# IMPACT OF SELLER OFFER PRIORITIZATION ON CUSTOMER SATISFACTION

## OLIST BRAZILIAN E-COMMERCE DATASET

## OBJECTIVE

This experiment evaluates whether a new seller offer prioritization algorithm improves customer satisfaction among single-order buyers on the Olist platform, aiming to convert single-order customers into repeat buyers, increase satisfaction, and ultimately drive higher revenue.

Hypothesis:

- Null (H0): Changing seller offer prioritization does not impact satisfaction among single-order buyers.
- Alternative (H1): Changing seller offer prioritization increases satisfaction among single-order buyers.

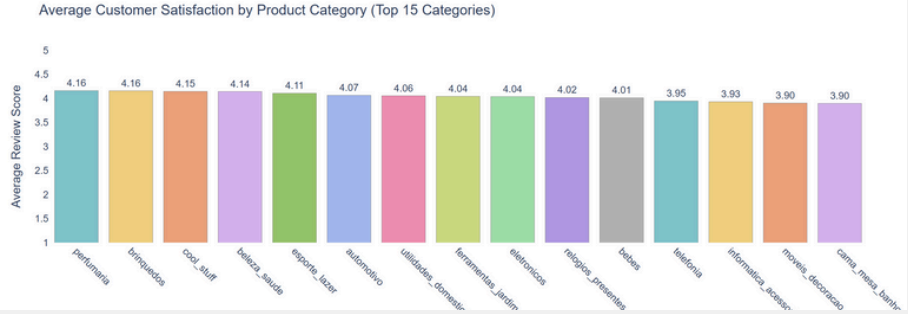
## TARGET AUDIENCE

The experiment focuses on single-order customers, who represent 93% of all buyers on the Olist platform, and targets the top five product categories by volume: cama_mesa_banho (bed, table and bath), beleza_saude (beauty and health), esporte_lazer (sports and leisure), moveis_decoracao (furniture and decoration), and informatica_acessorios (computers and accessories). To ensure data quality and relevance for measuring satisfaction, only delivered orders with valid review scores and known customer states are included. After applying these filters, the final dataset includes 37,770 orders from unique single-order customers. Repeat buyers are excluded, as they tend to report slightly higher satisfaction scores (4.04 vs. 3.99) and spend more (R$293.70 vs. R$145.05), which could bias results. This carefully defined audience allows the experiment to isolate the effect of the new seller offer prioritization on first-time customer satisfaction.

## JUSTIFICATION

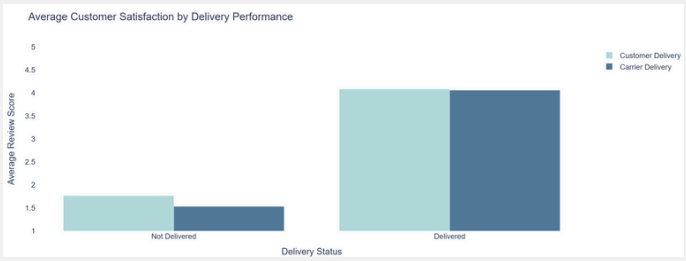
The decision to focus on single-order customers purchasing from the platform's top product categories is supported by several key findings:

Customer Satisfaction by Product Category: Perfumery and Toys lead with the highest average satisfaction scores of 4.16, closely followed by Beauty & Health (4.14) and Sports & Leisure (4.11). In contrast, Bed, Table & Bath products have a somewhat lower satisfaction average around 3.90. This indicates that focusing on the high-engagement categories will maximize the impact of the experiment on customer satisfaction.

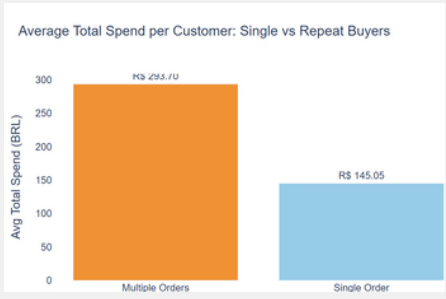



Geographic Distribution and Variation: Most customer orders are concentrated in São Paulo (46,427 delivered orders), Rio de Janeiro (14,054), and Minas Gerais (12,906). There are statistically significant differences in delivery rates by state (Chi-square = 154.13, df = 26, $p < 0.001$). Satisfaction also varies geographically, with the highest average review scores in Tocantins (4.13), São Paulo (4.10), and Paraná (4.08), while Bahia (3.77) and Rio de Janeiro (3.76) are on the lower end. This regional variation makes it necessary to stratify the experiment by customer state to control for geographic effects.

