

FIRST BUSINESS INTELLIGENCE ANALYTICS PROJECT

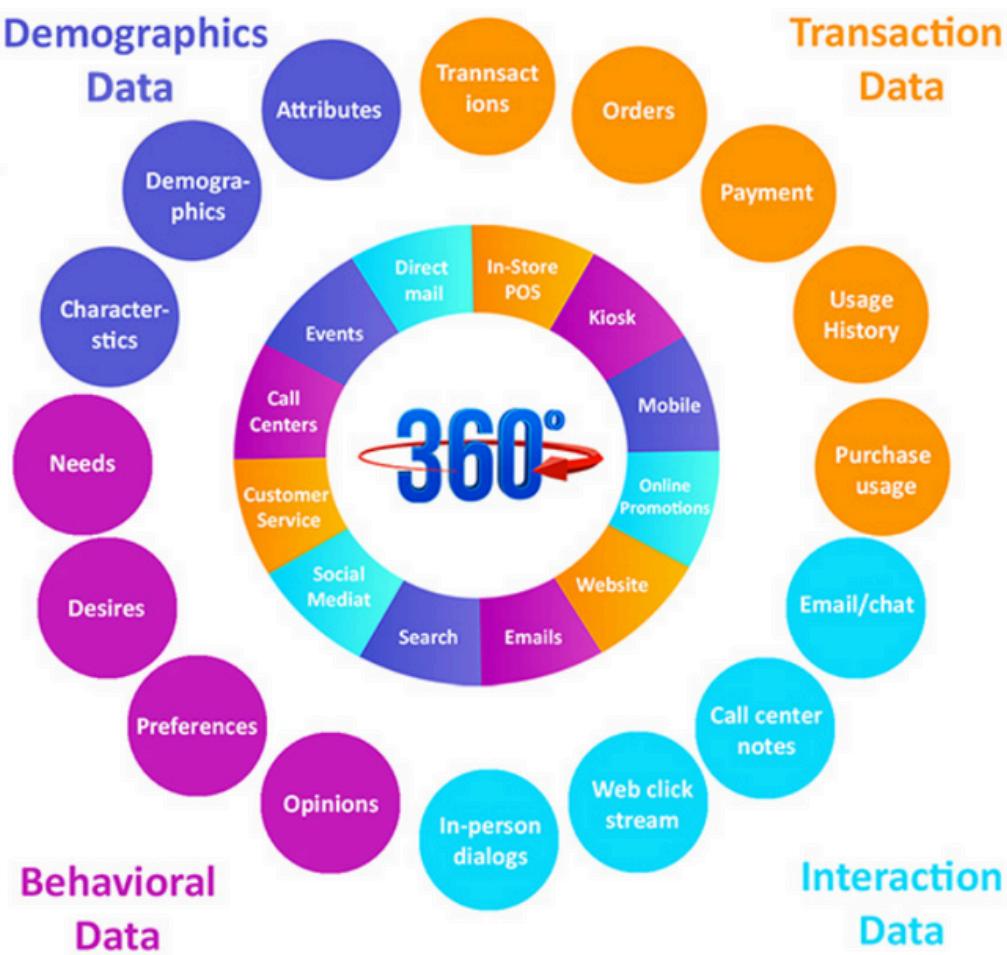
Report by **Trần Hoài Nam**

9.12.2024

TỔNG QUAN – CUSTOMER360 LÀ GÌ?

ĐỊNH NGHĨA

Customer360 là một chiến lược quản lý và phân tích khách hàng của doanh nghiệp, trong đó, mọi thông tin của từng khách hàng được thu thập vào một nền tảng chung. Nền tảng đó **chứa toàn bộ dữ liệu về các khía cạnh của khách hàng** đó, từ thông tin về nhân khẩu học, phương thức liên lạc, cho tới những quan hệ của khách hàng với doanh nghiệp.



Demographics Data – dữ liệu nhân khẩu học: thông tin cơ bản về khách hàng như tên, tuổi, giới tính, địa chỉ, nghề nghiệp, thu nhập, trình độ học vấn, ... giúp doanh nghiệp hiểu rõ hơn về khách hàng, phân chia khách hàng thành các nhóm đối tượng cụ thể để cá nhân hóa chiến lược tiếp thị, bán hàng và dịch vụ khách hàng.

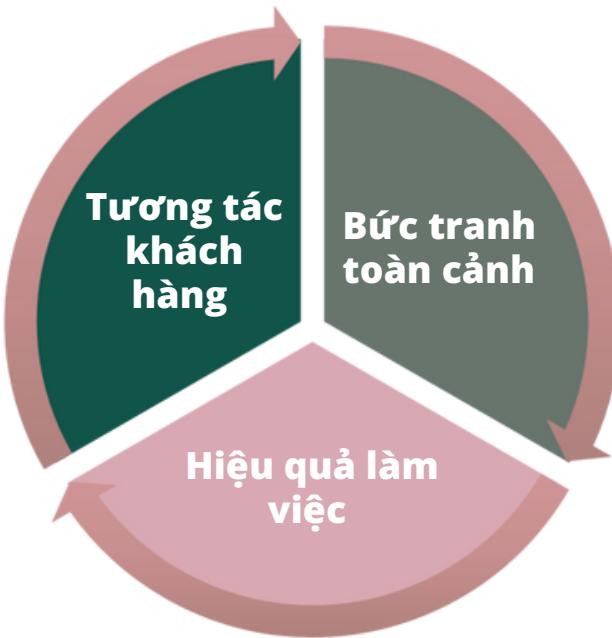
Transaction Data – dữ liệu giao dịch: Bao gồm thông tin về các giao dịch mua hàng của khách hàng như sản phẩm đã mua, giá cả, thời gian mua, phương thức thanh toán,... giúp doanh nghiệp theo dõi hành vi mua hàng của khách hàng, xác định xu hướng mua sắm, đánh giá hiệu quả của các chiến dịch marketing,...

Interaction Data – dữ liệu tương tác: Bao gồm thông tin về các tương tác giữa khách hàng và bộ phận chăm sóc khách hàng, chẳng hạn như: cuộc gọi điện thoại, email, tin nhắn, ... Qua đó, doanh nghiệp có thể đánh giá chất lượng dịch vụ khách hàng, xác định các vấn đề thường gặp, cải thiện quy trình chăm sóc khách hàng, ...

Behavioral Data – dữ liệu hành vi: thông tin về nhu cầu, sở thích, thói quen mua hàng, tần suất truy cập,... cung cấp hiểu biết về hành vi tiêu dùng hằng ngày của khách hàng, từ đó tối ưu hóa trải nghiệm khách hàng, cung cấp những chương trình tiếp thị mang tính cá nhân hóa, phù hợp với hành vi cụ thể của từng khách hàng.

Mô hình **Customer360** giúp doanh nghiệp hiểu rõ hơn về khách hàng, từ đó cung cấp được những trải mang tính cá nhân hóa, cải thiện được sự hài lòng và tăng lòng trung thành từ khách hàng. Đây là mục tiêu quan trọng để giúp doanh nghiệp cải thiện doanh thu và giá trị đem lại cho khách hàng.

LỢI ÍCH CỦA NỀN TẢNG CUSTOMER360



Bức tranh toàn cảnh

Customer360 cho doanh nghiệp một bức tranh toàn cảnh về khách hàng, tránh việc hồ sơ tương tác của khách hàng và doanh nghiệp bị mất kết nối và chia thành nhiều mảng nhỏ và qua các phòng ban tách biệt.

