

Analyzing Business Problem

Specialization: Payments/Monetization project

Task requirements

You should use tables from olist_db to come up with an analysis on your own. You will have to write SQL queries to extract needed data, then analyze and make a visualization using your preferred data visualization tool (Google Sheets / Tableau / Looker Studio). After all this, you will have to come up with a presentation oriented to executives from your company. Have in mind your audience for the presentation, these will be business-oriented people who are not data analysts and may not have studied statistics before.

- Write SQL queries to extract needed data, these queries should be well documented, with indentation and provided in separate file.
- Provide your analysis in preferred data visualization tool (Google Sheets / Tableau / Looker Studio), have in mind that this file should contain all the things you tried to look at even if they did not bring any useful insights.
- Provide presentation with around 10 slides which would be business focused and would present only the most interesting 3-5 insights.
- See whether you can apply at least 2 techniques learned in this or previous sprints.
- Explore the data. See whether there are interesting data points that can give more insights to your presentation.
- Provide analytical insights, what are the drawbacks of this analysis, what further analysis could you recommend?
- Do not forget how you should structure your presentation which you learnt during the first sprint of the first module.

Introduction

In the dynamic landscape of e-commerce, where swift and efficient operations can make all the difference, Olist stands as a pivotal player in the Brazilian market. As the driving force behind an online platform connecting merchants and their products to the nation's primary marketplaces, Olist empowers small businesses to expand their market reach across Brazil. To delve into the intricacies of Olist's operational efficiency, it becomes imperative to explore key metrics such as delivery times, revenue per product, and the geographical destinations to which sellers dispatch their products.

By scrutinizing these vital aspects, Olist can uncover insights that not only enhance its understanding of the company's overall performance but also pave the way for strategic decisions that optimize delivery processes, bolster revenue streams, and facilitate more targeted market expansion. This exploration into the intricacies of delivery times, revenue generation, and seller destinations becomes a linchpin in Olist's quest for sustained growth and market dominance in the ever-evolving e-commerce realm.

In this analysis, we'll break down our methods and share insights on how fast sellers deliver, the money they make, the products they sell, and where those products are heading.

1. Main Questions

Understanding sellers' delivery times, revenue, product offerings, and market destinations is crucial for Olist's success in the dynamic e-commerce landscape. Efficient deliveries, robust revenue streams, diverse product portfolios, and strategic seller outreach are pivotal factors. This analysis aims to uncover insights that will inform strategic decisions, strengthen support for sellers, and contribute to Olist's sustained growth in the competitive market.

To enhance the business's understanding of seller behavior, delivery times, delivery locations, products, and revenue, I raised specific questions aimed at gaining insights from both a delivery time and revenue perspective.

Main questions:

1. What is the average order value and how does this vary by product category or payment method?
2. Which product categories have the highest revenue on Olist, and how can the company increase revenue across different categories?
3. How profitable is Olist customers portfolio
4. How do customer ratings affect product sales performance on Olist?

2. Prepare & Process

In the second phase, I conducted a thorough review of the data sources, systematically assessing the type of data at my disposal. Within the provided Olist datasets, I examined each table to discern the diverse information available, ranging from customer details, order specifics, and order items to seller information, delivery destinations, and the frequency and punctuality of product deliveries.

Throughout this evaluation of the data, I deepened my comprehension of the key data fields essential for extracting information crucial to analysis. Additionally, I ensured that the available dataset have information required to address the specific questions that had been raised.

Furthermore, I crafted a query to extract the preliminary information essential for my analysis. In this process, I also generated several additional fields derived from the extracted data to enhance the depth and breadth of insights that could be derived from the dataset.

I have created such queries to get necessary data from dataset:

- Average order value by product category or payment method
- Best performing product categories by revenue (Pareto rule verification)

- Customers segmentation and RFM analysis
- Ratings distribution and sales performance

Furthermore, after extracting the necessary data, I developed a Looker Studio dashboard. This dashboard is designed to aid business owners and other departments in addressing key questions more efficiently.

Link to the Looker Studio dashboard: https://lookerstudio.google.com/u/0/reporting/dc499b1c-8559-448e-a227-8761c6288e4c/page/p_tphw12h3dd

3. Analyze & Share

For the primary analysis, I have created a Looker Studio Dashboard that enables users to interact effortlessly. By simply clicking on specific time intervals, selected regions, or choosing a particular date, delivery status or seller's segments, end users can filter their results to address various inquiries related to customer deliveries and revenue questions. This user-friendly interface simplifies the process of exploring and comprehending the data.

