

# BÁO CÁO

# Gian lận mã giảm giá của người bán và người mua

từ 30/10/2019 đến 2/11/2019

By Nguyễn Doãn Hoàng

## MỤC LỤC

Đặt vấn đề	2
Dữ liệu	2
Mục tiêu	2
1. Tổng quan	3
2. Người mua có khả năng gian lận	4
2.1. Seeding bằng đơn hàng ảo, đơn có giá trị thấp	4
2.2. Lạm dụng mã giảm giá, mua hàng với mục đích khác	7
3. Người bán có khả năng gian lận	8
3.1. Sự tương tác cố định giữa shop với người dùng	9
3.2. Shop có doanh thu, số đơn hàng và số người mua cao nhất	9
3.3. Seeding bằng đơn ảo, đơn có giá trị thấp	10

#### Đặt vấn đề

Shopee đưa ra chương trình khuyến mại 30% tổng hóa đơn từ ngày 30/10/2019 đến hết 2/11/2019, tối đa là 20000 VND cho mỗi hóa đơn. Có thông tin rằng có vài khách hàng đang liên kết với người bán để lạm dụng chương trình này. Điều này có thể gây ảnh hưởng đến uy tín của sàn giao dịch shopee và những người mua/ bán hàng thực sự. Báo cáo đưa ra các thông tin về những người mua và người bán có tiềm năng gian lận và nguyên nhân

#### Dữ liệu

Tập tin Transaction\_Data.csv chứa dữ liệu giao dịch của khách hàng.

#### Mục tiêu

Đưa ra 1 báo cáo trả lời các thông tin:

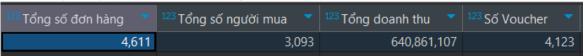
Khoảng tổng hóa đơn bao nhiều thì có số lượng giao dịch nhiều nhất

Đâu là 3 người bán có khả năng gian lận

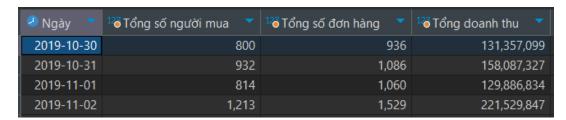
Đâu là 3 người mua có khả năng gian lận

# 1. Tổng quan

Trong 4 ngày diễn tra chương trình khuyến mãi, các chỉ số cơ bản như sau:

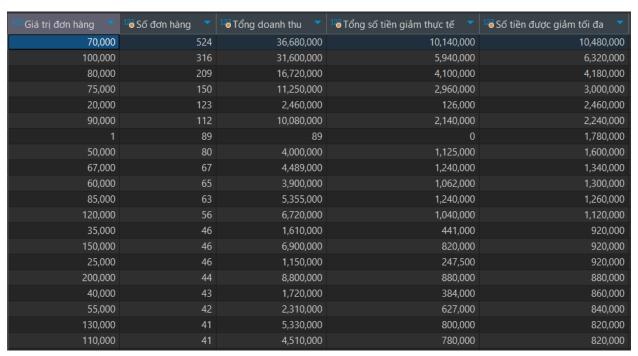


Theo thời gian, số khách hàng và số đơn hàng, doanh thu lần lượt là:



Giá trị đơn hàng ở mức 70K thì được giao dịch nhiều nhất trong chương trình khuyến mãi này, với tổng doanh thu lên tới 36 triệu và số lượng đơn cao nhất. Do voucher giảm 30% tối đa 20K, nên để tối ưu số tiền được giảm, người mua sẽ mua những đơn hàng có giá trị khoảng 70K trở lên..

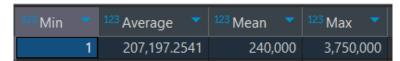
Bên cạnh đó, phát hiện số lượng đơn có giá trị bằng 1 khá cao, đây có thể sẽ bao gồm nhiều đơn hàng gian lận.



Số tiền được giảm tối đa nhỏ hơn số tiền giảm thực tế chứng tỏ không có gian lận trong việc dùng nhiều voucher 1 lúc hoặc lỗi hệ thống nào đó. Tuy nhiên, điều này cũng nói lên rằng có vẫn có những người mua không sử dụng voucher hoặc không biết đến chương trình khuyến mãi này

