TRƯỜNG ĐẠI HỌC MỞ THÀNH PHỐ HỒ CHÍ MINH

**KHOA CÔNG NGHỆ THÔNG TIN**



**BÁO CÁO MÔN HỌC**

**ĐỀ TÀI: PHÂN TÍCH PHÂN KHÚC KHÁCH HÀNG**

**Môn học: Phân tích dữ liệu**

**GVHD: ThS. Hồ Hướng Thiên**

**Lớp: IM2201**

**SINH VIÊN THỰC HIỆN:**

**Lê Thị Ngọc Hân – 2254050012**

**Lưu Thị Phương Thảo – 2254050061**

**Nguyễn Lê Huệ Tiên - 2254050069**

# **PHẦN MỞ ĐẦU**

# **Lý do chọn đề tài**

Trong bối cảnh chuyển đổi số và sự phát triển mạnh mẽ của thương mại điện tử, việc nắm bắt hành vi và nhu cầu của khách hàng ngày càng trở nên quan trọng đối với sự phát triển của các doanh nghiệp. Đặc biệt, với sự gia tăng các giao dịch mua sắm trực tuyến, dữ liệu khách hàng ngày càng phong phú và đa dạng. Những dữ liệu này không chỉ giúp doanh ngiệp hiểu rõ hơn về khách hàng của mình mà còn là cơ sở để xây dựng các chiến lược kinh doanh và marketing hiệu quả, tối ưu hóa trải nghiệm khách hàng và nâng cao hiệu quả kinh doanh.

Trung tâm Kinh doanh Hàng thời trang Merriman, thuộc Tổng Công ty Cổ phần Dệt may Hòa Thọ, chuyên kinh doanh các sản phẩm thời trang công sở nam và nữ thông qua các kênh bán hàng trực tuyến như website và các hệ thống cửa hàng trên toàn quốc. Báo cáo kinh doanh năm 2022 và nửa đầu năm 2023 cho thấy doanh thu từ thương mại điện tử và website chiếm 48,7% tổng doanh thu của trung tâm, điều này chứng tỏ vai trò quan trọng của khách hàng trực tuyến đối với hoạt động kinh doanh của công ty. Để duy trì và phát triển trên thị trường thương mại điện tử, việc phân loại và hiểu rõ hành vi của khách hàng trực tuyến là rất cần thiết để tối ưu hóa chiến lược marketing và phát triển bền vững.

Vì vậy, chúng tôi đã ứng dụng phương pháp phân khúc khách hàng bằng RFM kết hợp thuật toán phân cụm K-Means cho Trung tâm Kinh doanh Hàng thời trang Merriman – Tổng Công ty Cổ phần Dệt may Hòa Thọ. Dữ liệu sử dụng trong nghiên cứu này được trích xuất trực tiếp từ hoạt động kinh doanh của trung tâm, nhằm cung cấp những phân tích chi tiết và đề xuất giải pháp chiến lược giúp doanh nghiệp nâng cao hiệu quả kinh doanh, tối ưu hóa chiến lược tiếp cận khách hàng và gia tăng lợi thế cạnh tranh.

# **Đối tượng và phạm vi nghiên cứu**

Khách hàng đã mua sắn trên website Merriman.com của TTKDTT Merriman trong khoảng thời gian từ tháng 01/2022 đến 06/2023

**MỤC LỤC**

[**PHẦN MỞ ĐẦU** 2](#_Toc185865575)

[**1.** **Lý do chọn đề tài** 2](#_Toc185865576)

[**2.** **Đối tượng và phạm vi nghiên cứu** 2](#_Toc185865577)

[**CHƯƠNG 1. ỨNG DỤNG MÔ HÌNH PHÂN KHÚC KHÁCH HÀNG BẰNG RFM KẾT HỢP THUẬT TOÁN PHÂN CỤM K-MEANS** 3](#_Toc185865578)

[**1. Tiền xử lý dữ liệu.** 3](#_Toc185865579)

[**1.1. Xử lý từng bảng dữ liệu.** 3](#_Toc185865580)

[**1.2. Tiền xử lý bộ dữ liệu.** 5](#_Toc185865581)

[**1.3. Mô tả bộ dữ liệu.** 6](#_Toc185865582)

[**2. Phân tích tổng quát tình hình hoạt động kinh doanh trên Website.** 7](#_Toc185865583)

[**3. Xây dựng mô hình RFM** 9](#_Toc185865584)

[**3.1. Tính điểm RFM tổng thể** 10](#_Toc185865585)

[**3.2. Phân khúc và gắn nhãn nhóm khách hàng** 13](#_Toc185865586)

[**4. Xây dựng mô hình với thuật toán K-means** 14](#_Toc185865587)

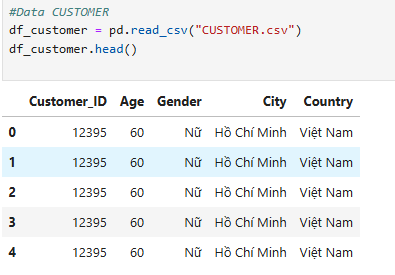
[**CHƯƠNG 2. ĐỀ XUẤT CHIẾN LƯỢC MARKETING** 18](#_Toc185865588)

# **CHƯƠNG 1.** **ỨNG DỤNG MÔ HÌNH PHÂN KHÚC KHÁCH HÀNG BẰNG RFM KẾT HỢP THUẬT TOÁN PHÂN CỤM K-MEANS**

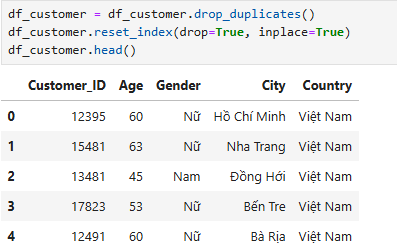
# **1. Tiền xử lý dữ liệu.**

## **1.1. Xử lý từng bảng dữ liệu.**

1. Bảng CUSTOMER

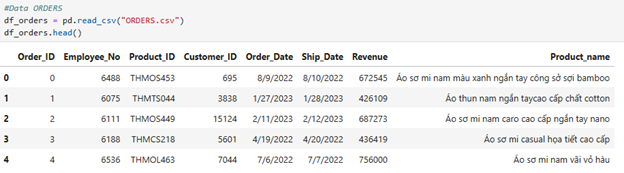


Dữ liệu trùng lặp nên tiến hành làm sạch dữ liệu bằng cách loại bỏ các bản ghi trùng lặp

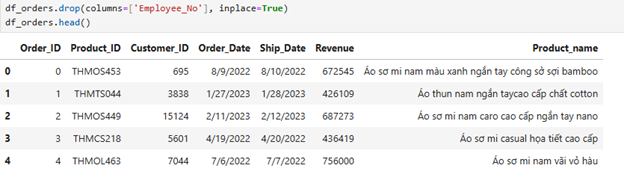


Ban đầu bảng có 121946 dòng và 5 cột. Sau khi tiến hành loại các bản ghi bị trùng, bảng Customer bây giờ có 18173 dòng và 5 cột, không có giá trị nào bị null hay bằng 0.

