E-commerce

PYTHON PROJECT

Case Study.

__by Thao Nguyen

Prompt

If the CEO of Olist company wants to **increase revenue by 10%** (The company has no control over the prices and shipping charges set by its partners).

Based on the given historical data attached, what insights and recommendations could I provide as a **data analyst** in this company.

Data source: Kaggle

Agenda Slide

01

Analysis Approach/ Framework

02

Business Opportunity 1:

Increase Customers

03

Business Opportunity 2:

Increase Order Frequency

04

Business Opportunity 3:

Increase Order Size

05

Key Recommend ations Slide

06

Q&A

AGENDA SLIDE

Structured framework

Step 1: Understand & Clarify the problem

The CPCC framework (Company, Product, Customer, Competitor)

Problem Formulation

Objectives:

Why: Why do we need to conduct this analysis?

The goal is to increase Olist's revenue by 10% within 6 months by driving consumer spending on the Olist platform.

Company

Business Lines & Revenue Segments

What: What industry does the company operate in?

The company provides an online platform for retailers to sell their products through e-commerce channels while also offering logistics services through delivery partners.

Which: What are the company's revenue sources?

- Commission fees from sales transactions.
- Shipping and delivery fees.
- Service fees from retailers (listing fees, advertising fees).

Data: freight_value, price, payment_value

Structured framework

Step 1: Understand & Clarify the problem

The CPCC framework (Company, Product, Customer, Competitor)

Company

Locations - Plant, Distribution

Where: Where does the Olist operate, and how is product distribution managed?
The company primarily operates in Brazil. The company does not own its own warehouses but collaborates with logistics partners to transport products.
Data: *Not available*

Key Partners:
Who: Who are the the company's key partners?
1) Merchants 2) Logistics companies 3) Payment gateway 4) IT and Cloud service
Data: *seller_id (Merchants)*

Product

SKUs - Types, Variants, Sizes

What: What industry does the company operate in?
The company provides an online platform for retailers to sell their products through e-commerce channels while also offering logistics services through delivery partners.
Who: Who decides which products are available on Olist?
Retailers or sellers determine the products and variations they provide through the Olist platform.
Data: *product_category_name*, *weight*, *length*, *height*, *width*

Structured framework

Step 1: Understand & Clarify the problem

The CPCC framework (Company, Product, Customer, Competitor)

Customer

Key customer segments

Who: Who are Olist's customers?
The company serves online shoppers in Brazil.
Demographics: Age, occupation, marital status, address,...
Data: *customer_city*, *customer_state*

When: When do customers search for products on Olist?
Data: *order_purchase_timestamp*

How: How do customers shop?
Browse products → Add to cart → Checkout → Order confirmation → Delivery → Feedback and review
Data: *payment_type*, *order_approved_at*, *order_delivered_carrier_date*, *order_delivered_customer_date*, *order_estimated_delivery_date*

Pain point:
What: What are the main issues faced by customers?
Customers may experience problems such as slow delivery, inconsistent product quality, or poor customer service.
Data: *review_score*, *review_comment_message*

Competitor

Who: Who are Olist's competitors in Brazil's e-commerce market? What is their market share?
Data: *Not available*

Structured framework

Step 2: Structuring the Problem

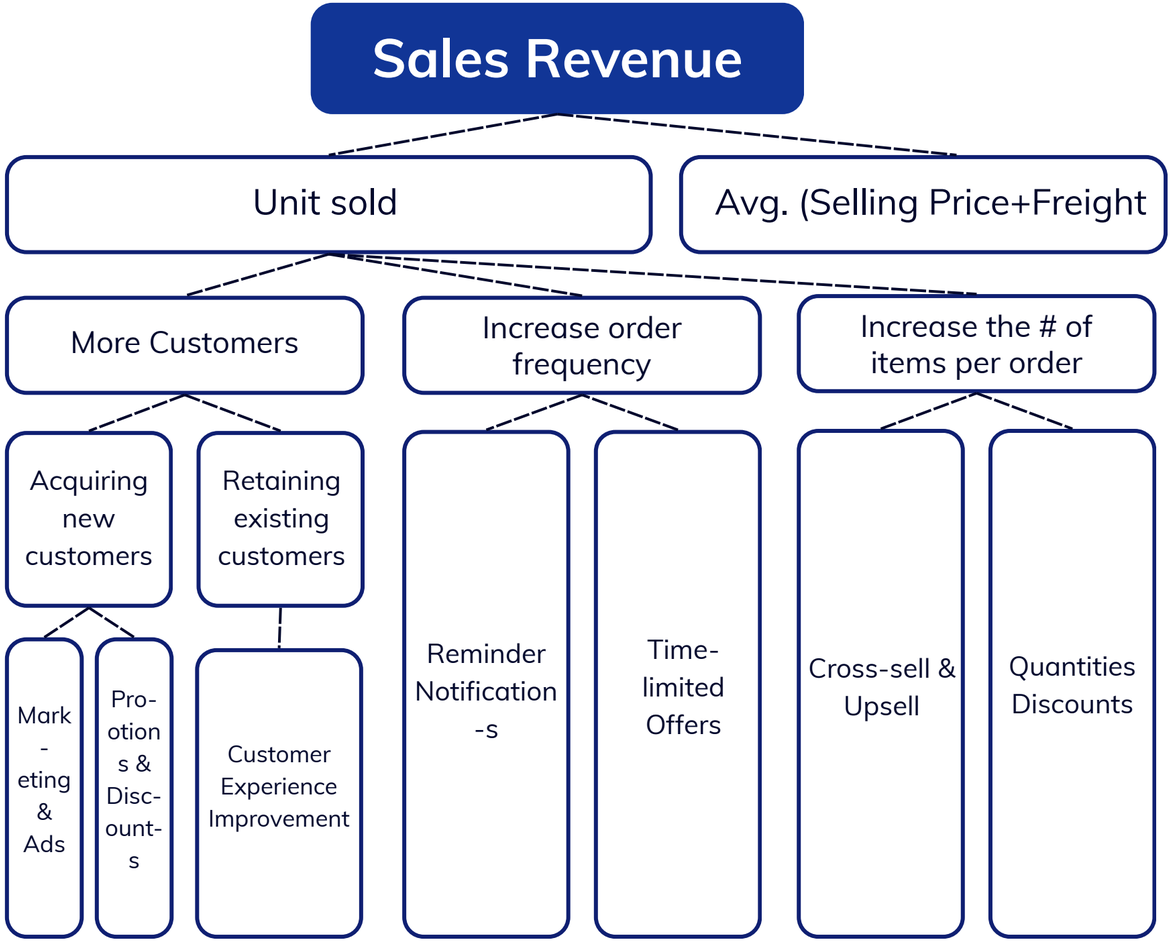
Sales Revenue = Units Sold x Avg.(Selling Price + Freight)

Apply the 80/20 rule

Focusing on the 20% factors that drive 80% of the impact

Apply MECE principle (Mutually Exclusive, Collectively Exhaustive)

The analysis is **well - structured** - breaking down problem into **distinct, non-overlapping** categories while **covering all possible factors without redundancy.**



Structured framework

Step 3: Finding Key Drivers

Identify three key drivers to increase the number of unit sold

01.

