

# RFM Analysis

Phân tích hành vi khách hàng bằng RFM model bởi Python tool

Do Thi Minh Thuy

 [Project on my Github](#)



# Agenda

## Mục tiêu của project

Thấu hiểu nhu cầu của stakeholder

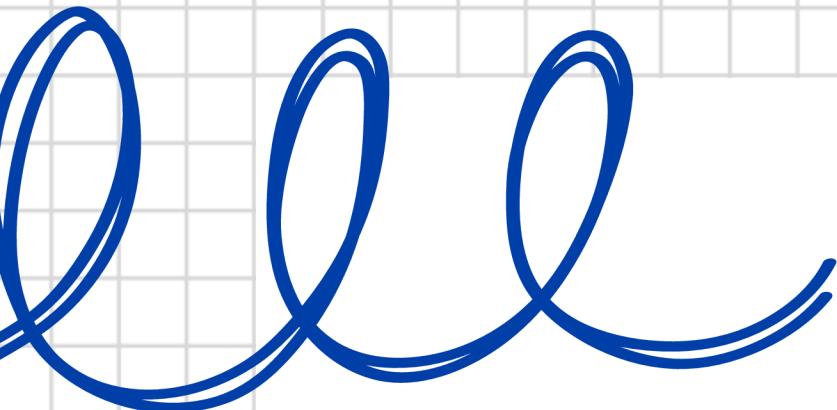
## Tư duy máy tính

Áp dụng Computational Thinking để đưa ra hướng tính toán các chỉ số Recency, Frequency, Monetary

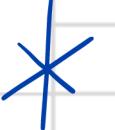
## Giải pháp

Từ bài phân tích bằng Python tool, đưa ra các giải pháp cho doanh nghiệp, cụ thể là team Marketing/ Sales



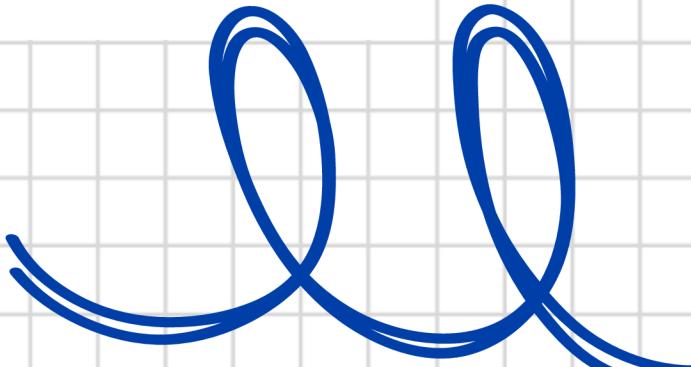


# Stakeholder đang cần gì?



**Case study:** Team Marketing/ Sales/ CS thường phải tự thực hiện các phân tích đơn giản để tìm ra hành vi khách hàng, từ đó có thể đưa ra các chiến dịch phù hợp. Tuy nhiên, vì quy mô doanh nghiệp ngày càng mở rộng và lượng khách hàng ngày càng tăng, stakeholders không thể tự xử lý các data này, nên cần DA có thể phân tích và phân loại khách hàng dựa vào các đặc điểm chung để có thể thực hiện các action phù hợp với từng phân khúc khách hàng.

**Thấu hiểu nhu cầu của stakeholder và bằng data doanh nghiệp đã thu nhập trong các năm qua, Data Analyst cần cung cấp phân tích Customer Behavior thông qua RFM Model và đưa ra được các Recommendation phù hợp cho stakeholders.**



# Computational Thinking

## Phân tách (Decomposition)

Phân tách vấn đề lớn thành các vấn đề nhỏ:

- RFM model bao gồm 3 chỉ số là Recency, Frequency, Monetary. Từ đó phân loại khách hàng thành các tệp
- Các steps thực hiện: import data - clean data - tính toán RFM score - visualization - recommendation

## Nhận dạng quy luật (Pattern Recognition)

Giải quyết các vấn đề mới bằng cách nhận dạng các quy luật chung với vấn đề cũ:

- Import data: import các data được chọn lọc trước đó để sử dụng tính toán RFM score
- Clean data: áp dụng EDA, check null, dtype, outliers,...
- Tính toán RFM score: xử lý tính toán group by ID của khách hàng,...
- Visualization: display segment
- Recommendation: đưa ra các giải pháp dựa vào đặc điểm của từng segment