1. Bảng ORDERS

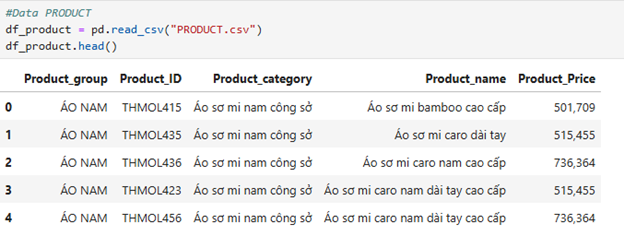


Loại bỏ cột Employee\_No vì nó không mang lại giá trị cho việc phân tích. Biến đổi kiểu dữ liệu datetime cho cột Order\_Date và Ship\_Date



Sau khi xử lý, bảng Orders có 142704 dòng và 8 cột, không có giá trị nào bị null hay bằng 0.

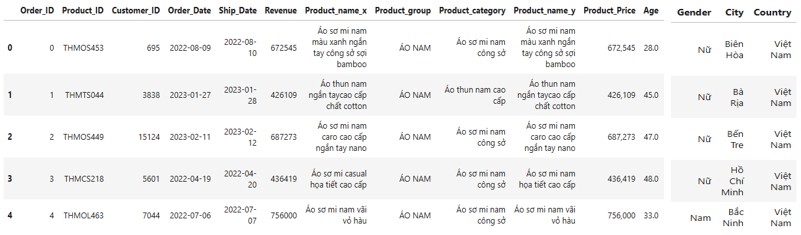
1. Bảng PRODUCT



Bảng Product có 195 dòng và 5 cột, không có giá trị nào bị null hay bằng 0.

## **1.2. Tiền xử lý bộ dữ liệu.**

Sau khi hoàn tất xử lý từng bảng dữ liệu riêng lẻ, tiến hành hợp nhất các bảng này thành một bảng dữ liệu tổng hợp FINAL, đồng thời loại bỏ các giá trị trùng lặp và cột không mang lại giá trị cho việc phân tích để đảm bảo tính chính xác và nhất quán. Bảng dữ liệu này sẽ được sử dụng làm cơ sở cho quá trình phân tích và thực hiện báo cáo.

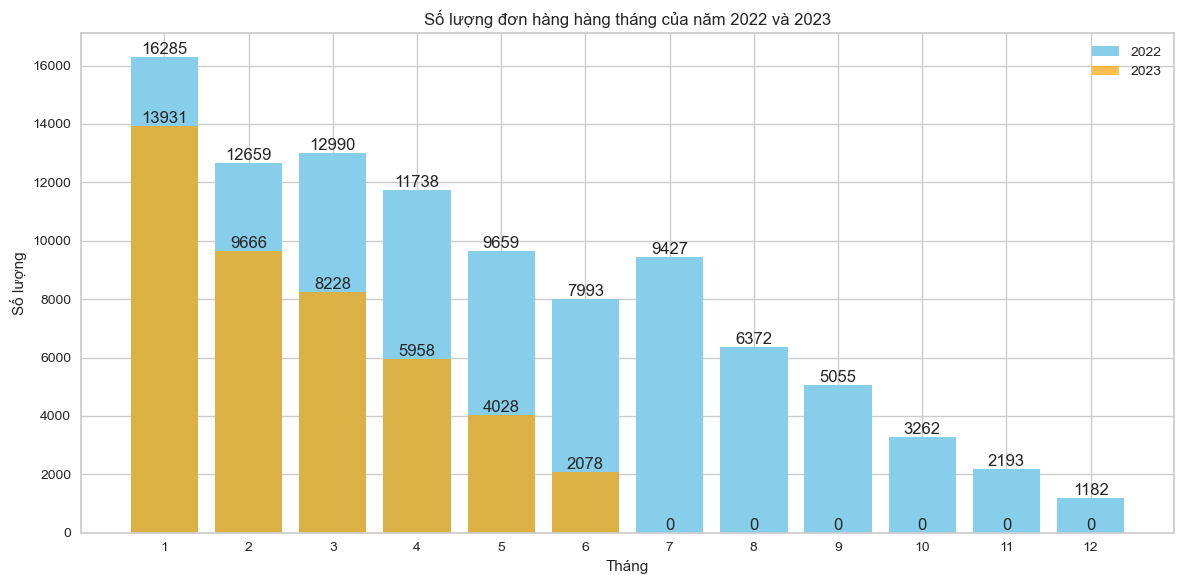


## **1.3. Mô tả bộ dữ liệu.**

Sau khi đã qua tiền xử lý và lọc những thông tin cần thiết cho việc phân tích thì dữ liệu có 142704 dòng và 14 cột.

|  |  |  |  |
| --- | --- | --- | --- |
| **STT** | **Tên cột** | **Datatype** | **Giải thích** |
| 1 | Order\_ID | Int64 | Mã đơn hàng |
| 2 | Product\_ID | String | Mã sản phẩm |
| 3 | Customer\_ID | Int64 | Mã khách hàng |
| 4 | Order\_Date | Datetime | Ngày đặt hàng |
| 5 | Ship\_Date | Datetime | Ngày giao hàng |
| 6 | Revenue | Int64 | Doanh thu |
| 7 | Product\_Name | String | Tên sản phẩm |
| 8 | Product\_group | String | Nhóm sản phẩm |
| 9 | Product\_category | String | Danh mục sản phẩm |
| 10 | Product\_Price | String | Giá sản phẩm |
| 11 | Age | Float64 | Tuổi khách hàng |
| 12 | Gender | String | Giới tính |
| 13 | City | String | Thành phố |
| 14 | Country | String | Quốc gia |