### 2. Người mua có khả năng gian lận

Tính toán tổng giá trị đơn hàng của người mua



Tính toán tổng số đơn hàng của người mua



Tính toán khoảng cách thời gian đặt hàng của người mua

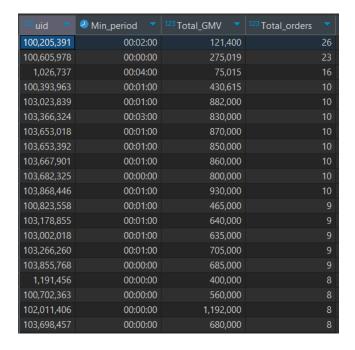


Trong đó 188379 giây tương đương với 52:19:39

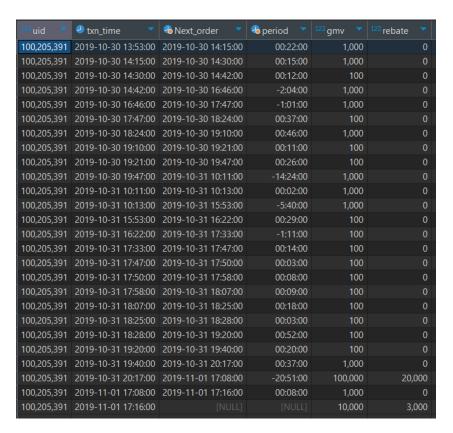
#### 2.1. Seeding bằng đơn hàng ảo, đơn có giá trị thấp

<u> </u>				•
<sup>123</sup> uid 🔻	¹ <mark>∛</mark> Số đơ	n hàng có	giá trị nhỏ hơr	1k 🔻
100,605,978				19
1,026,737				15
100,205,391				14
100,614,916				6
101,065,061				6
100,352,740				5
100,393,963				5
101,039,570				5
100,027,807				4
100,368,648				4

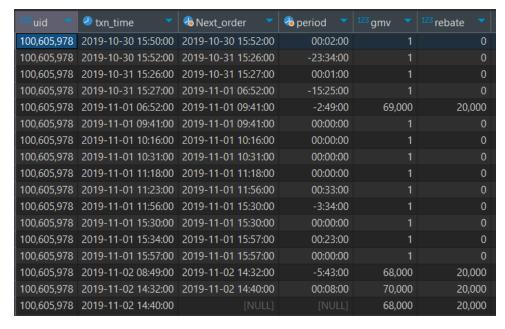
Các người mua có ID 10060978, 1026737, 100205391 có số lượng đơn hàng dưới 1K cao bất thường.



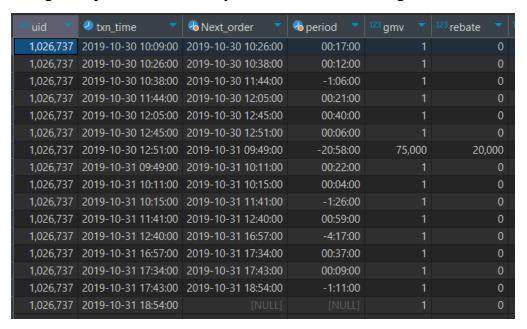
Đồng thời 3 người này cũng có tổng số đơn khá cao nhưng tổng giá trị các đơn hàng lại thấp hơn nhiều so với giá trị trung bình của toàn tập dữ liệu, thời gian giữa các lần mua khá ngắn. Chúng ta có thể thấy rằng đây là hành động có chủ ý nhằm chia đơn hàng lớn hơn thành những phần nhỏ hơn để thu lợi từ chương trình khuyến mại, hoặc tận dụng thời gian này để liên tục seeding các đơn ảo cho shop.



Người dùng 100205391 mua hàng liên tục trong 3 ngày, khoảng cách thời gian giữa các giao dịch khá sát, số lượng đơn 100đ và 1K chiếm đa số. Trong đó chỉ có 2 trên 26 đơn hàng dùng mã giảm giá



Người dùng 100605978 cũng tương tự khi người này cũng thực hiện mua hàng tương tự như người dùng đầu tiên, tổng GMV và tổng số đơn đặt hàng tương đối cao. Thậm chí người mua này còn có thời gian giữa các lần đặt hàng = 0. Điều này có thể xảy ra do người này đăng nhập nhiều thiết bị cùng lúc hoặc cố tình seeding cho 1 shop nào đó vì không cần quan tâm đến các yếu tố khác mà đặt hàng một cách nhanh vội.



Người dùng 1026737 cũng có các đặc điểm tương tự so với 2 người trước: có GMV thấp trong khi số đơn cao. Ngoài việc seeding đơn ảo, người này cũng có thể test hệ thống xem thử hệ thống có cho them voucher nếu họ đặt nhiều đơn hay không.

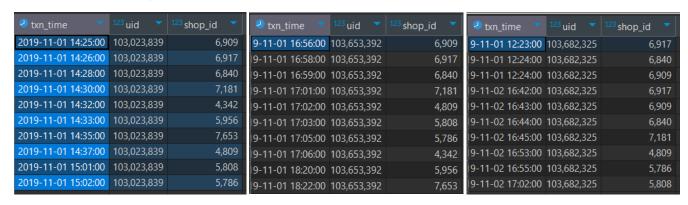
⇒ Từ đây, có thể nhận định rằng 3 người mua gian lận với ID lần lượt là 100205391, 100605978 và 1026737.