## Trừu tượng hóa (Abstraction)

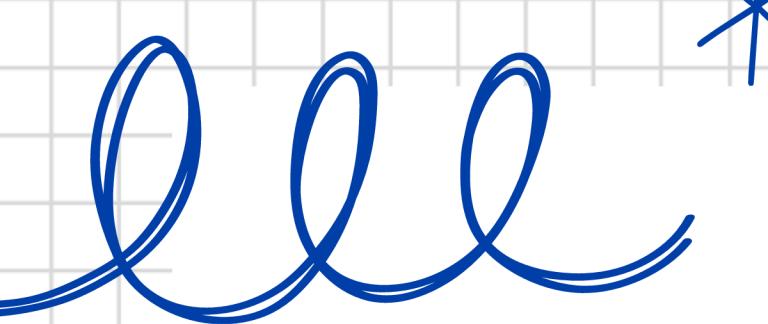
Tìm ra giải pháp chung để tính toán RFM score:

- Sử dụng qcut() để xác định điểm trong thang điểm 5 của 3 chỉ số R, F, M
- Cả 3 chỉ số đều sử dụng groupby theo CustomerID

## Tư duy thuật toán (Algorithm Design)

Xác định cách tính toán RFM score:

- Recency: số ngày kể từ ngày mua đơn hàng gần nhất -> càng lớn R score càng thấp
- Frequency: số đơn hàng đã mua được -> càng lớn F score càng cao
- Monetary: doanh thu từ khách hàng -> càng lớn M score càng cao
- RFM score: ghép 3 R,F,M score lại (ở dạng str)



# Tình trạng & Giải pháp

Nhóm khách hàng	Tình trạng	Giải pháp
Champions	<ul style="list-style-type: none"><li>- Đây là tệp khách hàng có số lượng khách hàng cao nhất trong tổng số khách hàng của doanh nghiệp (gần 800 khách hàng) nhưng tỉ trọng chiếm phần lớn trong tổng doanh số của doanh nghiệp (cao gấp 6 lần nhóm thứ 2).</li><li>- Đây là tệp khách hàng được ưu tiên hàng đầu của doanh nghiệp, với cả 3 chỉ số RFM đều cao nhất và không có chỉ số nào dưới 4. Thường đây là lượng khách hàng trung thành với thương hiệu; các đại lý, bán sỉ và bán lẻ; đối tác làm ăn quen thuộc với doanh nghiệp.</li></ul>	<ul style="list-style-type: none"><li>- Việc Champions có chỉ số Average Revenue Per User (ARPU) cao nhất trong các tệp khách hàng đang thể hiện tín hiệu rất lạc quan đối với doanh nghiệp. Đây là tệp khách hàng cần được tiếp tục care mạnh và mở rộng vì chiếm tỉ trọng doanh số rất cao trong doanh nghiệp.</li><li>- Với số lượng khách hàng lớn và đã có thông tin khách hàng trong data, việc Marketing đến tệp khách hàng này dễ truyền tải đến đúng đối tượng. Team Marketing/Sale nên tập trung các chương trình khuyến mãi/ chăm sóc khách hàng trung thành, các đặc quyền riêng với nhóm khách hàng này và có động thái hỏi thăm ngay khi thấy các khách hàng trong tệp này có dấu hiệu giảm 1 trong 3 chỉ số RFM.</li></ul>
Loyal	<ul style="list-style-type: none"><li>- Là nhóm đứng từ 4 trong tổng số các khách hàng (hơn 400 khách hàng), nhưng đây là nhóm chiếm tỉ trọng cao thứ 2 sau Champions trong tổng doanh số của doanh nghiệp.</li><li>- Đây là nhóm khách hàng trung thành của thương hiệu với 3 chỉ số cao hơn trung bình.</li></ul>	<ul style="list-style-type: none"><li>- Đây là nhóm khách hàng đã quen thuộc với thương hiệu và có vai trò PR cho doanh nghiệp rất tốt.</li><li>- Team Marketing/ CS cần có chính sách chăm sóc khách hàng tốt sẽ níu giữ sự trung thành của nhóm khách hàng này. Đây là nhóm khách hàng cần duy trì và tập trung mở rộng.</li></ul>