# **2. Phân tích tổng quát tình hình hoạt động kinh doanh trên Website.**

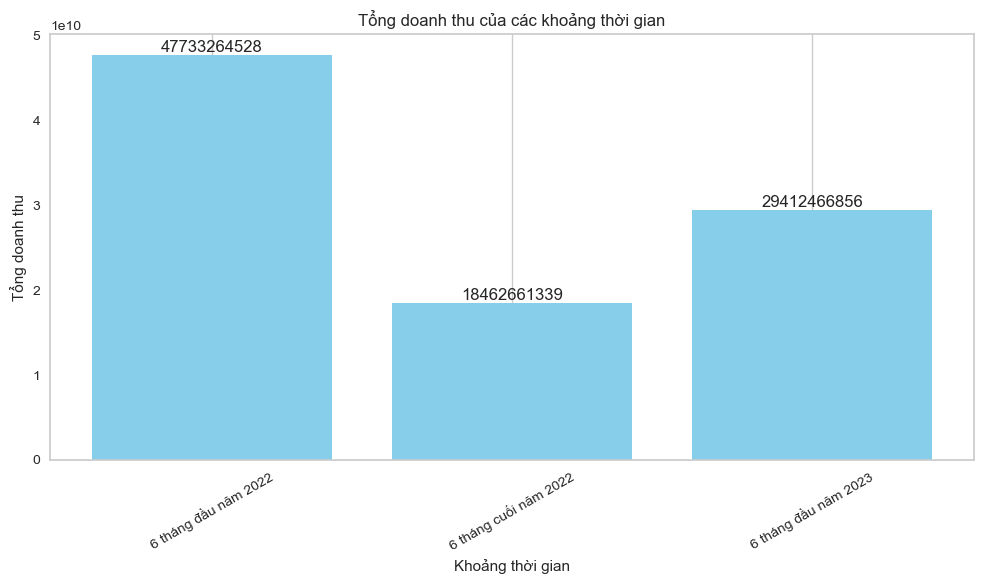
****

**Hình 2.1. Số lượng đơn hàng hàng tháng năm 2022 và 2023**

Số lượng đơn hàng qua các tháng trong hai năm 2022 và 2023 có sự biến động rõ rệt. Cả hai năm đều thể hiện xu hướng giảm dần số lượng đơn hàng từ đầu năm đến cuối năm. Đáng chú ý, trong năm 2022, tháng 1 và tháng 2 ghi nhận số lượng đơn hàng cao nhất, đặc biệt là tháng 1 – thời điểm cao điểm mua sắm sau kỳ nghỉ lễ và các chương trình khuyến mãi đầu năm. Từ tháng 3 trở đi, số lượng đơn hàng giảm dần, cho thấy nhu cầu tiêu dùng đã ổn định hơn.

Trong khi đó, dữ liệu năm 2023 chỉ cung cấp đến tháng 6, nhưng số lượng đơn hàng 6 tháng đầu năm 2023 thấp hơn rõ rệt so với cùng kỳ năm 2022. Điều này có thể xuất phát từ các yếu tố như sức mua giảm, chiến lược kinh doanh chưa hiệu quả hoặc những tác động khách quan từ thị trường. Sự chênh lệch này dẫn đến ảnh hưởng trực tiếp đến doanh thu, khi doanh số trong năm 2023 cũng thấp hơn so với năm trước.

Từ những dữ liệu này, doanh nghiệp có thể cân nhắc xây dựng chiến lược để thúc đẩy mua sắm vào các tháng thấp điểm, phân tích thêm nguyên nhân sụt giảm để đưa ra các biện pháp cải thiện, đồng thời dự báo và lên kế hoạch phù hợp nhằm tăng trưởng bền vững trong tương lai.

****

**Hình 2.2. Tổng doanh thu của các khoảng thời gian**

Doanh thu trong hai năm 2022 và 2023 có sự biến động rõ rệt. Năm 2022, 6 tháng đầu năm đạt gần 48 tỷ nhưng 6 tháng cuối năm giảm hơn một nửa, chỉ còn khoảng 18.5 tỷ, cho thấy sự sụt giảm đáng kể do yếu tố mùa vụ hoặc sức mua giảm.

Sang năm 2023, doanh thu 6 tháng đầu năm tăng trở lại lên 29 tỷ, cao hơn 6 tháng cuối năm 2022 khoảng 61%, thể hiện tín hiệu phục hồi tích cực. Tuy nhiên, mức này vẫn thấp hơn so với cùng kỳ năm 2022, cho thấy doanh nghiệp cần phân tích và cải thiện chiến lược kinh doanh để duy trì đà tăng trưởng ổn định và vượt qua các yếu tố tác động từ thị trường.

|  |  |
| --- | --- |
| https://lh7-rt.googleusercontent.com/docsz/AD_4nXcitFQ6L3YarSKhKvvTCqACqx5MNCk59ojCb9AqE24AJBgZlGmfGBrGNknO9YXdNodpdzkEcVSZOIjfTmINksfGXDIVhAkSEafgjAalE88l6EgiC3L1c5mN2faVFG1l_rxLSHvqwA?key=tA_l-wytks9cSb-5w8ELWrWt  **Hình 2.3. Tỉ lệ số lượng theo giới tính** | https://lh7-rt.googleusercontent.com/docsz/AD_4nXeTPz7Z2dqD0uk-RxCxqPa9cR2XjA6DNrSRIxVz71I3h7-IBH8m4JPcm1QKRLW0yx3gpeWYIebxNKp9YH-kaG8ExpG_S8zIp9XPVMmXgcYBWCDu8Ek3l4V7hGTi4LaCToQ6JfjZ0Q?key=tA_l-wytks9cSb-5w8ELWrWt  **Hình 2.4. Phân phối các nhóm sản phẩm** |

Dựa vào biểu đồ tỉ lệ giới tính cho thấy khách hàng nam chiếm 50,7% chỉ nhỉnh hơn khách hàng nữ 49,3%, thể hiện sự cân bằng về giới tính trong tập khách hàng

Tuy nhiên, khi phân tích theo nhóm các sản phẩm, có thể thấy khách hàng nam mua nhiều hơn ở hầu hết các nhóm, đặc biệt là các sản phẩm thuộc “Áo nam” (72,3%) và “Quần nam”(12,8%). Điều này khẳng định rằng nhóm khách hàng nam là đối tượng tiêu dùng chính của cửa hàng

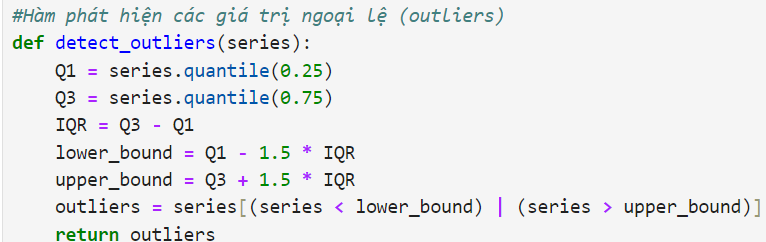
# **3. Xây dựng mô hình RFM**

Mô hình RFM là một mô hình thuộc Marketing Analysis và dùng để phân tích dữ liệu về Customer Value (giá trị khách hàng). Nhờ đó mà doanh nghiệp có thể rút ra chỉ số quan trọng liên quan tới từng nhóm khách hàng mà mình đang phục vụ. Qua đó sẽ có những chiến lược chăm sóc đặt biệt hay chương trình khuyến mãi cho từng đối tượng tiêu dùng. Mô hình RFM đánh giá các khía cạnh quan trọng của hành vi mua hàng của khách hàng dựa trên 3 yếu tố