#### 2.2. Lạm dụng mã giảm giá, mua hàng với mục đích khác

123 uid	<sup>12</sup> 6 Số lần dùng voucher	
103,366,324		10
103,653,018		10
103,682,325		10
103,653,392		10
103,667,901		10
103,868,446		10
103,023,839		10
103,002,018		9
103,178,855		9
103,266,260		9
103,855,768		9
100,702,363		8
103,044,301		8
103,044,781		8
103,650,095		7

Ta thấy có một vài người mua được phép sử dụng số lượng voucher có giá trị 20K nhiều lần. trong đó đáng chú ý là các người mua được sử dụng tới 10 lần. Dù chương trình chỉ diễn ra trong vòng 4 ngày nhưng số lần dùng của những người này lại rất cao và khác biệt so với số đông.

Tiến hành kiểm tra lịch sử giao dịch của 7 người có số lần dùng voucher nhiều nhất: (ví dụ 3 người mua):



#### Ta có nhận xét:

- Các người mua này đều mua đơn có giá trị lớn hơn 70K để tối đa hóa số tiền được giảm.
- Khoảng cách giữa các lần giao dịch rất ngắn, thậm chí người mua có ID 103622325 có thể đặt vài đơn trong cùng 1 mốc thời gian.
- Các shop mà những người này mua có các ID khác nhau nhưng lại mua hàng hóa giống với những người khác. Cụ thể, 7 người này mua 10 đơn hàng, các shop giống y hệt nhau với các ID shop là: 6909, 6917, 6840, 7181, 4342, 5956, 7653, 4809, 5808, 5786
- ➡ Điều này chứng tỏ rằng nhóm người này có thể được các shop này thuê để mua hàng đồng thời có thể đã sử dụng một phần mềm chuyên dụng để gian lận chiếm đoạt mã giảm giá của Shopee. Shopee nên có các hành động ngăn chặn và cảnh báo đối với cả người mua và người bán trong trường hợp này

Tiến hành kiểm tra lịch sử của các nhóm người mua có số lượng giao dịch là 9,8,7,6,5,4 lần, ta không thu được kết quả trùng lặp đáng ngờ như nhóm người mua dùng voucher 10 lần.

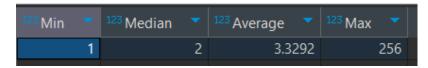
➡ Từ đây ta có thể nói rằng: nhóm 7 người mua sử dụng 10 lần voucher đều là nhóm tài khoản gian lận. Shopee nên có hình thức xử lí phù hợp đối với các nhóm này

#### 3. Người bán có khả năng gian lận

Doanh thu của các shop trong đợt khuyến mãi này:



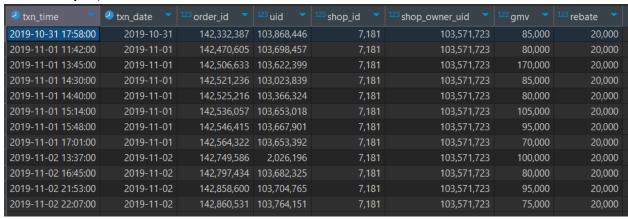
Số lượng đơn hàng của các shop trong đợt khuyến mãi:



Số lượng khách của 1 shop trong đợt này:



Trong phần phân tích trên, ta đã thu được 7 người đã có tình trạng lạm dụng mã giảm giá và mua ở các shop giống nhau. Ta tiến hành kiểm tra giao dịch của các shop này và không thu được kết quả đáng nghi về giá trị các đơn hàng, người mua cố định,...



Hình minh họa giao dịch của 1 shop bất kì trong nhóm 10 shop đáng nghi ở phần 2.2

#### 3.1. Sự tương tác cố định giữa shop với người dùng

•	•	<u> </u>	
123 shop_id	<sup>123</sup> uid ▼	¹ã Số giao dịch	
29,905	100,823,558		
10,979	1,191,456		8
10,979	103,282,624		7
10,979	103,086,848		
10,979	102,696,939		
10,873	103,993,346		
10,979	102,849,459		6
10,979	103,514,096		
10,979	103,005,918		
4,574	103,410,928		
5,643	102,892,016		5
790	102,001,686		5
29,950	1,168,734		5
29,769	103,274,327		5
5,454	100,151,179		5

Ta nhận thấy shop 10979 là shop có số đơn, người mua, doanh thu cao nên khó có thể là shop gian lận. Ta xét 3 shop có khả năng liên kết với người bán để gian lận đó là 29905, 4574 và 5643

Dù số lần tương tác giữa shop và các người dùng này cao nhưng giá trị đơn hàng, khoảng các giữa các lần đặt hàng và tổng số lượng đơn hàng đều không có gì đáng nghi. Mặc dù xuất hiện 1 đơn có giá trị 100đ nhưng chỉ từ 1 người mua có duy nhất 1 giao dịch.