# Tình trạng & Giải pháp

Nhóm khách hàng	Tình trạng	Giải pháp
Hibernating Customer	<ul style="list-style-type: none"><li>- Đây là tệp khách hàng "ngủ đông", với cả 3 chỉ số đều từ trung bình trở xuống. Số khách hàng trong nhóm này đang chiếm tỉ trọng khá cao trong doanh nghiệp, cao thứ nhì sau Champions (gần 700 khách hàng). Tỉ trọng doanh số của nhóm khách hàng này khá thấp, đứng thứ 5 trong các nhóm khách hàng.</li><li>- Thường nhóm khách hàng này là khách hàng đã ngưng sử dụng sản phẩm của doanh nghiệp trong một thời gian hoặc các khách hàng trải nghiệm sản phẩm.</li></ul>	<ul style="list-style-type: none"><li>- Tuy không gánh nhiều doanh số cho doanh nghiệp nhưng đây là tệp khách hàng phản ánh lên các phản hồi tiêu cực từ khách hàng từng sử dụng sản phẩm.</li><li>- Nhóm Marketing/ CS/ Sale nên tìm hiểu và thu thập reviews từ nhóm khách hàng này, đào sâu nguyên nhân ngưng sử dụng sản phẩm hay do đối thủ đã làm tốt hơn doanh nghiệp ở khâu quảng bá/ chăm sóc khách hàng? Từ đó đưa ra các kế hoạch phát triển sản phẩm/ marketing phù hợp.</li></ul>
Lost Customer	<ul style="list-style-type: none"><li>- Tương tự với Hibernating Customer, đây là nhóm khách hàng đã rời bỏ doanh nghiệp. Nhóm khách hàng này đứng thứ 3 trong tổng số khách hàng, "gánh" doanh số không cao.</li><li>- Đây cũng thường là nhóm khách hàng trải nghiệm sản phẩm.</li></ul>	