Nâng cao hiệu quả làm việc

Việc tập trung các dữ liệu về chung một nền tảng giúp cho các phòng ban tiếp cận với dữ liệu khách hàng nhanh và tiện lợi hơn, phối hợp làm việc với nhau tốt hơn, đặc biệt nếu dữ liệu được lưu lại trên điện toán đám mây.

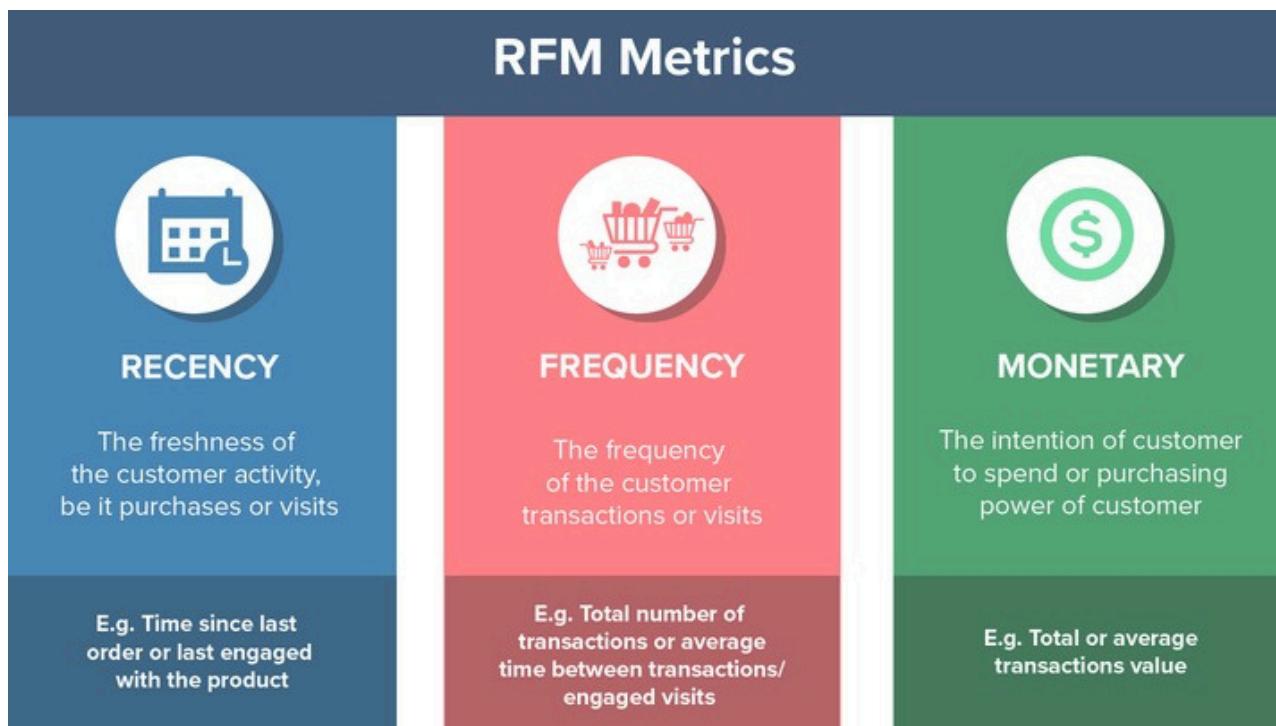
Tăng tính tương tác với khách hàng

Dựa vào dữ liệu đầy đủ, các phòng ban như marketing có thể phân loại khách hàng rõ ràng, dựa trên các tiêu chí ràng mạch, giúp tạo nên những trải nghiệm được cá nhân hóa dành cho từng тип khách hàng khác nhau.

MÔ HÌNH RFM

KHÁI NIỆM

RFM (Recency, Frequency, Monetary) là một mô hình phân tích hành vi khách hàng giúp doanh nghiệp xác định các nhóm khách hàng có giá trị cao, những khách hàng có khả năng rời bỏ và những khách hàng cần được chăm sóc đặc biệt. Dựa trên kết quả phân tích của RFM, doanh nghiệp có thể triển khai các chiến lược tiếp thị được cá nhân hóa và tối ưu hóa nguồn lực chăm sóc khách hàng.



LỢI ÍCH CỦA MÔ HÌNH RFM

Mô hình RFM giúp doanh nghiệp hiểu và phân loại khách hàng dựa trên ba yếu tố Recency, Frequency và Monetary. Qua đó, mô hình RFM đem lại nhiều lợi ích cho doanh nghiệp như:

Segmentation Phân loại

- Các phân loại khách hàng khác nhau ở điểm nào?
- Ai là những khách hàng tiềm năng cao nhất?

Targeting Nhắm mục tiêu

- Khách hàng tương tác với thương hiệu ở điểm nào?
- Đâu là thời gian, địa điểm, loại hình, định dạng tối ưu để nói với khách hàng về thương hiệu?

Positioning Định vị thương hiệu

- Loại thông điệp nào làm tăng niềm tin của khách hàng về thương hiệu
- Loại thông điệp nào sẽ làm khách mua một món hàng

Mô hình RFM cho phép các doanh nghiệp có được **insights** quan trọng về khách hàng mà từ đó họ có thể tác động được, qua đó **định hình chiến lược kinh doanh** lấy những insights đó làm trọng tâm. Mô hình này cho phép doanh nghiệp hiểu được **ý nghĩa của thương hiệu đối với khách hàng** hiện tại, giúp doanh nghiệp quản lý nhận thức của khách hàng và chuyển cảm xúc tích cực thành hành động mua hàng.

PHÂN LOẠI KHÁCH HÀNG CỤ THỂ THEO THANG ĐIỂM RFM

Phân khúc khách hàng là kết quả của việc tính toán các chỉ số RFM.
Mỗi khách hàng sẽ được tính điểm về 3 yếu tố R, F, M.

R,F,M có giá trị từ 1 đến 4, do đó ta sẽ có tổng cộng : $4 \times 4 \times 4 = 64$ tổ hợp khác nhau.
Sau đó, 3 số điểm này sẽ được gộp lại thành một chỉ số RFM duy nhất dùng để phân loại khách hàng vào từng nhóm:

Trong đó: R,F,M có các giá trị từ 1 đến 4

R có giá trị càng nhỏ thì càng được đánh giá cao

R = 1 : Khách hàng đã phát sinh giao dịch trong 31 ngày gần nhất

R = 2 : Khách hàng đã phát sinh giao dịch trong khoảng 32 đến 62 ngày gần nhất

R = 3 : Khách hàng đã phát sinh giao dịch trong khoảng 63 đến 92 ngày gần nhất

R = 4 : Khách hàng đã phát sinh giao dịch trong khoảng lớn hơn 92 ngày

F và M có giá trị càng lớn thì càng được đánh giá cao

F = 1: Khách hàng có tần suất mua hàng thấp nhất

F = 2: Khách hàng có tần suất mua hàng trung bình

F = 3: Khách hàng có tần suất mua hàng cao

F = 4: Khách hàng có tần suất mua hàng rất cao

M = 1: Đơn hàng có giá trị thấp nhất

M = 2: Đơn hàng có giá trị trung bình

M = 3: Đơn hàng có giá trị cao

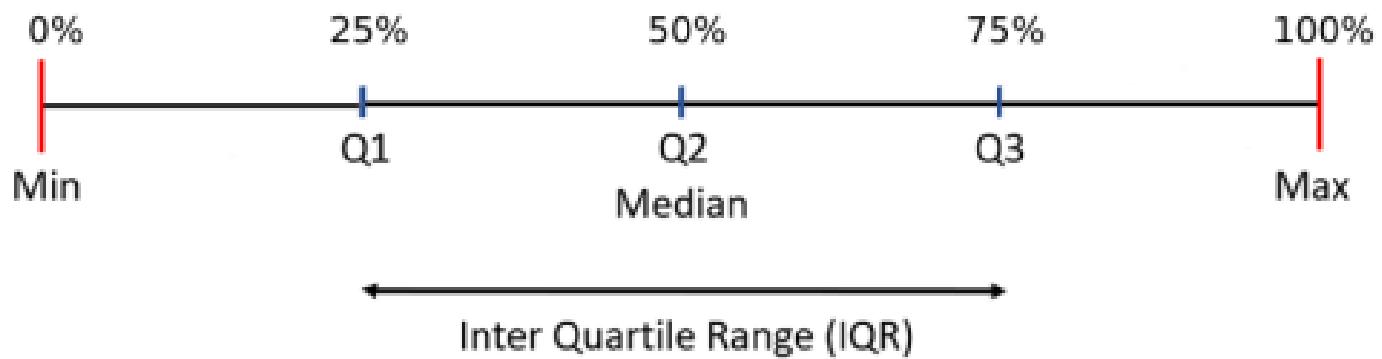
M = 4: Đơn hàng có giá trị rất cao

TỨ PHÂN VỊ - INTERQUARTILE RANGE

Tứ phân vị (IQR) là một phương pháp thống kê dùng để chia một tập dữ liệu đã được sắp xếp theo thứ tự thành bốn phần (tứ phân vị) có kích thước bằng nhau. Mỗi phần đại diện cho 25% dữ liệu trong tập dữ liệu

Tứ phân vị được xác định như sau:

- Sắp xếp các số theo thứ tự tăng dần
- Cắt dãy số thành 4 phần bằng nhau
- Tứ phân vị là các giá trị tại vị trí cắt



Với n là số quan sát:

$$Q1 = 25\% n;$$

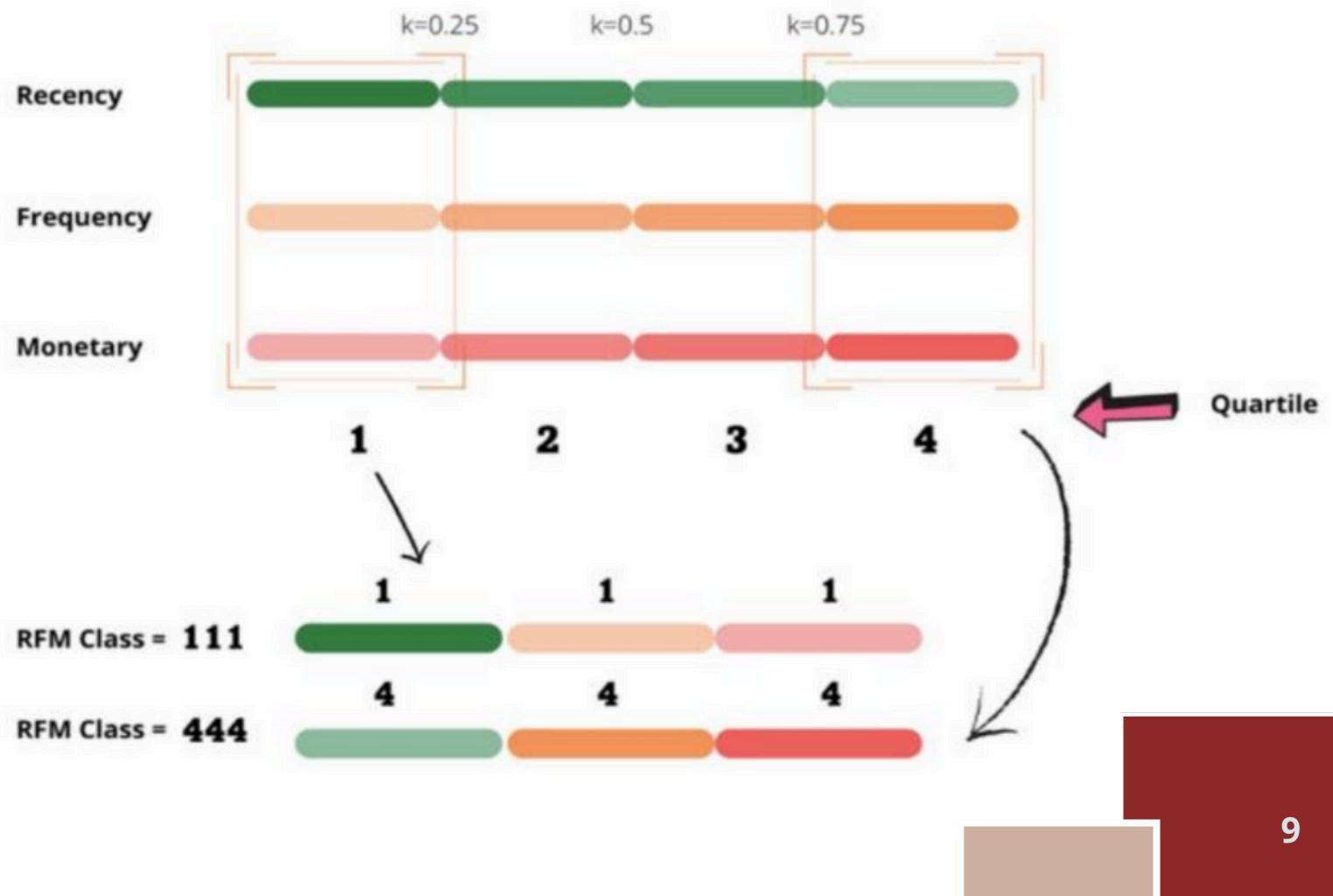
$$Q2 = 50\% n \text{ (} = \text{median) ;}$$

$$Q3 = 75\% n$$

Sử dụng tứ phân vị để đánh giá điểm cho 3 chỉ số R,F,M với thang điểm từ 1 – 4, cụ thể:

- Recency: được xếp thứ tự tăng dần tương ứng với điểm tăng dần từ 1 đến 4
- Frequency, Monetary: được xếp theo thứ tự tăng dần tương ứng với số điểm giảm dần từ 4 đến 1

Quy tắc tính điểm được làm rõ bằng hình dưới đây:

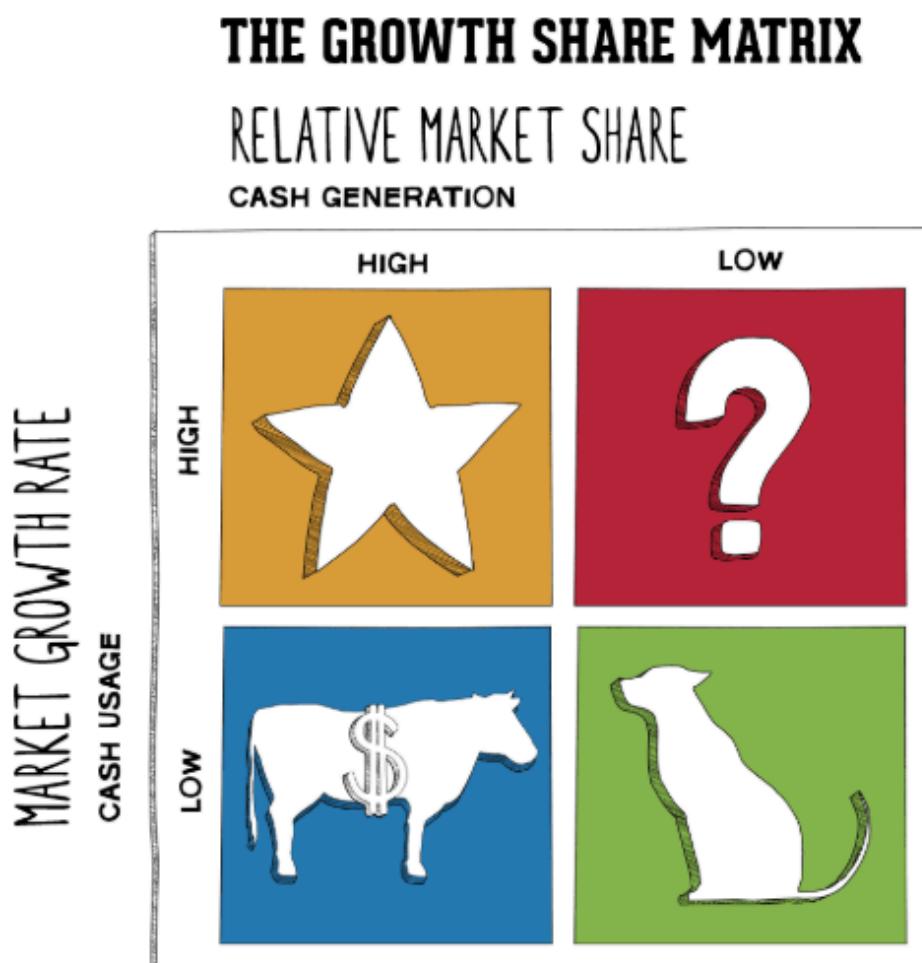


BCG MATRIX (Boston Consulting Group Matrix)

Ma trận BCG được hình thành để giúp các công ty xác định chiến lược tăng trưởng thị phần bằng cách phân tích danh mục sản phẩm của họ, từ đó đưa ra quyết định tiếp tục đầu tư hay từ bỏ sản phẩm.

Ma trận BCG dựa trên hai yếu tố chính:

- **Tốc độ tăng trưởng thị trường**: Đo lường tốc độ tăng trưởng của thị trường mà sản phẩm hoặc dịch vụ đang hoạt động.
- **Thị phần tương đối**: Đo lường thị phần của sản phẩm hoặc dịch vụ so với đối thủ cạnh tranh trong cùng thị trường.



Dựa trên hai yếu tố này, ma trận BCG chia các sản phẩm hoặc dịch vụ thành bốn nhóm chính:

- Ngôi sao (Stars): Thị phần lớn, tăng trưởng thị trường cao.
- Bò sữa (Cash Cows): Thị phần lớn, tăng trưởng thị trường thấp.
- Dấu hỏi (Question Marks): Thị phần nhỏ, tăng trưởng thị trường cao.
- Chó (Dogs): Thị phần nhỏ, tăng trưởng thị trường thấp.

Kết hợp mô hình phân khúc khách hàng RFM và ma trận BCG giúp doanh nghiệp định hình các chiến lược sản phẩm tương ứng với phân khúc khách hàng, từ đó tối ưu hóa chiến lược tiếp thị và chăm sóc khách hàng

Tổng quan về dữ liệu:

Dữ liệu khách hàng được thu thập bởi hệ thống nội bộ, thời gian dữ liệu được lưu từ ngày

1/6/2022 đến 31/8/2022. Bao gồm 2 bảng dữ liệu là Customer_Register và Customer_Transaction

- Dim table: Customer_Register – Bảng chứa thông tin khi đăng ký của user, bao gồm dữ liệu về:

Column name	Data type	Meaning
ID	Bigint	Mã khách hàng
Contract	Varchar	Mã hợp đồng
LocationID	Int	Mã vị trí
BranchCode	Tinyint	Mã chi nhánh
Status	Tinyint	Trạng thái
Created_date	Datetime	Ngày đăng ký
Stopdate	Datetime	Ngày hủy

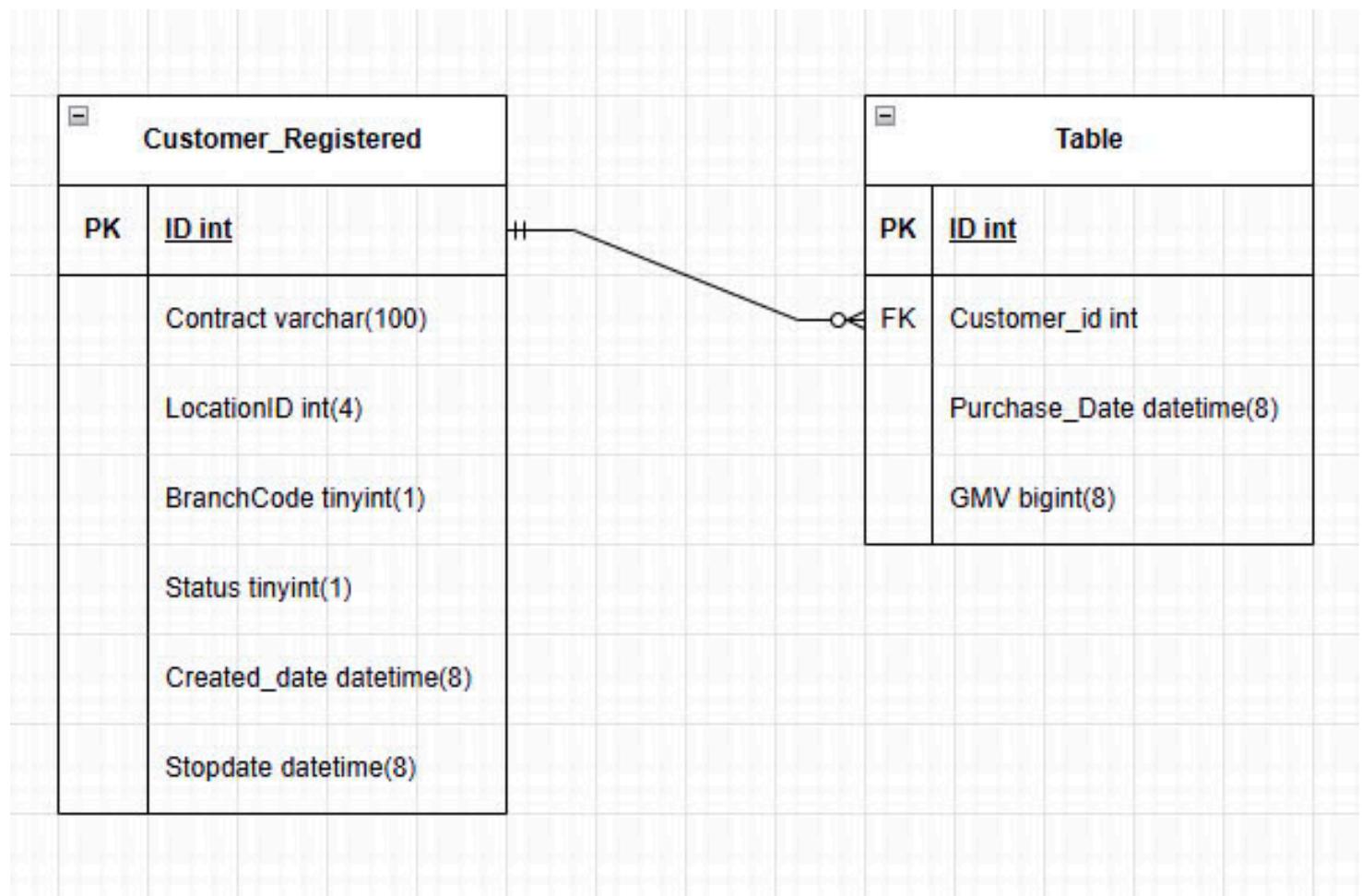
ID	Contract	LocationID	BranchCode	Status	created_date	stopdate
0	SGDN00215	8	1	0	2011-11-25 00:00:00.000	2012-01-05 00:00:00.000
1	SGDN00214	8	1	0	2012-06-14 00:00:00.000	NULL
2	SGD374348	8	1	0	2012-11-01 00:00:00.000	NULL
3	SGD022064	8	1	2	2011-06-22 00:00:00.000	2013-05-29 00:00:00.000
4	SGD041015	8	5	2	2011-12-17 00:00:00.000	2014-11-11 00:00:00.000
5	SGDN00211	NULL	NULL	2	2015-06-09 00:00:00.000	2015-09-09 00:00:00.000
6	SGD374348	8	1	3	2012-11-26 00:00:00.000	2012-12-13 00:00:00.000
7	SGDN00013	0	0	1	2017-12-11 00:00:00.000	NULL
8	BEAAA1809	75	0	1	2022-04-04 00:00:00.000	NULL
9	BNAAA4298	241	2	1	2022-04-19 00:00:00.000	NULL
10	SGAAH8487	8	3	1	2022-04-25 00:00:00.000	NULL
11	SGDN00017	0	0	2	2017-12-04 00:00:00.000	2018-04-27 00:00:00.000