⇒ Chứng tỏ không có sự tương tác cố định nào giữa người mua và người bán nhằm trục lợi từ chường trình khuyến mãi này.

#### 3.2. Shop có doanh thu, số đơn hàng và số người mua cao nhất

		,		8	
123 shop_id	¹ã Doanh thu ▼	123 shop_id	¹% Số đơn ▼	123 shop_id	¹ã Số người mua ▼
10,979	14,746,549	10,979	256	10,979	94
5,741	11,261,000	4,673	50	4,673	41
4,673	6,334,000	4,809	31	9,256	27
6,788	5,655,000	5,786	30	9,446	24
4,164	4,459,000	9,256	29	3,476	24
10,220	4,127,000	5,808	29	4,809	23
6,072	4,002,000	8,576	28	9,222	23
29,463	3,795,000	9,127	27	5,786	23
10,343	3,785,900	6,072	27	8,224	23
1,369	3,428,000	8,437	26	6,666	23
347	3,317,000	9,446	25	8,437	23
10,837	3,170,000	3,476	25	7,873	23
9,446	3,086,000	8,591	24	9,127	22
5,231	3,040,000	8,224	23	5,808	22
631	3,027,320	6,666	23	6,072	20
10,523	2,926,000	7,873	23	10,220	18
7,155	2,838,000	9,222	23	8,576	17
8,936	2,711,000	5,643	22	1,369	17
7,499	2,701,000	6,788	19	9,909	16
4,952	2,700,000	7,702	19	935	16

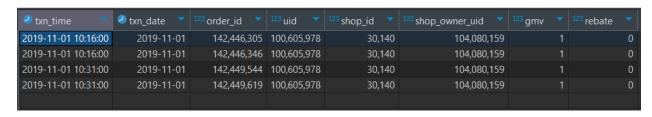
Ta nhận thấy những shop này, có thể có doanh thu vượt trội so với các shop khác, số đơn cao hơn hẳn so với toàn sàn hay số người mua cao là những shop ít có khả năng gian lận. Những giá trị cá biệt này có thể do loại hình sản phẩm mà các shop này kinh doanh.

➡ Tuy nhiên, Shopee cũng nên có một số hành động nhằm kiểm tra và xét duyệt từ đó quản lí các shop này kịp thời hơn.

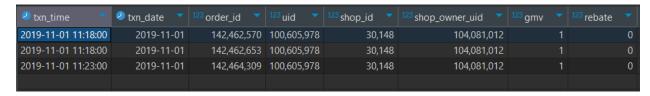
#### 3.3. Seeding bằng đơn ảo, đơn có giá trị thấp

	0 0			
123 shop_id	¹ã Số người mua		¹ã Số đơn <1k	
30,140		1		4
30,148		1		
30,185		1		3
29,380		1		2
10,564		1		2
29,712		1		2
29,728		1		2
29,831				2
8,947		1		2
30,223		1		2

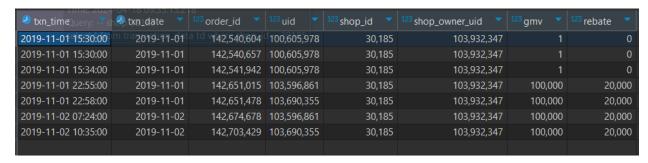
Một điều bất thường là dù các đơn hàng này chỉ dưới 1K nhưng số người mua chỉ có 1. Đây có thể là nick ảo của chủ shop nhằm seeding cho shop mình. Ta sẽ kiểm tra giao dịch của 3 shop đầu tiên



Shop 30140 chỉ bán những đơn hàng dưới 1K và người mua chỉ từ 1 tài khoản, vì vậy có thể nói shop này chắc chắn gian lận bằng cách set đơn ảo nhằm tang số lượng đơn hàng, từ đó sẽ cung cấp các đánh giá ảo và số lượng mua khống. Các đơn hàng của shop này cũng được thực hiện ở các khoảng thời gian khá gần nhau.



Shop 30148 cũng có tình trạng tương tự shop trên nên ta cũng có thể nói rằng đây là 1 shop đang thực hiện hành vi gian lận.



Dù shop 30185 cũng có các đơn có giá trị tương đối cao nhưng cũng xuất hiện 3 giao dịch tương tự với 2 shop trên.

⇒ Từ đây, ta có thể kết luận rằng, các shop có khả năng gian lận cao bằng cách seeding đơn ảo lần lượt có ID là 30140, 30148, 30185.