**Presentation of Findings on the Operations of Olist**

**Submitted by:**

**Group 4**

**Afsaneh Talebi**

**Karleng Lim**

**Maia Calvo**

**Table of Contents:**

1. **Overview**
2. **Presentation of Findings:**
   1. Customers
   2. Sales
   3. Reviews and Polarity Scores
   4. Delivery Performance
3. **Summary and Recommendations**

**I. Overview**

The analysts studied the problem by planning and strategizing how to best approach the project from the point of view of customers, sales, reviews and delivery performance. The team used the project proposal document and wrote down possible solutions in detail. The document served as the team’s guide throughout the project.

The team looked at the schema and inspected the contents. The team made a table in Excel containing the list of all variables of each table in the schema and put in notes on what these variables meant. The team proceeded to identify the variables to use to answer each question.

The team encountered some issues with the data. The geographic coordinates are showing up in other countries such as Italy, Spain and Mexico to name a few. This could be due to the use of VPN which could change a customer’s location, or the customer ID and records were registered in Brazil, but the customers shopped physically in another country. Reviews and comments were showing special characters. These have been cleaned for the project.

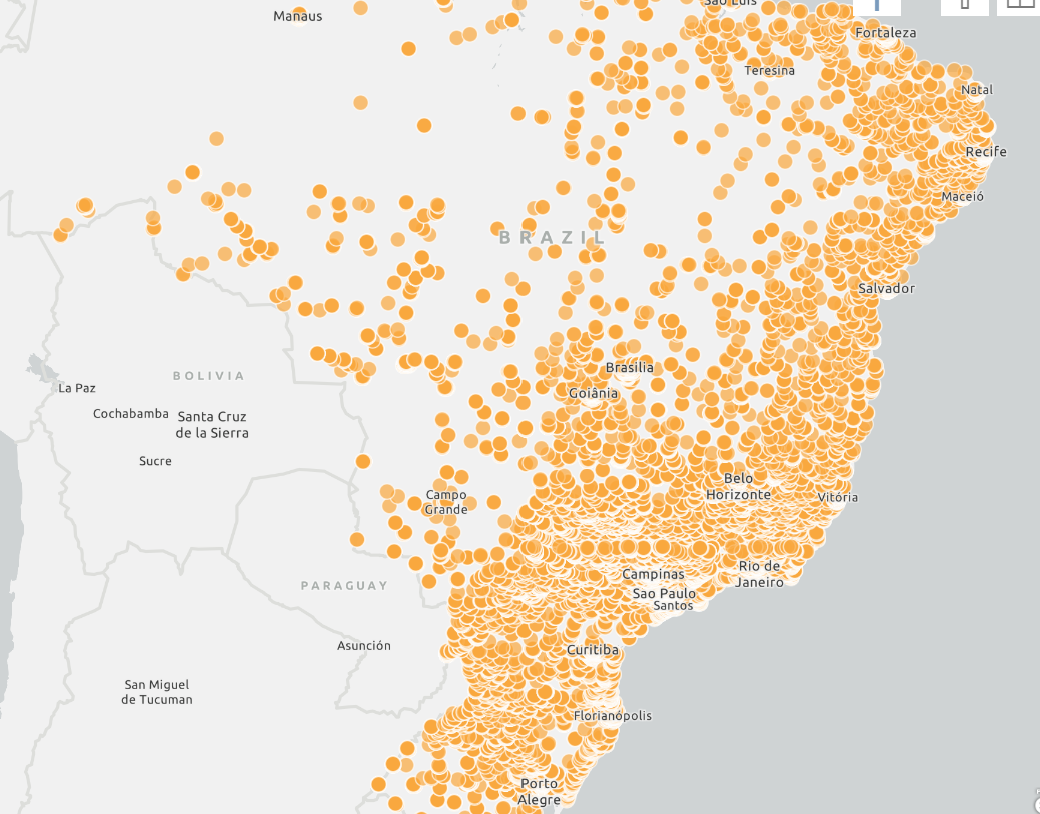
The tables in the schema were connected via the Power BI model view to ensure that the relationships between variables are used correctly.

The report will be presented in 4 parts. First, the customers and their shopping habits. Second, the sales figures and top selling product categories. Third, reviews and feedback of the customers. Fourth, the sellers and their delivery performance. Lastly, the team will present a summary and recommendations for Olist.

**II. Presentation of Findings**

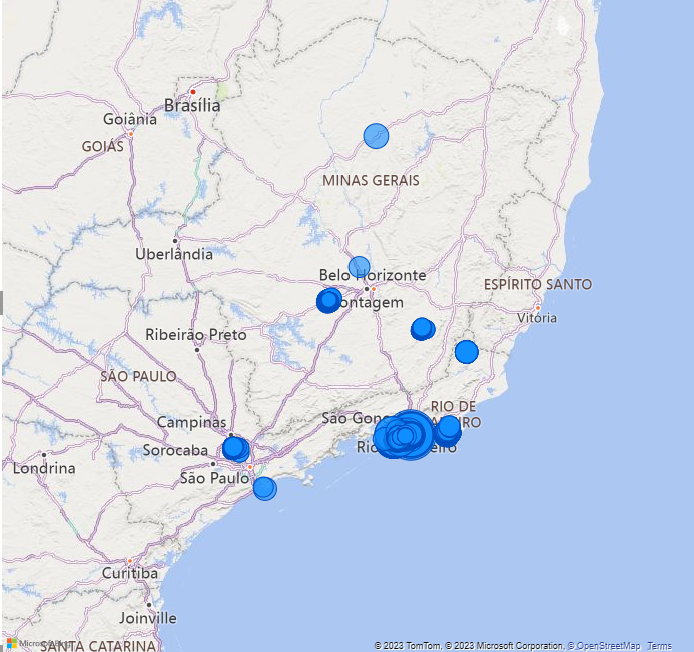
1. **Customers**

Graph 1. Location of customers in Brazil



Olist customers are highly concentrated on the south and eastern coast of Brazil. The company has fewer customers in the inland and border states.

Graph 2. Top 5 states with highest sales value



The top 5 states with the highest sales are Sao Paulo, Rio de Janeiro, Minas Gerais, Rio Grande do Sul and Parana. These 5 states also top the number of transactions.