– Fact table: Customer_Transaction – Bảng chứa thông tin các giao dịch của user, bao gồm dữ liệu về:

Column name	Data type	Meaning
ID	Bigint	Mã giao dịch
CustomerID	Varchar	Mã khách hàng
Purchase_date	Datetime	Ngày giao dịch
GMV	Bigint	Giá trị giao dịch

ID	CustomerID	Purchase_Date	GMV
0	1327813	2022-06-01 00:00:00.000	95000
1	1157830	2022-06-01 00:00:00.000	75000
2	873915	2022-07-01 00:00:00.000	95000
3	3505071	2022-07-01 00:00:00.000	90000
4	2930918	2022-07-01 00:00:00.000	109091
5	899882	2022-06-01 00:00:00.000	105000
6	2248818	2022-06-01 00:00:00.000	75000
7	3331485	2022-06-01 00:00:00.000	90000
8	3497579	2022-06-01 00:00:00.000	60000
9	2636994	2022-07-01 00:00:00.000	75000
10	3285167	2022-07-01 00:00:00.000	90000
11	3725376	2022-07-01 00:00:00.000	70000
12	1112691	2022-08-01 00:00:00.000	75000
13	1425581	2022-08-01 00:00:00.000	95000

Sơ đồ ERD



Xác định nhiệm vụ:

- Xác định được các chỉ số Recency (R), Frequency (F), Monetary (M) dựa trên dữ liệu về transaction của user.
- Thực hiện đánh giá điểm bằng tứ phân vị trong thống kê nhằm phân loại nhóm khách hàng để có những chính sách tiếp thị phù hợp cho từng đối tượng.
- Dựa vào dữ liệu ban đầu và dữ liệu sau khi phân tích, visualize và phân tích chi tiết dữ liệu.

Làm rõ khái niệm:

RFM Metrics:

- Recency: xác định ngày giao dịch gần nhất của user.

→ Được tính thông qua số ngày chênh lệch giữa ngày giao dịch gần nhất và mốc thời gian '2022-09-01' (giả định thời điểm thực hiện report)

- Frequency: tần suất sử dụng của user.

→ Xác định bằng số lượng giao dịch của user. Tuy nhiên, mỗi user sẽ có mốc thời gian đăng ký là khác nhau. Đồng nghĩa rằng khoảng thời gian đăng ký dịch vụ là khác nhau (độ tuổi hợp đồng khác nhau)

→ Để có sự so sánh phù hợp giữa các khách hàng, Frequency nên được tính dựa trên độ tuổi hợp đồng, nghĩa là: **Frequency = Số lượng giao dịch/Số ngày chênh lệch giữa created_date và '2022-09-01'**

- Monetary: doanh thu thu được từ mỗi user.

→ Xác định bằng tổng GMV của từng user. Tương tự như Frequency, Monetary nên được tính dựa trên độ tuổi hợp đồng, nghĩa là: **Frequency = Tổng GMV/Số ngày chênh lệch giữa created_date và '2022-09-01'**

Phân tích kết quả:

114.08K

Số lượng khách hàng

61.25

Average of recency

0.22

Average of Frequency

21.58K

Average of monetary

Từ dữ liệu phân tích cho thấy, hiện tại doanh nghiệp có khoảng 114.08 nghìn user thực hiện giao dịch, trong đó:

- Khoảng thời gian trung bình mà khách hàng truy cập lần cuối cùng cách hiện tại khoảng 61.25 ngày.
→ Các chương trình marketing hiện tại chưa thật sự hiệu quả với một bộ phận khách hàng.
- Tần suất khách hàng truy cập khoảng 0.22 lần/năm. Điều này chứng tỏ khách hàng rất hạn chế truy cập, tác động tiêu cực đến số lượng giao dịch được tiến hành.
→ Cần có chương trình thu hút khách hàng để họ truy cập thường xuyên hơn, tạo động lực thực hiện giao dịch nhiều hơn
- Chi tiêu trung bình của mỗi khách hàng khoảng 21.580/năm

Sử dụng **phương pháp IQR** để phân khoảng và tính điểm R – F – M, ta được bảng sau:

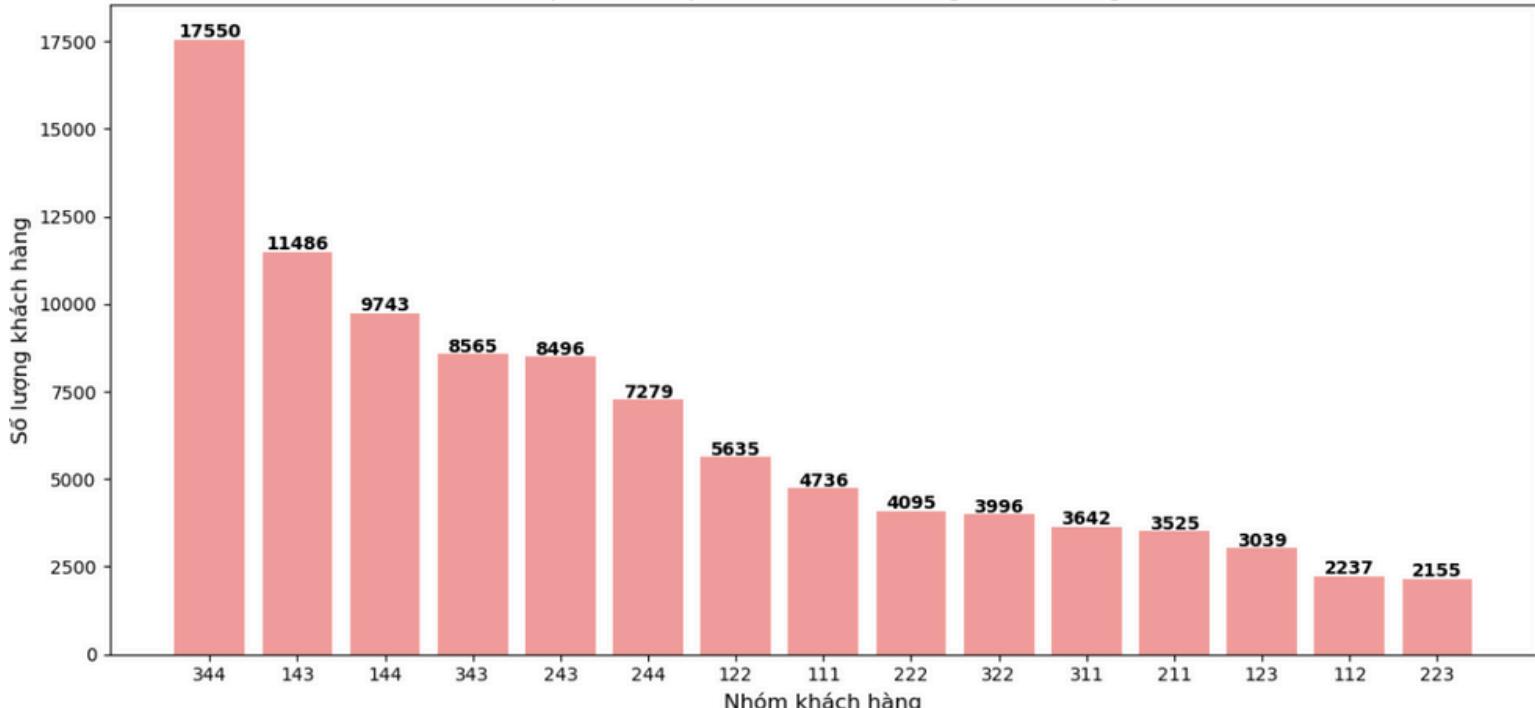
Điểm	Recency	Frequency	Monetary
1	> 92	0,14 - 0,2	0 - 15000
2	63 - 92	0,2 - 0,25	15000 - 18750
3	32 - 62	0,25 - 0,25	18750 - 23750
4	1 - 31	0,25 - 0,75	23750 - 110000

Bằng việc sử dụng **ma trận BCG**, tổ hợp RFM được phân bổ theo nhóm khách hàng như sau:

Segmentation	Mô tả	Tổ hợp
Stars (KH VIP)	Khách hàng thường xuyên truy cập và đóng góp lượng doanh thu lớn	444, 443, 434, 433, 344, 343
Cash Cows (KH trung thành)	Nhóm khách hàng này mặc dù không đóng góp doanh thu cao như Stars nhưng ổn định	442, 441, 432, 423, 342, 334, 333, 332, 331, 324, 323, 322, 244, 243, 234
Question Marks (KH tiềm năng)	Khách hàng không truy cập thường xuyên nhưng đóng góp doanh thu đáng kể	431, 424, 422, 421, 414, 413, 412, 411, 341, 321, 314, 313, 312, 311
Dogs	Khách hàng tần suất truy cập và doanh thu đóng góp thấp	242, 241, 233, 232, 231, 224, 223, 222, 221, 144, 143, 142, 141, 134, 133, 132, 131, 114, 113, 112, 111

CustomerID	GMV	recency	Frequency	Monetary	m_recency	m_frequency	m_monetary	R	F	M	RFM	seg
167518	0	13	140.000.000.000	0	180	11	1	4	1	1	411	TIEM NANG
180763	0	14	140.000.000.000	0	209	12	2	4	1	1	411	TIEM NANG
85902	0	16	140.000.000.000	0	247	13	3	4	1	1	411	TIEM NANG
184355	0	20	140.000.000.000	0	319	17	4	4	1	1	411	TIEM NANG
89489	0	10	140.000.000.000	0	135	8	5	4	1	1	411	TIEM NANG
75284	0	31	140.000.000.000	0	11057	22	6	3	1	1	311	TIEM NANG
180255	0	31	140.000.000.000	0	20959	613	7	3	1	1	311	TIEM NANG
180274	0	31	140.000.000.000	0	20960	614	8	3	1	1	311	TIEM NANG
142118	0	39	140.000.000.000	0	42454	1279	9	3	1	1	311	TIEM NANG
170846	0	46	140.000.000.000	0	42530	1281	10	3	1	1	311	TIEM NANG
153894	0	62	140.000.000.000	0	45940	2131	11	2	1	1	211	VANGLAI
180279	0	62	140.000.000.000	0	45368	2218	12	2	1	1	211	VANGLAI
180314	0	62	140.000.000.000	0	45383	2233	13	2	1	1	211	VANGLAI
178884	0	67	140.000.000.000	0	74128	2534	14	2	1	1	211	VANGLAI
156338	0	70	140.000.000.000	0	74153	2536	15	2	1	1	211	VANGLAI
156041	0	92	140.000.000.000	0	97454	3450	16	1	1	1	111	VANGLAI
180263	0	92	140.000.000.000	0	97635	3631	17	1	1	1	111	VANGLAI
125201	0	92	140.000.000.000	0	99217	4102	18	1	1	1	111	VANGLAI
206944	0	92	170.000.000.000	0	98170	4498	19	1	1	1	111	VANGLAI
195143	0	92	170.000.000.000	0	98349	4677	20	1	1	1	111	VANGLAI
209608	0	92	170.000.000.000	0	98718	5046	21	1	1	1	111	VANGLAI

Top 15 Tổ Hợp RFM Theo Số Lượng Khách Hàng



- Nhóm 344 có số lượng khách hàng cao nhất (17,550 khách hàng), vượt xa các nhóm còn lại.
- Nhóm 143 và 144 lần lượt đứng thứ hai và thứ ba với 11,486 và 9,743 khách hàng, cho thấy đây là những nhóm có sự đóng góp lớn.

Những khách hàng đang đứng top 5 về cả 3 tiêu chí:

CustomerID	recency	Frequency	Monetary	R	F	M
647145	30	0,25	24	4	4	4
748672	30	0,25	24	4	4	4
739785	29	0,25	25	4	4	4
758262	29	0,25	39,772	4	4	4
952984	29	0,25	39,772	4	4	4

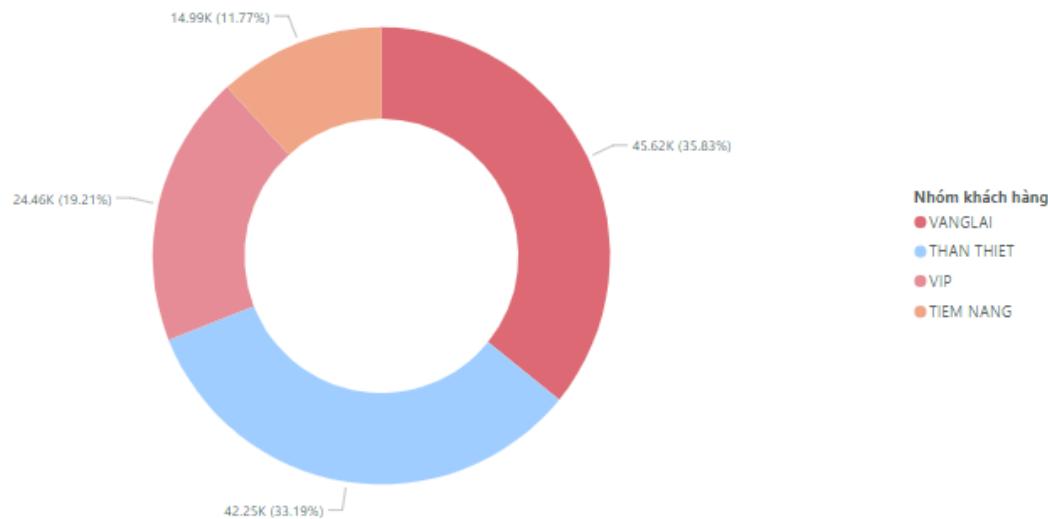
Top 5 khách hàng có tỉ lệ rời bỏ cao:

CustomerID	recency	Frequency	Monetary	R	F	M
125201	92	0,14	0	1	1	1
180263	92	0,14	0	1	1	1
156041	92	0,14	0	1	1	1
181534	92	0,14	2,023	1	1	1
94252	92	0,14	3,619	1	1	1

Top 10 khách hàng có tần suất mua hàng thấp nhưng có chi tiêu cao:

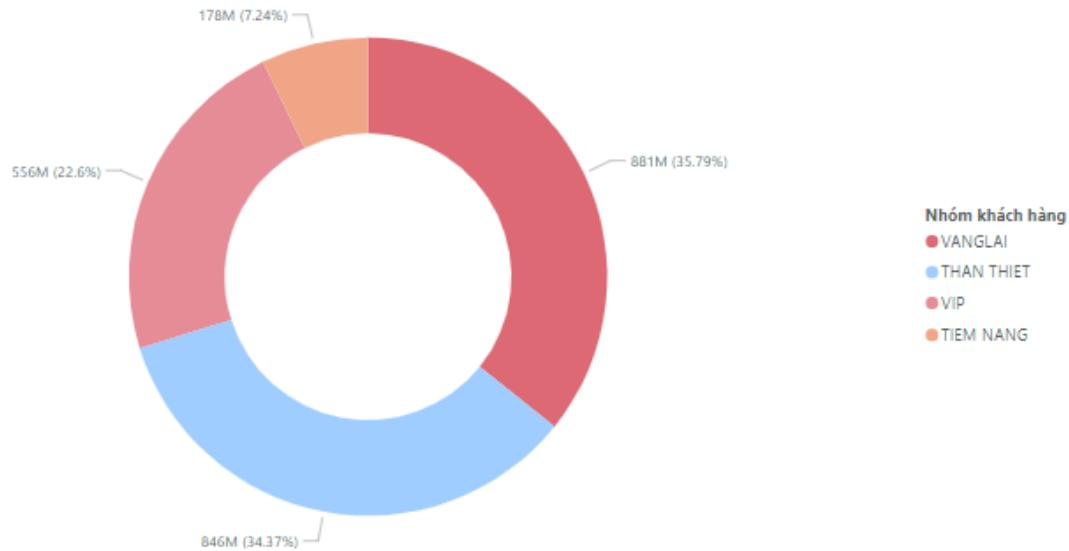
CustomerID	recency	Frequency	Monetary	R	F	M
436169	62	0,2	55.000	2	2	4
461285	87	0,2	42.000	2	2	4
608647	26	0,2	42.000	4	2	4
374744	92	0,2	41.000	1	2	4
603370	92	0,2	37.000	1	2	4
502693	31	0,2	37.000	3	2	4
518880	31	0,2	37.000	3	2	4
564806	31	0,2	37.000	3	2	4
463655	62	0,2	37.000	2	2	4
531703	62	0,2	37.000	2	2	4

Biểu đồ tỉ lệ nhóm khách hàng



- Nhóm VÃNG LAI có số lượng áp đảo, cho thấy doanh nghiệp có nhiều khách hàng chưa gắn bó lâu dài hoặc thường xuyên sử dụng dịch vụ. Doanh nghiệp cần tập trung chiến lược để chuyển đổi một phần khách hàng từ nhóm này sang các nhóm trung thành hoặc VIP.
- Nhóm VIP và TRUNG THÀNH chiếm hơn 42% tổng số khách hàng, điều này thể hiện doanh nghiệp đã xây dựng được một phần mối quan hệ bền vững với khách hàng.
- Nhóm TIỀM NĂNG chiếm tỷ lệ nhỏ nhất, cho thấy số lượng khách hàng đang trong giai đoạn "khám phá" sản phẩm/dịch vụ chưa cao.

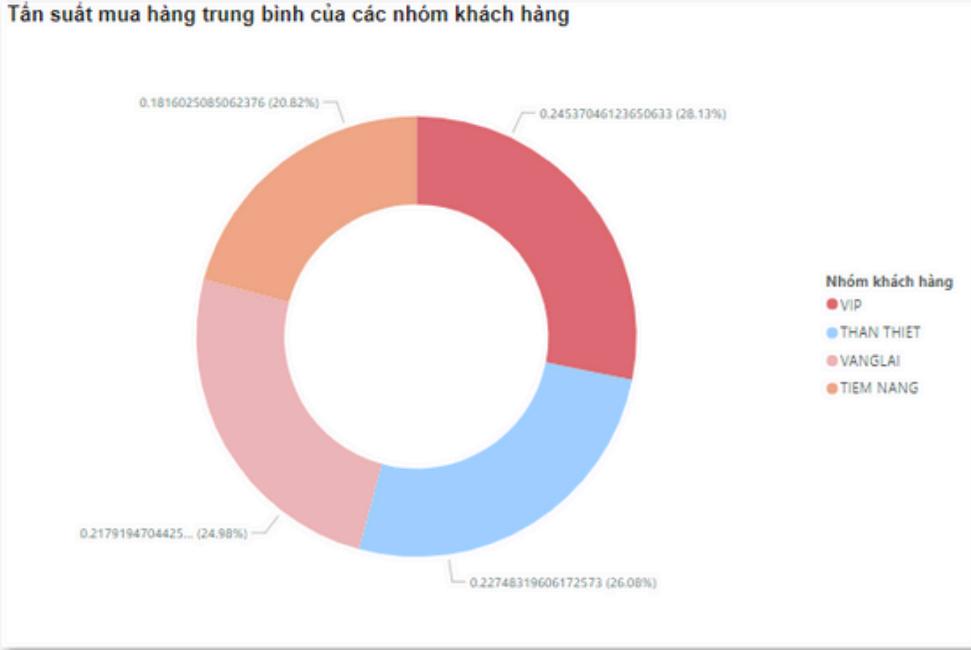
Tổng doanh thu theo từng nhóm khách hàng



Sự tương đương về tỷ lệ giữa số lượng khách hàng trong từng nhóm và phần trăm doanh thu mà họ mang lại là một yếu tố quan trọng cần xem xét. Trong đó, nhóm khách hàng Vǎng Lai chiếm tỷ lệ cao nhất với 35,8% tổng số khách hàng, đồng thời đóng góp khoảng 36% vào tổng doanh thu của công ty.

Tuy nhiên, nhóm khách hàng Tiềm Năng, dù chiếm 11,77% tổng số khách hàng, lại chỉ mang lại 7,24% doanh thu. Điều này cho thấy tiềm năng phát triển doanh thu từ nhóm khách hàng này chưa được khai thác đầy đủ. Vì vậy, công ty nên cân nhắc triển khai các chương trình ưu đãi hoặc chiến lược tiếp thị phù hợp nhằm khuyến khích nhóm khách hàng này gia tăng tần suất giao dịch và giá trị mua hàng. Việc này không chỉ giúp tăng doanh thu mà còn thúc đẩy sự gắn kết lâu dài của nhóm khách hàng tiềm năng với công ty.