Increase more customers (Acquiring new customers and retaining existing ones)

02.

Increasing order frequency (RFM analysis: Recency, Frequency and Monetary value)

03.

Increasing the number of items per order

Step 4: Asking the Right Questions

For each key areas, I applied a structured approach by asking 4 questions

01.

What do I need to know?

02.

Why it is important ?

03.

What actions can be taken based on insights?

04.

How can it be measured with data?

Opportunity 1: More customers

Acquiring new customers → (Promote Marketing & Ads)

What do I want to know?

1, Which regions have the highest performance?
2, Which products are the best-selling?
3, Which top merchants have the highest performance?

What do I want to know?

1, **Top-performing regions** → Identify where potential customers are concentrated.
2, **Best-selling products** → Determine which products are currently attracting the most customer interest.
3, **Top-performing merchants** → Identify the best-selling partners.

What actions can be taken based on insights?

1, Develop marketing strategies targeting these regions, such as advertising campaigns or promotions for specific cities.
2, Increase promotions for these products, apply discount campaigns for new product lines, or expand product offerings to attract more new customers.
3, Identifying top-selling merchants allows the company to focus marketing resources and support them in boosting sales. Additionally, Olist can learn from their strategies to help other merchants improve or attract new merchants.

How can it be measured with data?

High-performing regions : Number of customers per region.
Best-selling products: Number of orders per product & Average rating and number of reviews per product.
Top-performing merchants: 1) Number of orders 2) Revenue 3) Delivery time 4) Customer reviews.

Opportunity 1: More customers

Retaining Existing Customers → Customer Experience Improvement

What do I want to know?

Retention rate (How are customers returning to buy, and how is it generally assessed?)
1, Which top merchants are performing poorly?
2, Which payment methods have low success rates?
3, Delivery time

What do I want to know?

1, **Understanding why some merchants are underperforming** → affects customer experience with the product, making them perceive the product poorly.
2, **A low payment success rate** → customers fail to complete their purchases → leading to lost revenue.
3, **Long delivery times** → customers may feel dissatisfied, leading to reduced repeat purchases in the future.

What actions can be taken based on insights?

1, If Olist identifies merchants with low performance, they can implement support strategies such as training, improving operations, or adjusting policies to encourage merchants to improve their product and service quality.
2, If payment methods have low success rates → Olist can work on improving or optimizing those payment systems.
3, Understanding delivery times helps Olist find ways to optimize logistics processes to shorten delivery times and enhance customer experience.

How can it be measured with data?

Which top merchants are performing poorly: Number of returned orders & Customer rating and feedback.
Which payment methods have low success rates: Success rate for each payment method.
Delivery time (focus on slow orders):
Average delivery time (from when a customer places an order until they receive the product)
Delivery time by region or by merchant & On-time delivery rate.

Opportunity 2: Increase Order Frequency

Email/push notifications to encourage purchases

What do I want to know?

1, Time periods when order volume tends to be high.
2, Customer segmentation based on RFM.

Why do I want to know this?

1, To determine when customers are most likely to shop
2, **Customer segmentation using RFM** → Identify new customers, loyal customers, and inactive customers

What actions can be taken based on insights

1, Sending emails/push notifications at optimal times to encourage purchases.
- Similar to how Doordash sends push notifications about meal delivery promotions during dinner hours or reminds customers about their abandoned carts, Olist can use a similar strategy to increase order volume during key time periods.
- Additionally, optimizing customer support teams for peak hours is essential.

2, Knowing customer segmentation allows personalized email/push notification content.
For example, **loyal customers** may receive exclusive offers/gifts, while **customers needing incentives** might be targeted with discounts or free shipping.

How can it be measured with data?

1, RFM Model
- **Recency (R):** How long ago the customer last made a purchase.
- **Frequency (F):** Number of purchases within a specific time frame.
- **Monetary (M):** Total spending within a specific time frame.

2, Shopping trends based on time: Shopping patterns by day, week, and time of day.

Opportunity 3: Increase the # of Items per Order / Order Size

Cross-sell/ Up-sell

What do I want to know?

1, Average order size
2, Which products are frequently purchased together?

Why do I want to know this?

1, To identify customer groups with below-average order sizes
2, To determine which products are commonly bought together

What actions can be taken based on insights?

1, **Average order size:**
- Encourage customers to increase their spending to receive benefits (e.g., free shipping).
- Offer product bundles such as “game day special” deals (e.g., buy 3 items, get a discount).

2, Provide recommendations or create suitable product bundles:
Suggest complementary products to encourage additional purchases when a customer selects a specific product.

How can it be measured with data?

1, Average order size
- The average number of items per order.
- The average spending per order.

