụ thường xuyên trong quá khứ, tuy nhiên đã 1 thời gian rồi chưa sử dụng dịch vụ	244, 234, 243, 233, 224, 214, 213, 134, 144, 143, 133
Lost	Khách hàng đã sử dụng dịch vụ tuy nhiên không còn quay lại nữa	111, 112, 113, 114, 121, 122, 123, 221, 211, 222

*Bảng 3: Phân khúc khách hàng*

### III. QUY TRÌNH THỰC HIỆN:

- Đầu tiên, chúng ta có 2 bảng chứa thông tin chính:
- Bảng Customer\_Register chứa thông tin của khách hàng:

STT	Tên	Mô tả
1	ID	Mã khách hàng
2	Contract	Mã hợp đồng
3	LocationID	Mã vị trí
4	BranchCode	Mã chi nhánh
5	Status	Trạng thái
6	Created_date	Ngày đăng ký
7	Stop_date	Ngày hủy

*Bảng 4: Customer\_Register*

- Bảng Customer\_Transaction chứa thông tin về giao dịch của khách hàng:

STT	Tên	Mô tả
1	ID	Mã giao dịch
2	CustomerID	Mã khách hàng
3	Purchase_Date	Ngày thanh toán
4	GMV	Gross Monetary Value

*Bảng 5: Customer\_Transaction*

- Sau khi xác định bảng xong, ta nhận thấy rằng trong bảng dữ liệu còn một số những dữ liệu xấu, như những khách hàng có ID = 0 là của hệ thống, những ô không có giá trị hoặc hiển thị “Null”, hay là những ô ngày tháng còn để sai cách. Vì vậy cần làm sạch dữ liệu có trong bảng bằng những cách:

- Loại bỏ những ô có dữ liệu sai hoặc không có dữ liệu
- Chỉnh lại ngày tháng thành yyyy – mm – dd

- Tiếp theo, chúng ta sẽ chạy code SQL để có thể tính toán được R, F, M, sắp xếp lại và lấy được những số liệu cần thiết ( Code SQL sẽ để dưới phần IV )

- Dựa vào cách tính tứ phân vị, ta chia được các điểm R, F, M như bảng sau:

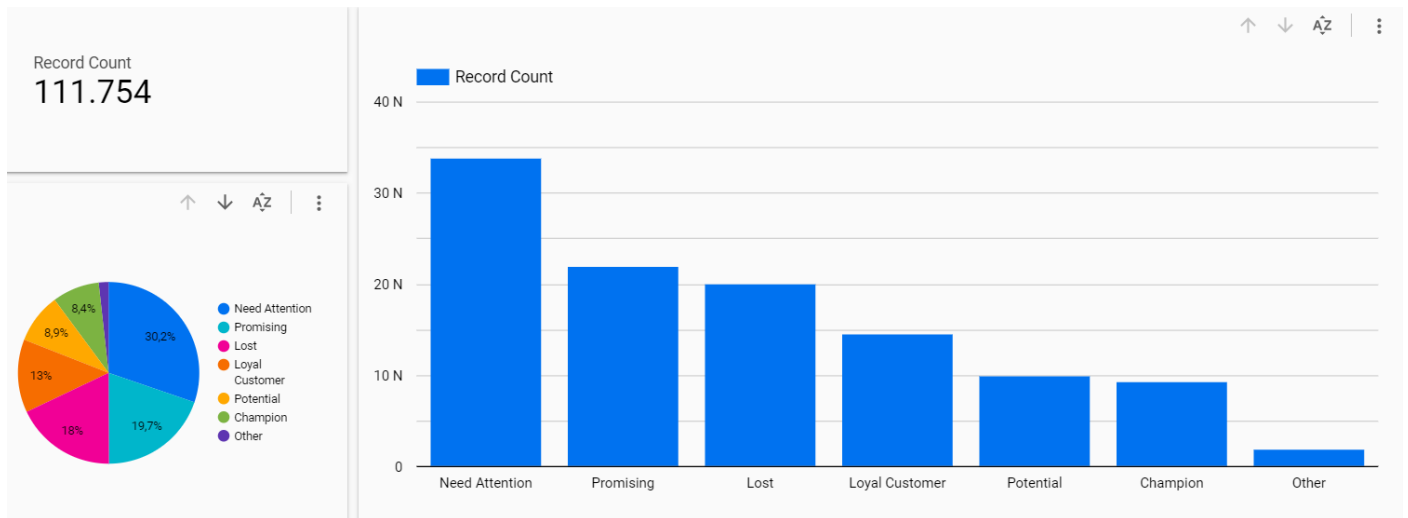
Điểm	Recency	Frequency	Monetary
1	Trên 89 ngày	0 – 0,5	0 – 20000
2	63 – 88 ngày	Trên 0,5 – 1	Trên 20000 - 50000
3	31 – 62 ngày	Trên 1 – 1.66667	Trên 50000 – 116000
4	Dưới 30 ngày	Trên 1,66667 - 2	Trên 116000

*Bảng 6: Bảng phân bố R, F, M*

- Cuối cùng là dùng Locker Studio, một công cụ trực quan hóa dữ liệu của Google để có thể tạo ra



bảng và phân tích:



Hình 4: Biểu đồ phân tích khách hàng

- Dữ liệu được tổng hợp từ 111.754 khách hàng tham gia vào giao dịch
- Qua biểu đồ, ta thấy được rằng:
  - + Tỷ lệ khách hàng thuộc diện **“Need Attention”** và **“Lost”** đang chiếm số lượng cực lớn ( 30,2% và 18%) , cho thấy rằng đông đảo khách hàng đã từng giao dịch nhưng hiện tại đã một thời gian không sử dụng, hoặc không còn sử dụng nữa . Điều này cho thấy rằng sản phẩm đang bị mất khách hàng trầm trọng. Nguyên nhân có thể phát sinh ra ở nhiều khía cạnh như chất lượng sản phẩm đi xuống, sản phẩm đối thủ cạnh tranh ưu thế hơn. Vì vậy, cần cải thiện chất lượng sản phẩm, đồng thời liên hệ, tương tác lại với khách hàng bằng những ưu đãi và khuyến mại khác nhau.

