ECommerce

Olist connects small businesses to Brazil's top marketplaces without contracts or hassles.
Sellers list products on the Olist Store and ship directly with Olist's logistics partners.

Objectives:

Explore key e-commerce patterns:

- Payment types
- Geographic reach
- Product categories
- Peak order times

Build a regression model to predict freight costs based on order weight.

Dataset includes:

Brazilian e-commerce order data from Olist Store, with 100K+ orders from 2016-2018 across multiple marketplaces.

It covers:

- Order status, price, payment, and freight details
- Customer locations and product attributes
- Customer reviews and satisfaction
- Geolocation data linking ZIP codes to lat/lng coordinates

Data Limitations

- Missing geolocation data led to approximation using nearest valid location
- 676 delivered orders had no review scores
- Up to 830 records had missing values – likely from incomplete or failed orders

ECommerce

Introduction | Payment Type | Geographical Reach | Item Categories | Freight | Busy Hours | Conclusion & Next Steps

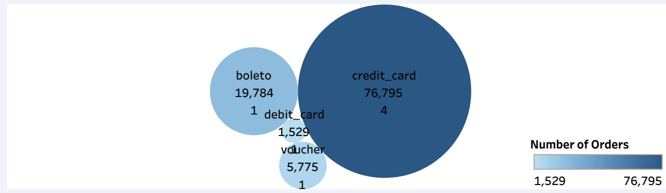
Insights & Opportunity

Credit card is the most common payment method
Installments are exclusively made via credit card, averaging **4 payments**

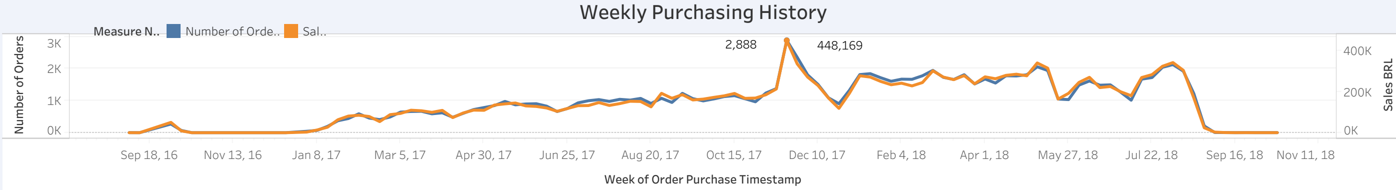
Opportunity: Launch an **Olist-branded credit card** with installment plans

Seasonal Insight
Notable sales spike during **Black Friday (Nov 2017)**:
• **448K BRL** in revenue..

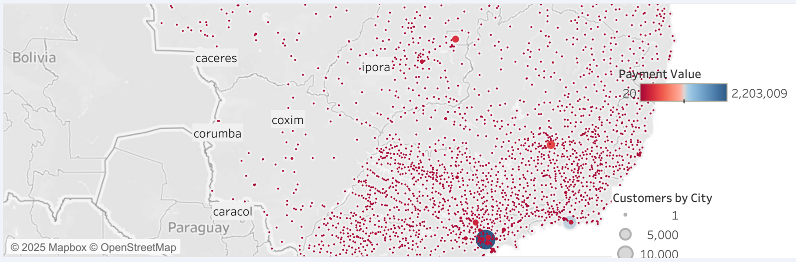
Payment Types



Weekly Purchasing History



ECommerce



Interactive Map Insight
The map above shows **customer cities**, highlighting **revenue** and **order size** across Brazil.

Strategic Recommendations

- **Enlist sellers** in underserved inland and northern regions
- **Add a distribution center** to manage high-demand inventory in those areas
- **Inform sellers** about local warehouse needs based on order volume..

Customer Distribution
São Paulo leads with **43K customers**, followed by Minas Gerais with **14K**. Other states have up to **7K**, reflecting Brazil's population distribution.

Seller Locations
Most sellers are concentrated in **São Paulo (73K)**. Several states show **little to no seller presence**, especially those in dark blue.
..

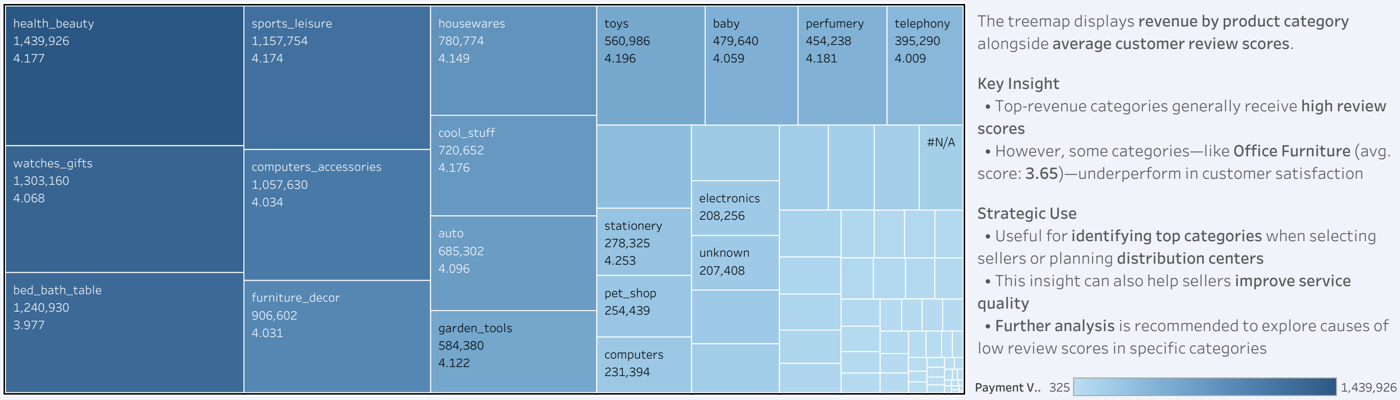
Revenue by State
Total sales value is highest in **São Paulo (~5M BRL)** and **Minas Gerais (>3M BRL)**, indicating strong customer bases and purchasing activity.