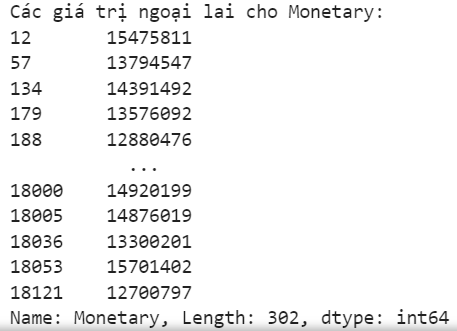
* Recency: Khoảng thời gian mà khách hàng mua hàng gần đây nhất
* Frequency:Tần suất mua hàng của khách hàng
* Monetary: Giá trị mỗi lần mua hàng của khách hàng

## **3.1. Tính điểm RFM tổng thể**

Đầu tiên cần tính các giá trị ngoại lai (outliner) để kiểm tra xem các giá trị này có phù hợp với thực tế hay không và nó sẽ tác động ảnh hưởng đến RFM ra sao



|  |  |
| --- | --- |
| https://lh7-rt.googleusercontent.com/docsz/AD_4nXcAbvtwTI20-6OO2ubkJDdI2x4Kx078z6c4Q9B5uL91umx3FbUfyq0wutnToszPOCcpYqiekMD3rO-KbC2QI4viafw2ipDfGBmnavVbbu2Ptn_Q-aJ1zsbUIuDIhrTt7ZdqUUwfxw?key=tA_l-wytks9cSb-5w8ELWrWt | https://lh7-rt.googleusercontent.com/docsz/AD_4nXefzM8hyRyo0OCo3UuHmKt7fyKcV9gYl-x_wk2oj8yyDNTP42La4oUU2FU3Am5PrxAnwQGeji2vYNHxwoj3OBfiK0K-N1LYSkp1JiUbQsHLNkLvZqXnuH4jQdTJ5QoZUl2HtzgGCQ?key=tA_l-wytks9cSb-5w8ELWrWt |

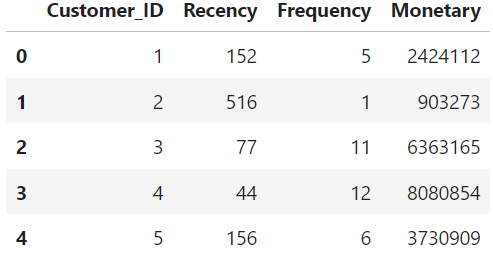


Vì các giá trị ngoại lai này phù hợp với thực tế và không làm ảnh hưởng đáng kể đến kết quả phân tích, nên chúng vẫn được giữ lại trong quá trình phân tích thống kê.

Recency: Biểu thị số ngày từ lần mua hàng gần nhất của mỗi khách hàng đến ngày xác định là sau ngày mua hàng gần nhất trong toàn bộ bộ dữ liệu. Điều này được thực hiện bởi vì bộ dữ liệu chỉ cung cấp thông tin đến ngày 29/06/2023, không phải đến thời điểm hiện tại.

Frequency: Dữ liệu sẽ được nhóm theo 'Customer\_ID', tức là tất cả các giao dịch liên quan đến mỗi khách hàng sẽ được tổng hợp lại. Sau đó, số lần mua hàng của mỗi khách hàng sẽ được đếm để tính tần suất mua sắm của họ.

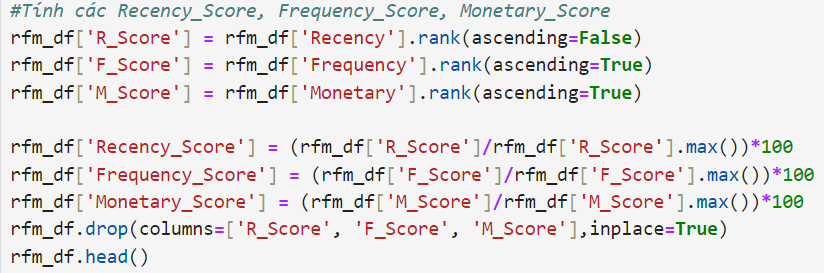
Monetary: Dữ liệu sẽ được nhóm theo 'Customer\_ID' và tính tổng giá trị của cột 'Revenue' cho tất cả các lần mua hàng của mỗi khách hàng, giúp xác định tổng chi tiêu của từng khách hàng.

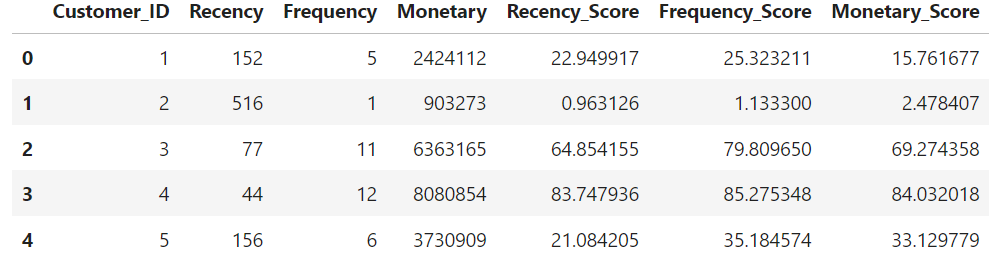
**

Dữ liệu cho thấy sự phân hóa rõ rệt giữa các khách hàng về tần suất mua hàng, giá trị chi tiêu và thời gian kể từ lần mua gần nhất. Một số khách hàng có xu hướng mua hàng rất thường xuyên và chi tiêu nhiều, trong khi những khách hàng khác ít khi mua hàng và chi tiêu thấp.

Tiếp theo, xếp hạng khách hàng, chúng tôi tiến hành phân loại khách hàng dựa trên thang điểm 0 - 100. Trong đó

* Đối với khách hàng có số điểm gần với 100 sẽ có xu hướng quay lại mua hàng, tần suất giao dịch cao, chi tiêu nhiều.
* Đối với khách hàng có số điểm gần với 0 sẽ có khả năng không quay lại, tần suất giao dịch ít, chi tiêu ít.

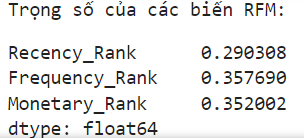




Theo phân tích tổng quát tình hình công ty, việc doanh thu 6 tháng đầu năm 2023 giảm đáng kể so với cùng kỳ năm 2022 là điều đáng chú ý, mục tiêu tối đa hóa doanh thu từ khách hàng trở nên quan trọng hơn hết.