Main Information

First of all, I created a few key metrics to provide a quick overview of the business's main information, capturing attention effectively.

Number of orders	Total revenue	Product categories with 80 % revenue	Average order value
100.2K	\$20.05M	16	\$165

How many orders do we have?

We have 100,2 K orders.

How much revenue did the seller generate?

Sellers generated around \$20,05 M with average order value of \$165

-

How many product categories generates 80% of the revenue ?

19 product categories generates 80% of revenue.

Main Business questions

1. What is the average order value and how does this vary by product category or payment method?

The average order value is 165\$ and it varies depending on product category and payment method

By product category

Categories such as computers, fixed telephony, and small appliances (home oven and coffee) generate high revenue per order, indicating they consist of higher-priced or premium products.

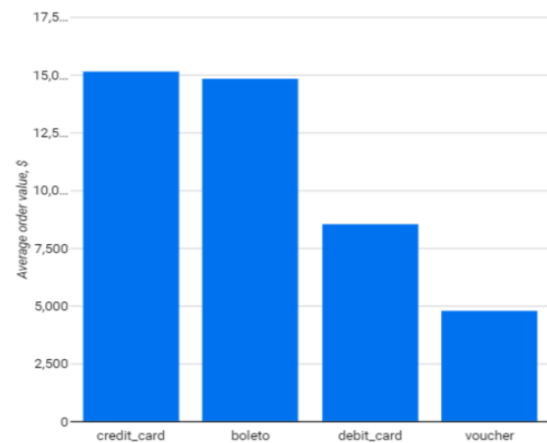
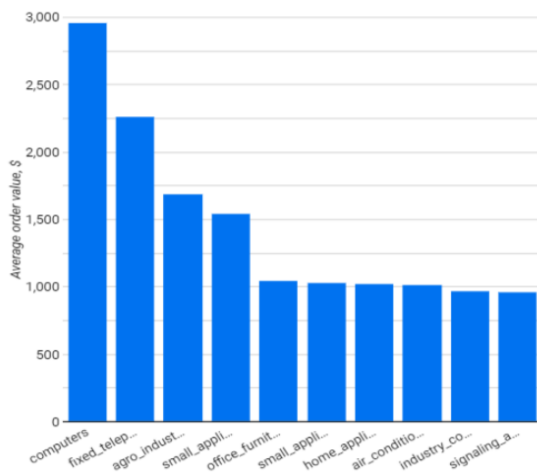
By payment method

"Credit card payments have the average order value (AOV) relatively high, suggesting that customers paying with credit cards tend to spend more per transaction.

Boleto payments have the AOV comparable to credit card payments, suggesting similar spending patterns among customers who choose this payment method.

Debit card payments AOV is lower, indicating that customers paying with debit cards tend to spend less per transaction. This payment method may be less popular or used for smaller purchases.

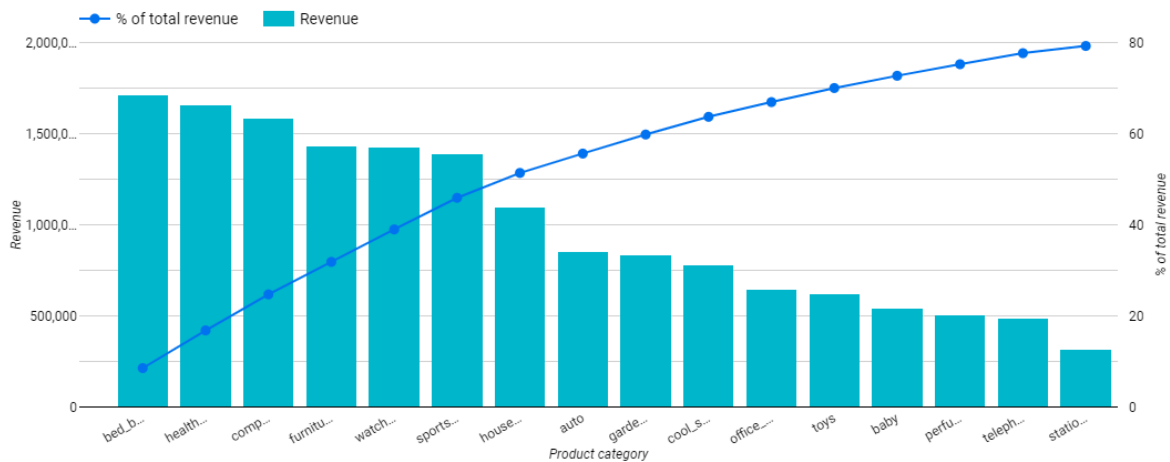
Voucher payments AOV is significantly lower compared to other payment methods, indicating that customers using vouchers tend to make smaller purchases. Vouchers may be used for discounts or promotions, attracting price-sensitive customers."



