

CUSTOMER ANALYTICS REPORT



PREPARED BY
NGUYEN ANH NGUYEN

Table of Contents

I. Giới thiệu về Customer 360.....	2
II. Mô hình RFM	2
III. BCG Matrix	3
IV. Cách làm.....	4
1. Data preparation.....	4
2. Phân tích chi tiết	5
V. Code xử lý.....	7

I. Giới thiệu về Customer 360



Customer 360 là cái nhìn toàn diện về khách hàng, được xây dựng từ bốn nhóm dữ liệu chính:

- Demographic (thông tin nhân khẩu học): Tuổi, giới tính, đặc điểm cá nhân.
- Behavioral (hành vi): Sở thích, mong muốn, ý kiến.
- Transaction (giao dịch): Lịch sử mua hàng, thanh toán, sử dụng sản phẩm.
- Interaction (tương tác): Email, mạng xã hội, chăm sóc khách hàng, truy cập web.

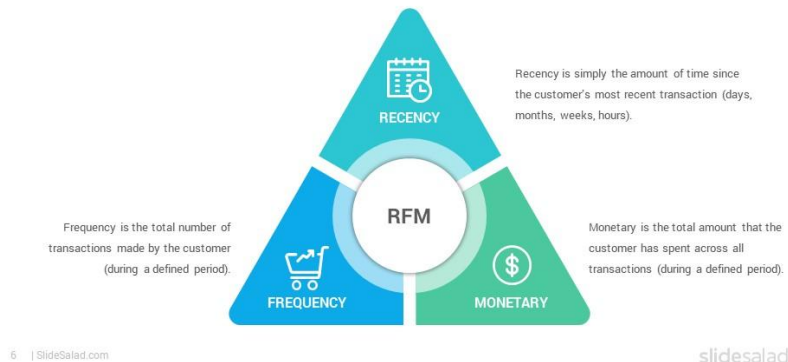
Tất cả được tổng hợp thành một hồ sơ duy nhất, giúp doanh nghiệp hiểu rõ khách hàng để phục vụ tốt hơn và cá nhân hóa trải nghiệm hiệu quả hơn.

Bản báo cáo này tập trung sử dụng và phân tích khách hàng dựa trên Transaction data thu thập trong 3 tháng (01/06/2022 – 31/08/2022) thông qua mô hình phân tích RFM (Recency, Frequency, Monetary).

II. Mô hình RFM

RFM Customer Segmentation Model

RFM Customer Segmentation Model



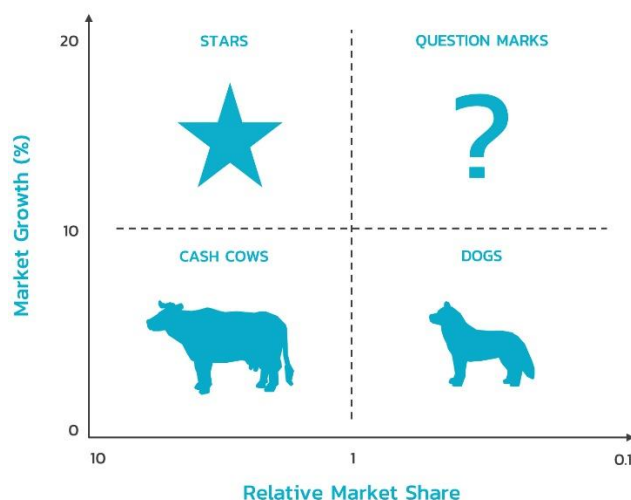
Mô hình RFM là một cách đơn giản nhưng hiệu quả để doanh nghiệp phân loại và đánh giá khách hàng dựa trên 3 yếu tố:

- **Recency** (Gần đây): Lần mua hàng gần nhất là khi nào
- **Frequency** (Tần suất): Mua hàng bao nhiêu lần
- **Monetary** (Giá trị): Đã chi tiêu bao nhiêu tiền

Bằng cách xem xét ba yếu tố này, doanh nghiệp có thể nhận biết ai là khách hàng trung thành, ai có nguy cơ rời bỏ, từ đó đưa ra những ưu đãi hoặc chiến lược chăm sóc phù hợp để giữ chân và tăng giá trị từ khách hàng đó.

