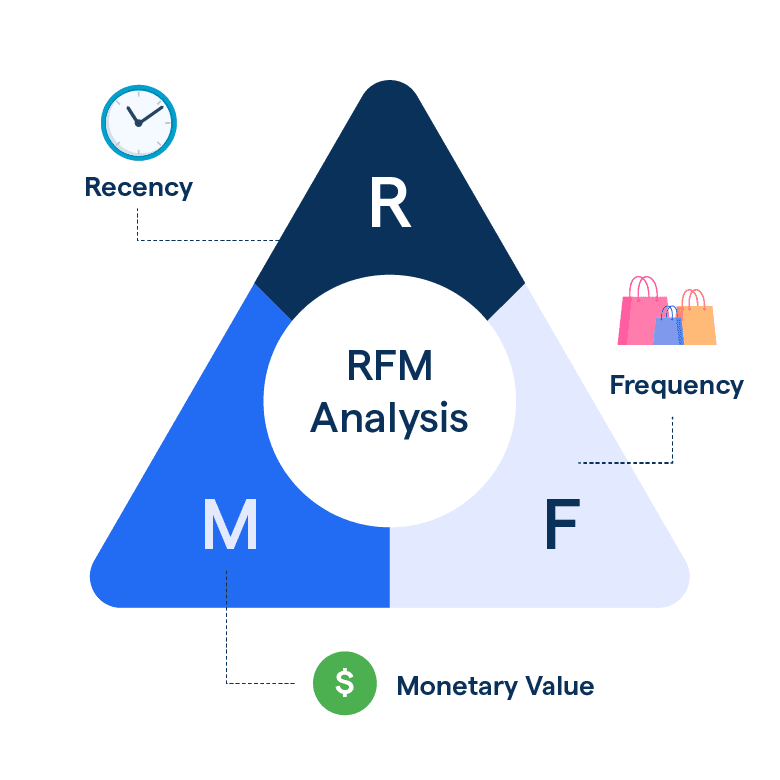
X

CUSTOMER

ANALYTICAL

**REPORT**

Phan Quốc Thắng

**Mục Lục**

[**I.** **Tổng quát:** 1](#_Toc174218523)

[**1.** **Giới thiệu dự án:** 1](#_Toc174218524)

[**2.** **Mô hình RFM** 2](#_Toc174218525)

[**3.** **BCG Matrix** 4](#_Toc174218526)

[**II.** **Xử lý dữ liệu và tính toán:** 6](#_Toc174218527)

[**1.** **Mô tả dữ liệu** 6](#_Toc174218528)

[**2.** **Xử lý dữ liệu:** 7](#_Toc174218529)

[**3.** **Phân nhóm khách hàng theo RFM** 8](#_Toc174218530)

[**III.** **Phân tích dữ liệu** 9](#_Toc174218531)

[**1.** **Tổng quan** 9](#_Toc174218532)

[**2.** **Phân tích theo từng phân nhóm khách hàng** 9](#_Toc174218533)

[**3.** **Phân bổ doanh thu theo từng năm tuổi của khách hàng** 14](#_Toc174218534)

[**4.** **Pie and Donut chart phân bổ doanh thu theo từng điểm số R, F, M** 15](#_Toc174218535)

[**5.** **Boxplot chart R, F, M** 16](#_Toc174218536)

**Danh mục hình ảnh**

[Figure 1 Customer\_Transaction: Chi tiết các giao dịch cuả khách hàng. 6](#_Toc174218080)

[Figure 2 Customer\_Registered: Thông tin chi tiết đăng ký của khách hàng 6](#_Toc174218081)

[Figure 3 Xử lý và tính toán dữ liệu 7](#_Toc174218082)

[Figure 4 Tạo cột dữ liệu R, F, M 7](#_Toc174218083)

[Figure 5 Tổng quan dữ liệu 9](#_Toc174218084)

[Figure 6 Biểu đồ phân tích theo nhóm khách hàng 9](#_Toc174218085)

[Figure 7 Biểu đồ số lượng khách hàng theo từng nhóm 10](#_Toc174218086)

[Figure 8 Biểu đồ trung bình doanh thu trên người theo từng nhóm nhóm khách hàng 11](#_Toc174218087)