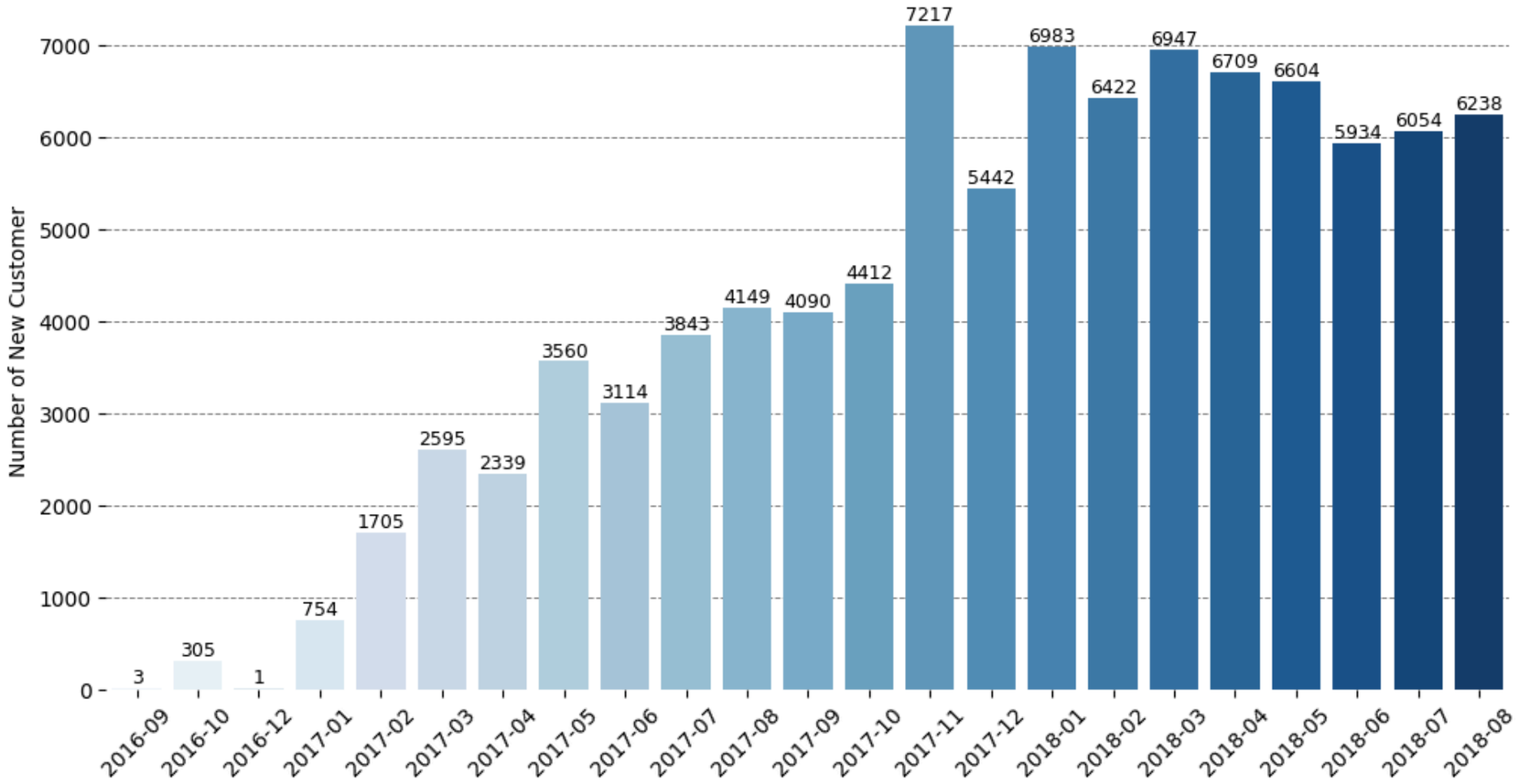
2, Which products are frequently purchased together?

Opportunity 1

MORE CUSTOMERS

New customers + Retain Customers

The number of new customers grew strongly and reached its peak in December 2017, and then remained stable.



Total New Customers by Months

From January 2017 to December 2017, the number of **new customers grew steadily**, with a sharp spike in November 2017 (possibly due to Black Friday).

From January 2018 to August 2018, the number of new customers slightly decreased but remained stable at around

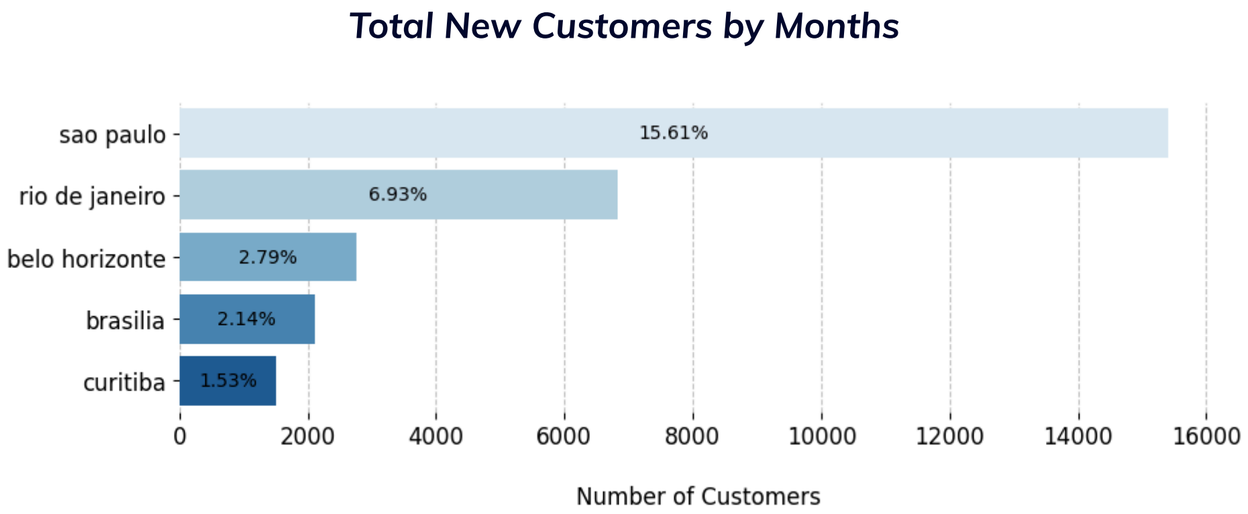
6,000 to 7,000 customers.