Qua bản báo cáo này, chúng ta sẽ chấm điểm và chia khách hàng theo 4 phân khúc dựa trên ma trận BCG.

III. BCG Matrix



Phân khúc khách hàng theo mô hình BCG Matrix giúp doanh nghiệp xác định đâu là nhóm khách hàng nên tập trung nguồn lực. Dựa trên hai tiêu chí là mức độ tăng trưởng (khách hàng có đang tăng chi tiêu hay tương tác không) và giá trị đóng góp (chi tiêu nhiều hay ít), khách hàng được chia thành 4 nhóm.

- **Khách Hàng VIP (Star):** Những khách hàng có giá trị cao và đang tăng trưởng mạnh, cần được ưu tiên chăm sóc và phát triển lâu dài.
- **Khách Hàng Trung Thành (Cash Cows):** Tuy không còn tăng trưởng nhiều nhưng vẫn mang lại doanh thu ổn định, nên duy trì mối quan hệ hiệu quả.
- **Khách Hàng Tiềm Năng (Question Marks):** Có tiềm năng nhưng chưa mang lại nhiều giá trị, cần được theo dõi và thử nghiệm để phát triển thêm.
- **Khách Hàng Vãng Lai (Dogs):** là những khách hàng ít giá trị và không tăng trưởng, có thể được loại khỏi các chiến dịch chính hoặc tái kích hoạt nếu phù hợp.

IV. Cách làm

1. Data preparation.

- Bảng Customer_Register

Name	Type	Definition
ID	bigint	Mã định danh duy nhất cho từng dòng dữ liệu
Contract	varchar	Mã hợp đồng hoặc chuỗi định danh liên quan đến hợp đồng
LocationID	int	Mã định danh vị trí hoặc khu vực
BranchCode	tinyint	Mã chi nhánh (số nguyên nhỏ, dùng cho chi nhánh cụ thể)
Status	tinyint	Trạng thái
created_date	datetime	Ngày giờ tạo bản ghi hoặc bắt đầu hợp đồng

- Bảng Customer_Transaction

Name	Type	Definition
ID	bigint	Mã định danh duy nhất cho mỗi bản ghi trong bảng
CustomerID	varchar	Mã khách hàng, chuỗi ký tự đại diện cho từng khách hàng (tối đa 200 ký tự)
Purchase_Date	datetime	Ngày và giờ khách hàng thực hiện mua hàng
GMV	bigint	Tổng giá trị đơn hàng (Gross Merchandise Value) của giao dịch đó

- Lọc khách hàng: Join 2 bảng và lọc bằng điều kiện customerIS != 0 và customerID is not null
- Sử dụng Interquartile để chia tỷ lệ dữ liệu RFM của khách hàng

→ Kết quả:

Result	1	2	3	4
Recency	≥ 92 ngày	62 – 91 ngày	32 – 61 ngày	1 – 31 ngày

Frequency	1 lần	2 lần	3 lần	4 lần
Monetary	$\leq 12k$	12k – 17k	17k – 20k	$\geq 20k$

→ Kết quả cho thấy dữ liệu thu thập từ 114 nghìn khách hàng được chia thành 47 tổ hợp