[Figure 9 Biểu đồ tổng doanh thu và doanh thu theo tháng của từng nhóm khách hàng 11](#_Toc174218088)

[Figure 10 Thông tin chi tiết các nhóm RFM trong nhóm At-Risk VIP customers 13](#_Toc174218089)

[Figure 11 Tree map doanh thu theo năm tuổi khách hàng ủng hộ công ty 14](#_Toc174218090)

[Figure 12 Pie and Donut chart phân bổ doanh thu theo từng điểm số R, F, M 15](#_Toc174218091)

[Figure 13 Boxplot chart R, F, M 16](#_Toc174218092)

**Danh mục hình ảnh**

[Table 1 Customer\_Transaction: Chi tiết các giao dịch cuả khách hàng. 6](#_Toc174218226)

[Table 2 Customer\_Registered: Thông tin chi tiết đăng ký của khách hàng 6](#_Toc174218227)

[Table 3 Phân nhóm khách hàng theo RFM 8](#_Toc174218228)

1. **Tổng quát:**
2. **Giới thiệu dự án:**

**Hiện trạng ngành hàng:**

Trong ngành bán lẻ, đặc biệt là lĩnh vực siêu thị, sự cạnh tranh ngày càng gay gắt từ các tiểu thương, siêu thị ngoại, và trung tâm thương mại tạo ra những thách thức lớn đối với các doanh nghiệp nội địa.

Theo thống kê gần đây, các siêu thị ngoại chiếm khoảng **40%** thị phần bán lẻ tại Việt Nam, trong khi tỷ lệ khách hàng trung thành của các doanh nghiệp nội địa chỉ đạt khoảng **30%**. Điều này nhấn mạnh sự cần thiết trong việc cải thiện chiến lược kinh doanh để duy trì và nâng cao sức cạnh tranh.

**Vấn đề của công ty:**

Công ty X, một siêu thị hàng đầu tại thị trường nội địa, đang đối mặt với thách thức trong việc hiểu rõ hơn về khách hàng của mình nhằm duy trì sức cạnh tranh và tiếp tục phát triển bền vững. Mặc dù Công ty X đã đạt được thành công nhất định, **tỷ lệ khách hàng rời bỏ vẫn đạt mức 20%**, ảnh hưởng đáng kể đến doanh thu và lợi nhuận của công ty. Công ty đã thu thập một lượng lớn dữ liệu khách hàng từ các giao dịch hàng ngày nhưng chưa thực sự khai thác hết tiềm năng của dữ liệu này để tối ưu hóa hoạt động kinh doanh.

**Giải pháp**

Để tận dụng triệt để giá trị của dữ liệu khách hàng, công ty X đã quyết định áp dụng mô hình RFM (Recency, Frequency, Monetary) nhờ vào tính đơn giản, dễ hiểu và khả năng áp dụng thực tế cao của nó.

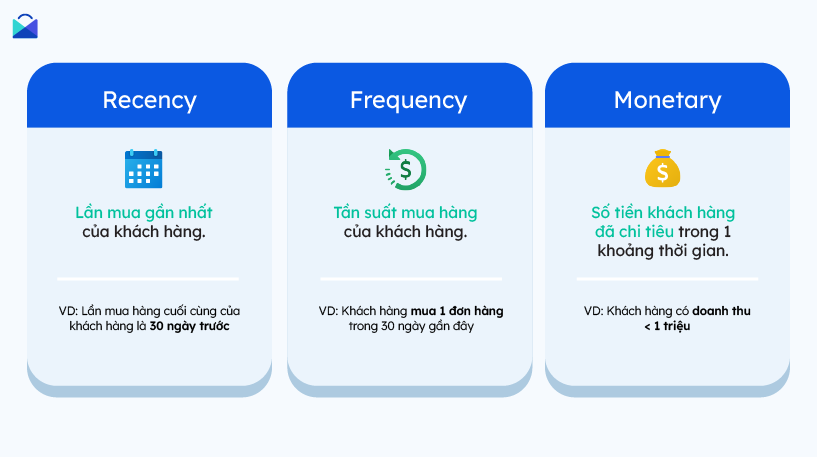
**Ngày thực hiện dự án: 2022-09-02**

**Mục đích của dự án**

* Phân nhóm khách hàng
* Mô tả đặc tính dữ liệu của từng nhóm.
* Đề xuất chiến lược tiếp cận và bán hàng phù hợp
* Nhận diện và giữ chân khách hàng có giá trị cao góp phần tăng tỉ lệ chuyển đổi mua hàng và tăng doanh thu