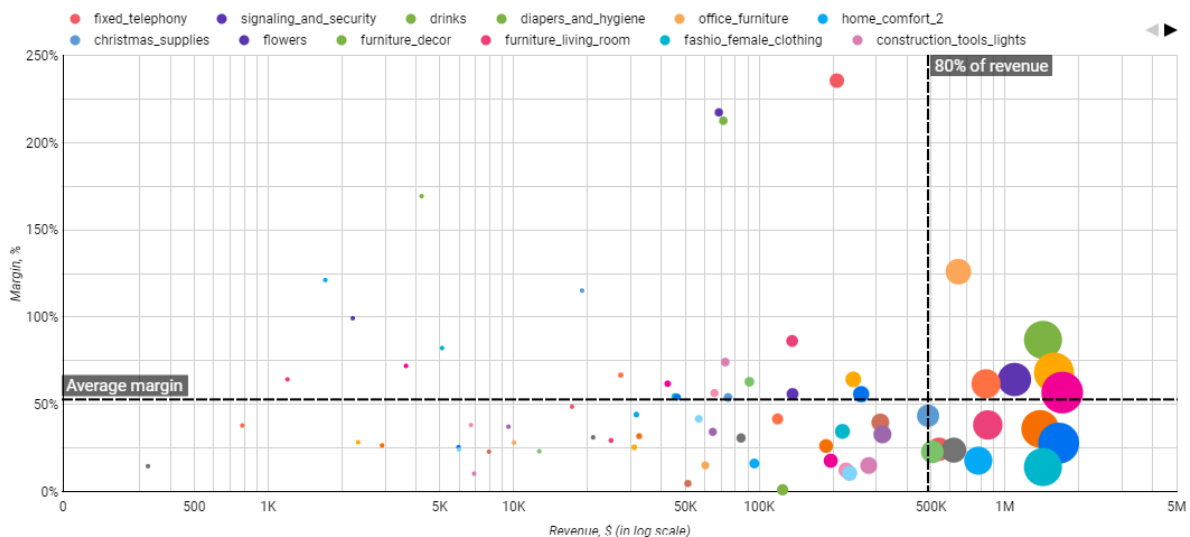
2. Which product categories have the highest revenue on Olist, and how can the company increase revenue across different categories?

From a product standpoint, we offer 71 unique products, but the Pareto chart reveals that only 19 products (25%) generate over 80% of our revenue. Focusing on promoting and advertising these top 19 products is recommended for optimizing our revenue.

Additionally, customers often purchase from multiple categories in one order. Utilizing popular product pairs can enhance the shopping experience by providing targeted suggestions when customers add items to their carts. For instance, suggesting another product from a popular pair when one is added can encourage complementary purchases.



Based on the provided data, the top product categories contributing to 80% of the total revenue include: health & beauty, watches & gifts, bed & bath table, sports & leisure, computers & accessories, furniture & decor, housewares, cool stuff, auto, and garden tools. For increasing revenue it is recommended to develop strategies to boost sales of TOP products by revenue



High Margin, High Revenue Quadrant (Top Right Quadrant):

Product Categories: "signaling_and_security", "food", "construction_tools_safety", "watches_gifts", "telephony"

Characteristics: These categories have both a relatively high margin percentage and contribute significantly to the total revenue. They are considered the most beneficial categories as they generate high profits while contributing substantially to the overall revenue.

High Margin, Low Revenue Quadrant (Top Left Quadrant):

Product Categories: "small_appliances_home_oven_and_coffee", "bed_bath_table", "computers", "fashion_shoes", "tablets_printing_image"

Characteristics: These categories have a relatively high margin percentage but contribute less to the total revenue. They might represent niche or specialized products with higher profit margins but lower sales volume compared to other categories.

Low Margin, High Revenue Quadrant (Bottom Right Quadrant):

Product Categories: "computers_accessories", "cool_stuff", "christmas_supplies", "small_appliances", "luggage_accessories"

Characteristics: These categories have a lower margin percentage but contribute significantly to the total revenue. They represent high-volume products or categories with lower profit margins but substantial sales.