Table 1. Top 5 highest states by sales value and number of transactions

| State name | Sales value | # of Transactions |
| --- | --- | --- |
| Sao Paulo | 5,998,227 | 41,746 |
| Rio de Janeiro | 2,144,380 | 12,852 |
| Minas Gerais | 1,872,257 | 11,635 |
| Rio Grande do Sul | 890,899 | 5,466 |
| Parana | 811,156 | 5,045 |

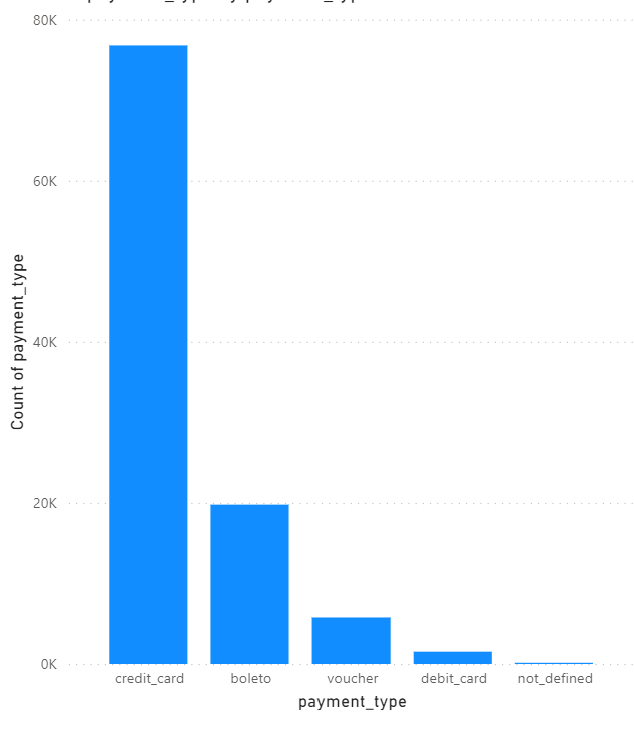
However, looking at the transaction per value tells a different story. The value per transaction per State is higher in Paraiba, Acre, Rondonia, Amapa and Alagoas.

Table 2. Top 5 States by Value per Transaction

| State name | Sales value | # of Transactions | Value/Transaction |
| --- | --- | --- | --- |
| Paraiba | 141,546 | 536 | 264 |
| Acre | 19,681 | 81 | 243 |
| Rondonia | 60,866 | 253 | 241 |
| Amapa | 16,263 | 68 | 239 |
| Alagoas | 96,962 | 413 | 235 |

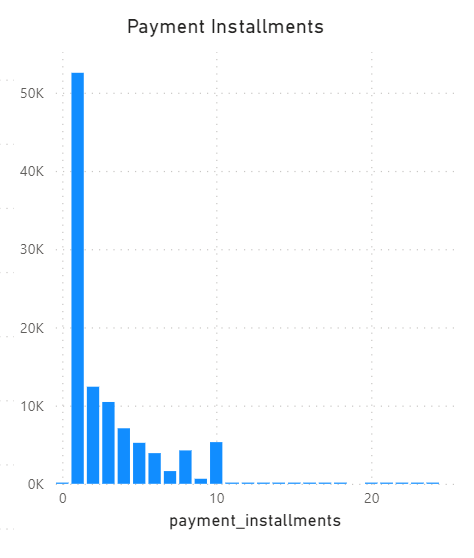
In terms of payment type, customers use credit cards the most, followed by boleto and voucher. Boleto is a popular, convenient and secure way to pay in Brazil. It behaves similarly to Bpay. The customer can pay through banks, post offices and other authorized locations.

Graph 3. Frequency by payment type



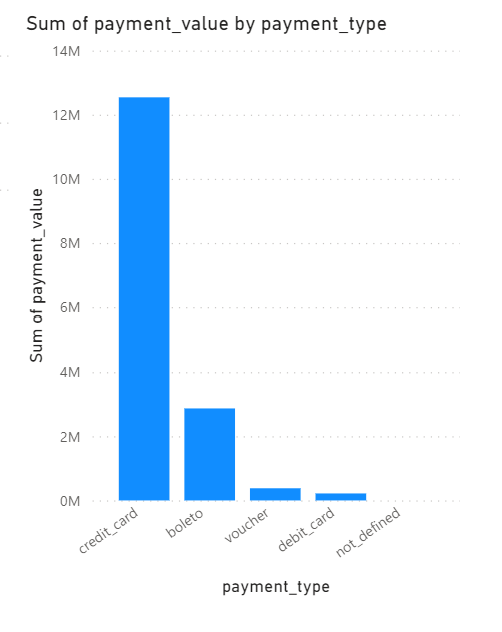
The majority of customers fall within the one to two month payment scheme.

Graph 4. Distribution of payment installments



The value of credit card payments amount to 12.54 million Real, the largest source of revenue of the company by payment type.