Olist needs to implement additional marketing strategies to attract more new customers in order to achieve a

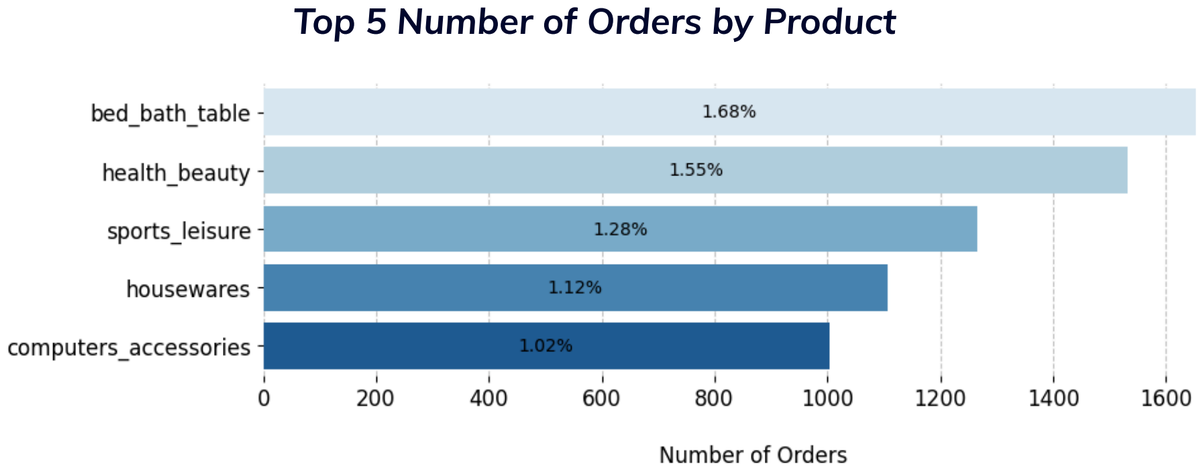
10%

revenue growth in the next year.

Strengthen marketing and promotional strategies in Sao Paulo (the area with the highest demand) for bath table and health beauty products to attract new customers.

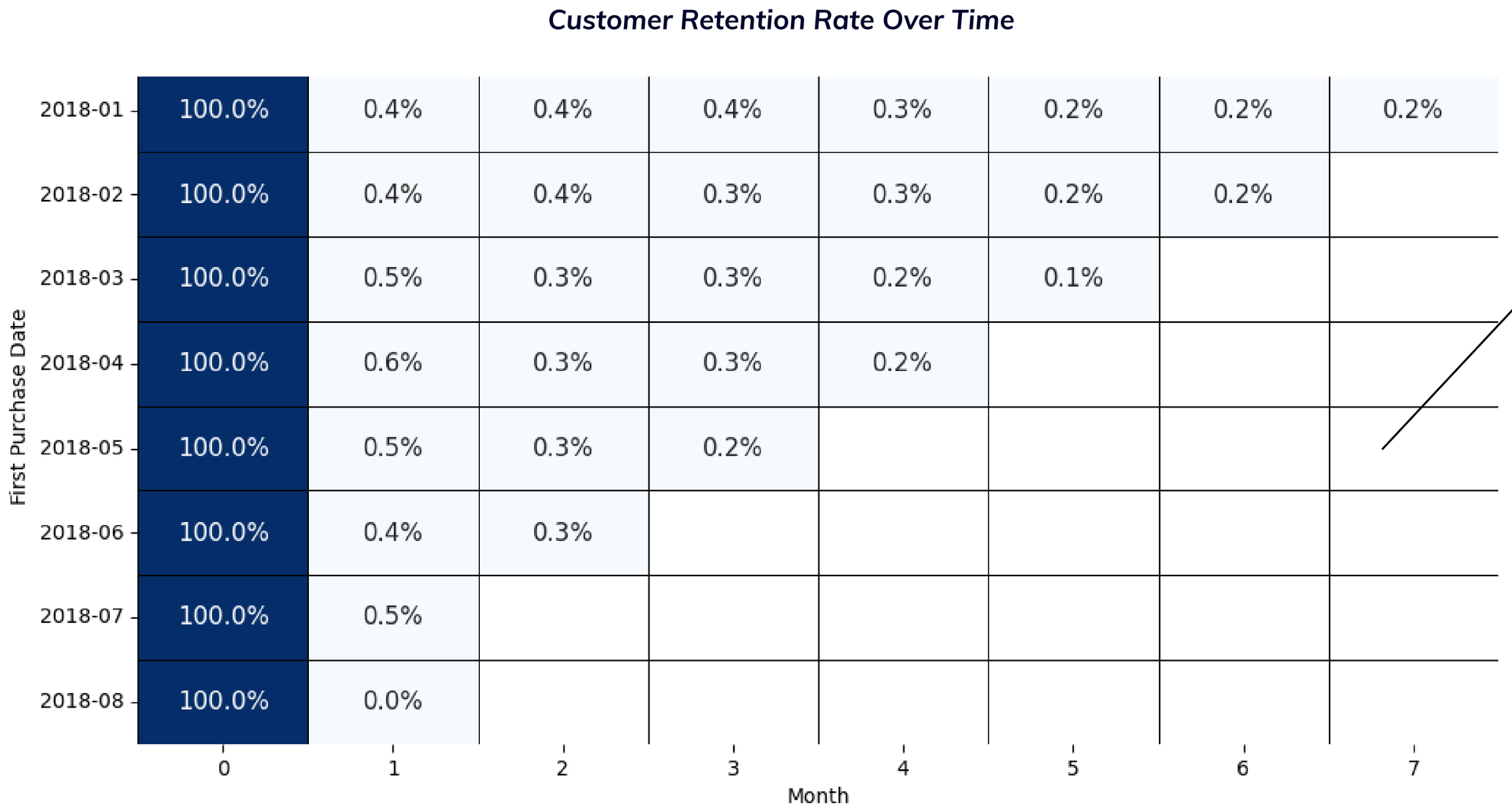


Sao Paulo is the city with the **highest concentration** of Olist customers. The product lines that are most in demand in this area are bed bath tables and health beauty products.



Since Sao Paulo is the area with the highest concentration of customers and strong demand for bed bath tables, it is important to **leverage the current momentum by developing a marketing/promotion strategy targeting Sao Paulo** with bath table and health beauty products to attract more new customers.

Only about 0.4% of customers want to return to make a purchase after their first purchase.

