

Báo Cáo Phân Tích Dữ Liệu Online Retail Dataset

Thực hiện: Lương Đức Thắng

Tóm tắt nội dung

Báo cáo phân tích tập dữ liệu *Online Retail* nhằm đánh giá hành vi mua hàng, hiệu suất sản phẩm và sự khác biệt giữa các thị trường. Nội dung tập trung vào việc nhận diện các xu hướng chính, xác định nhóm khách hàng có giá trị cao, phân tích sản phẩm chủ lực và đưa ra những đề xuất chiến lược dựa trên kết quả quan sát. Kết quả phân tích cho thấy doanh thu mang tính mùa vụ mạnh, nhóm khách hàng VIP đóng góp phần lớn doanh thu và một số thị trường châu Âu có tiềm năng tăng trưởng rõ rệt.

1 Mở đầu

Dữ liệu bán lẻ trực tuyến đang trở thành nền tảng quan trọng trong việc hiểu hành vi người tiêu dùng và tối ưu chiến lược kinh doanh. Tuy nhiên, không phải doanh nghiệp nào cũng khai thác đầy đủ giá trị từ nguồn dữ liệu này. Tập dữ liệu *Online Retail* phản ánh chi tiết các giao dịch trong giai đoạn gần hai năm, bao gồm mã sản phẩm, khách hàng, đơn giá và quốc gia phát sinh giao dịch. Việc phân tích giúp nhận diện xu hướng mua sắm, sự khác biệt giữa các nhóm khách hàng và thị trường, từ đó xây dựng chiến lược nhằm cải thiện doanh thu và nâng cao hiệu quả vận hành.

Mục tiêu phân tích:

- Khám phá xu hướng doanh thu theo thời gian.
- Đánh giá mức độ đóng góp doanh thu theo quốc gia.
- Nhận diện các phân khúc khách hàng dựa trên hành vi.
- Xác định sản phẩm mang lại doanh thu cao.
- Đề xuất chiến lược kinh doanh với nền tảng từ dữ liệu.

2 Mô tả dữ liệu

Bộ dữ liệu bao gồm các giao dịch được ghi nhận trong giai đoạn:

- **01/01/2010 – 09/12/2011**

Tổng quan:

- **Tổng doanh thu:** 15.99M GBP.
- **Số lượng bán:** 10M đơn vị.
- **Đơn hàng:** 42.98K hóa đơn.
- **Khách hàng:** 5,842.
- **Sản phẩm:** 5,157.
- **Giá trung bình:** 3.69 GBP.

Dữ liệu được chia thành 4 nhóm:

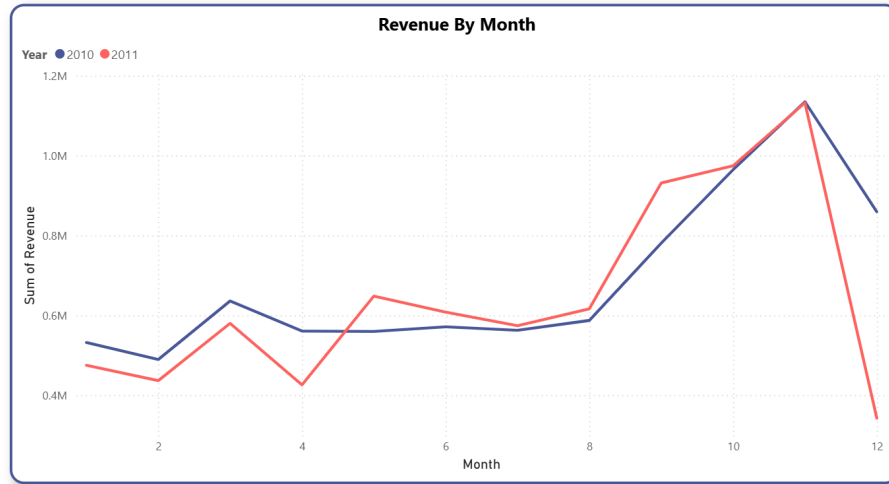
1. Thông tin giao dịch.
2. Danh mục sản phẩm.
3. Hồ sơ khách hàng.
4. Quốc gia phát sinh đơn hàng.