Graph Payment value by payment type

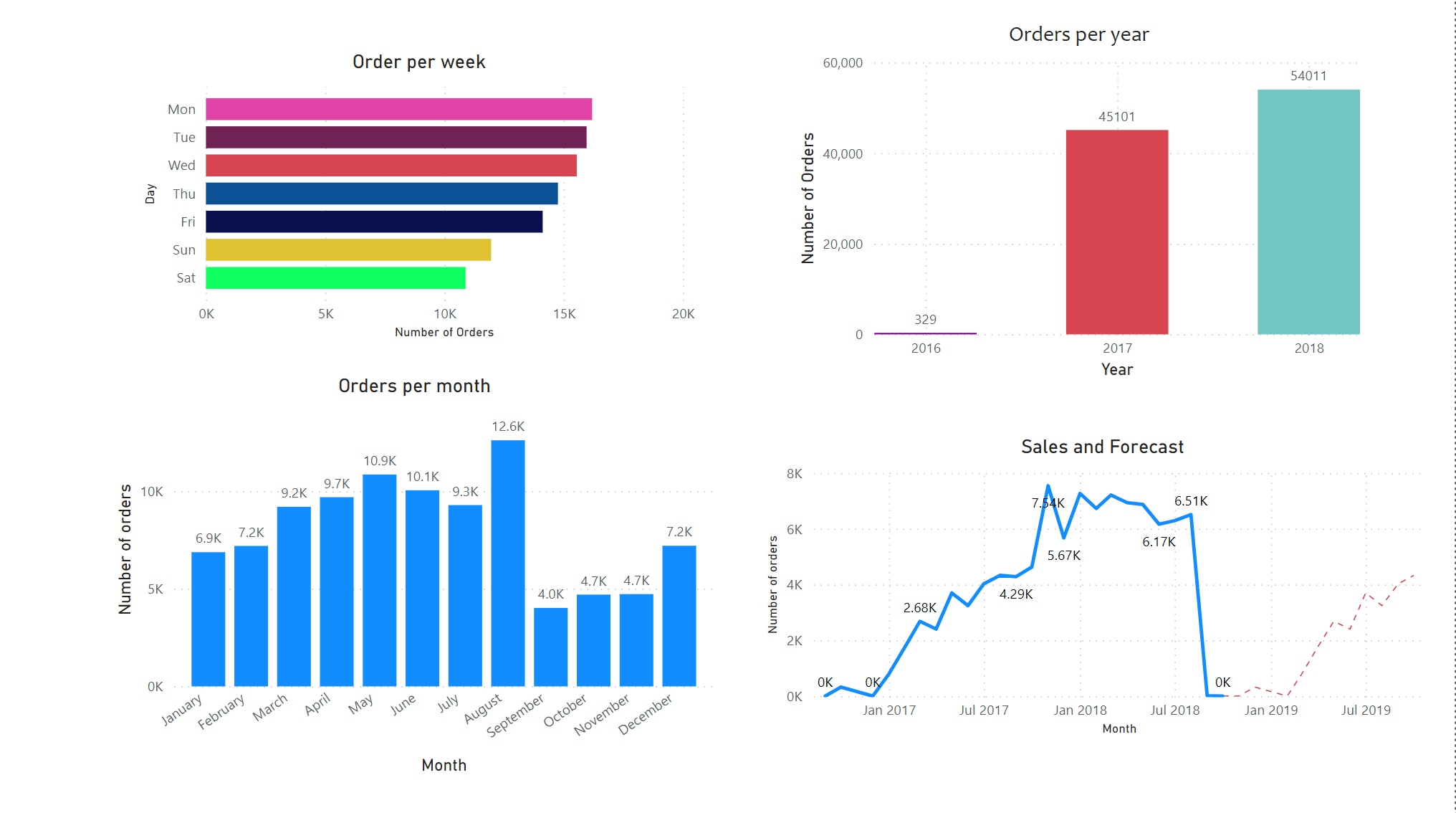


The use of credit cards and installment plans come as no surprise. While Brazil may be the largest economy in Latin America, the income per capita sits at USD 8,800 in 2022. This is around USD 730 per month or AUD 1,100 monthly.

Due to this income level and bouts of high inflation that Brazil is known for, it is crucial for the company to offer small monthly payment options to make consumption viable and maintain sales figures even in unstable economic scenarios. Businesses have to make it convenient for customers to pay as well.

1. **Sales**

Graphs 5. Sales trends according to week, month, year and forecasts

****

These charts highlight trends in the sales data by number of orders. The top left chart shows the number of orders on a daily basis. Most of the orders are done on the weekday and slow down on the weekend.

For the yearly chart, sales took off in 2017 and 2018.

In the monthly chart, there is a spike in online shopping in August and a slowdown in “Ber” months or the last quarter of the year.

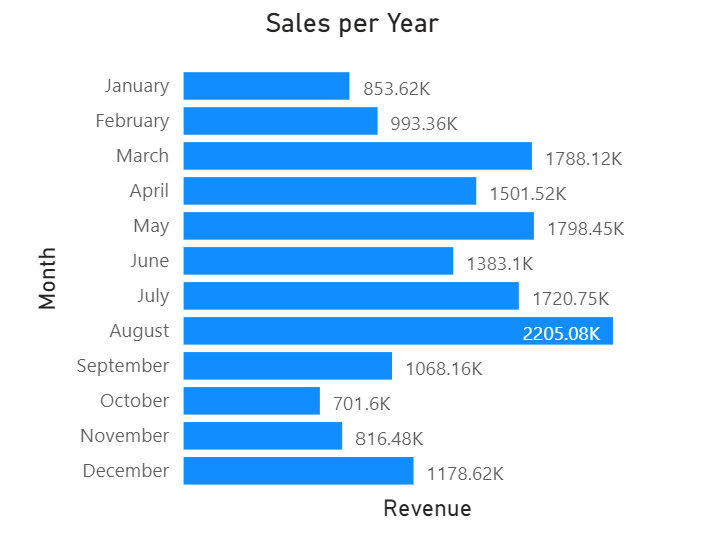
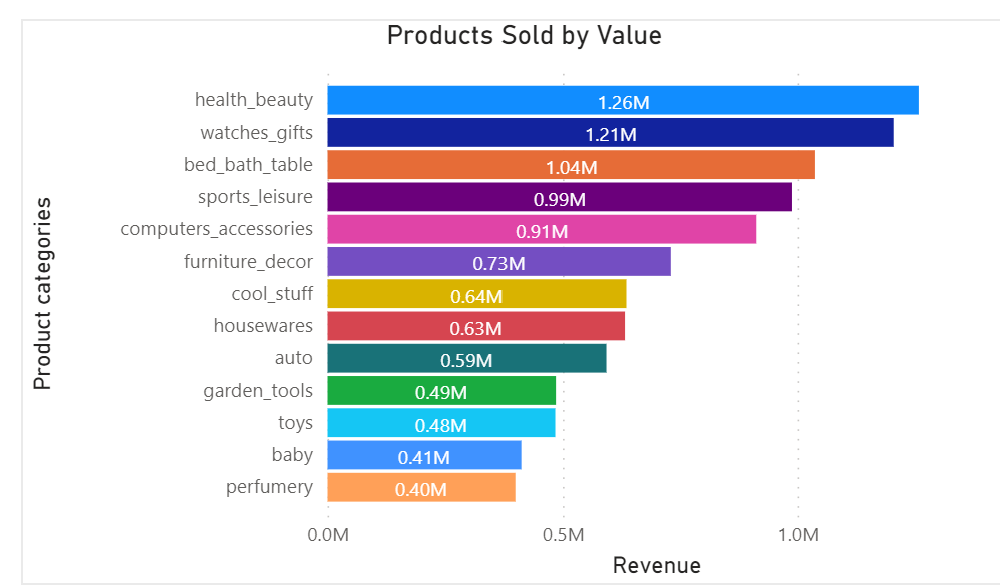
August in Brazil is a big month for sales (similar to mid-year or Black Friday sales). This explains the spike in this particular month. The months following August show slower sales due to buyers taking a break from purchases in August. Online shopping picks up again in December for the Christmas holidays.