1. **Mô hình RFM**

Mô hình RFM là một kỹ thuật phân tích dữ liệu khách hàng phổ biến trong lĩnh vực marketing, giúp doanh nghiệp phân loại và hiểu rõ hơn về khách hàng của mình dựa trên ba tiêu chí chính:



* **Recency (R):** Đo lường thời gian gần nhất kể từ lần mua hàng cuối cùng của khách hàng. Khách hàng mua gần đây hơn được coi là có giá trị hơn.
* **Frequency (F):** Đo lường tần suất mua hàng của khách hàng trong một khoảng thời gian cụ thể. Khách hàng mua hàng thường xuyên hơn được coi là có giá trị hơn.
* **Monetary (M):** Đo lường tổng giá trị tiền tệ mà khách hàng đã chi tiêu trong một khoảng thời gian cụ thể. Khách hàng chi tiêu nhiều hơn được coi là có giá trị hơn.

**Cách Tính RFM Áp Dụng tại Công ty X**

Công ty X đã áp dụng mô hình RFM để phân loại khách hàng bằng cách sử dụng phân vị tứ phần (Interquartile Range - IQR).

A diagram of a number of blue squares

Description automatically generated

**Đây là cách tính cụ thể:**

Đối với mỗi cột:

|  |  |  |
| --- | --- | --- |
| **Cột** | **Mô tả** | **Thứ tự sắp xếp** |
| n\_GMV\_per\_month | Tổng doanh số theo tháng | Tăng dần |
| n\_FREQUENCY\_per\_month | Tần suất mua hàng theo tháng | Tăng dần |
| Days\_from\_latest\_purchase | Số ngày kể từ lần mua cuối cùng | Giảm dần |

Ta sẽ sắp xếp lại mỗi cột theo thứ tự như trên và chia ra thành 4 phần với mỗi phần có số lượng khách hàng bằng nhau. Mỗi phần tương ứng với một điểm.

Cột F, M sẽ được sắp xếp theo thứ tự tăng dần

**Recency (Thời gian từ lần mua gần nhất - số ngày):**

* Điểm 1: Từ 0 đến Q1 (phân vị thứ nhất) – Khách hàng đã lâu không mua hàng.
* Điểm 2: Từ Q1 đến Q2 (phân vị thứ hai) – Khách hàng mua hàng cách đây một thời gian.
* Điểm 3: Từ Q2 đến Q3 (phân vị thứ ba) – Khách hàng mua hàng gần đây.
* Điểm 4: Từ Q3 trở đi (trên phân vị thứ ba) – Khách hàng vừa mới mua hàng.

**Frequency (Số lần mua hàng trong 1 tháng):**

* Điểm 1: Từ 0 đến Q1 (phân vị thứ nhất) – Khách hàng ít mua hàng.
* Điểm 2: Từ Q1 đến Q2 (phân vị thứ hai) – Khách hàng có tần suất mua hàng trung bình.
* Điểm 3: Từ Q2 đến Q3 (phân vị thứ ba) – Khách hàng mua hàng nhiều lần.
* Điểm 4: Từ Q3 trở đi (trên phân vị thứ ba) – Khách hàng mua hàng rất thường xuyên.

**Monetary (Chi tiêu hàng tháng):**

* Điểm 1: Từ 0 đến Q1 – Khách hàng chi tiêu thấp nhất.
* Điểm 2: Từ Q1 đến Q2 – Khách hàng có mức chi tiêu thấp trung bình.
* Điểm 3: Từ Q2 đến Q3 – Khách hàng có mức chi tiêu cao trung bình.
* Điểm 4: Từ Q3 trở đi – Khách hàng chi tiêu cao nhất.

Phương pháp phân loại này giúp Công ty X dễ dàng nhận diện và tập trung vào các nhóm khách hàng có giá trị cao nhất, từ đó xây dựng các chiến lược phù hợp để giữ chân và phát triển mối quan hệ với họ.

1. **BCG Matrix**

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.



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.