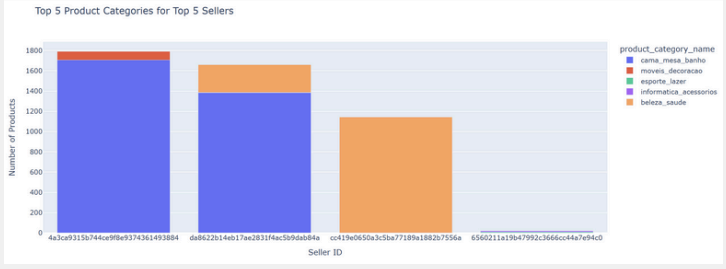
Impact of Delivery Performance: Delivery status strongly affects satisfaction scores. Orders confirmed as delivered by customers average 4.08, and those confirmed by carriers average 4.06. Meanwhile, undelivered orders have dramatically lower scores – 1.76 and 1.53 respectively. This large difference justifies including only delivered orders in the experiment to ensure satisfaction scores reflect actual customer experience.




Focus on Single-Order Customers: Single-order customers account for approximately 93% of buyers, while repeat customers make up about 7%. Repeat buyers have a higher average satisfaction (4.04 vs. 3.99) and nearly double the average spending (R$293.70 vs. R$145.05) compared to single-order buyers. Including repeat customers could introduce bias because their prior experience may influence satisfaction differently. Thus, focusing on single-order buyers isolates the effect of the seller offer prioritization on first-time purchase satisfaction.




Seller Diversity and Product Offering: Sellers with more products tend to have higher average satisfaction scores. For example, sellers with around 24 products average 4.03 in satisfaction, while those with only one product show more variability. The highest average product prices among sellers reach approximately R$895. This supports stratifying the audience by seller ID to maintain a balanced representation of seller diversity and product range.



## AUDIENCE CHARACTERISTICS

The selected customers come from all regions of Brazil, with the largest groups in São Paulo (SP), Minas Gerais (MG), Rio de Janeiro (RJ), and Rio Grande do Sul (RS), which represent Olist's main markets. The product categories are well balanced, with each of the top five categories making up roughly 16% to 26% of the total sample. This ensures fair representation across key product lines. In terms of sellers, the dataset includes a wide variety—while the top 10 sellers handle a notable share of transactions, they do not dominate the sample. This broad seller diversity helps maintain balanced and unbiased distribution between the test and control groups.

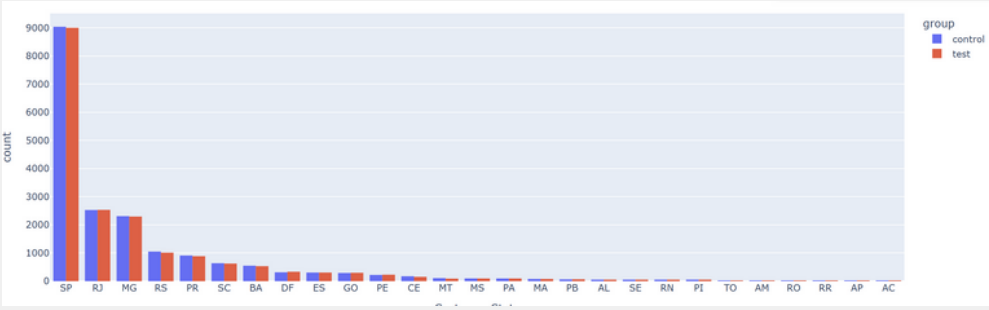
## EXPERIMENTAL DESIGN AND STRATIFICATION

To ensure a fair comparison, customers were grouped (stratified) by state, product category, and seller ID. Only groups with at least two customers were used so they could be evenly split. This gave a final sample of 31,140 customers.
Half (15,570) were randomly assigned to the control group with the current seller offer system, and the other half to the test group with the new seller offer prioritization.