On the sales and sales forecast chart, there is a big drop in the number of orders. The team investigated the data and it was a big event that took place in September that caused security issues in Brazil when the presidential frontrunner Bolsonaro was stabbed. Elections were heated and huge protests and demonstrations from anti-Bolsonaro filled 27 states while pro-Bolsanaro occupied 27 states.

The forecast shows the number of orders picking up after a drop in September and October.

Due to the limitations of PowerBi in making forecasts, the team could have deleted the last data point. However, the September and October figures of 2018 play an important role as these help understand how large of an impact political instability has on sales. The team recommends using a more sophisticated forecasting tool where variables can be used to capture these events and form part of statistical regression techniques should the seasonal index fail in mimicking and forecasting historical sales.

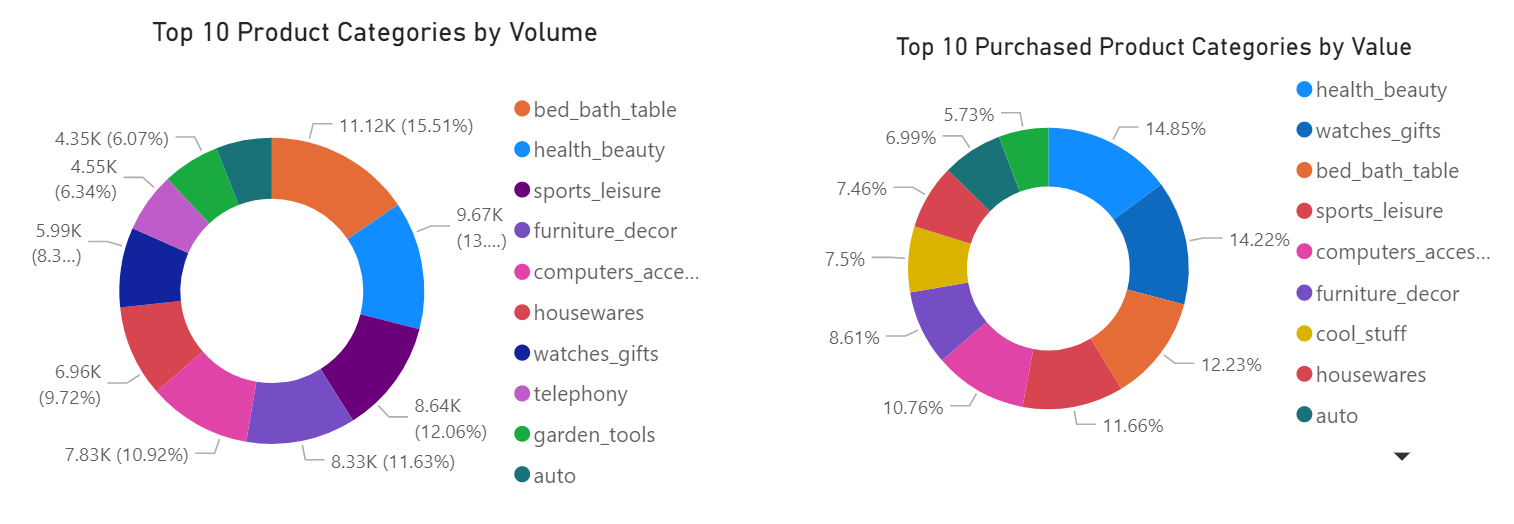
Graph 6. Top Products sold by value and Monthly Sales

****

The majority of customers of Olist shop for health and beauty, watches and gifts, bed, bath and table, sports and leisure, and computer accessories. These categories top the orders in terms of volume and value.

Out of the 23 states, 6 states show that health and beauty are the top sellers. These states are Ceara, Distrito Federal, Maranhao, Mato Grosso do Sul, Rio de Janeiro and Santa Catarina.

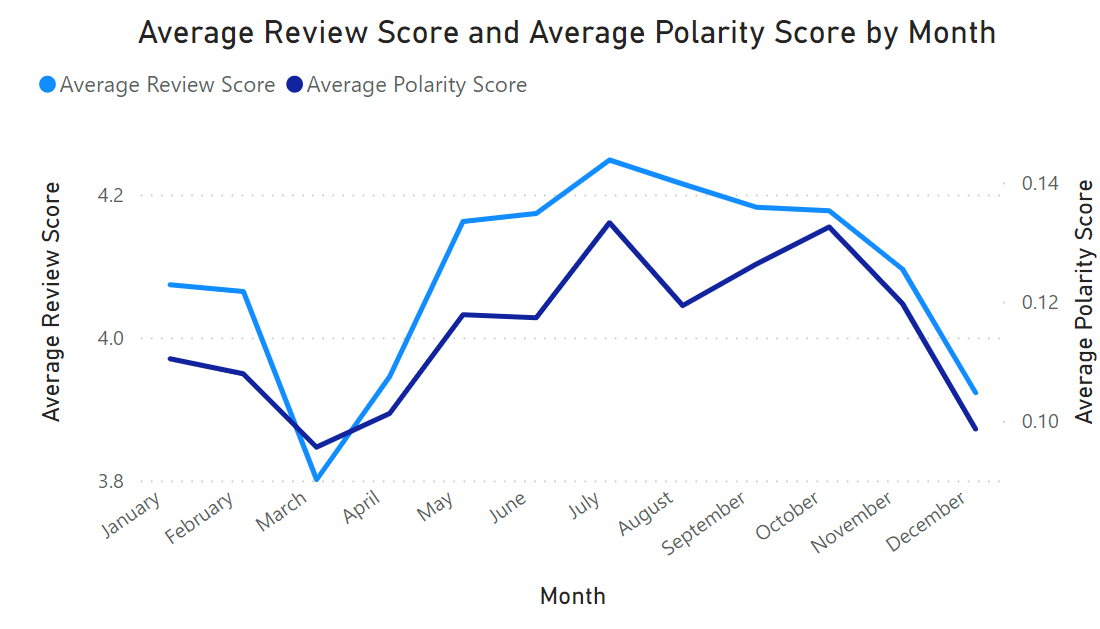
Graph 7. Top 10 product categories by volume and value