1. **Xử lý dữ liệu và tính toán:**
2. **Mô tả dữ liệu**

Table 1 Customer\_Transaction: Chi tiết các giao dịch cuả khách hàng.

|  |  |  |
| --- | --- | --- |
| **STT** | **Trường** | **Mô tả** |
| 1 | ID | TransactionID |
| 2 | CustomerID | Mã khách hàng |
| 3 | PurchaseDate | Ngày mua hàng |
| 4 | GMV | Gross Money Value |

A screenshot of a computer

Description automatically generated

Figure 1 Customer\_Transaction: Chi tiết các giao dịch cuả khách hàng.

Table 2 Customer\_Registered: Thông tin chi tiết đăng ký của khách hàng

|  |  |  |
| --- | --- | --- |
| **STT** | **Trường** | **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 |

A screenshot of a computer

Description automatically generated

Figure 2 Customer\_Registered: Thông tin chi tiết đăng ký của khách hàng

1. **Xử lý dữ liệu:**

Biến đổi cột stopdate thành dạng “YYYY-mm-dd hh:mm:ss”

Xóa dữ liệu có stopdate khác Null vì đó là những khách hàng đã ngưng sử dụng.

Nối 2 bảng Customer\_Transaction và Customer\_Registered thông qua cột CustomerID và ID.

**Tính toán:**

* Tổng doanh số trên từng khách hàng: total\_order\_value
* Tổng doanh số theo từng tháng: n\_GMV\_per\_month
* Tần suất mua hàng theo tháng: n\_FREQUENCY\_per\_month
* Trung bình giá trị mỗi lần mua: average\_order\_value
* Số tháng kể từ ngày tạo khách hàng: Months\_Customer\_Lifetime
* Số ngày kể từ lần mua hàng gần nhất: days\_from\_latest\_purchase

A screenshot of a computer program

Description automatically generated

Figure 3 Xử lý và tính toán dữ liệu

* A computer screen shot of text

  Description automatically generatedCác chỉ số R, F, M tính tới thời điểm '2022-09-02'

Figure 4 Tạo cột dữ liệu R, F, M

1. **Phân nhóm khách hàng theo RFM**

Table 3 Phân nhóm khách hàng theo RFM

|  |  |  |
| --- | --- | --- |
| **Cus\_cluster** | **Description** | **RFM** |
| VIP Customers | Tần suất cao, doanh thu trên tháng cao, khách mua gần đây | '444', '443', '434', '433', '344', '343', '334', '333' |
| At-Risk VIP Customers | Tần suất cao, doanh thu trên tháng cao, khách mua lâu rồi | '144', '143', '134', '133', '244', '243', '234', '233' |
| Regular Loyal Customers | Tần suất cao, doanh thu trên tháng thấp, khách mua gần đây | '441', '442', '431', '432', '341', '342', '331', '332' |
| At-Risk Regular Customers | Tần suất cao, doanh thu trên tháng thấp, khách mua lâu rồi | '141', '142', '131', '132', '241', '242', '231', '232' |
| Occasional VIP Customers | Tần suất thấp, doanh thu trên tháng cao, khách mua gần đây | '414', '413', '424', '423', '314', '313', '324', '323' |
| At-Risk Occasional VIP Customers | Tần suất thấp, doanh thu trên tháng cao, khách mua lâu rồi | '114', '113', '124', '123', '214', '213', '224', '223' |
| New or Casual Customers | Tần suất thấp, doanh thu trên tháng thấp, khách mua gần đây | '411', '412', '421', '422', '311', '312', '321', '322' |
| Lost Customers | Tần suất thấp, doanh thu trên tháng thấp, khách mua lâu rồi | '111', '112', '121', '122', '211', '212', '221', '222' |

1. **Phân tích dữ liệu**
2. **Tổng quan**

Tập dữ liệu gồm **931 nghìn khách hàng** với tổng doanh thu hằng tháng đạt **7,27 tỷ đồng**. Trung bình mỗi khách hàng sẽ chi tiêu **88 nghìn đồng** với khoảng **90,8 nghìn lượt mua hàng** mỗi tháng. Trung bình mỗi khách hàng đã mua 1,11 lần và thời gian trung bình khách hàng gắn bó với công ty là **38 tháng** tính đến thời điểm “02-09-2022”

A close-up of a number

Description automatically generated

Figure 5 Tổng quan dữ liệu

1. **Phân tích theo từng phân nhóm khách hàng**

A screenshot of a graph

Description automatically generated

Figure 6 Biểu đồ phân tích theo nhóm khách hàng

A screenshot of a graph

Description automatically generated

Figure 7 Biểu đồ số lượng khách hàng theo từng nhóm

Nhìn chung các nhóm khách hàng chủ yếu của công ty X là Lost Customers, VIP Customers, At-Risk VIP Customers, New or Casual Customers chiếm tới 88% tổng lượt khách hàng của công ty X.