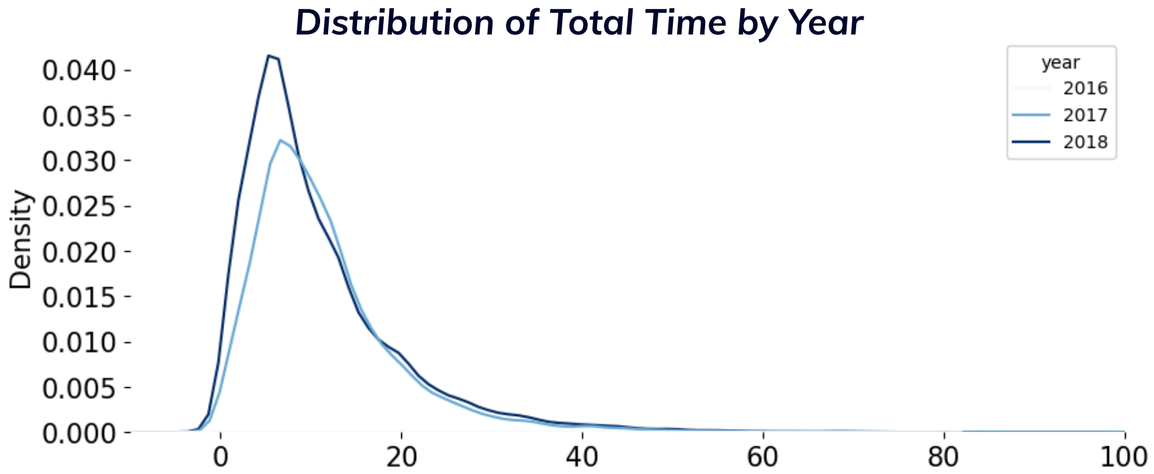


Olist is not doing well in retaining customers and needs to take measures to improve the customer experience, including:

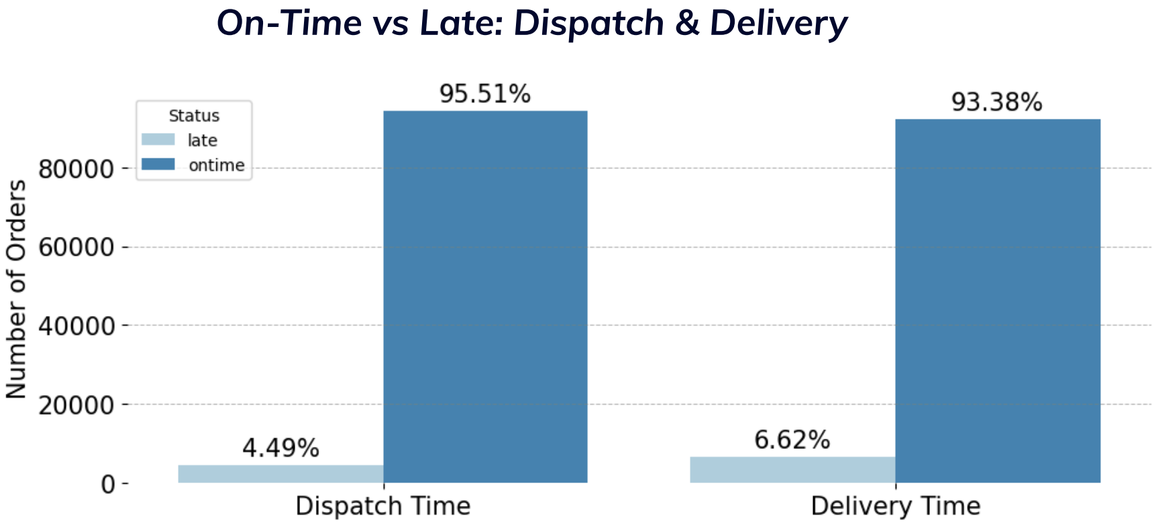
Dispatch time & Delivery time

Payment process

Olist's average dispatch & delivery time has improved but still takes 10 days, which does not meet customer expectations.



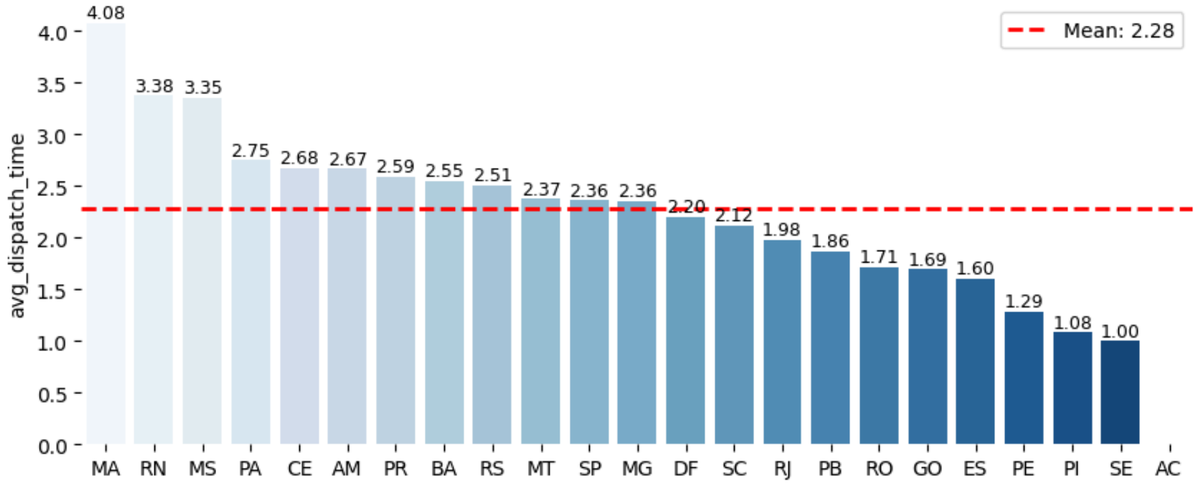
From 2016 to 2018, Olist made improvements in total time (dispatch time & delivery time), with a relatively low rate of late deliveries compared to expectations. However, the average delivery time remains as long as 10 days per order, falling short of customer expectations for speed.



To enhance customer satisfaction, Olist needs to further reduce order processing time by:
- Optimizing the order preparation process.
- Improving delivery efficiency.

Analyze the reasons for the long order dispatch times in MA, RN, and MS, and issue warnings to high-revenue merchants with a high rate of late orders to reduce order dispatch time.

Average Dispatch Time by Seller State



Top 5 Sellers with Highest Late Rate and Orders

seller_id	count_order	revenue	late_rate
835f0f7810c76831d6c7d24c7a646d4d	44	6287.570000	59.090909
54965bbe3e4f07ae045b90b0b8541f52	78	13531.060000	42.307692
1da366cade6d8276e7d8beea7af5d4bf	43	8722.760000	41.860465
a49928bcdf77c55c6d6e05e09a9b4ca5	98	10664.580000	40.816327
a7f13822ceb966b076af67121f87b063	75	15595.120000	40.000000

The three states with long order dispatch times are MA, RN, and MS(because some orders take 10-20 days to prepare).