****

1. **Reviews and Polarity Scores**

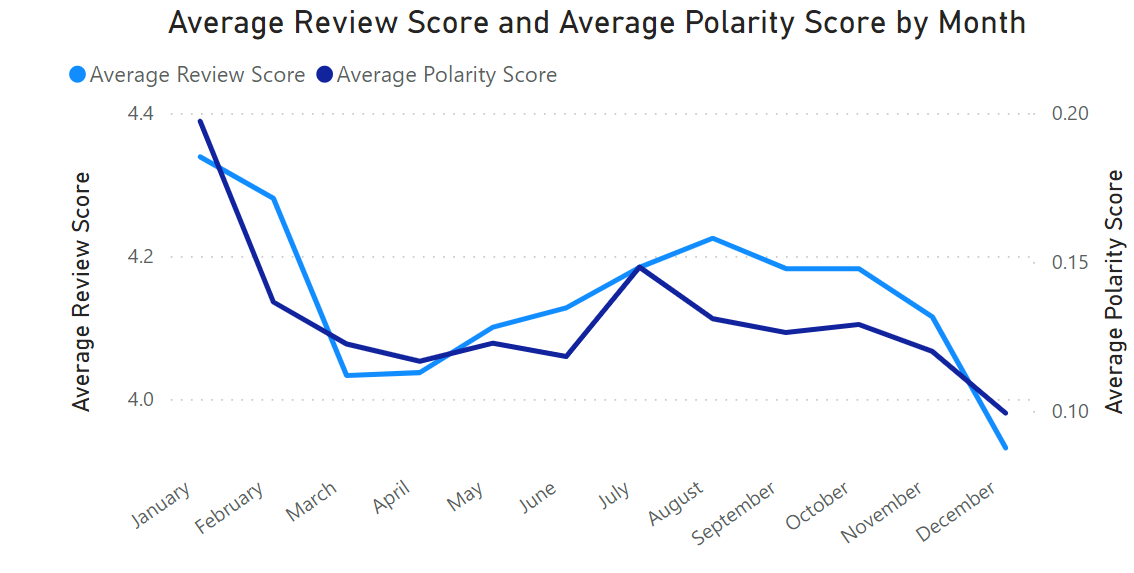
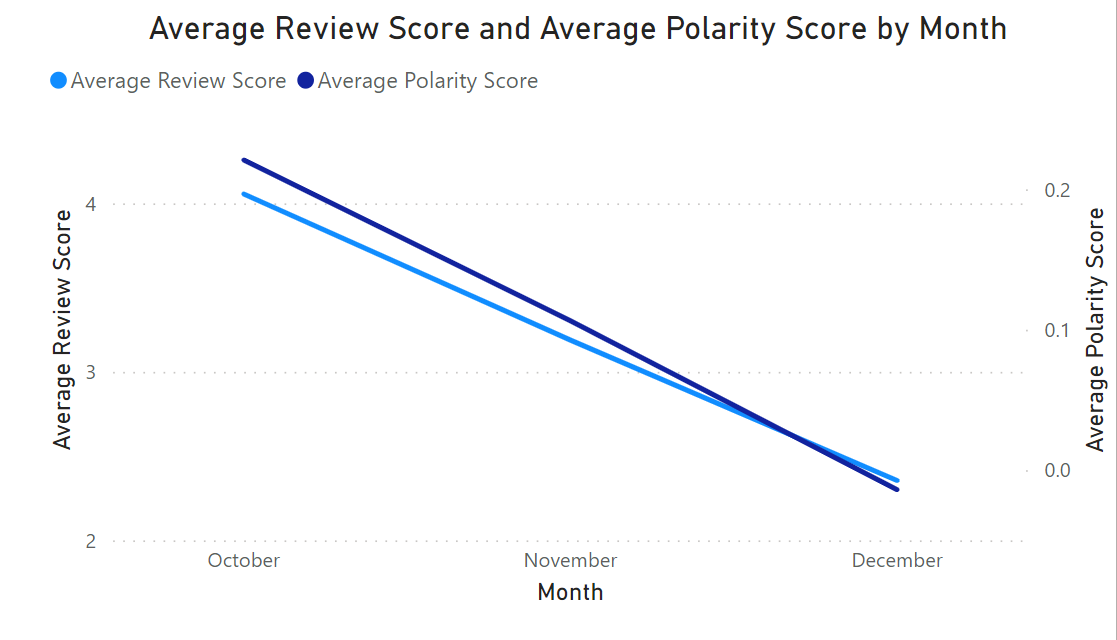
The reviews and comments of Olist customers were in Portuguese. The team had to translate these into English and use the NLTK library to assess the polarity scores.

Graph 8. Average review score and polarity score by month



The polarity score and average review scores by customers move in the same direction. These scores tend to be high in January, before dipping in March. The second quarter of the year shows a rise in scores and drops down after August. The drop coincides with August peak sales and customer complaints with late deliveries and issues with sellers such as delivering the wrong item or informing the customer late that the product they ordered is not available.