- + Tiếp theo là những khách hàng thuộc diện **“Promising”** và **“Loyal Customer”** đang ở vị trí thứ 2 và thứ 4, với 19,7% và 13%. Đó là những khách hàng đang giao dịch khá thường xuyên và những khách hàng trung thành. Lý do có thể rằng những sản phẩm của ta đang đáp ứng được nhu cầu nhất thời của họ. Đây là những khách hàng thật sự quan trọng, nhất là trong bối cảnh ngày càng nhiều những khách hàng đang rời bỏ việc sử dụng sản phẩm. Do đó, cần đặc biệt quan tâm và chăm sóc những khách hàng này, để tránh việc họ rơi vào những khách hàng ở trên.

- + Trong khi đó , một điều đáng mừng những khách hàng tiềm năng và khách hàng trung thành vẫn giữ một tỉ lệ khá ổn, đó là **“ Potential”** và **“ Champion”** với 8,9% và 8,4%. Đó chứng tỏ sản phẩm vẫn được lòng một vài tệp khách hàng.

Qua một số những insight thu được ở trên, ta có thể rút ra một cái nhìn tổng quát về sản phẩm đang có:

- ➔ Cần nâng cao và cải thiện sản phẩm để có thể giữ chân được tệp khách hàng trung thành, đồng thời cũng không quên tung ra những chiến lược marketing khác nhau như khuyến mãi, giảm giá để thu hút khách hàng
- ➔ Với một số khách hàng đã lâu không sử dụng sản phẩm, có thể giữ liên lạc với họ, hoặc gửi những phiếu đánh giá sản phẩm, để có thể biết được rằng sản phẩm không vừa lòng khách hàng tại đâu, cần khắc phục những khuyết điểm nào và cải thiện.
- ➔ Tạo lợi thế cạnh tranh so với những công ty khác bằng cách ra mắt nhiều sản phẩm với mẫu mã đa dạng khác nhau.

## IV. SQL CODE QUERRY:

```
with cte as (  
select contract,
```

```

datediff('2022-09-01', max(purchase_date)) as Recency ,
sum(GMV)/timestampdiff(year, max(created_date), '2022-09-01') as Monetary ,
count(customerid)*1.0/ timestampdiff(year, max(created_date), '2022-09-01') as
Frequency ,
row_number() over(order by datediff('2022-09-01', max(purchase_date))) as
rn_recency ,
row_number() over(order by sum(GMV)/timestampdiff(year, max(created_date), '2022-
09-01') desc ) as rn_monetary ,
row_number() over(order by count(customerid)*1.0/ timestampdiff(year,
max(created_date), '2022-09-01') desc ) as rn_frequency
from customer_registered cr
join customer_transaction ct on cr.id = ct.customerid
where ct.customerid <> 0 and cr.stopdate = 0
group by contract
order by count(customerid)*1.0/ timestampdiff(year, max(created_date), '2022-09-
01') desc
),
customer_statistics as(
select Contract, Recency, Frequency, Monetary ,
case when rn_recency >= (select min(rn_recency) from cte) and rn_recency < (select
count(rn_recency)*0.25 from cte) then 4
    when rn_recency >= (select count(rn_recency)*0.25 from cte) and rn_recency <
(select count(rn_recency)*0.5 from cte) then 3
    when rn_recency >= (select count(rn_recency)*0.5 from cte) and rn_recency <
(select count(rn_recency)*0.75 from cte) then 2
    else 1 end as R,
case when rn_frequency >= (select min(rn_frequency) from cte) and rn_frequency <
(select count(rn_frequency)*0.25 from cte) then 1
    when rn_frequency >= (select count(rn_frequency)*0.25 from cte) and
rn_frequency < (select count(rn_frequency)*0.5 from cte) then 2
    when rn_frequency >= (select count(rn_frequency)*0.5 from cte) and
rn_frequency < (select count(rn_frequency)*0.75 from cte) then 3
    else 4 end as F,
case when rn_monetary >= (select min(rn_monetary) from cte) and rn_monetary <
(select count(rn_monetary)*0.25 from cte) then 1
    when rn_monetary >= (select count(rn_monetary)*0.25 from cte) and rn_monetary
< (select count(rn_monetary)*0.5 from cte) then 2
    when rn_monetary >= (select count(rn_monetary)*0.5 from cte) and rn_monetary
< (select count(rn_monetary)*0.75 from cte) then 3
    else 4 end as M
from cte
where recency is not null
),
customer_RFM as (
select *, concat(R,F,M) as RFM
from customer_statistics
)
select *,
case when rfm in (444, 443, 434, 344) then 'Champion'
    when rfm in (442, 441, 432, 431, 433, 343, 342, 341) then 'Loyal Customers'
    when rfm in (424, 423, 324, 323, 413, 414, 343, 334) then 'Potential'
    when rfm in (333, 332, 331, 313) then 'Promising'
    when rfm in (244, 234, 243, 233, 224, 214, 213, 134, 144, 143, 133) then
'Need Attention'
    when rfm in (111, 112, 113, 114, 121, 122, 123, 221, 211, 222) then 'Lost'
    else 'Other'
end as Cus_type
from customer_RFM

```