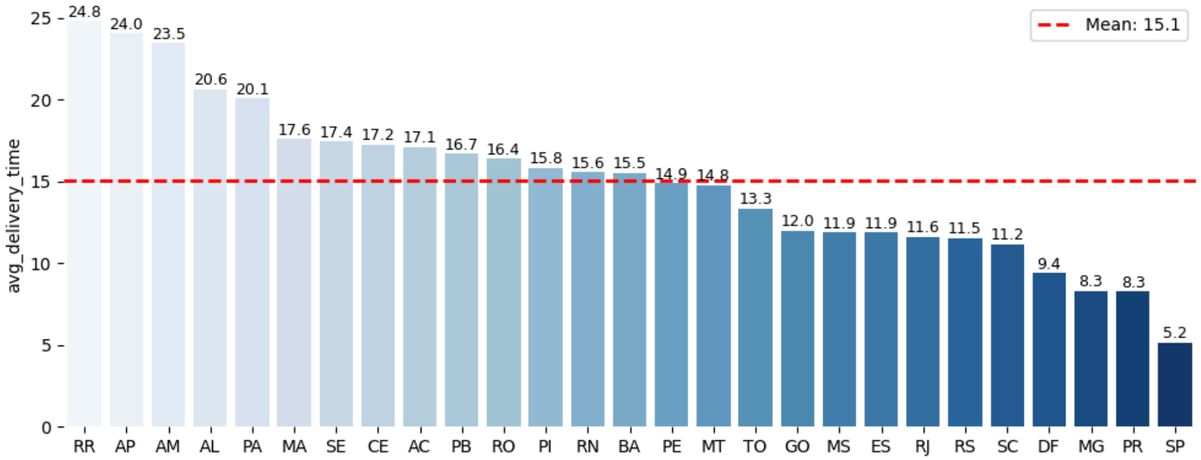
Some merchants have good revenue and order volume but have **30-60% of their orders** with dispatch times longer than expected.

In some states with order dispatch times longer than the average, particularly MA, RN, and MS, Olist should investigate the reasons behind this to implement measures to reduce dispatch times.

Additionally, for merchants with good revenue and order volumes but a high rate of delayed order preparation, Olist should issue warnings to help them improve customer satisfaction.

Prioritize allocating shipper resources to areas with long delivery times (RR, AP, AM, AL, PA) and high late delivery rates (CR, PI, SE, MA, AL) to reduce delivery times and improve transportation efficiency.

Average Delivery Time by Customer State



The average delivery time for Olist is 15 days, with five states having long delivery times: RR, AP, AM, AL, and PA.

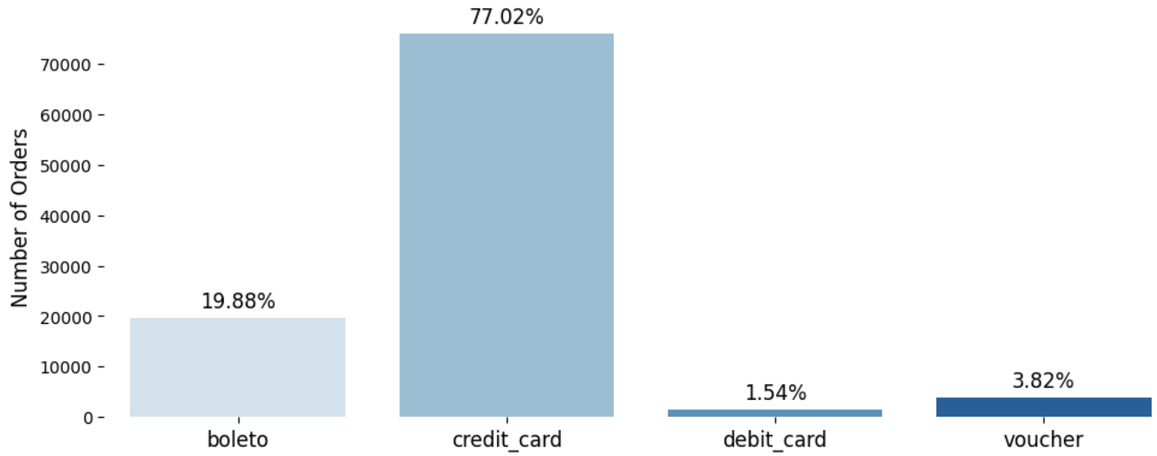
Some states have a late delivery rate of up to 20.5% compared to the expected time: CR, PI, SE, MA, and AL.

Proportion of Late and On-time Deliveries by Customer State

In regions with long delivery times and high late delivery rates, Olist needs to prioritize allocating shipper resources to reduce delivery times and enhance delivery efficiency.

Optimize the credit card payment method, the most preferred option, to ensure a smooth experience and minimize errors to below 0.5%.

Number of Orders by Payment Type



Most orders are paid for with credit cards, and the second most used payment method is boleto, which has a relatively low error rate, below