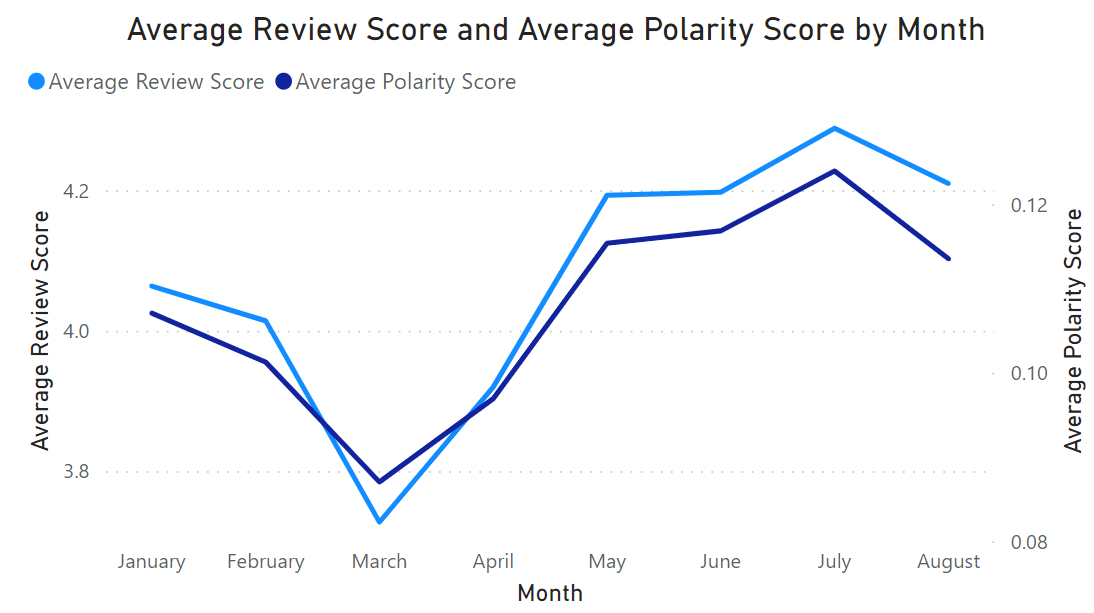
Graph 9. Review and Polarity Score in 2016 (left) and in 2017 (right)



The first year of operations of Olist showed a downward trend in reviews. This is due to the company navigating its operations in the first few months. Reviews have shown complaints in delays in delivery and negligence from the seller in shipping the correct product or shipping fragile items.

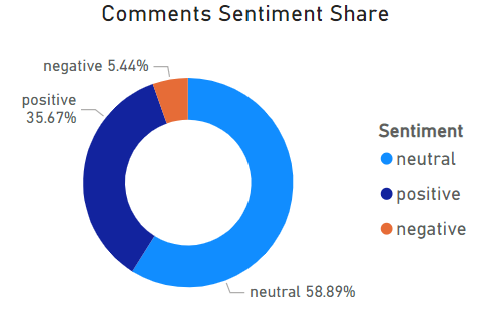
There were more positive reviews in 2017. However, the reviews again have dropped in the first quarter, picked up in the second quarter, and then dropped again in the last two quarters. The issues were the same as in 2017: delivery delays, incorrect product, broken products and lack of timely response from the sellers.

Graph 10. Average review score and polarity score in 2018



The reviews in 2018 performed similarly to 2017.

Graph 11. Sentiment share of reviews (2016-2018)



The review scores of the customers generally fall under the neutral category at 58.89%, followed by positive scores at 35.67% and then negative at 5.44%.

The average rating of customers from 2016 to 2018 stands at 4.09. The customers rated Olist at 3.55 in 2016, 4.11 in 2017 and 4.07 in 2018.

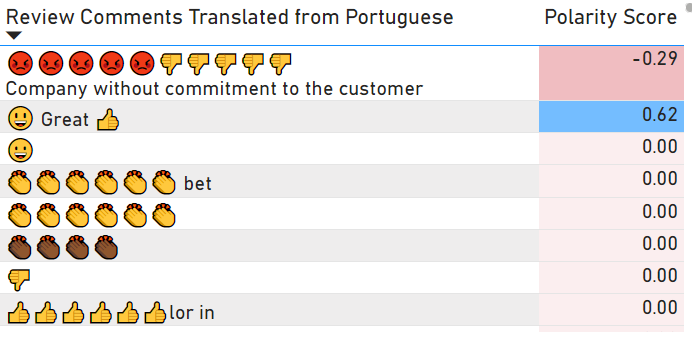
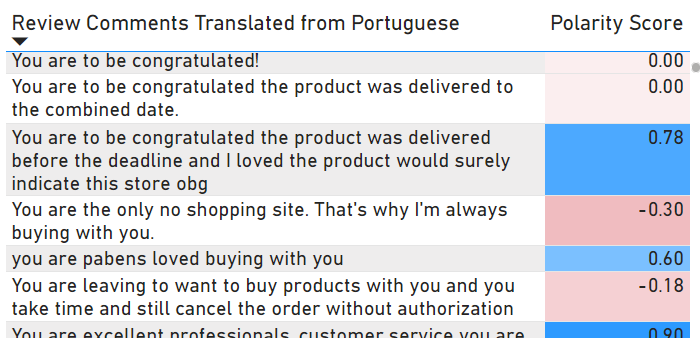
Graph 12. Key words appearing in the reviews



The following chart shows high frequency words appearing in the reviews. The larger the text, the more often the word appears. So far, positive words have come up in the study.

The team would like to highlight that there are limitations to the NLTK library in assessing polarity scores. There were comments that were positive that were rated neutral. Negative comments were also rated neutral.