Tần suất mua hàng trung bình của các nhóm khách hàng



Biểu đồ trên cho thấy tần suất mua hàng trung bình của các nhóm khách hàng. Nhóm "VIP" dẫn đầu với 28.13%, cho thấy họ có thói quen mua sắm thường xuyên và có khả năng chi tiêu cao. Nhóm "THAN THIET" theo sau với 26.08%, cho thấy sự quan tâm đáng kể từ nhóm này. Trong khi đó, nhóm "TIEM NANG" và "VANG LAI" có tần suất mua hàng thấp hơn, lần lượt là 26.08% và 20.82%. Điều này cho thấy sự khác biệt rõ rệt về mức độ tham gia của từng nhóm khách hàng vào hoạt động mua sắm.

Kết luận:

1. Nhóm khách hàng VIP

- Đặc điểm:
 - Có mức độ tương tác cao với tần suất mua trung bình là 0,25.
 - Đóng góp doanh thu cao nhất, bao gồm cả những khách hàng chi tiêu lớn, trong đó có một trường hợp ngoại lệ với giá trị giao dịch gần 556 triệu đồng.
- Chiến lược:
 - Duy trì lòng trung thành thông qua các ưu đãi cá nhân hóa.
 - Áp dụng chế độ đai ngộ VIP đặc biệt để củng cố mối quan hệ, đảm bảo khách hàng tiếp tục đóng góp ở mức cao và giảm thiểu rủi ro mất khách hàng.

2. Nhóm khách hàng thân thiết

- Đặc điểm:
 - Tần suất mua trung bình cao (0,23) với tổng giá trị giao dịch đạt 846 triệu đồng.
 - Tuy nhiên, mức độ tương tác gần đây giảm, cho thấy họ trở nên ít hoạt động hơn trong thời gian gần đây.
- Chiến lược:
 - Nhanh chóng triển khai các chiến dịch tiếp thị được cá nhân hóa nhằm khơi dậy lại sự hứng thú và tăng cường tương tác.
 - Cung cấp các chương trình khuyến mãi đặc biệt để duy trì lòng trung thành và cải thiện mức độ hoạt động.

3. Nhóm khách hàng tiềm năng

- Đặc điểm:
 - Giá trị giao dịch trung bình đạt 178 triệu đồng với tần suất mua 0,22.
 - Đại diện cho cơ hội tăng trưởng tiềm năng lớn nếu được khai thác đúng cách.
- Chiến lược:
 - Tập trung vào các chiến lược tiếp thị phù hợp để tăng mức độ tương tác và giá trị chi tiêu.
 - Hướng tới việc chuyển đổi nhóm này thành khách hàng trung thành hoặc VIP, góp phần gia tăng giá trị dài hạn.

4. Nhóm khách hàng vãng lai

- Đặc điểm:
 - Đóng góp doanh thu đáng kể (881 triệu đồng), nhưng tần suất mua chỉ đạt 0,18.
 - Mức độ hoạt động gần đây cao cho thấy rủi ro mất khách hàng nếu không có hành động kịp thời.
- Chiến lược:
 - Triển khai ngay các chiến dịch kích hoạt lại có mục tiêu, cá nhân hóa để thu hút sự quan tâm của nhóm này.
 - Tăng cường liên lạc qua các kênh phù hợp và cung cấp các ưu đãi hấp dẫn để giữ chân khách hàng

Tóm lại

- Nhóm khách hàng VIP và thân thiết: Đây là nền tảng doanh thu quan trọng, cần được đầu tư và nuôi dưỡng để duy trì lòng trung thành và giá trị hiện có.
- Nhóm khách hàng tiềm năng: Mang lại cơ hội tăng trưởng lớn nhất, cần các chiến lược thúc đẩy chi tiêu và tăng mức độ tương tác.
- Nhóm khách hàng vãng lai: Đòi hỏi các hành động tức thời để cứu vãn và giảm thiểu nguy cơ mất doanh thu.

Việc tập trung vào các nhóm khách hàng này với chiến lược phù hợp sẽ giúp doanh nghiệp đảm bảo sự phát triển bền vững và gia tăng lợi nhuận lâu dài.

SQL query:

```
• select CustomerID ,GMV ,DATEDIFF(day, max(Purchase_Date), '2022-09-01') as recency ,
       round(1.0* count(ct.CustomerID)/DATEDIFF(year, created_date, '2022-09-01'),2) as Frequency,
       round(1.0* sum(GMV)/DATEDIFF(year, created_date, '2022-09-01'),2) as Monetary,
       ROW_NUMBER() over (order by DATEDIFF(day, max(Purchase_Date), '2022-09-01')) as rn_recency,
       ROW_NUMBER() over (order by round(1.0* count(ct.CustomerID)/DATEDIFF(year, created_date, '2022-09-01'),2)) as rn_frequency,
       ROW_NUMBER() over (order by round(1.0* sum(GMV)/DATEDIFF(year, 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 CustomerID != 0
 group by CustomerID, created_date, GMV
```

```
• SELECT *, CONCAT(R, F, M) as RFM,
CASE
WHEN CONCAT(R, F, M) IN ('444', '443', '434', '433', '344', '343') THEN 'VIP'
WHEN CONCAT(R, F, M) IN ('442', '441', '432', '423', '342', '334', '333', '332', '331', '324', '323', '322', '244', '243', '234' ) THEN 'TRUNG THANH'
WHEN CONCAT(R, F, M) IN ('431', '424', '422', '421', '414', '413', '412', '411', '341', '321', '314', '313', '312', '311') THEN 'TIEM NANG'
ELSE 'VANGLAI'
END AS seg
   FROM ##RFM
```

```

• 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 ##RFM
from ##calculation

```

•SELECT

```
percentile_cont(0.25) WITHIN GROUP (ORDER BY recency) OVER () AS quartile_1,  
percentile_cont(0.50) WITHIN GROUP (ORDER BY recency) OVER () AS quartile_2,  
percentile_cont(0.75) WITHIN GROUP (ORDER BY recency) OVER () AS quartile_3,  
MIN(recency) OVER () AS min_recency,  
MAX(recency) OVER () AS max_recency  
FROM ##calculation;
```

•SELECT

```
percentile_cont(0.25) WITHIN GROUP (ORDER BY Frequency) OVER () AS quartile_1,  
percentile_cont(0.50) WITHIN GROUP (ORDER BY Frequency) OVER () AS quartile_2,  
percentile_cont(0.75) WITHIN GROUP (ORDER BY Frequency) OVER () AS quartile_3,  
min (Frequency) OVER () as min_frequency,  
max (Frequency) OVER () as max_frequency  
FROM ##calculation;
```

•SELECT

```
percentile_cont(0.25) WITHIN GROUP (ORDER BY Monetary) OVER () AS quartile_1,  
percentile_cont(0.50) WITHIN GROUP (ORDER BY Monetary) OVER () AS quartile_2,  
percentile_cont(0.75) WITHIN GROUP (ORDER BY Monetary) OVER () AS quartile_3,  
min (Monetary) OVER () as min_monterey,  
max (Monetary) OVER () as max_monterey  
FROM ##calculation;
```