114.08K

Total customer

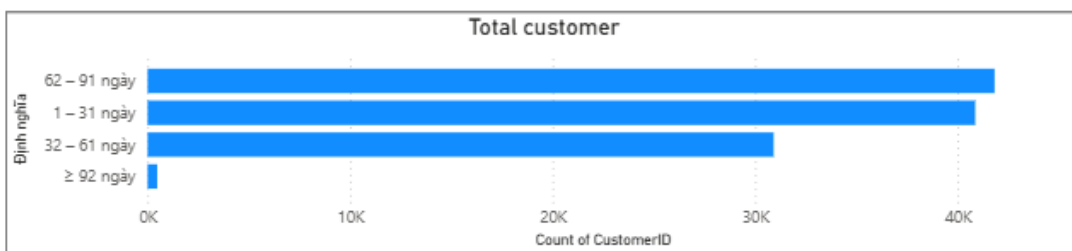
47

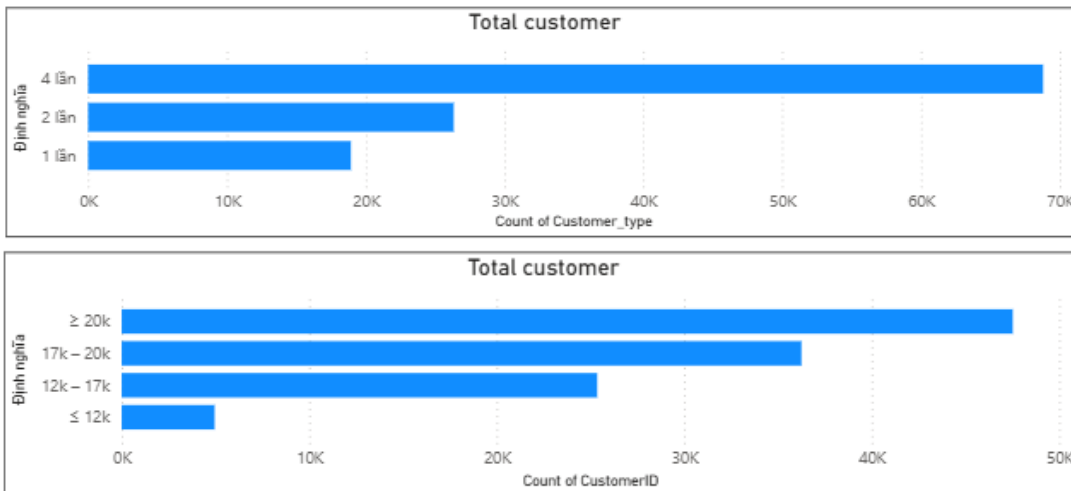
Total Group of RFM

- Chia tổ hợp khách hàng theo BCG matrix

Nhóm khách hàng	Mô tả	Tổ hợp
Khách hàng VIP	Khách hàng có mức chi tiêu lớn, thương xuyên sử dụng dịch vụ và mới gần đây	444, 443, 343, 344, 333, 334, 244, 243
Khách hàng Trung Thành	Khách hàng thường xuyên sử dụng dịch vụ, mới đây tuy nhiên mức chi tiêu không nhiều như khách hàng VIP	133, 134, 143, 144, 233, 234, 234, 244
Khách hàng Tiềm Năng	Khách hàng mới sử dụng dịch vụ, chưa nhiều nhưng mức chi tiêu đáng kể	311, 312, 313, 314, 321, 322, 323, 324, 332, 334, 412, 413, 414, 422, 423, 424, 432, 442
Khách hàng Vãng Lai	Khách hàng mới sử dụng dịch vụ, chưa nhiều và chi tiêu chưa đáng chú ý hoặc từ lâu chưa sử dụng dịch vụ.	121, 122, 123, 124, 131, 132, 133, 134, 141, 142, 143, 144, 311, 312, 411, 412.

2. Phân tích chi tiết





- Recency (Thời gian mua hàng gần nhất):

Phần lớn khách hàng thực hiện giao dịch trong vòng 1–31 ngày và 62–91 ngày, mỗi nhóm đều có hơn 40.000 khách hàng. Nhóm từ 32–61 ngày có số lượng thấp hơn nhưng vẫn duy trì ở mức đáng kể. Đáng chú ý, nhóm khách hàng mua hàng cách đây trên 92 ngày gần như rất ít, cho thấy phần lớn khách hàng vẫn có tương tác gần đây và khả năng rời bỏ chưa quá cao.