Table 3. Sample reviews and equivalent scores based on NLTK sentiment analysis

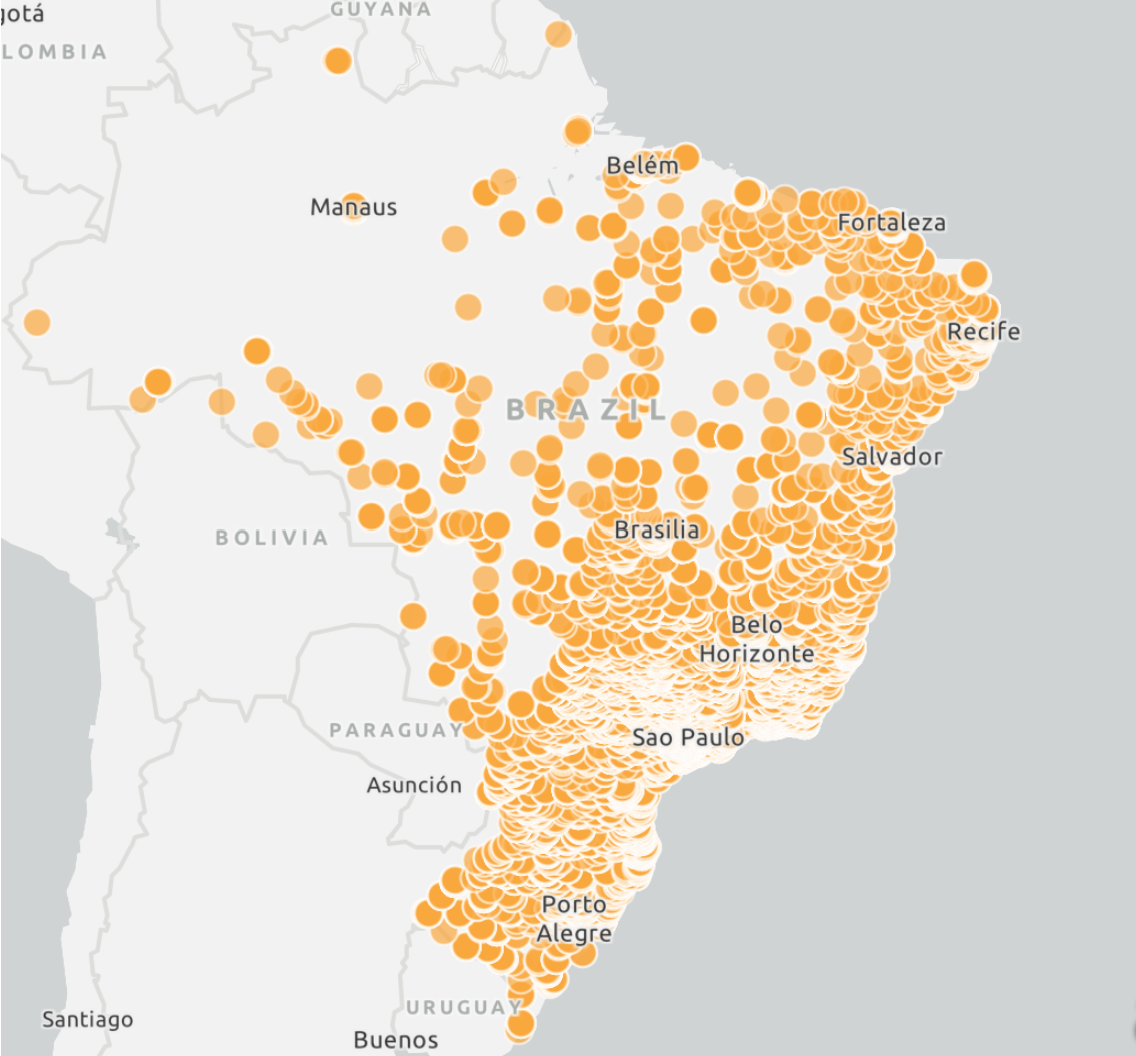


Even with the limitations of NLTK, the team would like to highlight that the polarity scores calculated with the NLTK library follows the average scores of customers as seen in Graphs 8 to 10.

1. **Delivery Performance**

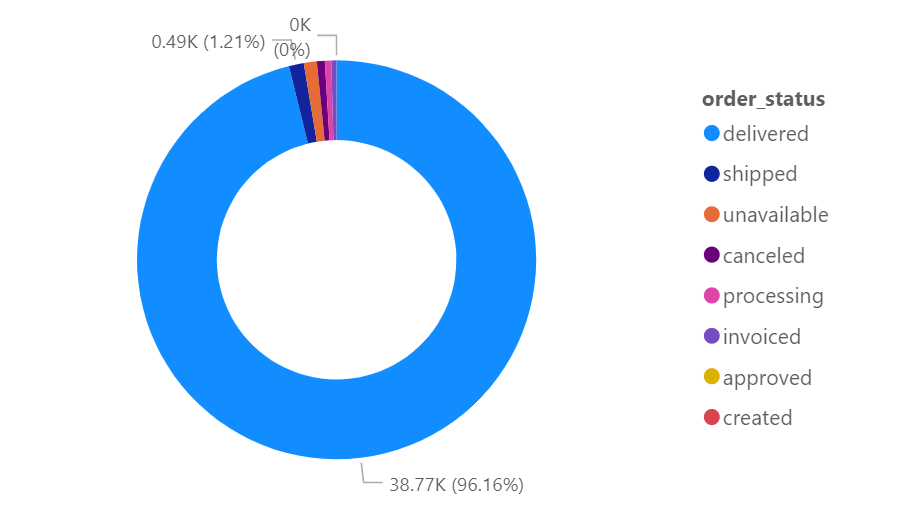
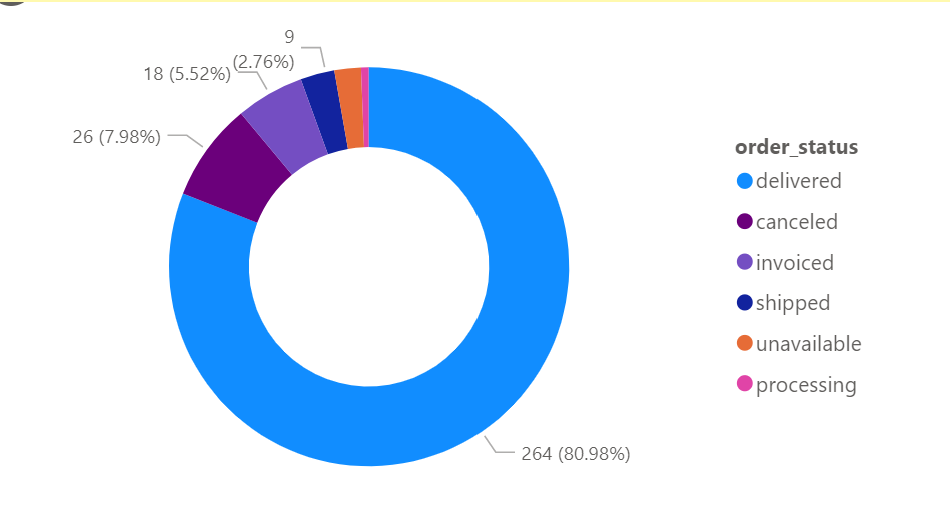
There is a good ratio of sellers to customers. These sellers are geographically well distributed to meet the needs of the customers.