# Tình trạng & Giải pháp

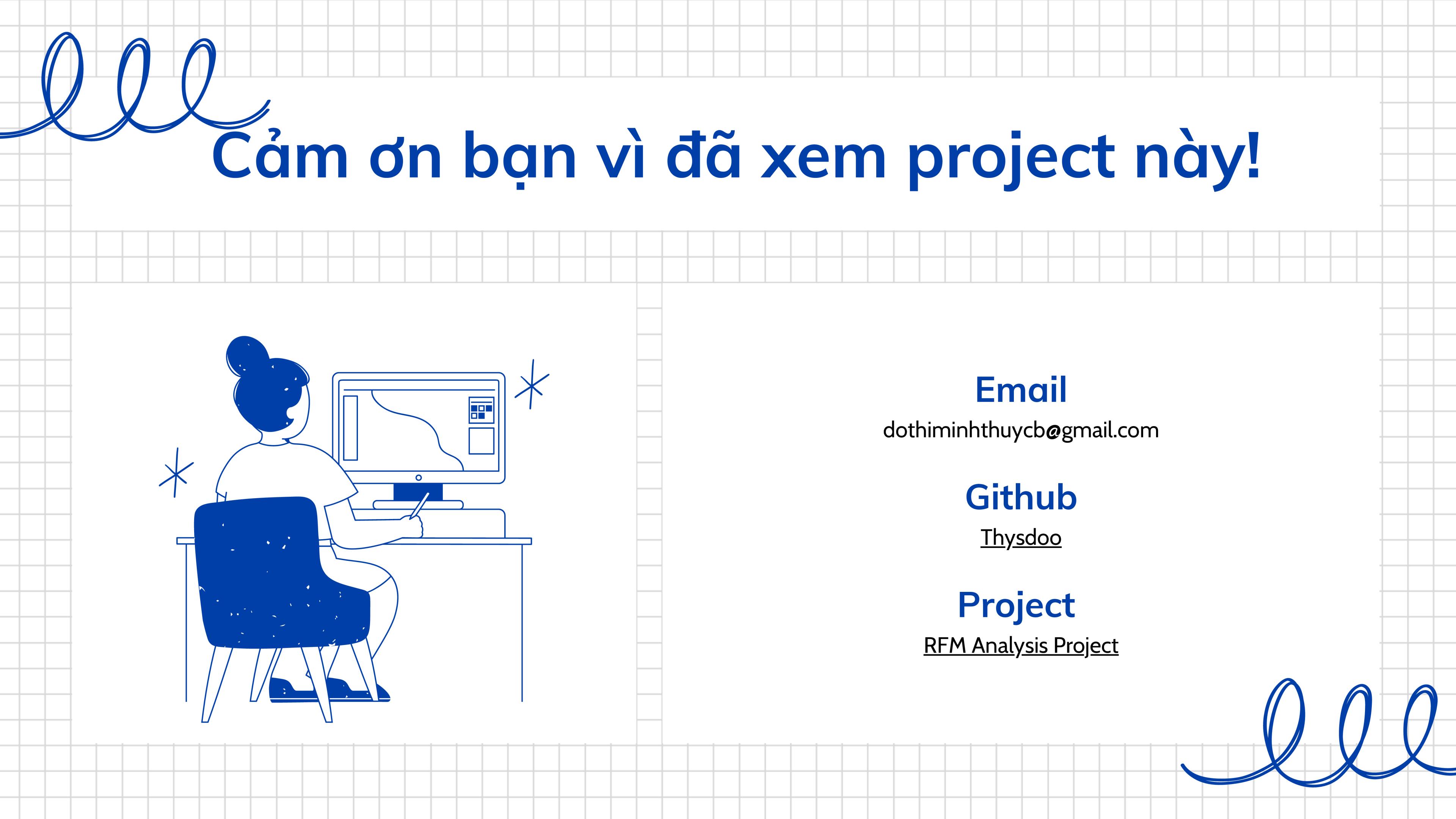
Nhóm khách hàng	Tình trạng	Giải pháp
At Risk	<ul style="list-style-type: none"><li>- Đây là nhóm khách hàng đông thứ 5 (gần 400 khách hàng) trong các nhóm nhưng tỉ trọng doanh số đứng thứ 3 và không quá kém cạnh Loyal</li><li>- Nhóm khách hàng này đã mua hàng thường xuyên và chi tiêu một số tiền lớn nhưng không mua hàng gần đây. Điều này cho thấy khả năng họ rời bỏ doanh nghiệp cao, có thể do họ không còn nhu cầu với sản phẩm hoặc tệ hơn là họ đã tìm ra sản phẩm thay thế, hoặc chuyển sang ủng hộ đối thủ.</li></ul>	<ul style="list-style-type: none"><li>- Với nhóm khách hàng này, doanh nghiệp cần gửi cho họ các chiến dịch tái kích hoạt cá nhân hóa để tái thiết lập kết nối và tư vấn, chăm sóc để cung cấp cho họ các sản phẩm phù hợp để khuyến khích mua hàng tiếp theo.</li></ul>



# Tổng kết

Nhìn chung, nhóm khách hàng Champions với cả 3 chỉ số R,F,M đều cao hơn trung bình đang chiếm tỉ lệ cao nhất phản ánh **tình hình kinh doanh tích cực của doanh nghiệp**. Tuy nhiên, ngoài việc tập trung vào nhóm khách hàng “gánh” doanh số cao này, doanh nghiệp cũng nên phân bổ nguồn lực vào thu hút khách hàng mới bằng cách nghiên cứu lại sản phẩm/dịch vụ của mình thông qua các phản hồi của tệp khách hàng Hibernating Customer và Lost Customer, vì 2 tệp khách hàng này cũng chiếm tỉ trọng cao không kém Champions trong tổng số lượng khách hàng. Việc cạnh tranh đường dài của doanh nghiệp không chỉ tập trung vào doanh số vì việc lắng nghe đánh giá tiêu cực của khách hàng, nghiên cứu phát triển sản phẩm và cải thiện chất lượng dịch vụ là yếu tố cốt lõi giúp doanh nghiệp có được niềm tin từ khách hàng.





# Cảm ơn bạn vì đã xem project này!



## Email

[dothiminhthuycb@gmail.com](mailto:dothiminhthuycb@gmail.com)

## Github

[Thysdoo](#)

## Project

[RFM Analysis Project](#)