Low Margin, Low Revenue Quadrant (Bottom Left Quadrant):

Product Categories: "construction_tools_construction", "home_appliances", "market_place", "furniture_bedroom", "music"

Characteristics: These categories have both a relatively low margin percentage and contribute less to the total revenue. They might include products with low-profit margins and lower sales volume compared to other categories.

3. How profitable is Olist customers portfolio

Best customers These customers have made recent purchases (average recency: 104.8 days) with frequent purchases (average frequency: 2.03) and a high monetary value (average monetary value: \$775.06). The recommendation is to reward these customers with exclusive offers, loyalty programs, or personalized discounts to encourage continued engagement and foster loyalty.

Can't loose them: These customers have a longer average recency (439.4 days) but still hold a high monetary value (average monetary value: \$585.36). They make frequent purchases (average frequency: 2.01). Despite their past spending, their lack of recent activity suggests they may be losing interest or engaging with competitors. Targeted re-engagement efforts are crucial to prevent churn and retain their business.

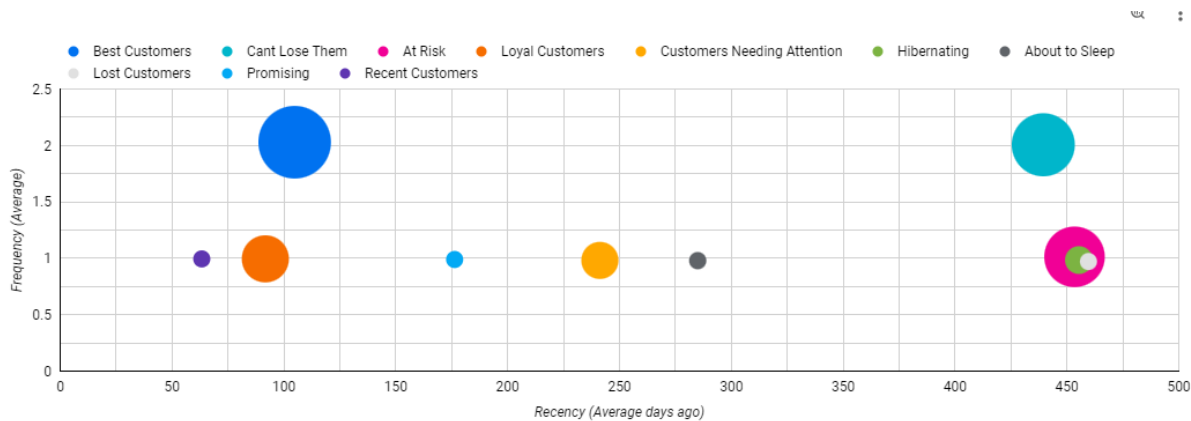
At risk These customers have a longer average recency (453.2 days) and have a moderate monetary value (average monetary value: \$537.60). Despite their past spending, their lack of recent activity suggests they may be losing interest or engaging with competitors.

Loyal Customers: These customers have made recent purchases (average recency: 91.7 days) and contribute a moderate monetary value (average monetary value: \$325.91). These are the backbone of the business, demonstrating strong loyalty and consistent engagement. Maintaining personalized communication and offering exclusive rewards can further solidify their loyalty.

Customers Needing Attention: These customers have a moderate average recency (241.2 days) and monetary value is relatively lower (average monetary value: \$202.01). While they have not completely disengaged, their recent activity suggests a potential decline. Proactive outreach and tailored offers can help reignite their interest and prevent further disengagement.

Hibernating Customers: These customers have a longer average recency (455.4 days) and their monetary value is the lowest among the segments (average monetary value: \$113.86)., indicating they have been inactive for a significant period and have low spending habits. These customers are at high risk of churn and may require targeted

win-back campaigns to re-engage them. Understanding the reasons for their inactivity is essential for effective reactivation efforts.



RFM segment	# of customers	% of customers	Recency (Average days ago)	Frequency (Average)	Monetary (Average)
1. Best Customers	384	0.4%	105	2.03	775.06
2. Cant Lose Them	168	0.2%	439	2.01	585.36
3. At Risk	6087	6.2%	453	1.01	537.6
4. Loyal Customers	24329	24.7%	92	1	325.91
5. Customers Needing Attention	30651	31.2%	241	0.98	202.01
6. Hibernating	12139	12.3%	455	0.98	113.86
7. About to Sleep	6020	6.1%	285	0.98	43.44
8. Lost Customers	6331	6.4%	460	0.97	43.2
9. Promising	6173	6.3%	176	0.99	43.12
10. Recent Customers	6072	6.2%	63	0.99	43.01

About to Sleep Customers: These customers have a moderate average recency (284.9 days) and their monetary value is the lowest among the segments (average monetary value: \$43.44). Despite recent activity, their low spending suggests a lack of engagement. Targeted promotions or product recommendations may encourage additional purchases before they become completely inactive.