0.5%

Payment Method Failure Rate

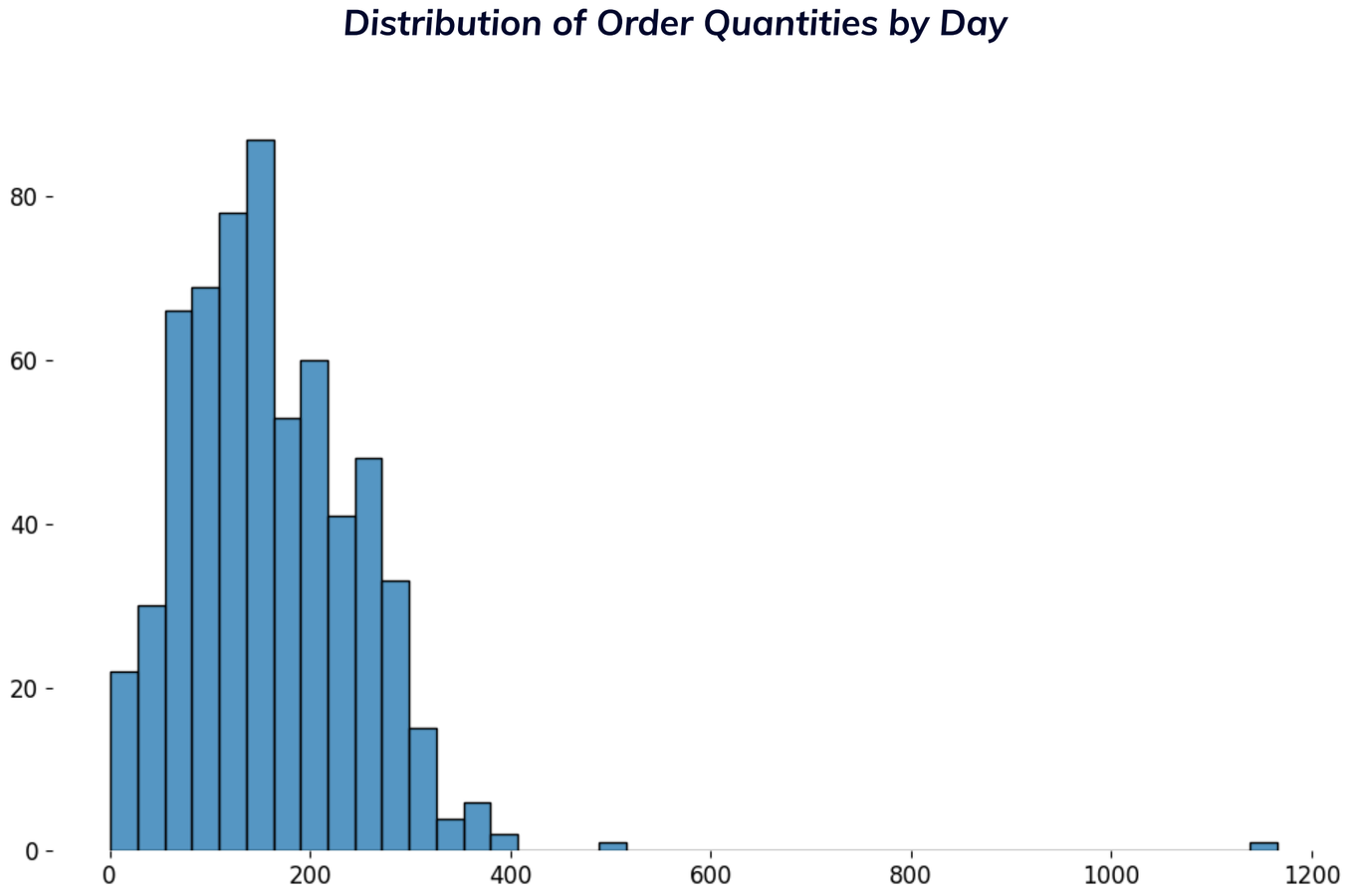
payment_type	count_order	revenue	failure_rate
boleto	19614	2842240.160000	0.402774
credit_card	75991	12434000.780000	0.496111
debit_card	1521	215105.730000	0.394477
voucher	3766	337388.700000	0.557621

Since Credit Card is the most preferred payment method, Olist should focus on optimizing this payment option to ensure customers can use it easily and with minimal errors.

Opportunity 2

INCREASE ORDER FREQUENCY

The number of orders per day mainly ranges from 50 to 200 orders. Improvements are needed to achieve a 10% revenue growth.

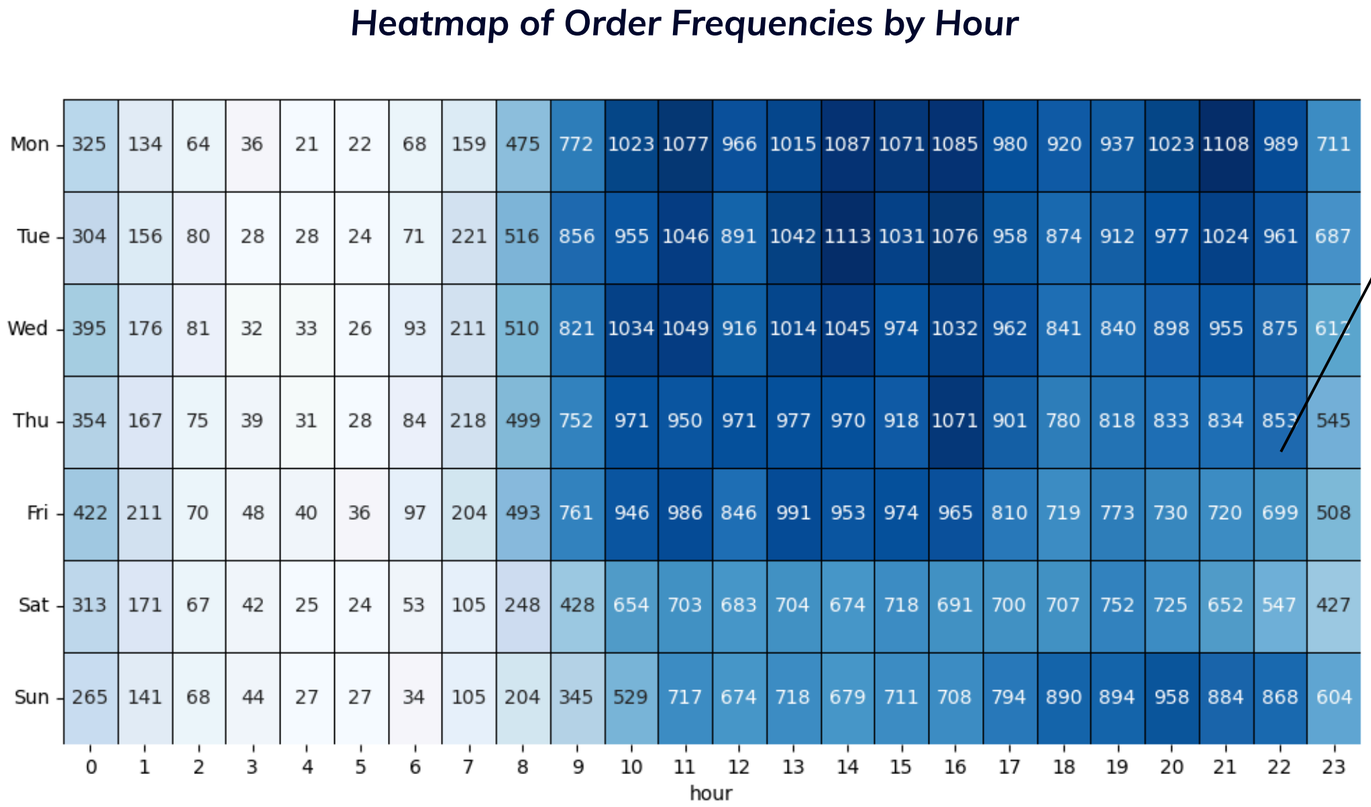


The number of orders per day ranges from 0 to 1,200 orders, but the focus is mainly on the range of 50 to 200 orders.

