

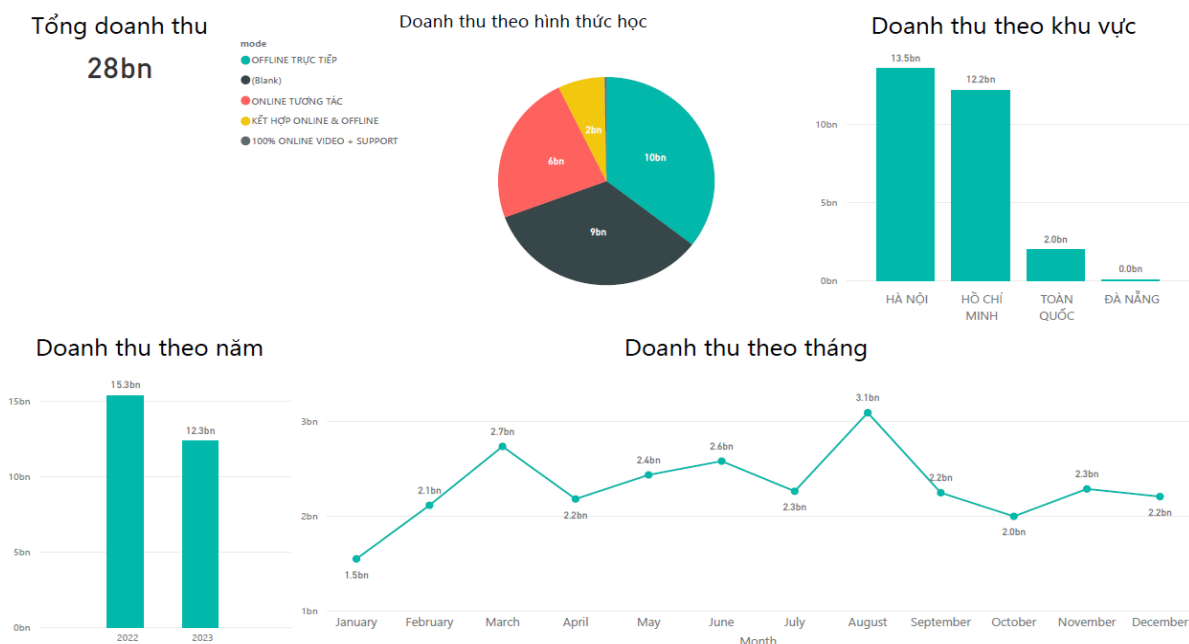
Phần I: Xử lý dữ liệu

Dữ liệu mẫu đã cho có nhiều lỗi

- Cột `cus_mobi` dữ liệu được thêm vào không theo quy tắc, dẫn đến sự không đồng nhất trong dữ liệu cần phải làm sạch và đưa về cùng 1 dạng trước khi phân tích.
- Cột `course_mode_id` bị thiếu rất nhiều dữ liệu, 1/3 số dòng không có giá trị id nên chúng ta không thể xóa những dòng bị thiếu như vậy sẽ gây nên sai lệch lớn trong quá trình phân tích. Giải pháp ở đây là nên thay những dòng “blank” thành giá trị 0.
- Cột `cus_location_id` cũng có 1 số giá trị “blank” nhưng không đáng kể nên ta có thể thực hiện xóa hoàn toàn những ô đó để tiện cho công việc phân tích về sau.

Phần II: Phân tích dữ liệu

1. Xây dựng dashboard và đưa ra nhận xét, phân tích chung về tình hình kinh doanh



- Có sự sụt giảm khá rõ trong doanh thu của năm 2023 so với 2022.
- Theo những biểu đồ trên ta có thể thấy tháng 3 và tháng 8 có doanh thu vượt trội so với các tháng còn lại .
- Phần lớn doanh thu đến từ những học viên từ Hà Nội và Hồ Chí Minh, đa phần đều học theo hình thức offline.

Year	Month	Sum of total_fee	Revenue MoM%	
2022	January	671M		
2022	February	963M	↑	43.46%
2022	March	1154M	↑	19.78%
2022	April	1103M	↓	-4.44%
2022	May	1254M	↑	13.70%
2022	June	1674M	↑	33.53%
2022	July	1328M	↓	-20.64%
2022	August	1725M	↑	29.84%
2022	September	1373M	↓	-20.42%
2022	October	1248M	↓	-9.08%
2022	November	1524M	↑	22.12%
2022	December	1322M	↓	-13.26%
2023	January	873M	↓	-33.94%
2023	February	1148M	↑	31.47%
2023	March	1581M	↑	37.65%
2023	April	1076M	↓	-31.94%
2023	May	1180M	↑	9.69%
2023	June	905M	↓	-23.29%
2023	July	932M	↑	2.99%
2023	August	1367M	↑	46.62%
2023	September	872M	↓	-36.20%
2023	October	747M	↓	-14.35%
2023	November	761M	↑	1.86%
2023	December	882M	↑	15.94%
Total		27661M		3.29%

- Tốc độ tăng trưởng theo tháng là 3,29%. Trong kỳ báo cáo, 50% số tháng ghi nhận sự suy giảm doanh số so với các tháng trước, trong khi 13 tháng còn lại cho thấy xu hướng tăng trưởng doanh số so với các giai đoạn trước đó.