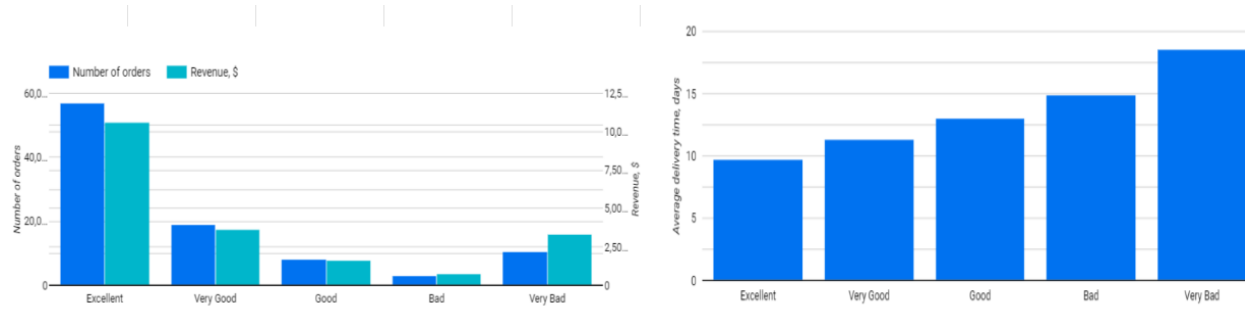
Lost Customers: These customers have the longest average recency (459.6 days), indicating they have likely churned or lost interest, and monetary value is also low (average monetary value: \$43.20). Re-engaging these customers may be challenging, but understanding the reasons for their disengagement and offering attractive incentives or personalized messages may help win them back.

Promising Customers: These customers have made recent purchases (average recency: 176.3 days) and their monetary value is the lowest among the segments (average monetary value: \$43.12), indicating recent activity but with minimal spending. While they have shown recent interest, their low spending suggests they may still be evaluating the products or services. Providing personalized recommendations or offers can help convert their interest into purchases.

Recent Customers: These customers have made very recent purchases (average recency: 63.3 days) and monetary value is the lowest among the segments (average monetary value: \$43.01), indicating recent but minimal engagement. These customers are new or have recently re-engaged, presenting an opportunity to nurture their

relationship early on. Tailored follow-ups and incentives can encourage repeat purchases and build long-term loyalty.

4. How do customer ratings affect product sales performance on Olist?



Number of Orders: The number of orders decreases as the rating goes from "Excellent" to "Very Bad." This indicates a correlation between customer satisfaction and the volume of orders. Though there are significant amount of orders ended up with low rating

Revenue: Despite the decrease in the number of orders, the revenue generated doesn't follow a strictly decreasing trend from "Excellent" to "Very Bad." There's a slight increase in revenue from "Excellent" to "Very Good," followed by a decrease for "Good" and "Bad," and then an increase again for "Very Bad."

Delivery time: The higher delivery the worst rating customers gives for the products. It might be a factor to investigate and improve.

5. Key Insights and recommendations

As from the analysis we now have a better picture of our sellers performance from revenue and delivery times and their behavior from various perspectives as: Segmentation, product categories and delivery status.

1. TOP performing product categories by average order value are computers, fixed telephony, and small appliances. It is recommended to ensure wider choice of those products to increase number of orders.

2. The top product categories contributing to 80% of the total revenue (Pareto rule) include: health & beauty, watches & gifts, bed & bath table, sports & leisure, computers & accessories, furniture & decor, housewares, cool

stuff, auto, and garden tools. Identify cross-selling and upselling opportunities within the top-performing product categories to increase average order value (AOV) and maximize revenue per customer.

3. Customers RFM analysis shows that most of the customers purchase only once on Olist marketplace. It also shows that 37% of customers delivering high monetary values to the company are either at risk to churn or need attention to bring them back. It is recommended to build customer re-engagement strategy and work more on user experience.

4. Product rating distribution and its impact on sales performance from one side indicates a correlation between customer satisfaction and the volume of orders but it also shows that some products with higher number of orders and revenue generates quite many negative reviews. Relationship between rating and average delivery time shows that faster delivery times typically correlate positively with positive outcomes, while slower delivery times correlate negatively with achieving desired results.