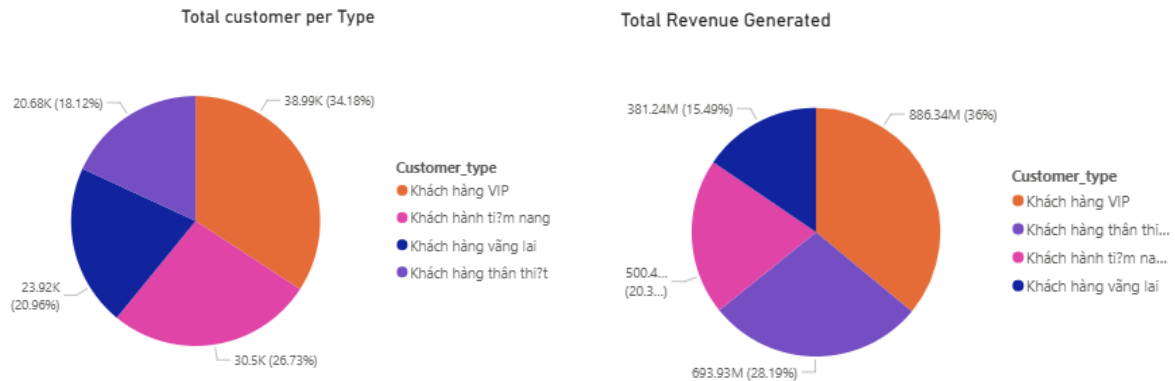
- Frequency (Tần suất mua hàng):

Nhóm khách hàng mua hàng 4 lần chiếm số lượng lớn nhất với gần 70.000 người, cho thấy một tệp khách hàng trung thành đáng kể. Nhóm mua 2 lần có khoảng 30.000 người, trong khi nhóm mua 1 lần chiếm tỷ trọng thấp nhất. Điều này phản ánh rằng chương trình chăm sóc sau mua đang hoạt động hiệu quả và có thể tiếp tục khai thác nhóm mua nhiều lần để tăng doanh thu.

- Monetary (Giá trị chi tiêu):

Nhóm khách hàng có chi tiêu từ 20K trở lên là đông nhất, với gần 50.000 khách hàng, tiếp theo là nhóm chi tiêu từ 17K–20K. Nhóm chi tiêu 12K–17K chiếm số lượng ít hơn, trong khi nhóm chi dưới 12K là thấp nhất. Điều này cho thấy phần lớn doanh thu đến từ nhóm có sức mua cao, và nhóm có chi tiêu thấp nên được cân nhắc nuôi dưỡng hoặc loại bỏ khỏi các chiến dịch đầu tư cao.

→ Doanh nghiệp nên tập trung vào việc **giữ chân và gia tăng giá trị từ nhóm mua nhiều lần và chi tiêu cao**, đồng thời xây dựng các chiến lược **kích hoạt lại nhóm khách hàng lâu chưa mua và chuyển đổi nhóm mua ít hoặc chi tiêu thấp sang nhóm có giá trị cao hơn**.



Phân tích cơ cấu khách hàng cho thấy nhóm **“Khách hàng VIP”** chiếm tỷ trọng lớn nhất về số lượng với **34.18%** (38.9K khách hàng), tiếp theo là **“Khách hàng tiềm năng”** (26.73%), **“Khách hàng vãng lai”** (20.96%) và cuối cùng là **“Khách hàng thân thiết”** (18.12%). Tuy nhiên, khi so sánh về tổng doanh thu đóng góp, nhóm **“Khách hàng VIP”** tiếp tục giữ vai trò chủ lực khi tạo ra **36% tổng doanh thu** (886.34M), dù chỉ chiếm 1/3 về số lượng. Đáng chú ý, nhóm **“Khách hàng thân thiết”** tuy có số lượng thấp nhất nhưng vẫn đóng góp **28.19% doanh thu**, vượt cả nhóm **“Khách hàng tiềm năng”** (20.3%) và **“Khách hàng vãng lai”** (15.49%). Kết quả này cho thấy **giá trị đóng góp không hoàn toàn tỷ lệ thuận với số lượng**, và nhóm **“Khách hàng thân thiết”** và **“VIP”** mang lại hiệu quả cao nhất.

Do đó, doanh nghiệp nên tập trung nhiều hơn vào việc **chăm sóc nhóm khách hàng này**, đồng thời **xây dựng chiến lược chuyển đổi nhóm vãng lai và tiềm năng thành nhóm trung thành** nhằm tối ưu hóa doanh thu dài hạn.