Average Spend per Customer
When normalized by customer count, **smaller states show higher average spend**. This may suggest demand for **high-ticket items** not readily available locally...

ECommerce

Introduction | Payment Type | Geographical Reach | Item Categories | Freight | Busy Hours | Conclusion & Next Steps

Which category items sell more? What is their review score?



Category Performance Treemap

The treemap displays **revenue by product category** alongside **average customer review scores**.

Key Insight

- Top-revenue categories generally receive **high review scores**
- However, some categories—like **Office Furniture** (avg. score: **3.65**)—underperform in customer satisfaction

Strategic Use

- Useful for **identifying top categories** when selecting sellers or planning **distribution centers**
- This insight can also help sellers **improve service quality**
- **Further analysis** is recommended to explore causes of low review scores in specific categories

Freight Cost vs. Product Weight

Key Findings

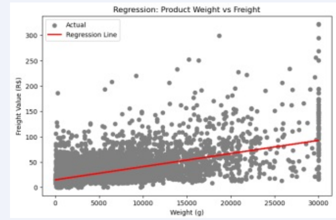
- A **linear regression model** was built to predict freight cost based on product weight
- **Slope:** R$0.0026 per gram
- **RMSE:** ~R$12.50 — average prediction error
- **R^2 Score:** ~0.38 — **38% of variation** in freight cost is explained by weight

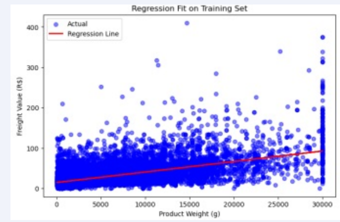
Model Performance

- **Test RMSE:** R$12.69 (slightly higher, shows minor generalization error)
- **R^2 Consistency:** 0.38 (train) vs. 0.39 (test) — indicates **good generalization**
- **Conclusion:** Weight impacts freight, but **other factors** (e.g., size, distance, carrier) matter too

..

Metric	Training Set	Test Set
Slope	0.0026	0.0026
RMSE (R$)	12.50	12.69
R^2 Score	0.382	0.385



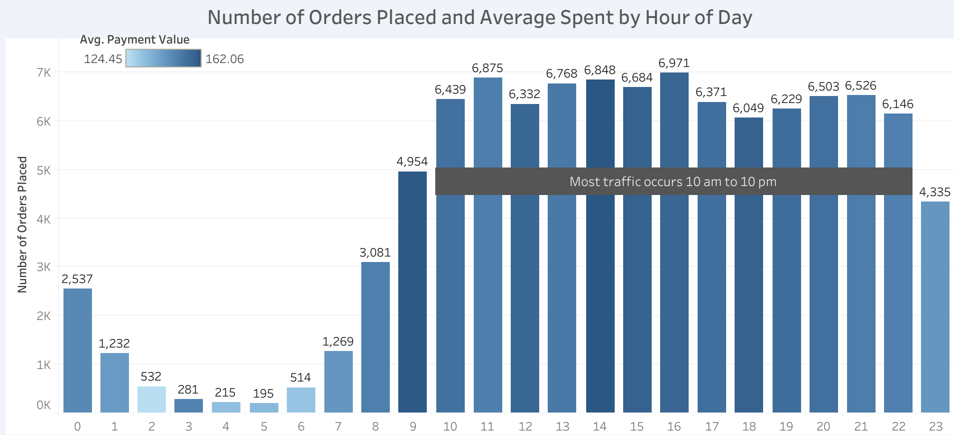


ECommerce

Introduction	Payment Type	Geographical Reach	Item Categories	Freight	Busy Hours	Conclusion & Next Steps

Order Timing & Revenue Insight

- Most orders are placed between **10 AM and 10 PM**, indicating a strong **marketing opportunity window** for sellers during this time.
- The **bar chart** also shows **average revenue per order by hour**, ranging from **124 to 162 BRL**—a **narrow gap**, suggesting **purchase amount is consistent throughout the day**.
- Slight peaks at **9 AM and 2 PM** show the **highest average order values**, which could be leveraged for **targeted promotions**.



ECommerce

Conclusion

Insights from Olist's data reveal key trends in customer behavior, product performance, and regional demand.

A regression model showed a moderate relationship between freight cost and product weight, suggesting room for improvement by incorporating additional features like volume and shipping location.

Opportunities include:

• Launching an Olist credit card
• Expanding seller presence in underserved regions
• Improving freight prediction with more features

Altogether, these findings provide a strong foundation for data-driven decisions to improve customer satisfaction, seller performance, and overall operational efficiency.
..

Next Steps

Enhance Freight Model
Include product dimensions, shipping distance, and delivery method to improve prediction accuracy.

Seller Strategy
Identify high-potential regions with low seller presence and recommend seller onboarding or distribution hubs.

Customer Segmentation
Cluster customers by behavior (e.g., spend, location, review score) to tailor marketing and service.

Category Deep-Dive
Analyze low-performing categories (e.g., Office Furniture) to uncover drivers of poor reviews.

Credit Strategy..