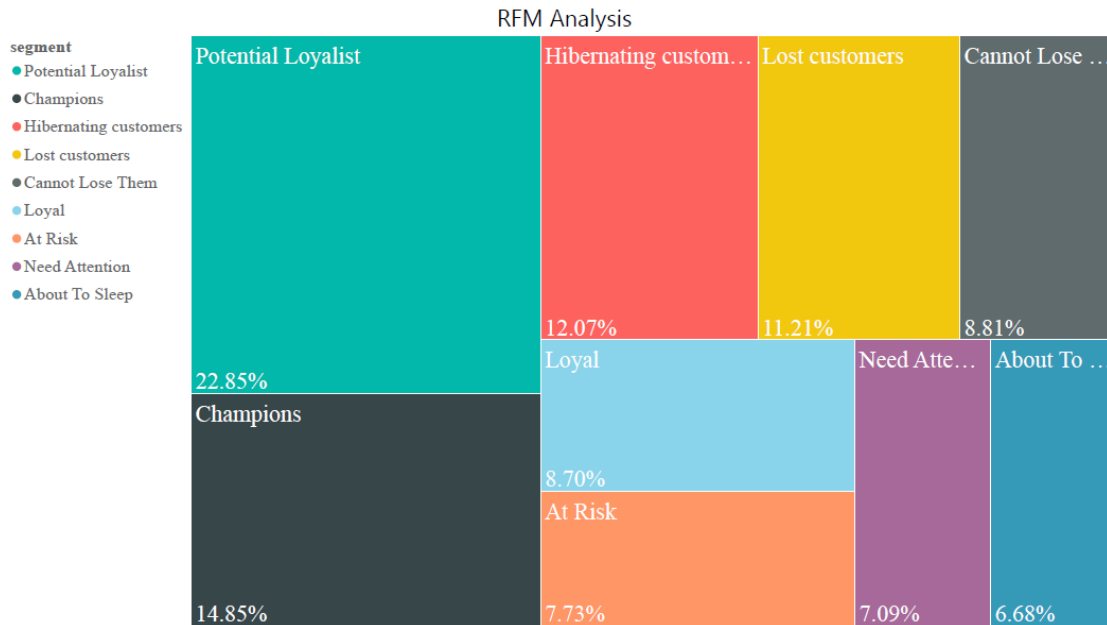
2. Sử dụng mô hình RFM để phân khúc khách hàng

- Xây dựng mô hình RFM bằng Sql bằng cách phân khúc khách hàng thành 11 cấp độ cùng với cách tính điểm như trong bảng sau

Customer Segment	Activity	Actionable Tip
Champions	Bought recently, buy often and spend the most!	Reward them. Can be early adopters for new products. Will promote your brand.
Loyal Customers	Spend good money with us often. Responsive to promotions.	Upsell higher value products. Ask for reviews. Engage them.
Potential Loyalist	Recent customers, but spent a good amount and bought more than once.	Offer membership / loyalty program, recommend other products.
Recent Customers	Bought most recently, but not often.	Provide on-boarding support, give them early success, start building relationship.
Promising	Recent shoppers, but haven't spent much.	Create brand awareness, offer free trials
Customers Needing Attention	Above average recency, frequency and monetary values. May not have bought very recently though.	Make limited time offers, Recommend based on past purchases. Reactivate them.
About To Sleep	Below average recency, frequency and monetary values. Will lose them if not reactivated.	Share valuable resources, recommend popular products / renewals at discount, reconnect with them.
At Risk	Spent big money and purchased often. But long time ago. Need to bring them back!	Send personalized emails to reconnect, offer renewals, provide helpful resources.
Can't Lose Them	Made biggest purchases, and often. But haven't returned for a long time.	Win them back via renewals or newer products, don't lose them to competition, talk to them.
Hibernating	Last purchase was long back, low spenders and low number of orders.	Offer other relevant products and special discounts. Recreate brand value.
Lost	Lowest recency, frequency and monetary scores.	Revive interest with reach out campaign, ignore otherwise.

Segment	Scores
Champions	555, 554, 544, 545, 454, 455, 445
Loyal	543, 444, 435, 355, 354, 345, 344, 335
Potential Loyalist	553, 551, 552, 541, 542, 533, 532, 531, 452, 451, 442, 441, 431, 453, 433, 432, 423, 353, 352, 351, 342, 341, 333, 323
New Customers	512, 511, 422, 421, 412, 411, 311
Promising	525, 524, 523, 522, 521, 515, 514, 513, 425, 424, 413, 414, 415, 315, 314, 313
Need Attention	535, 534, 443, 434, 343, 334, 325, 324
About To Sleep	331, 321, 312, 221, 213, 231, 241, 251
At Risk	255, 254, 245, 244, 253, 252, 243, 242, 235, 234, 225, 224, 153, 152, 145, 143, 142, 135, 134, 133, 125, 124
Cannot Lose Them	155, 154, 144, 214, 215, 115, 114, 113
Hibernating customers	332, 322, 231, 241, 251, 233, 232, 223, 222, 132, 123, 122, 212, 211
Lost customers	111, 112, 121, 131, 141, 151

- Sau khi thực hiện tính toán trong Sql, nhập dữ liệu vào Power BI và thực hiện trực quan hóa dữ liệu



- Phân khúc khách hàng hàng đầu (Champions và Cannot lose them): Cần duy trì sự trung thành và mức chi tiêu lớn. Họ là doanh thu chính và giúp quảng bá thương hiệu
- Khách hàng tiềm năng và gần đây (Potential Loyalist & một phần của Loyal): Xây dựng mối quan hệ lâu dài bằng ưu đãi (free trials, dịch vụ đi kèm, ...) để họ có thể trở thành khách hàng hàng đầu trong tương lai.
- Khách hàng cần quan tâm (Need Attention & About To Sleep): Kết nối lại bằng các ưu đãi, gợi ý mua hàng dựa trên lịch sử và thu thập thông tin về những khó khăn hay những điều không hài lòng khi sử dụng sản phẩm.

- Khách hàng gần như đã mất và rời bỏ (At Risk, Need Attention mức độ cao hơn, Lost Customers): gợi ý sản phẩm phù hợp, thu thập phản hồi để cải thiện sản phẩm.