Lượng khách hàng At-Risk (đã lâu không mua hàng) và mua hàng gần đây khá là tương đồng nhau. Đồng thời nhóm khách hàng Lost Customers rất cao có thể phản ánh rằng công ty không giữ chân khách hàng tốt hoặc ngành hàng bán lẻ là ngành hàng có tỉ lệ rời bỏ cao bởi xu hướng thích sự đổi mới, tiện lợi và ngẫu hứng của khách hàng.

A graph of a number of blue bars

Description automatically generated with medium confidence

Figure 8 Biểu đồ trung bình doanh thu trên người theo từng nhóm nhóm khách hàng

Trung bình giá trị mỗi khách hàng mang lại ở từng nhóm khách hàng là khá tương đồng với nhau. Cao nhất là nhóm khách hàng Occasional VIP Customers với 115k/ người và thấp nhất là At-Risk Regular Customers với 57k/ khách hàng. Bởi nhóm Occasional VIP Customer mua hàng với tần suất thấp và số lượng lớn, doanh thu nhiều nên việc giá trị đơn hàng của họ cao hoàn toàn dễ hiểu, họ có thể là nhà phân phối nhỏ lẻ hoặc doanh nghiệp. Điều tương tự cũng xảy ra với At-Risk Regular Customers.

A close-up of a card

Description automatically generated

Figure 9 Biểu đồ tổng doanh thu (bên trái) và doanh thu theo tháng của từng nhóm khách hàng

**Xét về tổng doanh thu mang lại:**

Nhóm VIP customer và At-risk Customers chiếm tới gần **50% tổng doanh số** từ khi công ty hoạt động cho đến nay.

Tổng doanh số mang lại của kênh khách hành New or Casual Customers và Lost customers cũng rất cần được để tâm khi chiếm gần **40% doanh thu công ty.**

Vấn đề cần phải quan tâm ở đây là nhóm Lost customers khi họ chiếm tới gần **20,77% doanh thu công ty**. Vậy điều này có đáng quan ngại không?

**Xét về doanh thu trung bình theo tháng:**

Quan sát biểu đồ doanh thu theo tháng thì nhóm Lost Customers chỉ đống góp **300 triệu đồng** trên tháng tức là **4% doanh thu hằng tháng**. Điều này chứng tỏ nhóm này không còn tác động đáng kể đến doanh thu cửa hàng trong những tháng gần đây do đó vấn đề này không quá nghiêm trọng.

Vấn đề tiếp theo nằm ở việc nhóm **At-Risk VIP customers và VIP customers** chiếm tỉ trọng quá lớn trong tổng doanh thu hằng tháng. Do đó công ty cần phải chăm sóc đặc biệt cẩn trọng đối với 2 nhóm khách hàng này. Bên cạnh đó nhóm **At-Risk VIP customers** đang là nhóm chiếm tới **47% doanh thu tháng** nên ta cần xác định nguyên nhân tại sao lâu rồi họ không mua hàng để kịp thời chăm sóc và tránh làm mất khách hàng. Nếu không công ty sẽ thiệt hại rất đáng kể.

**Một điểm sáng** đó là tổng doanh thu của **nhóm New or Casual Customers** khá ổn định chứng tỏ công ty vẫn có khả năng thu hút khách hàng mới hoặc khách vãng lai. Đây cũng là kênh tiềm năng để tung các chương trình quảng cáo nhằm tăng doanh thu cho công ty.