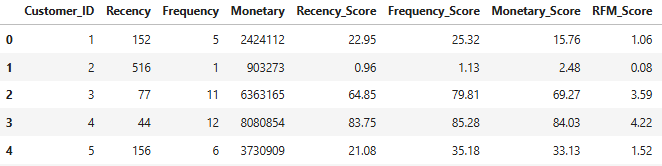
'Monetary\_Score' đại diện cho tổng số tiền mà khách hàng đã chi tiêu, phản ánh trực tiếp vào doanh thu và lợi nhuận của công ty. Một khách hàng có giá trị mua hàng cao ('Monetary\_Score' cao) có thể quan trọng hơn về mặt tài chính so với một khách hàng mua hàng thường xuyên (có 'Frequency\_Score' cao) hoặc một khách hàng gần đây mới mua hàng (có 'Recency\_Score' cao) nhưng chi tiêu ít hơn. Điều này có thể khiến cho 'Monetary\_Score' trở thành yếu tố quyết định khi xem xét giá trị dài hạn của một khách hàng, và do đó, nó nhận được trọng số cao hơn trong việc tính điểm tổng hợp RFM.

Điểm RFM được tính toán dựa trên ba yếu tố chính: Recency (Khoảng thời gian mà khách hàng mua hàng gần đây nhất), Frequency (Tần suất mua hàng của khách hàng), và Monetary (Giá trị mỗi lần mua hàng của khách hàng). Mỗi yếu tố này sẽ được chuẩn hóa theo thứ hạng và sau đó kết hợp lại để tính ra điểm RFM tổng hợp cho mỗi khách hàng. Mức độ quan trọng của từng yếu tố trong công thức này được xác định bằng các trọng số khác nhau.



Công thức tính điểm RFM được áp dụng như sau:

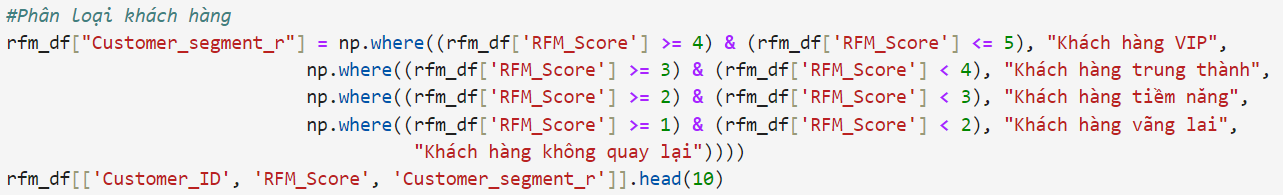
|  |
| --- |
| 0.29\*Recency + 0.36\*Frequency + 0.35 \*Monetary |

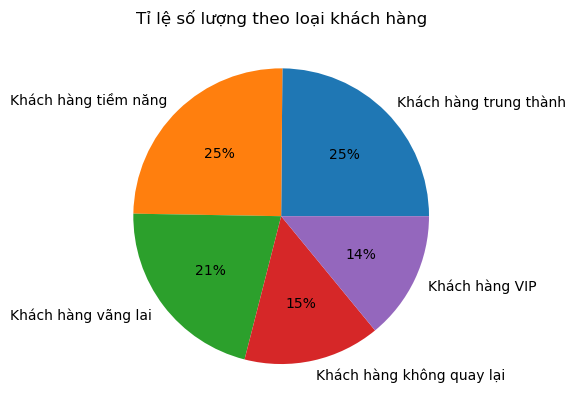


## **3.2. Phân khúc và gắn nhãn nhóm khách hàng**

Sau khi tính các giá trị RFM, chúng ta tiếp tục phân khoảng giá trị này và gán chúng cho một mức điểm số trên thang điểm 1-5, với 1 điểm là giá trị thấp nhất và 5 điểm là giá trị cao nhất. Ta sử dụng phương pháp chia khoảng đơn giản. Trong đó

* Điểm RFM >= 4: Khách hàng VIP. Đây là nhóm khách hàng mua gần đây, mua hàng thường xuyên và chi tiêu nhiều nhất
* 4 > Điểm RFM >= 3: Khách hàng trung thành. Đây là nhóm khách hàng có phản ứng tích cực với các chương trình khuyến mãi của khách hàng
* 3 > Điểm RFM >= 2: Khách hàng tiềm năng. Đây là nhóm khách hàng mới mua gần đây, nhưng đã chi tiêu một khoản lớn và mua nhiều hơn một lần
* 2 > Điểm RFM >= 1: Khách hàng vãng lai. Đây là nhóm khách hàng có tần suất mua ít, đã lâu không mua hàng và số tiền chi trả cho một đơn hàng không nhiều
* Điểm RFM < 1: Khách hàng không quay lại. Đây là nhóm khách hàng có tần suất mua rất ít, đã hơn 6 tháng khách hàng không phát sinh đơn hàng và số tiền chi trả cho một đơn hàng rất ít