## VALIDATION OF GROUP BALANCE

To ensure a fair and unbiased comparison, stratified random sampling was applied based on three key variables: customer state, product category, and seller ID. This process created two balanced groups, with 15,570 customers in both the control and test groups.

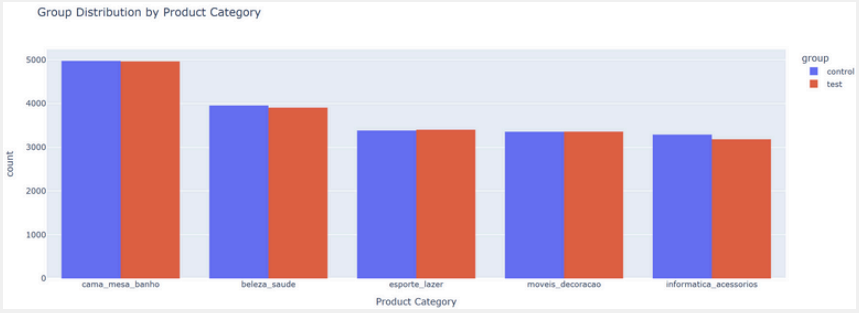
Geographic Balance: Customers were distributed across all Brazilian states. A chi-square test for independence was conducted to compare the geographic distribution between the test and control groups. The test yielded a chi-square statistic of 9.28 with 26 degrees of freedom and a p-value of 0.9989. This high p-value indicates no statistically significant difference in geographic distribution between the two groups, confirming that the groups are geographically balanced.



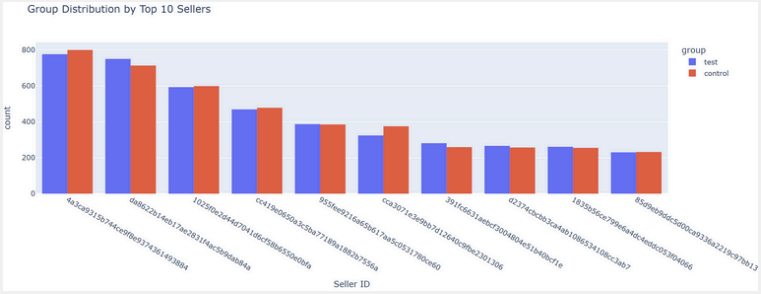

Product Category Balance: The proportions of customers in each product category were nearly identical across groups:

- cama_mesa_banho (bed, table & bath): 26.39% vs. 26.24%
- beleza_saude (beauty & health): 20.76% vs. 20.86%
- esporte_lazer (sports & leisure): 18.08% vs. 17.84%

This ensures both groups have equal representation across high-demand product segments.




Seller Representation Balance: Seller representation was similarly balanced. The top 10 sellers accounted for a significant share of transactions, and customer distribution among these sellers varied by no more than ±2% between the test and control groups. This indicates that no individual seller was disproportionately represented in either group, supporting a fair comparison.




## CUSTOMER SATISFACTION RESULTS

Customer satisfaction was defined as giving a review score of 4 or higher. The control group showed a satisfaction rate of 75.44%, while the test group had a rate of 74.84%. A chi-square test found no statistically significant difference between the groups ($\chi^2 = 1.79$, $p = 0.1812$), indicating similar satisfaction levels before the treatment.

## REPEAT PURCHASE BEHAVIOR ANALYSIS

An initial analysis of repeat purchase rates after the experiment start date showed no significant difference between groups ($p = 1.0$). Since this analysis is based on historical data without actual seller prioritization changes, the result reflects no observed impact within the dataset timeframe. It is recommended that Olist implement and monitor a real seller offer prioritization focused on single sellers within the top five product categories to better assess its potential effect on repeat purchase behavior over time.

## CONCLUSION

The stratification process successfully balanced test and control groups by geography, product category, and seller representation, eliminating sampling bias. Statistical tests confirmed no significant differences between groups at baseline, providing a reliable foundation for evaluating the seller offer prioritization's effect on customer satisfaction. While no immediate improvement in satisfaction was detected, these findings support continued experimentation and future assessment of longer-term impacts on customer behavior.