**Đi sâu hơn và vấn đề cần giải quyết: nhóm khách hàng At-Risk VIP customers**

A screenshot of a computer

Description automatically generated

Figure 10 Thông tin chi tiết các nhóm RFM trong nhóm At-Risk VIP customers

Có thể thấy phần lớn **doanh thu** tập trung ở **nhóm RFM 144 và 244** đó là các nhóm khách hàng đã lâu không mua hàng nhưng trước đó họ mua thường xuyên và doanh thu lớn. Họ chiếm tới **43% tổng doanh thu hằng tháng** của công ty. Nếu phát hiện được nguyên nhân khiến họ không hài lòng sẽ giúp ích đáng kể cho công ty.

1. **Phân bổ doanh thu theo từng năm tuổi của khách hàng tính đến ngày ‘2022-09-02’:**

A screenshot of a computer screen

Description automatically generated

Figure 11 Tree map doanh thu theo năm tuổi tính từ ngày nhập thông tin khách hàng lên hệ thống

**Về phân bổ doanh thu theo từng năm tuổi của khách hàng tính đến ngày ‘2022-09-02’:**

* Nhóm khách hàng từ 0 đến 1 năm chiếm **19 tỷ đồng với 23% tổng doanh thu**
* Các nhóm khách hàng 2,3,4 năm tuổi khá tương đồng nhau về doanh thu đều là **12, 13 tỷ đồng** chiếm **15-16% tổng doanh thu.**
* Có thể thấy công ty nên tập trung chăm sóc vào giai đoạn đầu của tuổi hợp đồng khách hàng. Thời gian khách hàng gắn bó với công ty càng lâu thì số tiền họ chi tiêu có xu hướng giảm dần.

**Một số gợi ý triển khai chiến dịch:**

* **Ưu Đãi Cao Tháng Đầu:** Cung cấp giảm giá đặc biệt hoặc sản phẩm miễn phí trong tháng đầu để khuyến khích mua sắm sớm.
* **Hỗ Trợ Cá Nhân Hóa:** Tư vấn và gợi ý sản phẩm phù hợp với nhu cầu khách hàng ngay từ đầu để tăng chi tiêu.
* **Lợi Ích Tăng Dần:** Triển khai ưu đãi tăng theo tháng, khuyến khích khách hàng mua nhiều hơn trong giai đoạn đầu.

1. **Pie and Donut chart phân bổ doanh thu theo từng điểm số R, F, M**

A group of pie charts

Description automatically generated

Figure 12 Pie and Donut chart phân bổ doanh thu theo từng điểm số R, F, M

Donut chart (3 cái nằm trên) M, F đều thể hiện sự vượt trội trong % doanh thu của mức điểm 4 trong khi bên R thì những khách hàng mua gần đây tức thang điểm 3,4 mang lại hiệu quả khá tốt với **50% tổng doanh thu công ty**. Điều này chứng tỏ công ty đang làm ăn tốt bởi những khách hàng mua đã lâu thì trước đó họ đã có thời gian tích lũy mua hàng rất lớn nên % tổng doanh số đóng góp cho công ty cao.

3 pie chart phía dưới là % doanh thu theo tháng của các mức điểm R, F, M và nó rất cân bằng với nhau ở cả 4 mức điểm trong cả 3 biểu đồ.

1. **Boxplot chart R, F, M**

A screenshot of a diagram

Description automatically generated

Figure 13 Boxplot chart R, F, M

**Số ngày kể từ lần mua cuối cùng** phân bố với tần số khá đều nhau với **25% dữ liệu dưới** mức **25 ngày** và **75% dữ liệu dưới** mức **71 ngày.**

**Tần số mua hàng trên tháng của khách hàng** có một số điểm ngoại lai tuy nhiên cũng phân bổ khá tốt với **50% dữ liệu trải dài từ 0.12 lần đến 0.55 lần**

**Doanh số trên tháng** thì rất nhiều điểm ngoại lai và các điểm ngoại lai này mang lại doanh thu cho công ty rất rất lớn so với các khách hàng còn lại. Đây vừa là thuận lợi vừa là thách thức với công ty bởi lẽ cần phải giữ chân nhóm khách hàng này, chăm sóc họ thật tốt, tìm ra cách khiến họ chi tiền mạnh từ đó áp dụng cho các nhóm khách hàng khác để tăng doanh thu. Nếu họ rời bỏ công ty thì sẽ gây thiệt hại không thể đo đếm được và tệ hơn nữa khi họ chọn đối thủ của công ty X chúng ta là đối tác mới.