

# Customer 360 Analysis

## 1. Tổng quan dự án

Mục tiêu của dự án là xây dựng **Customer 360 View** từ dữ liệu Olist, cung cấp cái nhìn tổng quan về vòng đời khách hàng, hành vi mua sắm, chi tiêu, đánh giá và nguy cơ rời bỏ doanh nghiệp (churn). Dự án kết hợp phân tích dữ liệu, trực quan hóa, mô hình RFM/Clustering, dự đoán churn và NLP để phân tích sentiment bình luận khách hàng, đồng thời triển khai các công cụ tương tác dữ liệu:

- **Streamlit App:** cho phép người dùng chọn và xem chi tiết vòng đời từng khách hàng.
- **n8n Chatbot:** hỗ trợ chủ doanh nghiệp với **khuyến nghị hành động kinh doanh, cảnh báo tự động** và gợi ý quyết định dựa trên dữ liệu.
- **PowerBI:** dashboard trực quan tổng hợp các insight về hành vi, phân nhóm, RFM, churn và đánh giá sentiment từ bình luận.

## 2. Dataset

- **Olist Brazilian E-Commerce dataset:** Bộ dữ liệu Olist là dữ liệu thực tế từ sàn thương mại điện tử Brazil, bao gồm đầy đủ thông tin về khách hàng, đơn hàng, sản phẩm, seller, thanh toán và đánh giá. Đây là bộ dữ liệu được nhiều nghiên cứu và dự án phân tích khách hàng sử dụng, có thể trực tiếp áp dụng để xây dựng Customer360, phân nhóm khách hàng và phân tích hành vi mua sắm.

- `customers.csv`: Thông tin khách hàng.
- `orders.csv`: Thông tin đơn hàng.
- `order\_items.csv`: Chi tiết sản phẩm trong đơn.
- `products.csv`: Thông tin sản phẩm.
- `sellers.csv`: Thông tin người bán.
- `order\_payments.csv`: Phương thức thanh toán, giá trị.
- `order\_reviews.csv`: Đánh giá và feedback khách hàng.
- `geolocation.csv`: Thông tin địa lý.
- `product\_category\_name\_translation.csv`: Dịch tên danh mục sản phẩm.

**- Một số bài báo đã sử dụng bộ data Olist:**

- + [118366.pdf](#): Bài báo sử dụng dữ liệu Olist để phân tích hoạt động thương mại điện tử, phát hiện doanh số theo mùa, thời gian xác nhận đơn hàng, mức độ hài lòng và tỉ lệ mua lại của khách, đồng thời đưa ra gợi ý cải thiện giữ chân khách hàng.
- + [Implementation Of ETL E-Commerce For Customer Clustering Using RFM And K-Means Clustering | Jurnal Harian Regional](#): Bài báo rút ra rằng phân khúc khách hàng bằng RFM + K-Means giúp xác định rõ nhóm VIP, trung bình và at-risk, từ đó hỗ trợ doanh nghiệp tối ưu chiến lược giữ chân và tăng loyalty.

### **3. Câu hỏi nghiên cứu (Research Questions)**

#### **3.1. Tổng quan khách hàng**

- + Khách hàng tập trung ở vùng nào và đóng góp bao nhiêu vào doanh thu?
- + Tổng số đơn hàng, tổng chi tiêu trung bình theo vùng hoặc danh mục sản phẩm.

#### **3.2. Hành vi mua hàng (RFM)**

- + Khách hàng nào là VIP, khách hàng at-risk, khách mới có tiềm năng?
- + Có outlier về tần suất mua hoặc giá trị đơn hàng cần lưu ý không?

#### **3.3. Lịch sử đơn hàng & sản phẩm**

- + Khách hàng thường mua sản phẩm nào?
- + Sản phẩm/danh mục nào bán chạy hoặc có biến động bất thường?
- + Sản phẩm nào có nguy cơ hết hàng (dựa trên bán trung bình và tồn kho)?

#### **3.4. Đánh giá & Feedback**

- + Khách hàng nào đánh giá thấp hoặc có sentiment tiêu cực?
- + Thời gian giao hàng ảnh hưởng đến rating và khả năng churn như thế nào?

#### **3.5. Phân nhóm khách hàng & risk**

- + Cluster khách hàng thành nhóm VIP, trung bình, at-risk — nhóm nào quan trọng nhất?
- + Nhóm khách nào dễ churn hoặc cần chiến dịch giữ chân?

### **3.6. Customer Churn & Retention**

- + Khách hàng nào có nguy cơ rời đi cao?
- + Đặc điểm (recency, frequency, rating, delivery) nào liên quan mạnh đến churn?
- + Hệ thống có thể gợi ý chiến dịch retention hiệu quả cho từng nhóm khách.

### **3.7. Khuyến nghị hành động kinh doanh**

- + Nên gửi voucher, flash sale hay chiến dịch marketing cho nhóm nào để tăng LTV?
- + Sản phẩm/danh mục nào cần nhập thêm, tăng quảng cáo hoặc điều chỉnh tồn kho?
- + Seller nào cần kiểm tra hoặc tối ưu vận hành để giảm giao trễ?

### **3.8. Tổng quan Customer 360**

- + Tạo bảng tổng hợp theo với RFM, lịch sử đơn, chi tiêu, đánh giá, cluster, churn score và các khuyến nghị hành động.
- + Cung cấp dữ liệu trực quan và actionable insights cho chủ doanh nghiệp.

## **4. Kết quả kỳ vọng**

### **- Streamlit App:**

- Cho phép người dùng chọn khách hàng và xem toàn bộ vòng đời, lịch sử giao dịch, RFM, cluster và churn score.
- Có nút "Gợi ý hành động" để gọi n8n Chatbot nhận khuyến nghị kinh doanh dựa trên dữ liệu và chiến lược do các đội nhóm trong một hệ thống kinh doanh.

### **- n8n Chatbot:**

- Hỗ trợ chủ doanh nghiệp với khuyến nghị hành động, cảnh báo tự động và gợi ý quyết định (ví dụ: restock, campaign, kiểm tra seller).
- Trả lời câu hỏi liên quan Customer360, KPI, trend và insight hành vi khách hàng.

**- PowerBI:**

- Dashboard tổng quan, trực quan hóa tất cả các thông số về khách hàng xuyên suốt từ lúc đến và thời điểm của bộ dữ liệu.