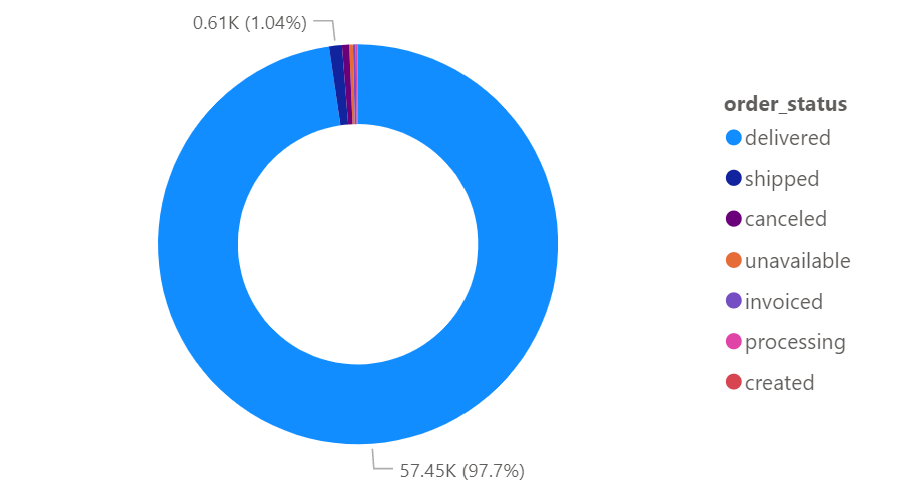
Graph 13. Seller location



There have been improvements in fulfilled delivery from 2016 to 2018.

Graphs 14. Order statuses between 2016 to 2018





However, the average number of days between expected delivery date and delivered dates have decreased. This indicates that the late delivery times have been rising across the years.

| Year | Average number of days between estimated delivery and delivered date | Longest days exceeding estimated delivery date |
| --- | --- | --- |
| 2016 | 36.29 | 38 |
| 2017 | 12.16 | 181 |
| 2018 | 11.57 | 188 |

From 2016 to 2018, the lowest average delivery days occur in March, April and December. Brazil has several public holidays in April (Good Friday, Easter, Tiradentes Day) and Christmas, New Year’s Eve in December. The rainiest months also happen in March and April.

The average difference between estimated delivery date and delivered date for all states from 2016 to 2018 is 13.33. The States that performed below average are the following:

| State Code | State | Average Days | Month of Concern |
| --- | --- | --- | --- |
| AL | Alagoas | 8.71 | October |
| BA | Bahia | 10.79 | April, May |
| CE | Ceara | 10.80 | March, April, December |
| DF | Distrito Federal | 12.05 | March, August |
| ES | Espirito Santo | 10.5 | March, April, December |
| GO | Goias | 12.19 | March, April, December |
| MA | Maranhao | 9.57 | March, April, December |
| MG | Minas Gerais | 13.24 | March, April, August, December |
| MS | Mato Grosso do Sul | 11.05 | March, April, August, December |
| PB | Paraiba | 13.26 | April, May, October |
| PE | Pernambuco | 13.29 | April, May, August |
| PI | Piaui | 11.31 | April, December |
| PR | Parana | 13.31 | April, August, December |
| RJ | Rio de Janeiro | 11.77 | January, March, April, December |
| SC | Santa Catarina | 11.51 | March, April, December |
| SE | Sergipe | 10.02 | January, April, May, June, December |
| SP | Sao Paulo | 11.08 | August, December |
| TO | Tocantins | 12.13 | March, April, December |

The analysts originally thought the States that would have late delivery issues would have been located in states farther from the capital cities, like Amazonas, but it is more the cosmopolitan and coastal areas that are encountering issues. This could be explained that in spite of a good ratio of sellers to customers, there are not enough delivery workers available to meet the high number of orders. There could be infrastructure issues and high traffic during delivery times as well.

Delivery issues tend to happen in April, August and December. With so many shoppers buying online in August, a strain in delivery can be expected. December, apart from being a shopping and holiday filled month due to Christmas, is a challenge for the company due to a mismatch between demand for orders and the number of delivery workers willing to work through the holidays.

April delivery issues can be explained with significant events happening in the country. In the same month in 2017, cities went into a partial shutdown for its first general strike in more than two decades. Brazil witnessed demonstrations taking place across the country. Private and public schools have closed and there were a lot of roadblocks and burning barricades in the Brasilia capital. These along with economic issues have impacted the ability of the company to deliver products on time.

These months of concern generally coincides with negative reviews as well.

**III. Summary and Recommendation**

Olist customers are value hunters. Given that the country has an average monthly income profile of USD 730 ( or AUD 1, 100) per month, it is consistent that these customers would shop during August sales and employ credit cards and installment plans.

Given this profile, the team recommends to continue offering installment plans and offer reward points to keep sales sustainable and viable. Since the customers are also large users of credit cards, the team recommends speaking to credit card companies to lower transaction fees to keep the overall shopping cost lower.

On slower months, the company could launch promotions, particularly around the “ber” months or the last quarter after the big August sale.

The negative reviews coincide with issues in deliveries and seller negligence. It is worth investigating the complaints of the customers regarding specific sellers. Also, there is a need for improvement in delivery time.

For delivery services, the team recommends examining further the coastal states that are encountering issues in deliveries. These could be infrastructure related issues where perhaps a better vehicle could be more suited to the terrain or traffic (for example, motorbikes are better at weaving through heavy traffic). It could also be a mismatch between the number of workers available to deliver packages and the number of orders made by the customers. In this case, better delivery worker planning is needed.

It is also worthwhile for the company to invest in customer support service and improve on seller response time. Better tracking delivery and notifications are key to ensure that deliveries are made on time to ensure customer satisfaction.

Having a key performance metric dashboard to monitor delivery performance, company sales and reviews is also key to assist the company in making data driven decisions and focus on areas that require support.