V. Code xử lý.

```
select ct.CustomerID, datediff(day,max(ct.Purchase_Date), '2022-09-01') as recency,
round (1.0 * COUNT(ct.ID)/datediff(year,cr.created_date, '2022-09-01'),2) as
frequency,
round (1.0 * sum(ct.GMV)/datediff(year,cr.created_date, '2022-09-01'),2) as monetary,
row_number() over (order by datediff(day,max(ct.Purchase_Date), '2022-09-01')) as
rn_recency,
row_number() over ( order by round (1.0 * COUNT
(ct.ID)/datediff(year,cr.created_date, '2022-09-01'),2)) as rn_frequency,
row_number() over ( order by round (1.0 * sum (ct.GMV)/datediff(year,cr.created_date,
'2022-09-01'),2)) as rn_monetary
into #calculation
from Customer_Transaction ct
join Customer_Registered cr on ct.CustomerID = cr.ID
WHERE ct.CustomerID != 0 and ct.CustomerID is not null
group by ct.CustomerID, cr.created_date

select * from #calculation1

select *,
```



```

case
    when recency < (select recency from #calculation
                     where rn_recency = (select cast(count(distinct(CustomerID)) *0.25 as int)
from #calculation)) then '4'
    when recency >= (select recency from #calculation
                     where rn_recency = (select cast(count(distinct(CustomerID)) *0.25 as int)
from #calculation))
        and recency < (select recency from #calculation
                       where rn_recency = (select cast(count(distinct(CustomerID)) *0.5 as int)
from #calculation)) then '3'
    when recency >= (select recency from #calculation
                     where rn_recency = (select cast(count(distinct(CustomerID)) *0.5 as int)
from #calculation))
        and recency < (select recency from #calculation
                       where rn_recency = (select cast(count(distinct(CustomerID)) *0.75 as int)
from #calculation)) then '2'
    else '1'
end as R,
case
    when frequency < (select frequency from #calculation
                      where rn_frequency = (select cast(count(distinct(CustomerID)) *0.25 as
int) from #calculation)) then '1'
    when frequency >= (select frequency from #calculation
                       where rn_frequency = (select cast(count(distinct(CustomerID)) *0.25 as
int) from #calculation))
        and frequency < (select frequency from #calculation
                          where rn_frequency = (select cast(count(distinct(CustomerID)) *0.5 as
int) from #calculation)) then '2'
    when frequency >= (select frequency from #calculation
                       where rn_frequency = (select cast(count(distinct(CustomerID)) *0.5 as
int) from #calculation))
        and frequency < (select frequency from #calculation
                          where rn_frequency = (select cast(count(distinct(CustomerID)) *0.75 as
int) from #calculation)) then '3'
    else '4'
end as F,
case
    when monetary < (select monetary from #calculation
                     where rn_monetary = (select cast(count(distinct(CustomerID)) *0.25 as
int) from #calculation)) then '1'
    when monetary >= (select monetary from #calculation
                      where rn_monetary = (select cast(count(distinct(CustomerID)) *0.25 as
int) from #calculation))
        and monetary < (select monetary from #calculation
                         where rn_monetary = (select cast(count(distinct(CustomerID)) *0.5 as int)
from #calculation)) then '2'
    when monetary >= (select monetary from #calculation
                      where rn_monetary = (select cast(count(distinct(CustomerID)) *0.5 as int)
from #calculation))
        and monetary < (select monetary from #calculation
                         where rn_monetary = (select cast(count(distinct(CustomerID)) *0.75 as
int) from #calculation)) then '3'
    else '4'
end as M
into #newcalculation

```

```
from #calculation
```

```
select *,  
concat(R,F,M) as RFM  
from #newcalculation
```

```
select *,  
CASE  
    when "Group" in ('444', '443', '343', '344', '333', '334', '244', '243') then  
'Khách hàng VIP'  
    when "Group" in  
( '311', '312', '313', '314', '321', '322', '323', '324', '332', '334', '412', '413', '414', '422',  
'423', '424', '432', '442') then 'Khách hàng tiềm năng'  
    when "Group" in ('133', '134', '143', '144', '233', '234', '234', '244') then 'Khách  
hàng thân thiết'  
    else 'Khách hàng vắng lai'  
    END as "Customer_type"  
from dbo.RFM_Result rr
```