



# **CUSTOMER ANALYTICAL PROJECT**

TÁC GIẢ  
THÙY TRANG

# Table of CONTENTS

I

**Giới thiệu về Customer 360**

II

**Xác định vấn đề**

III

**Dữ liệu**

IV

**Phương pháp phân tích**

4.1.1 Khái niệm về mô hình RFM

4.1.2 Ứng dụng mô hình RFM trong bài:

4.1.3 Mô hình IQR:

V

**Phân loại khách hàng và Giải pháp**



# I. GIỚI THIỆU VỀ CUSTOMER 360

Customer 360 là một nền tảng phân tích xoay quanh các khía cạnh về khách hàng và nền tảng này chứa toàn bộ dữ liệu về khách hàng. Dữ liệu này bao gồm 4 khía cạnh chính: nhân khẩu học, **dữ liệu về giao dịch**, dữ liệu về tương tác và dữ liệu về hành vi. Trong bài phân tích này tôi sẽ sử dụng dữ liệu về giao dịch.



## II. XÁC ĐỊNH VẤN ĐỀ

Hiện tại để tìm giải pháp thúc đẩy việc tăng doanh thu cho doanh nghiệp tôi tiến hành phân tích về những giao dịch mua hàng của những khách hàng đã từng mua thông qua bảng 3.2 và 3.4 (ở mục III) bằng phương pháp RFM. Từ đó sẽ lên những chiến dịch chăm sóc khách hàng phù hợp với mỗi phân khúc khách hàng.



### III. DỮ LIỆU

Ở bảng Customer\_transaction (Bảng 3.2) chứa thông tin giao dịch của khách hàng, bảng Customer\_registered (Bảng 3.4) chứa thông tin đăng kí hợp đồng của khách hàng. Từ dữ liệu ở 2 bảng này tôi sẽ phân tích để đưa ra những phân khúc khách hàng khác nhau dựa trên phương pháp phân tích là RFM và IQR.

**Bảng 3.1. Customer\_transaction**

<u>Tên cột</u>	<u>Mô tả</u>
<u>TransactionID</u>	<u>Mã giao dịch</u>
<u>CustomerID</u>	<u>Mã Khách hàng</u>
<u>Purchase Date</u>	<u>Ngày mua hàng</u>
<u>GMV</u>	<u>Gross Monetary Value (doanh thu)</u>

**Bảng 3.2. Dữ liệu mẫu của Bảng 3.1**

123 Transaction_ID ▼	123 CustomerID ▼	ABC Purchase_Date ▼	123 GMV ▼
0	1,327,813	2022-06-01	95,000
1	1,157,830	2022-06-01	75,000
2	873,915	2022-07-01	95,000
3	3,505,071	2022-07-01	90,000
4	2,930,918	2022-07-01	109,091
5	899,882	2022-06-01	105,000
6	2,248,818	2022-06-01	75,000
7	3,331,485	2022-06-01	90,000
8	3,497,579	2022-06-01	60,000
9	2,636,994	2022-07-01	75,000

**Bảng 3.3 Customer\_registered**

<u>Tên cột</u>	<u>Mô tả</u>
ID	Mã Khách hàng
Contract	Mã hợp đồng
Location ID	Mã vị trí
Branch Code	Mã chi nhánh
Status	Trạng thái
Created_date	Ngày đăng ký
Stopdate	Ngày hủy hợp đồng

**Bảng 3.4 Dữ liệu mẫu của bảng 3.3**

<small>123</small> ID	<small>ASC</small> Contract	<small>123</small> LocationID	<small>123</small> BranchCode	<small>123</small> Status	<small>ASC</small> created_date	<small>ASC</small> stopdate
1	SGDN00215	8	1	0	2011-11-25	2012-05-01
2	SGDN00214	8	1	0	2012-06-14	
3	SGD374348	8	1	0	2012-01-11	
4	SGD022064	8	1	2	2011-06-22	2013-05-29
5	SGD041015	8	5	2	2011-12-17	2014-11-11
6	SGDN00211	[NULL]	[NULL]	2	2015-09-06	2015-09-09
7	SGD374348	8	1	3	2012-11-26	2012-12-13
8	SGDN00013	0	0	1	2017-11-12	
9	BEAAA1809	75	0	1	2022-04-04	
10	BNAAA4298	241	2	1	2022-04-19	