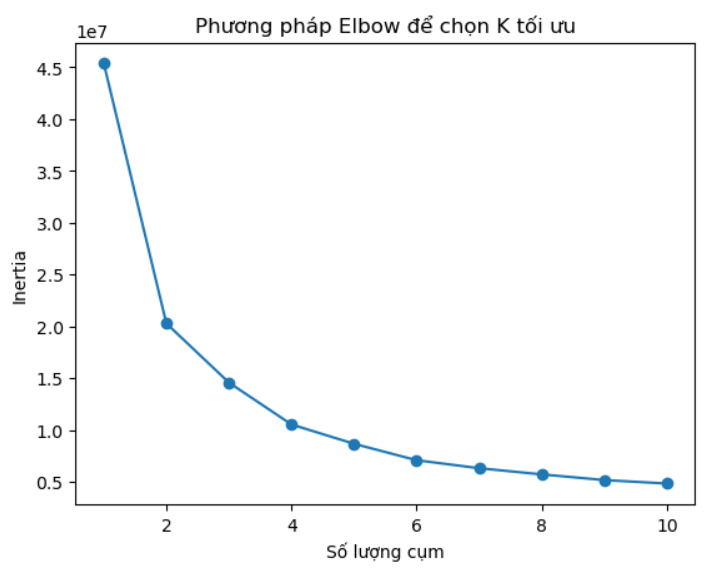


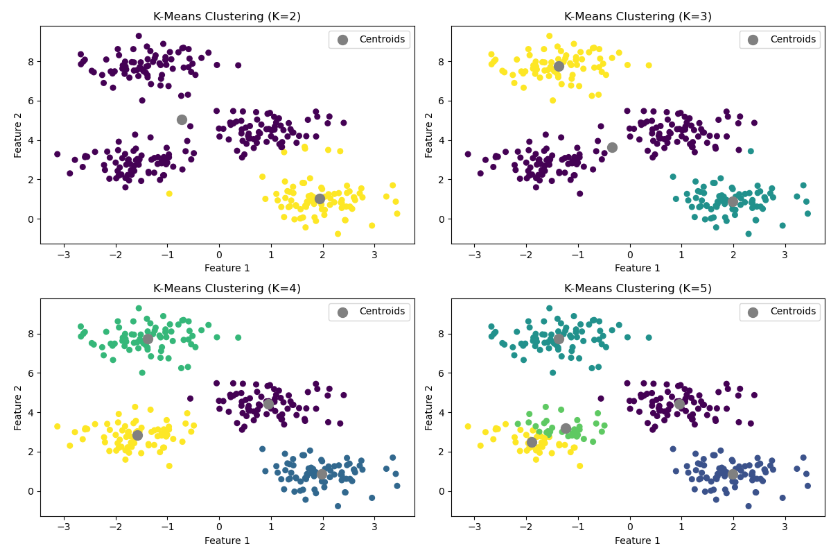


# **4. Xây dựng mô hình với thuật toán K-means**

Sử dụng thuật toán K-mean clustering để phân cụm khách hàng với đầu vào là giá trị của ba biến Recency, Frequency và Monetary đã chuẩn hóa trên mô hình RFM đã xây dựng.

Trước tiên, ta dùng phương pháp Elbow (phương pháp khuỷu tay) để xác định số lượng cụm tối ưu. Phương pháp này dựa vào mức độ thay đổi của tổng bình phương khoảng cách từ mỗi điểm dữ liệu đến tâm cụm gần nhất của nó (Within-Cluster Sum of Squares - WCSS) khi thay đổi số lượng cụm.



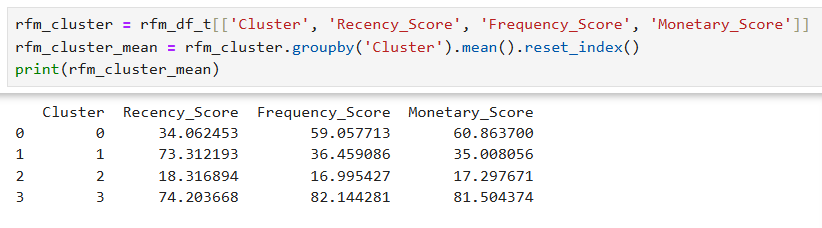


**Hình 4.1. Kết quả thực nghiệm được minh họa bằng Đồ thị Elbow (2a-hình bên trên) và bốn đồ thị phân cụm (2b-hình bên dưới)**

Trong đồ thị Elbow (Hình 2a), điểm K=4 có thể xem là điểm “khuỷu tay”trên đồ thị, nơi có sự thay đổi đáng kể về tốc độ giảm WCSS. Việc tăng số cụm từ 3 lên 4 mang lại cải thiện đáng kể, trong khi tăng từ 4 lên 5 (hoặc cao hơn) chỉ mang lại cải thiện nhỏ.

Bên cạnh đó, trong biểu đồ phân cụm (Hình 2b), với K=4 cũng cho thấy sự phân chia rõ ràng và có ý nghĩa nhất giữa các cụm.

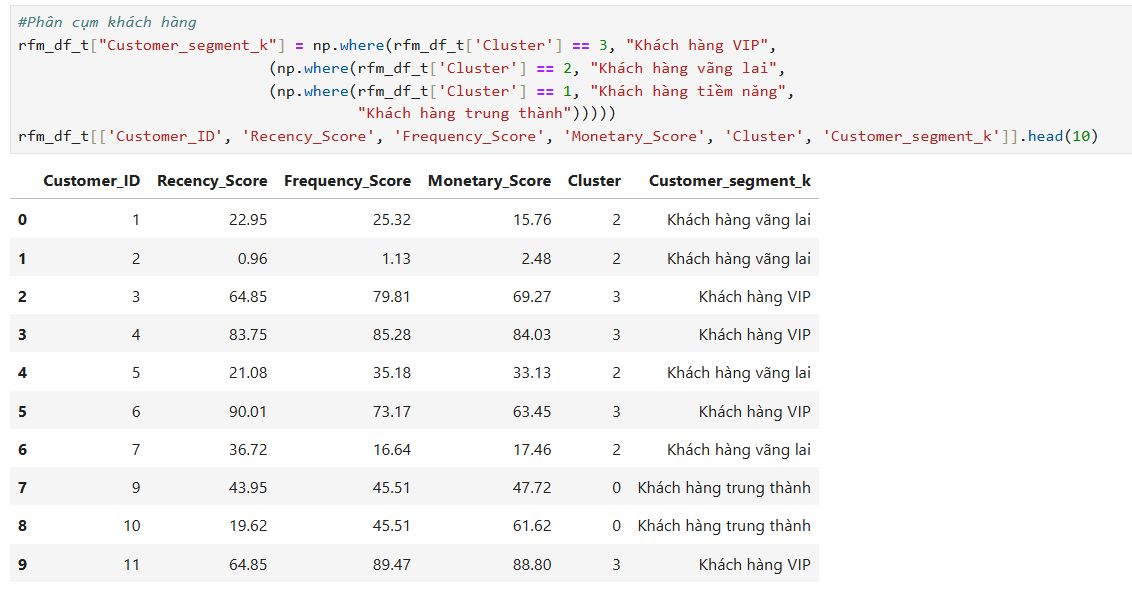
Kết hợp với kết quả mô hình RFM bên trên, ta quyết định chọn K=4 là số cụm tối ưu.



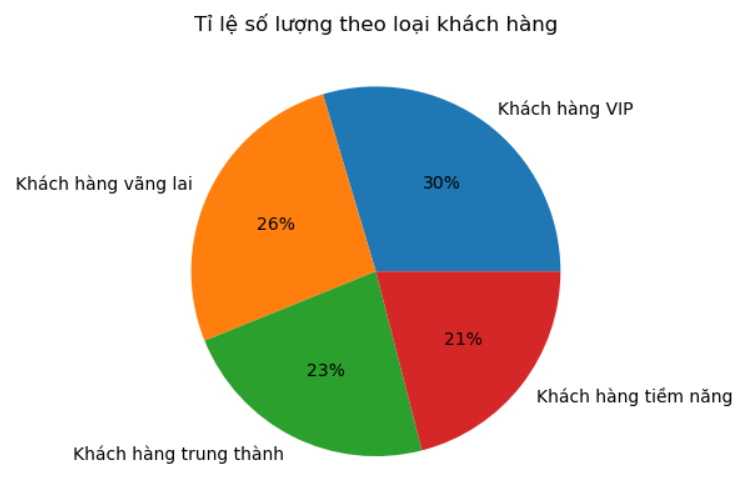
**Hình 4.2. Kết quả tính RFM trung bình của từng cụm**

