Tên người dùng, tuổi, tháng sinh, nghề nghiệp, giới tính, trung bình số tiền mỗi lần mua hàng 30 ngày gần nhất, số lần click vào sản phẩm này trong 30 ngày gần nhất,10 lịch sử tìm kiếm gần nhất, số lần mua sản phẩm này, tháng mua sản phẩm này, tổng số lượng sản phẩm mà người dùng mua trong 30 ngày gần nhất, tên sản phẩm vừa mua, khoản cách mua sản phẩm này gần nhất, số lần hủy sản phẩm này, số lần trả sản phẩm này, số lần thêm vào giỏ nhưng lại xóa của sản phẩm này, điểm của sản phẩm,10 bình luận về sản phẩm, giảm giá, tên sản phẩm, loại sản phẩm, mùa phát hành sản phẩm, màu sắc sản phẩm, thương hiệu sản phẩm, phong cách sản phẩm, size sản phẩm, lượt bán của sản phẩm trong 30 ngày gần nhất, giá tiền sản phẩm, mô tả sản phẩm, nhãn.

Mô tả các cột dữ liệu:

Toàn bộ đều là tiếng việt trừ các cột đặc thù như hãng sản phẩm

Dữ liệu về thời trang

Một bộ dữ liệu thực tế nhất có thể không cần cân bằng nhãn hay làm đẹp dữ liệu

Các cột cách nhau dấu gạch dưới (\_)

Tên người dùng : tên của người dùng gồm họ tên

Tuổi của người dùng : số tuổi

Ngày sinh, tháng sinh : ngày và tháng sinh của người dùng

Nghề nghiệp của người dùng: nghề nghiệp người dùng

Trung bình số tiền mỗi lần mua hàng: trung bình thanh toán các đơn hàng trong 30 ngày gần nhất

Số lần click vào sản phẩm này trong tháng gần nhất: số lần kích chuột vào sản phẩm trên web trong 30 ngày gần nhất

Lịch sử tìm kiếm sẽ là 10 tìm kiếm gần nhất của người dùng và cách nhau bởi dấu phẩy(,)

Số lần mua sản phẩm này của người dùng : Số lần mua sản phẩm này của người dùng

Ngày mua sản phẩm này, tháng mua sản phẩm này, năm mua sản phẩm này: thời gian mua sản phẩm

Tổng số lượng sản phẩm mà người dùng mua : Tổng số lượng sản phẩm mà người dùng mua trong 30 ngày gần nhất

Tên sản phẩm vừa mua : tên của sản phầm mua gần đây nhất

Khoản cách mua sản phẩm này gần nhất sẽ được tính theo ngày

Số lần hủy, trả, thêm vào giỏ nhưng lại xóa của sản phẩm này: số lần hủy, trả, thêm vào giỏ nhưng lại xóa sản phẩm này trong 30 ngày gần nhất

Điểm của sản phẩm: trung bình điểm số của sản phẩm từ người dùng từ 1-10

Bình luận về sản phẩm: lấy 10 bình luận gần nhất và các bình luận cách nhau dấu phẩy

Giảm giá: khi tạo dòng dữ liệu này thì sản phẩm có đang giảm giá không.

Tên sản phẩm: tên đầy đủ sản phẩm

Loại sản phẩm: thể loại của sản phẩm do đây là dữ liệu về thời trang nên loại sẽ có dạng: quần ào, phụ kiện, mũ, …

Mùa phát hành sản phẩm: Mùa phát hành sản phẩm

Màu sắc sản phẩm: màu sắc sản phẩm

Thương hiệu sản phẩm: Các thương hiệu thời trang như adidas, …

Phong cách sản phẩm: vd: trẻ trung năng động, trưởng thành, …

Size sản phẩm: vd: m-6xl cho áo, 32-33 cho quần,…

Lượt bán của sản phẩm: lượt bán của sản phẩm trong 30 ngày gần nhất

Giá tiền sản phẩm: giá tiền sản phẩm

Mô tả: mô tả về sản phẩm ít nhất 50 chữ tổng hợp lại các thông tin về sản phẩm giới thiệu ưu nhược điểm sản phẩm, nhu cầu, sự phù hợp, đối tượng mà sản phẩm nhắm đến

Đánh giá của người dùng về sản phẩm: đánh giá của người dùng về sản phẩm

Nhãn: 0-1 thể hiện xem người dùng có mua sản phẩm không