3 Phân tích thống kê mô tả

3.1 Doanh thu theo tháng

Kết quả phân tích theo từng tháng cho thấy doanh thu tăng mạnh từ tháng 9 và đạt mức cao nhất trong giai đoạn tháng 10–12. Đây là đặc điểm điển hình trong thương mại điện tử khi nhu cầu mua sắm cuối năm tăng lên, đặc biệt trong các dịp lễ hội. Mức doanh thu của năm 2011 cũng vượt trội hơn so với năm 2010, phản ánh sự mở rộng của thị trường và nhu cầu ngày càng cao.

Quan sát này khẳng định tầm quan trọng của việc chuẩn bị trước các chiến dịch marketing, tối ưu tồn kho và nâng cao năng lực vận hành trước mùa cao điểm.



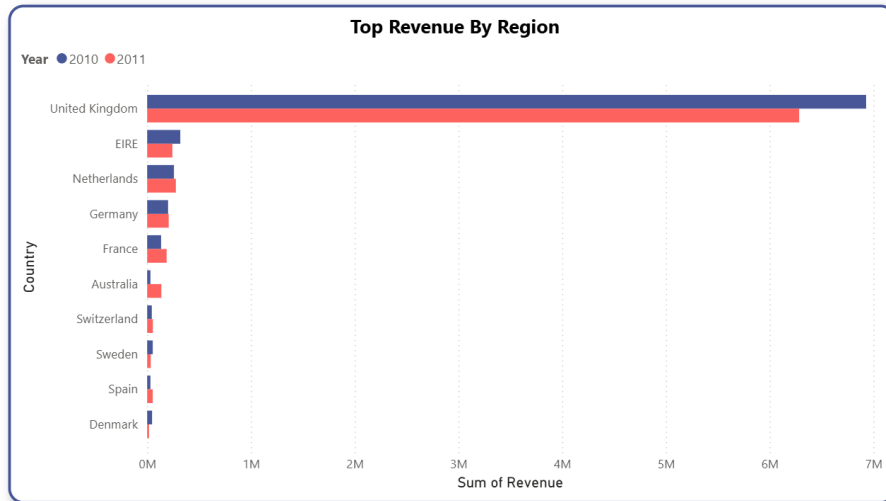
Hình 1: Doanh thu theo tháng (2010–2011)

3.2 Phân tích theo vùng địa lý

Dữ liệu cho thấy thị trường Vương quốc Anh chiếm tỷ trọng doanh thu cao nhất, vượt trội hoàn toàn so với các quốc gia khác. Điều này phù hợp với đặc điểm vận hành nội địa, chi phí logistics thấp và tệp khách hàng lớn.

Ngoài UK, các thị trường EIRE, Netherlands và Germany có mức doanh thu ổn định và tiềm năng. Ngược lại, Australia có doanh thu thấp hơn nhiều, cho thấy rào cản về vận chuyển hoặc nhu cầu hạn chế.

Kết quả này mở ra cơ hội nghiên cứu sâu hơn các thị trường châu Âu có khả năng mở rộng doanh thu trong tương lai.



Hình 2: Doanh thu theo thị trường

3.3 Tổng quan khách hàng

Khách hàng có tần suất mua trung bình 6.29 lần và chi tiêu trung bình 3.02K GBP. Tuy nhiên, chỉ số Recency trung bình 201 ngày cho thấy nhiều khách hàng đã không quay lại trong thời gian dài. Điều này phản ánh mức độ rời bỏ cao và sự không ổn định trong hành vi mua lặp lại.