Dựa vào điểm RFM, tiến hành phân chia cụm khách hàng có sự thay đổi như sau:

* Cluster = 3: Khách hàng VIP. Đây là nhóm khách hàng có chỉ số RFM cao thể hiện nhóm khách hàng mua gần đây, mua hàng thường xuyên và chi tiêu nhiều nhất
* Cluster = 2: Khách hàng vãng lai. Đây là nhóm khách hàng có chỉ số RFM thấp thể hiện khách hàng có tần suất mua ít, đã lâu không mua hàng và số tiền chi trả cho một đơn hàng không nhiều
* Cluster = 1: Khách hàng tiềm năng. Đây là nhóm khách hàng có Recency và Monetary thấp, nhưng Frequency ở mức trung bình thể hiện khách hàng mới mua gần đây, nhưng đã chi tiêu một khoản lớn và mua nhiều hơn một lần
* Cluster = 0: Khách hàng trung thành. Đây là nhóm khách hàng có chỉ số Frequency và Monetary cao, mặc dù Recency thấp thể hiện khách hàng có phản ứng tích cực với các chương trình khuyến mãi của khách hàng, mặc dù mua hàng gần đây ít nhưng có tần suất mua hàng và chi tiêu cao

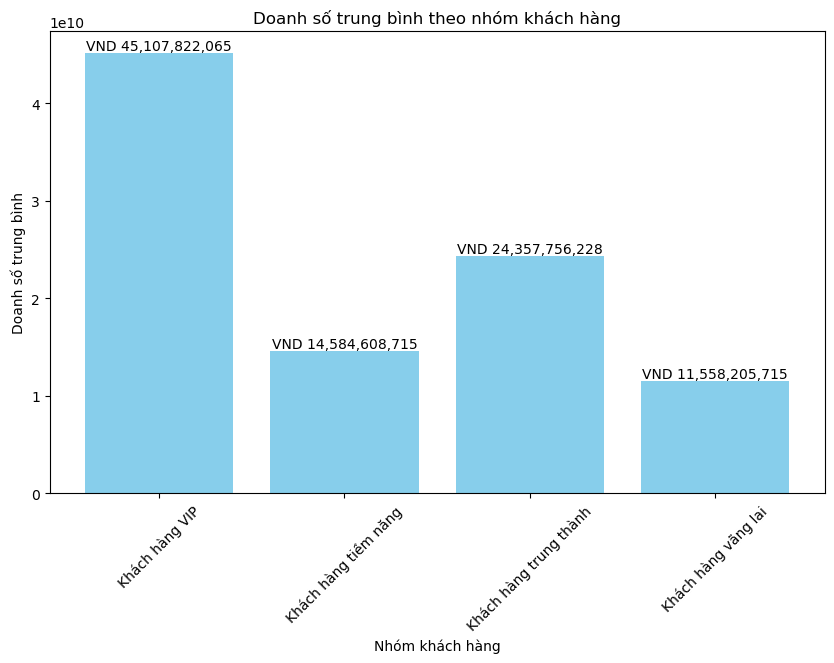


**Hình 4.3. Phân nhóm khách hàng theo cụm**



**Hình 4.4. Tỉ lệ số lượng khách hàng theo phân khúc**

# **CHƯƠNG 2. ĐỀ XUẤT CHIẾN LƯỢC MARKETING**



Dựa vào biểu đồ về doanh số trung bình theo nhóm khách hàng, có thể đề xuất một số chiến lược marketing như sau:

**1. Tập trung vào khách hàng VIP**

* Khách hàng VIP mang lại doanh số trung bình cao nhất (VND 45,116,540,120). Đây là nhóm có giá trị lớn nhất đối với doanh nghiệp, do đó:
  + Tạo các chương trình chăm sóc khách hàng đặc biệt như ưu đãi độc quyền, sự kiện riêng tư hoặc quà tặng cá nhân hóa.
  + Xây dựng hệ thống chăm sóc cá nhân hóa như gán nhân viên tư vấn riêng hoặc hỗ trợ ưu tiên.
  + Duy trì mối quan hệ bền chặt với nhóm này để tăng khả năng mua lại và lòng trung thành.

**2. Xây dựng lòng trung thành từ khách hàng trung thành**

* Khách hàng trung thành có doanh số trung bình tương đối cao (VND 24,365,724,473), chứng minh họ là nhóm đáng tin cậy:
  + Tăng cường chương trình tích điểm thưởng hoặc hoàn tiền để khuyến khích mua sắm thường xuyên.
  + Cung cấp các gói giảm giá dài hạn hoặc thẻ thành viên nâng cấp.
  + Khuyến khích khách hàng trung thành giới thiệu bạn bè, người thân qua các chương trình referral (giới thiệu) với phần thưởng hấp dẫn.

**3. Thúc đẩy chuyển đổi từ khách hàng tiềm năng**

* Khách hàng tiềm năng có doanh số trung bình thấp hơn nhưng vẫn đáng chú ý (VND 14,571,317,545):
  + Đầu tư vào các chiến dịch quảng cáo nhắm mục tiêu dựa trên sở thích, hành vi mua sắm của họ.
  + Tạo các ưu đãi dùng thử hoặc giảm giá cho lần mua đầu tiên để tăng tỷ lệ chuyển đổi.
  + Gửi email/ tin nhắn cá nhân hóa để nhắc nhở và khuyến khích họ thực hiện giao dịch.

**4. Tăng doanh số từ khách hàng vãng lai**

* Khách hàng vãng lai có doanh số trung bình thấp nhất (VND 11,554,810,585), tuy nhiên vẫn có tiềm năng:
  + Tập trung vào các chiến lược remarketing để nhắc nhở họ quay lại mua sắm.
  + Tạo các chương trình khuyến mãi ngắn hạn, giảm giá mạnh để thu hút sự chú ý và thúc đẩy mua hàng nhanh.
  + Gửi khảo sát hoặc ưu đãi để tìm hiểu lý do họ chưa trở thành khách hàng thường xuyên và cải thiện trải nghiệm.

**5. Phân bổ ngân sách quảng cáo hợp lý**

* Dựa trên giá trị trung bình từng nhóm, doanh nghiệp nên tập trung ngân sách vào khách hàng VIP và trung thành (vì đây là các nhóm có lợi nhuận cao nhất), đồng thời thiết lập các chiến dịch bổ sung để khai thác tiềm năng từ khách hàng tiềm năng và vãng lai.