## IV. PHƯƠNG PHÁP PHÂN TÍCH

### 4.1.1 Khái niệm về mô hình RFM

Để phân tích khách hàng dựa trên dữ liệu từ bảng 3.2 và bảng 3.4 tôi sử dụng mô hình RFM. Vậy RFM là gì? Bảng 4.1 sau sẽ giải thích về mô hình RFM

***Bảng 4.1.1 Giải thích về mô hình RFM***

<u>Thuật ngữ</u>	<u>Khái niệm</u>
Recency ( R )	<u>Lần gần nhất khách hàng mua</u>
Frequency (F)	<u>Tần suất mua hàng của khách hàng</u>
Monetary (M)	<u>Tổng số tiền khách đã bỏ ra để mua hàng</u>

### 4.1.2 Ứng dụng mô hình RFM trong bài:

Tôi tiến hành chạy Code SQL để tính ra được R, F, M. Báo cáo này tôi lấy mốc là ngày 1/9/2022 để tính cho toàn bộ dữ liệu trên.

Cụ thể:

Recency (R): Thời gian cuối cùng khách hàng mua hàng đến thời điểm 1/9/2022

Frequency (F): Tổng số lần khách hàng đã mua hàng tính đến thời điểm 1/9/2022

Monetary (M): Tổng số tiền khách đã bỏ ra để mua hàng tính đến thời điểm 1/9/2022

Để tránh trường hợp bị thiên vị thì tôi tính thêm tuổi khách hàng bằng cách lấy thời điểm 1/9/2022 trừ cho ngày mà khách hàng đăng ký giao dịch. Sau đó lấy 2 chỉ số Frequency (F), Monetary (M) lần lượt chia cho tuổi của khách hàng

**Bảng 4.1.2. Dữ liệu mẫu RFM sau khi chạy code**

ABC contract	123 monetary	123 frequency	123 recency	123
AGD000018	70	0	31	
AGD000237	40	0	31	
AGD001038	43	0	92	
AGD001899	55	0	92	
AGD001957	91	0	62	
AGD002195	30	0	92	
AGD002745	58	0	31	
AGD002866	38	0	31	
AGD003187	51	0	31	
AGD003414	42	0	92	
AGD003445	42	0	31	
AGD003467	50	0	62	

## 4.2 Mô hình IQR và ứng dụng:

Mô hình IQR sẽ được dùng chấm điểm ba chỉ số R, F, M.

IQR nghĩa là khoảng tứ phân vị. Từ mẫu dữ liệu R, F, M tôi đã tính ra, sắp xếp lại và chia thành 4 phần bằng nhau. Điểm càng lớn thì doanh thu, mức độ thường xuyên càng cao và khách mới vừa mua hàng. Trong bài này điểm số cao nhất là 4 và lần lượt 3,2,1 là số điểm sẽ giảm dần tương ứng.

**Bảng 4.2 Bảng điểm R, F, M**

ABC contract	123 monetary	123 frequency	123 recency	123 M	123 F	123 R
AGD000018	70	0	31	1	4	4
AGD000237	40	0	31	4	4	4
AGD001038	43	0	92	4	4	2
AGD001899	55	0	92	2	2	1
AGD001957	91	0	62	1	1	2
AGD002195	30	0	92	4	4	1
AGD002745	58	0	31	2	2	3
AGD002866	38	0	31	4	4	4
AGD003187	51	0	31	3	3	3
AGD003414	42	0	92	4	3	1
AGD003445	42	0	31	4	4	4
AGD003467	50	0	62	3	2	3
AGD003477	46	0	92	3	4	1



## V. Phân loại khách hàng và giải pháp

### 5.1. Phân loại khách hàng

Tôi chọn 10 nhóm khách hàng có số lượng khách hàng nhiều nhất tôi phân chia thành 4 nhóm khách hàng bao gồm: kim cương, vàng, bạc, đồng.