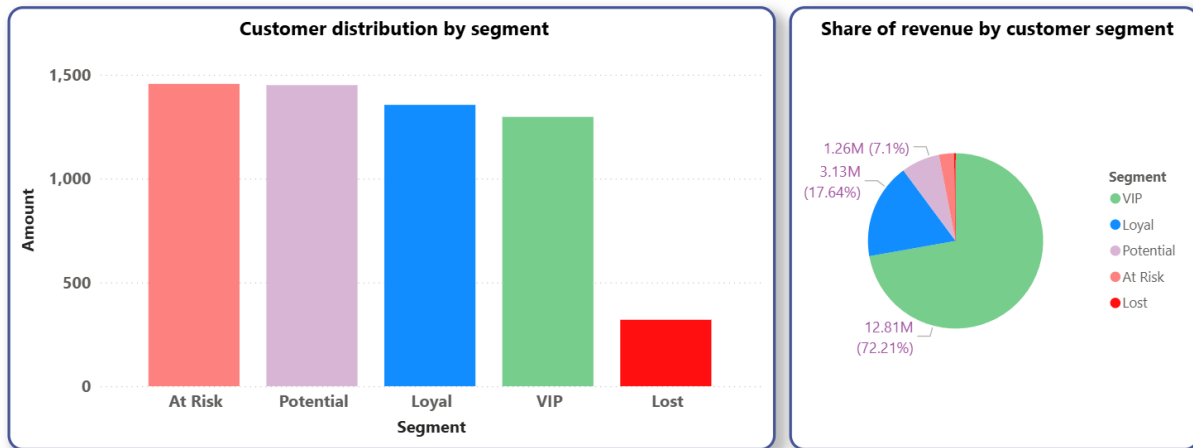
Doanh nghiệp cần tập trung nhiều hơn vào việc duy trì khách hàng hiện hữu thông qua các chương trình chăm sóc, nhắc nhở và thúc đẩy hành vi quay lại.

4 Phân tích RFM và hành vi khách hàng

4.1 Phân nhóm khách hàng theo RFM

Phân tích RFM cho thấy sự phân hóa rõ rệt giữa các nhóm khách hàng. Nhóm **VIP** đóng góp tới 72.21% tổng doanh thu mặc dù chiếm tỷ lệ nhỏ về số lượng. Đây là nhóm khách hàng đặc biệt quan trọng với giá trị mua hàng cao và tần suất ổn định.

Nhóm **At Risk** từng có mức chi tiêu tốt nhưng đang giảm mua đáng kể. Đây là nhóm cần được ưu tiên tiếp cận để ngăn ngừa rời bỏ. Nhóm **Lost** không còn mang lại giá trị và không phải đối tượng tối ưu để đầu tư marketing.

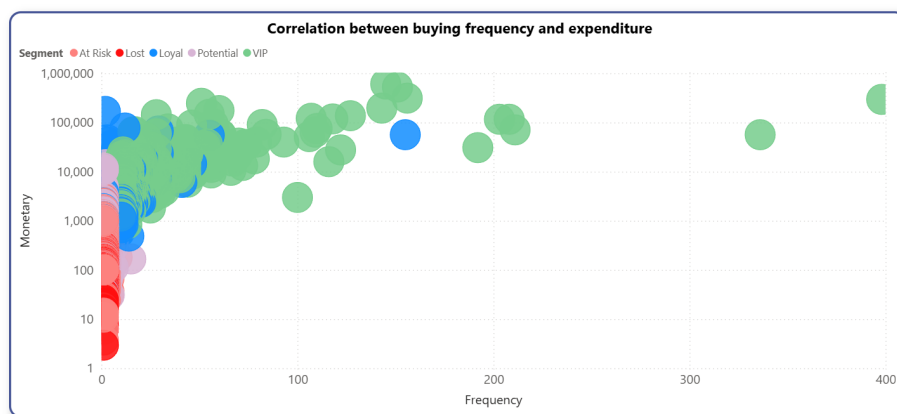


Hình 3: Phân khúc khách hàng theo RFM

4.2 Tương quan Frequency–Monetary

Mối quan hệ giữa tần suất mua và chi tiêu cho thấy khách hàng mua càng nhiều thì chi tiêu càng lớn, đặc biệt tập trung ở hai nhóm VIP và Loyal. Các nhóm có tần suất thấp (Lost, At Risk) nằm ở vùng chi tiêu thấp, phản ánh hành vi mua không liên tục hoặc giá trị giao dịch thấp.

Doanh nghiệp cần ưu tiên các chiến dịch cá nhân hóa nhằm duy trì nhóm có giá trị cao và tái kích hoạt nhóm từng có giá trị tốt.