**6. Ứng dụng công nghệ và dữ liệu**

* Sử dụng hệ thống CRM (Quản lý quan hệ khách hàng) để phân tích hành vi và tạo các chiến dịch cá nhân hóa hiệu quả hơn.
* Theo dõi và đánh giá hiệu quả từng chiến dịch để tối ưu hóa ngân sách marketing.

|  |  |
| --- | --- |
|  |  |
|  |  |

Dựa trên biểu đồ phân bố độ tuổi trong các cụm khách hàng, có thể đề xuất các chiến lược marketing phù hợp với đặc điểm từng nhóm tuổi. Phân tích cho thấy rằng đa số khách hàng tập trung trong độ tuổi 30-55, với tỷ lệ cao nhất trong tất cả các cụm, tiếp theo là độ tuổi 55+ và cuối cùng là nhóm 16-30.

**1. Nhóm khách hàng 30-55 (chiếm tỷ lệ cao nhất - khoảng 67-69% trong các cụm)**

Đặc điểm nhóm: Đây là nhóm khách hàng trung niên, thường có thu nhập ổn định, tiêu dùng hướng đến chất lượng, giá trị và sự tiện lợi.

* Chiến lược:
  + Tăng cường quảng bá chất lượng sản phẩm/dịch vụ: Nhấn mạnh vào độ bền, hiệu quả, và giá trị dài hạn của sản phẩm.
  + Ưu đãi giá trị: Cung cấp các gói khuyến mãi khi mua số lượng lớn hoặc các combo sản phẩm liên quan.
  + Chăm sóc khách hàng cá nhân hóa: Sử dụng email, tin nhắn, hoặc các chương trình khách hàng thân thiết để giữ chân nhóm này.
  + Nội dung marketing thực tế và chuyên sâu: Viết bài đánh giá sản phẩm hoặc video giới thiệu tính năng chi tiết, phù hợp với nhu cầu của họ.

**2. Nhóm khách hàng 55+ (chiếm khoảng 28-30% tùy cụm)**

Đặc điểm nhóm: Đây là nhóm khách hàng lớn tuổi, thường ưu tiên sản phẩm dễ sử dụng, đáng tin cậy, và có tính tiện ích cao.

* Chiến lược:
  + Quảng cáo tập trung vào giá trị sức khỏe và tiện ích: Nhấn mạnh các lợi ích sức khỏe, an toàn hoặc tính đơn giản của sản phẩm/dịch vụ.
  + Kênh marketing truyền thống: Sử dụng các kênh như báo in, truyền hình, hoặc các chiến dịch offline (bán hàng tại cửa hàng).
  + Hỗ trợ khách hàng: Cung cấp hotline hỗ trợ hoặc nhân viên tư vấn riêng để hướng dẫn sử dụng sản phẩm.
  + Chương trình ưu đãi cho người lớn tuổi: Cung cấp giảm giá hoặc các chương trình đặc biệt cho khách hàng từ 55 tuổi trở lên.

**3.Nhóm khách hàng 16-20 (chiếm tỷ lệ nhỏ nhất - dưới 2% trong các cụm)**

Đặc điểm nhóm: Đây là nhóm khách hàng trẻ, thích các sản phẩm hiện đại, thời trang, có tính cá nhân hóa cao. Họ cũng dễ bị thu hút bởi xu hướng và hoạt động trên mạng xã hội.

* Chiến lược:
  + Tận dụng mạng xã hội: Tăng cường quảng cáo trên Facebook, Instagram, TikTok, kết hợp với influencer marketing để thu hút sự chú ý.
  + Chiến dịch sáng tạo và giải trí: Tạo các nội dung viral, giveaway, hoặc minigame để tương tác với nhóm này.
  + Khuyến mãi nhanh: Áp dụng các chương trình giảm giá trong thời gian giới hạn (flash sale) để kích thích mua sắm.

|  |  |
| --- | --- |
|  |  |
|  |  |

Dựa vào biểu đồ tỷ lệ giới tính trong từng cụm khách hàng, dưới đây là các đề xuất chiến lược marketing phù hợp:

**1. Cụm 0: Nam chiếm 50,3%, Nữ chiếm 49,7%**

* Tỷ lệ giới tính gần như cân bằng, nên chiến lược marketing cần linh hoạt và phục vụ tốt cả hai giới.
* Chiến lược:
  + Cung cấp các sản phẩm/dịch vụ phổ biến với cả nam và nữ.
  + Thực hiện các chiến dịch quảng cáo trung lập, tập trung vào giá trị sản phẩm hơn là giới tính.
  + Tạo chương trình ưu đãi đôi, khuyến khích mua cặp (nam và nữ), ví dụ: "Mua 1 tặng 1 cho bạn đồng hành".

**2. Cụm 1: Nam chiếm 52,3%, Nữ chiếm 47,7%**

* Nam giới chiếm tỷ lệ cao hơn một chút, nên ưu tiên nội dung và sản phẩm hướng đến nhóm nam.
* Chiến lược:
  + Nhấn mạnh các sản phẩm/dịch vụ dành cho
  + Tạo nội dung quảng cáo mạnh mẽ, nhấn vào phong cách nam tính nhưng vẫn giữ sự cân bằng để không bỏ qua nhóm nữ.
  + Tổ chức sự kiện hoặc các chiến dịch liên quan đến thể thao hoặc công nghệ, thu hút nam giới tham gia.

**3. Cụm 2: Nam chiếm 50,9%, Nữ chiếm 49,1%**

* Tương tự cụm 0, tỷ lệ giới tính gần cân bằng.
* Chiến lược:
  + Tạo các sản phẩm/dịch vụ hướng đến cả hai giới, nhưng có thể ưu tiên nhẹ cho nhóm nam giới.
  + Phân phối các chiến dịch marketing trên các kênh trung lập, như Facebook, YouTube, hoặc quảng cáo theo chủ đề gia đình, cặp đôi.
  + Khuyến khích trải nghiệm mua sắm cho cả gia đình, với các gói combo phù hợp cho cả nam và nữ.

**4. Cụm 3: Nam chiếm 49,4%, Nữ chiếm 50,6%**

* Nữ giới chiếm tỷ lệ nhỉnh hơn, nên ưu tiên các sản phẩm/dịch vụ dành cho nữ nhưng vẫn không bỏ qua nam giới.
* Chiến lược:
  + Đẩy mạnh các sản phẩm liên quan đến thời trang, làm đẹp, và chăm sóc cá nhân dành cho nữ giới.
  + Tạo các chiến dịch quảng cáo nhẹ nhàng, tinh tế, nhắm vào sở thích và hành vi mua sắm của phụ nữ.
  + Tổ chức các chương trình dành riêng cho nữ như ưu đãi ngày Quốc tế Phụ nữ (8/3), Ngày Phụ nữ Việt Nam (20/10).