

An eCommerce Company Case Study

Bối cảnh

Bạn là một Data Analyst làm việc cho một công ty thương mại điện tử tên là X. Bạn được giao nhiệm vụ chuẩn bị một bài thuyết trình để trình bày tổng quan tình hình kinh doanh và vận hành của công ty tính đến thời điểm hiện tại cho Giám đốc bán hàng và Giám đốc vận hành. Bài thuyết trình tối thiểu phải bao gồm các thông tin sau: **tổng quan tình hình kinh doanh, mức độ hài lòng của khách hàng, và đề xuất 2 đến 3 lĩnh vực (areas) mà công ty có thể cải thiện.**

Một số thông tin bổ sung cho case study:

- Vì chỉ có dữ liệu đến năm 2018, nên ta sẽ giả sử hiện tại đang là tháng 9 năm 2018 (các dữ liệu sau tháng 9/2018 bạn có thể bỏ qua)
- Công ty có trụ sở tại Mỹ, tuy nhiên được thành lập ở Brazil (đó là lý do vì sao một số thông tin được viết bằng tiếng Bồ Đào Nha)

Dữ liệu Bộ dữ liệu bao gồm:

- **Orders dataset:** Cung cấp thông tin về các đơn hàng

order_id: unique ID của đơn hàng

customer_id: unique ID của khách hàng

order_status: trạng thái đơn hàng

order_purchase_timestamp: thời gian đơn hàng được đặt mua

order_approved_at: thời gian đơn hàng được phê duyệt

order_delivered_carrier_date: thời gian hàng được đưa đến cho đơn vị vận chuyển

order_delivered_customer_date: thời gian hàng được đưa đến khách hàng

order_estimated_delivery_date: thời gian dự kiến đơn hàng được đưa đến khách hàng

- **Order items dataset:** Cung cấp thông tin về từng món hàng trong đơn hàng và chi phí ship

order_id: unique ID của đơn hàng

order_item_id: ID của món hàng trong đơn hàng (món hàng số 1 có ID là 1, món hàng số 2 có ID là 2, v.v. Dựa vào đây ta cũng biết được mỗi đơn hàng có bao nhiêu món hàng)

product_id: unique ID của sản phẩm nằm trong đơn hàng

seller_id: unique ID của người bán hàng

price: giá của món hàng

freight_value: phí ship

- **Order payments dataset:**

order_id: unique ID của đơn hàng

payment_sequential: thứ tự của thanh toán

payment_type: hình thức thanh toán

payment_installments: thanh toán 1 lần ($\text{payment_installments} = 1$) hay trả góp ($\text{payment_installments} > 1$, khi đó số tiền thanh toán sẽ được trả dần thành nhiều lần. Tổng giá trị của các lần thanh toán đó bằng payment_value)

payment_value: giá trị của thanh toán

- **Product dataset**: Cung cấp thông tin về sản phẩm

product_id: unique ID của sản phẩm

product_category_name: Tên nhóm sản phẩm

product_name_lenght: Số kí tự (chữ, số) trong tên sản phẩm

product_description_lenght: Số kí tự trong phần mô tả sản phẩm

product_photos_qty: Số lượng ảnh mô tả sản phẩm

product_weight_g: Cân nặng của sản phẩm (g)

product_length_cm: Chiều dài sản phẩm (cm)

product_height_cm: Chiều cao sản phẩm (cm)

product_width_cm: Chiều rộng/sâu sản phẩm (cm)

- **Product category name translated dataset**: Dịch tên các product categories từ tiếng Bồ Đào Nha sang tiếng Anh.

- **Order reviews dataset**: Cung cấp thông tin review của mỗi đơn hàng

review_id: unique ID của review

order_id: unique ID của đơn hàng

review_score: Điểm khách hàng đánh giá

review_comment_title: Tiêu đề của review

review_comment_message: Nội dung của review

review_creation_date: Ngày tạo review

review_answer_timestamp: Ngày giờ review được trả lời

- **Customers dataset dataset:** Cung cấp thông tin về khách hàng.

customer_id: unique ID của khách hàng. Trường này dùng để link với trường customer_id ở bảng orders_dataset.

customer_unique_id: mã unique ID của khách hàng trong hệ thống quản lý thông tin khách hàng

customer_zip_code_prefix: zip code của khách hàng

customer_city: thành phố nơi khách hàng sống

customer_state: bang nơi khách hàng sống