Olist needs to encourage customers to make more purchases in order to achieve the goal of a

10% revenue growth in the coming year.

Enhance email/push notification campaigns during peak hours (10 AM - 4 PM and 8 PM - 9 PM) and optimize the customer support team to drive more orders.



The order volume tends to be higher during two time slots: 10 AM - 4 PM (slightly decreasing at 12 PM) and 8 PM - 9 PM.

Saturdays and Sundays have fewer orders compared to other days.

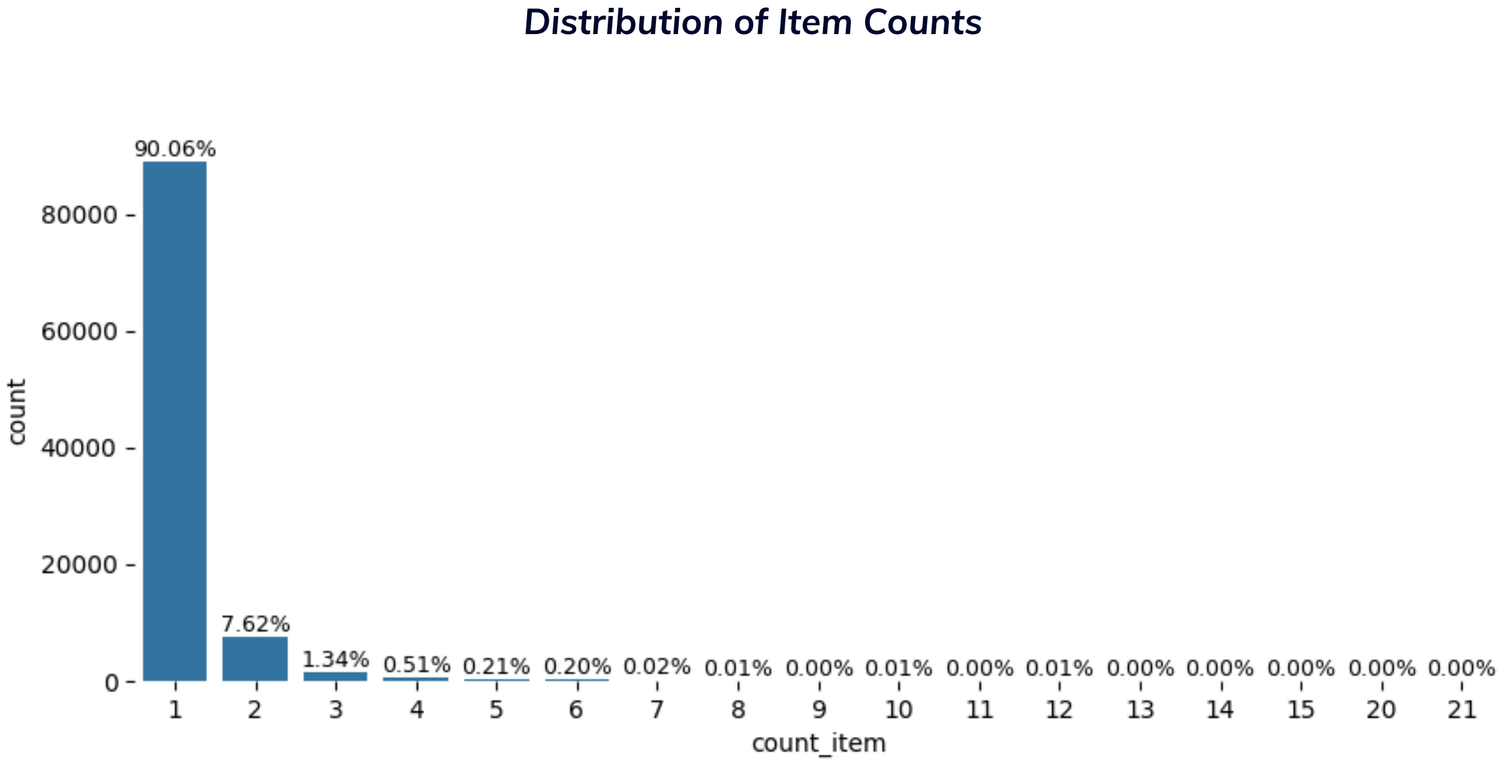
To take advantage of the time periods when customers tend to place the most orders, Olist should **send emails/push notifications during those times** to encourage users to make additional purchases.

Additionally, the customer support team should be optimized during peak hours.

Opportunity 3

INCREASE # ITEMS/ORDER

Encourage customers to purchase more products per order by offering volume discounts and suggesting relevant product bundles.



As 91%

of customers purchase fewer than 2 products with most users currently buying only 1 product per order

Olist should create incentives for customers to buy more items through offers and promotions to increase the number of items sold per order. For example, "Buy 3 products, get X$ off.

Additionally, product suggestions or bundles can be offered to encourage customers to purchase additional products once they have already chosen a specific item.

Key Recommendations

Increase Customers

- Strengthen marketing and promotional strategies in Sao Paulo (the area with the highest demand) for bath table and health beauty products to attract new customers.
- Analyze the reasons for long order dispatch times in MA, RN, and MS, and issue warnings to high-revenue merchants with a high rate of late orders to reduce dispatch times.
- Prioritize allocating shipper resources to regions with long delivery times (RR, AP, AM, AL, PA) and high late delivery rates (CR, PI, SE, MA, AL) to reduce delivery times and improve transportation efficiency.
- Optimize the credit card payment method, the most preferred option, to ensure a smooth experience and minimize errors to below 0.5%.

Increase Order Frequency

- Enhance email/push notification campaigns during peak hours (10 AM - 4 PM and 8 PM - 9 PM) and optimize the customer support team to drive more orders.

Increase # Item/Order

- Encourage customers to purchase more products per order by offering volume discounts and suggesting relevant product bundles.