Hình 4: Tương quan giữa tần suất mua và số tiền chi

5 Phân tích sản phẩm

5.1 Top sản phẩm theo doanh thu

Năm sản phẩm mang lại doanh thu cao nhất chủ yếu thuộc nhóm hàng trang trí và phụ kiện gia dụng. Những sản phẩm này thường có kích thước nhỏ, giá cả phải chăng và phù hợp để mua làm quà hoặc mua kèm nhiều lần. Hai sản phẩm có doanh thu cao nhất là REGISTRY CAKESTAND 3 TIER và WHITE HANGING HEART T-LIGHT HOLDER.

Những sản phẩm này đóng vai trò chủ lực và cần được ưu tiên trong quản lý tồn kho và chiến lược marketing.

Top Selling Products		
Description	Sum of Revenue Ordered	
REGENCY CAKESTAND 3 TIER	269,736.70	3659
WHITE HANGING HEART T-LIGHT HOLDER	226,282.65	4738
JUMBO BAG RED RETROSPOT	134,845.16	2683
ASSORTED COLOUR BIRD ORNAMENT	119,095.81	2571
POSTAGE	109,554.67	1931

Hình 5: Top 5 sản phẩm có doanh thu cao nhất

5.2 Xu hướng tiêu thụ theo tháng

Các sản phẩm top-selling có sự tăng trưởng doanh số rõ rệt từ tháng 9 trở đi. Một số sản phẩm thể hiện nhu cầu ổn định quanh năm, giúp giảm mức độ phụ thuộc vào mùa vụ. Điều này cho phép doanh nghiệp hoạch định tồn kho chính xác hơn và xây dựng các chiến dịch khuyến mãi theo mùa hiệu quả hơn.

6 Phân tích theo vùng thị trường

6.1 Doanh thu ngoài UK

Các quốc gia như EIRE và Netherlands cho thấy tiềm năng tăng trưởng nhờ mức doanh thu ổn định. Ngược lại, doanh thu tại Australia thấp, cho thấy nhu cầu hạn chế hoặc chi phí logistics cao làm giảm khả năng cạnh tranh.

6.2 Xu hướng tại UK

UK thể hiện sự tăng trưởng doanh thu mạnh vào cuối năm, phù hợp với xu hướng mua sắm trong mùa lễ hội. Đây là thị trường chủ lực và cần được tập trung nguồn lực để tối ưu.

7 Insight và đề xuất chiến lược

7.1 Insight tổng hợp

- Doanh thu có tính mùa vụ mạnh, đặc biệt vào quý IV.
- Nhóm khách hàng VIP mang lại phần lớn doanh thu.
- Một số thị trường EU có tiềm năng mở rộng.
- Các sản phẩm chủ lực giữ vai trò quan trọng trong tổng doanh thu.
- Tỷ lệ khách hàng không quay lại cao, cần chiến lược tái kích hoạt.

7.2 Đề xuất chiến lược

1. Tăng ngân sách marketing cho mùa cao điểm từ tháng 9–12.
2. Xây dựng chương trình chăm sóc khách hàng cao cấp cho nhóm VIP.
3. Triển khai chiến dịch tái kích hoạt cho nhóm At Risk.
4. Mở rộng thị trường tại Netherlands và EIRE với chiến dịch quảng bá phù hợp.
5. Tập trung vào nhóm sản phẩm top-selling và mở rộng các biến thể sản phẩm liên quan.

8 Kết luận

Báo cáo phân tích dữ liệu Online Retail cung cấp cái nhìn toàn diện về hành vi mua sắm của khách hàng, hiệu suất sản phẩm và sự khác biệt giữa các thị trường. Những insight thu được là nền tảng quan trọng giúp doanh nghiệp nâng cao hiệu quả quản lý khách hàng, tối ưu chiến lược bán hàng theo mùa vụ và mở rộng thị trường tiềm năng. Các đề xuất đưa ra góp phần hỗ trợ doanh nghiệp trong việc cải thiện doanh thu và củng cố lợi thế cạnh tranh trong bối cảnh thương mại điện tử ngày càng phát triển.