Segmentation	RFM_Group	Đặc điểm
kimcuong	444, 344, 244	Những khách hàng ghé mua hàng thường xuyên, bỏ tiền ra nhiều nhất để mua hàng
vang	333, 233, 322	Những khách hàng ghé mua hàng với tần suất, chi tiền mua hàng vừa phải
bac	411	Khách hàng mới ghé nhưng độ chi tiêu và mức độ ghé thường xuyên thấp
dong	211, 122, 111	Khách hàng mua hàng đã lâu, độ chi tiêu và mức độ ghé thường xuyên đều thấp

**Bảng 5.1. Bảng phân loại khách hàng**

ABC RFM_Group	123 total_client	ABC Segmentation
411	14,629	bac
322	8,990	vang
244	8,253	kimcuong
233	8,163	vang
122	6,837	dong
333	4,370	vang
111	4,209	dong
444	4,058	kimcuong
211	3,881	dong
344	3,386	kimcuong

## 5.2. Giải pháp

- Nhóm khách hàng kim cương: đưa ra những chương trình khách hàng thân thiết
- Khách hàng hạng Vàng: Đưa ra quà tặng miễn phí cho các giao dịch trên giá trị trung bình đơn hàng của phân khúc này
- Khách hàng hạng Bạc: Lấy ý kiến trải nghiệm về sản phẩm, tặng voucher cho những đơn hàng tiếp theo
- Khách hàng hạng Đồng: Cần thực hiện những khảo sát tìm ra nguyên nhân dẫn đến quyết định không quay lại và đưa ra hướng khắc phục

# Tài liệu code SQL

```
CREATE TABLE CTE AS
SELECT
    contract,
    CAST(SUM(GMV) / DATEDIFF('2022-09-01', MAX(created_date)) AS SIGNED INTEGER) AS
monetary,
    CAST(COUNT(ID) / DATEDIFF('2022-09-01', MAX(created_date)) AS SIGNED INTEGER) AS
frequency,
    CAST(DATEDIFF('2022-09-01', MAX(Purchase_Date)) AS SIGNED INTEGER) AS recency,
    CAST(ROW_NUMBER() OVER (ORDER BY SUM(GMV) / DATEDIFF('2022-09-01',
MAX(created_date)) DESC) AS SIGNED INTEGER) AS rn_monetary,
    CAST(ROW_NUMBER() OVER (ORDER BY COUNT(ID) / DATEDIFF('2022-09-01',
MAX(created_date)) DESC) AS SIGNED INTEGER) AS rn_frequency,
    CAST(ROW_NUMBER() OVER (ORDER BY DATEDIFF('2022-09-01', MAX(Purchase_Date)) ASC)
AS SIGNED INTEGER) AS rn_recency
FROM
    customer_registered
JOIN
    customer_transaction
ON
    customer_registered.id = customer_transaction.customerID
GROUP BY
    contract;

CREATE TABLE rfm as
(SELECT
    contract,
    monetary,
    frequency,
    recency,
    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,
```

**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\_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**

**FROM**

**CTE**

**ORDER BY**

**contract);**

**WITH A AS (**

**SELECT \*, CONCAT(R, F, M) AS RFM\_Group**

**FROM rfm)**

**SELECT**

**RFM\_Group,**

**COUNT(\*) AS total\_client,**

**CASE**

**WHEN RFM\_Group IN ('444', '244', '344') THEN 'kimcuong'**

**WHEN RFM\_Group IN ('333', '233', '322') THEN 'vang'**

**WHEN RFM\_Group = '411' THEN 'bac'**

**WHEN RFM\_Group IN ('211', '111', '122') THEN 'dong'**

**END AS Segmentation**

**FROM A**

**GROUP BY RFM\_Group**

**ORDER BY total\_client DESC;**