

SHOPEE FRAUD

FRAUD IN E-COMMERCE

ANH NGUYEN | 2024

Contents

[Đặt Vấn Đề 2](#_Toc168313168)

[Kiến thức (domain knowledge) & Tool 2](#_Toc168313169)

[Phân Tích Giao dịch 2](#_Toc168313170)

[Phân tích Người mua 5](#_Toc168313171)

[Phân tích Người bán 7](#_Toc168313172)

# Đặt Vấn Đề

Hiện nay, việc gian lận trên sàn thương mại điện tử đang diễn ra với thủ đoạn ngày càng tinh vi, thực hiện bởi cả người mua lẫn người bán. Gây khó khan cho doanh nghiệp vận hành như thuật toán sai lệch, gây mất doanh thu cho sàn, reviews sản phẩm ảo, seeding, lợi dụng voucher của sàn để kiếm lợi nhuận, vậy đâu là cách để nhìn ra sự bất thường này trong hàng triệu giao dịch mỗi ngày?

# Kiến thức (domain knowledge) & Tool

Báo cáo này sẽ sử dụng python cho phân tính và trực quan hóa dữ liệu.

1. Chỉ ra đâu là GMV có lượt mua cao nhất, tại sao?
2. Xác định cách thức gian lận của người mua, chỉ ra những người mua bất thường
3. Xác định cách thức gian lận của người bán, chỉ ra những người bán bất thường

# Phân Tích Giao dịch

Dữ liệu bao gồm mã đơn hàng, mã người mua, mã người bán, GMV, khuyến mãi (rebate) và thời gian giao dịch được thực hiện. Bước tiếp theo sẽ làm sạch và đưa dữ liệu về đúng định dạng cần có của nó.

A screenshot of a computer screen

Description automatically generated

Sử dụng dạng đồ thị phù hợp để xác định GMV có lượt giao dịch lớn nhất. Biểu đồ này có trục x là GMV, cột y là số lượng đơn đặt hàng. Nhìn vào chúng ta có thể nhận thấy mức GMV >= 70,000vnd và GMV < 80,000vnd có số lượng đơn cao nhất. Vì ở mức này là mức thấp nhất mà người mua sử dụng được voucher (thể hiện ở cột rebate) do sàn Shopee cung cấp.

A graph of a distribution of gross mechanise value

Description automatically generated

Chúng ta có thể dung một số kỹ thuận truy xuất dữ liệu cơ bản để thể hiện số liệu này:

A screenshot of a computer

Description automatically generated

Dùng kỹ thuật thống kê chúng ra sẽ phân loại giao dịch thành 3 loại như sau:

1. GMV >= 70,000 & GMV < 80,000: phân loại ‘70s’
2. GMV >= 60,000 & GMV < 70,000: phân loại ‘60s’
3. GMV >= 50,000 & GMV < 60,000: phân loại ‘50s’

A black screen with white text

Description automatically generated

Tại đây chúng ta có thể xác định được tính hiệu quả của việc khuyến mãi, voucher sẽ thúc đẩy tiêu dung như thế nào, từ đó sẽ có kế hoạch hợp lý cho việc thúc đẩy mua hàng thông qua voucher sản phẩm hoặc vận chuyển.

# Phân tích Người mua

Kiểm tra dấu hiệu bất thường của người mua: Trước hết sẽ xem xét về tổng lượt giao dịch của 1 người mua:

A screenshot of a computer

Description automatically generated

Lưu ý: Tập trung vào ‘Top server’ nhưng không có nghĩa rằng người mua với số lượng đơn lớn sẽ là người mua gian lận, cần đi vào chi tiết để cho thể biết gốc rễ vấn đề.

Đầu tiên chúng ta sẽ kiểm tra người mua 100605978 có lượt mua là 23 lượt. Ở cột GMV có thể nhận thấy rằng người mua này có thể mua được hàng mà không cần trả tiền với số lượng lớn ‘bất thường’. Câu hỏi đặt ra, tại sao người bình thường có thể săn được nhiều đơn 1đ tại các shop dễ dàng như vậy?  
Khi kiểm tra ID các shop giao dịch, số lượng lặp lại của các shop khá nhiều.

* Người mua này có hành vi khá tinh vi, có thể để qua mặt thuật toán, không đặt mua quá nhiều đơn 1đ cho một shop, mỗi shop sẽ thực hiện 3-4 đơn. Điều này có thể cho rằng người mua đang lợi dụng để ‘seeding ảo’ đưa ‘review’ ảo tăng tương tác cho shop.

A black screen with numbers

Description automatically generated

Tiếp theo chúng ta sẽ xem tiếp tục người mua 103653018.

Người mua này thoạt nhìn sẽ xảy ra các giao dịch với mức GMV bình thường thậm chí có những đơn hàng vượt quá 100,000vnd. Nhưng nếu nhìn vào cột Khuyến mãi và thời gian xảy ra giao dịch, Người mua này sử dụng được rất nhiều mã khuyến mãi trong cũng một ngày và thời điểm là 15h trở đi.   
A screenshot of a computer screen

Description automatically generated

* Đây là người mua có dấu hiêu gian lận về mã khuyến mãi do sàn cung cấp, chúng ta có thể nhìn vào ID shop, không bị trùng lặp, người bán có lẽ sẽ không có lỗi trong trường hợp này.

Tiếp theo là người mua 1026737, người mua này chỉ phát sinh giao dịch 1đ và có một giao dịch 75,000vnd nhưng giao dịch đó có mã, giao dịch liền trước là giao dịch 1đ trong cùng một shop có mã 102368516. Đây là người mua seeding ảo cho các shop nhưng thủ thuật tinh vi hơn là mỗi shop chỉ thực hiện 1, tối đa là 2 giao dịch.

A black screen with white numbers

Description automatically generated

# Phân tích Người bán

Xem xét hành vi người bán “Bất thường” thông qua GMV, giải thích cho việc tại sao không dung rebate để xem xét hành vi người bán vì mã voucher được sàn tặng cho những đơn hàng >=70,000vnd thì trong một ngày hoặc một thời gian, shop bán được rất nhiều đơn hàng và tất cả sử dụng voucher thì vẫn có thể là hợp lệ. Tuy nhiên GMV 1đ bao gồm seeding ảo và lợi dụng mã khuyến mãi của sàn.

Lọc những nhà bán hàng có tượng đơn với GMV < 10,000vnd:

Những nhà bán hàng top đầu bán được đơn với GMV dưới 10,000vnd gồm: 100205391, 100605978, 1026737.

Cần sử dụng báo cáo theo tuần, tháng, quý, để truy xuất được số lượng thường xuyên của các nhà bán hàng trên 3 đơn 1đ để xác định người bán có hành vi bất thường.

A screenshot of a black screen

Description automatically generated