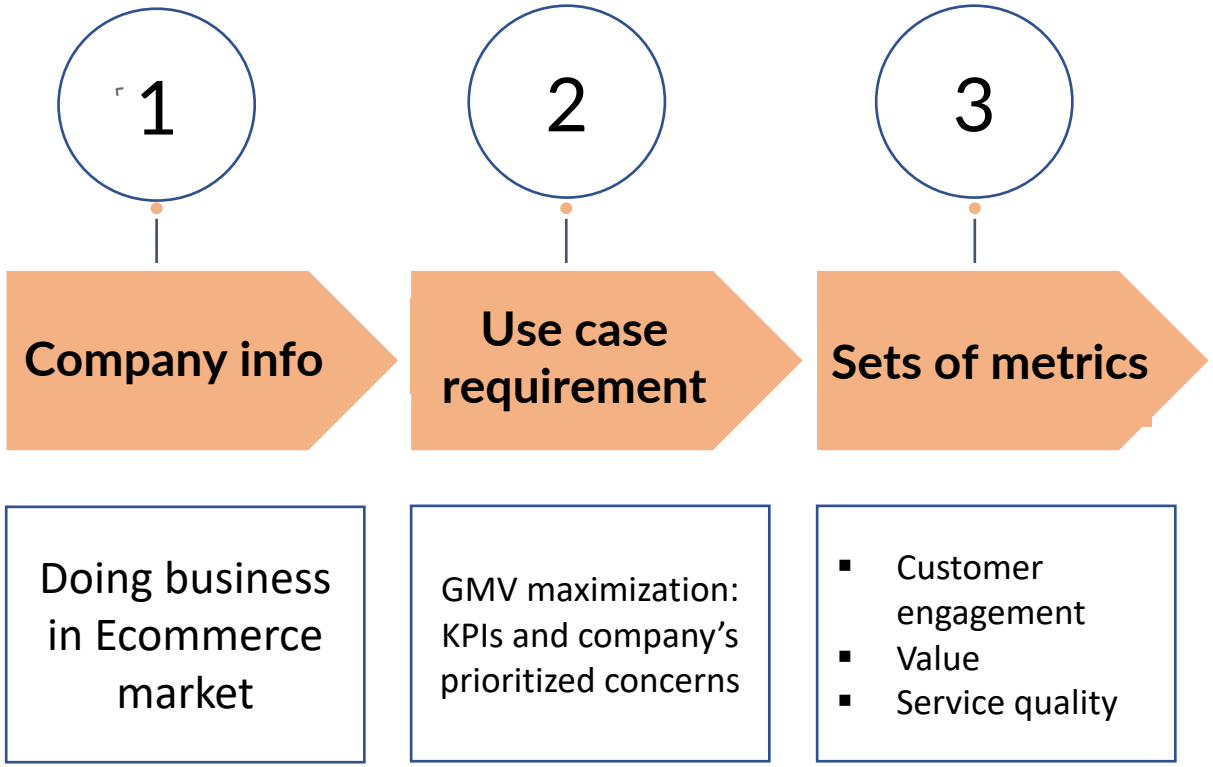
Business highlights: Brazillian ecommerce use case The focus should be customer interaction.



Customer engagement status

- It fluctuates
- Average number of customer: 4505
- Average growth rate in number of customer: 15.8%

Value

- It strongly correlates to number of customers or customer engagement status
- Average total value: 797224
- Average growth rate in value: 9%

Service quality

- In total, we have 100K reviews on 99K orders
- Average score is 4.
- Care more about logistics services.

GMV maximization: KPIs and business priority

The key concern is about sustainably competitive position

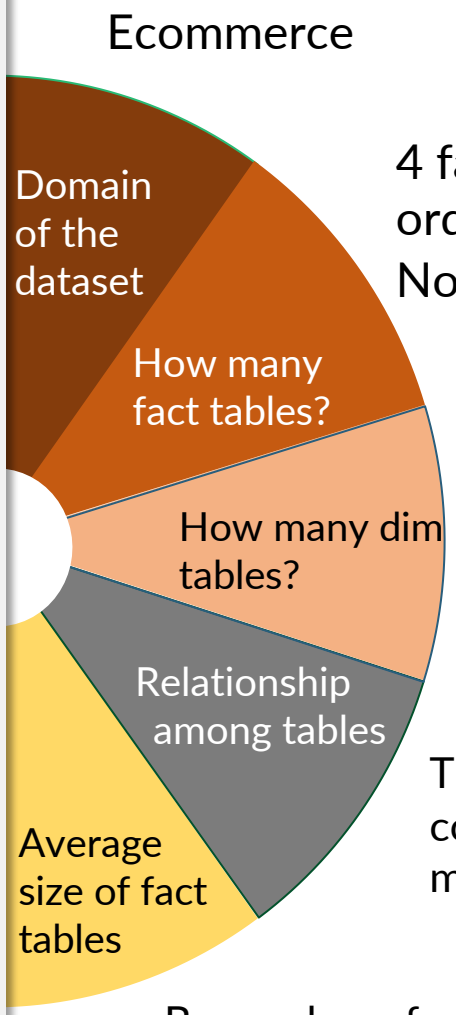
In a competitive market and customers are easy to switch like ecommerce, is the company maintaining its competitive advantage through ***growth metrics?***

Which are ***metrics help portrait company's health***?

- Customer engagement status
- Value
- Service quality

Company needs to build up a stable relationship with customers by ***understanding their customers and offering them matched products/services***, but not trying to sell as many as possible. The metrics ***tell company what shall be their first priority*** at the moment to maintain the interaction with their customers.

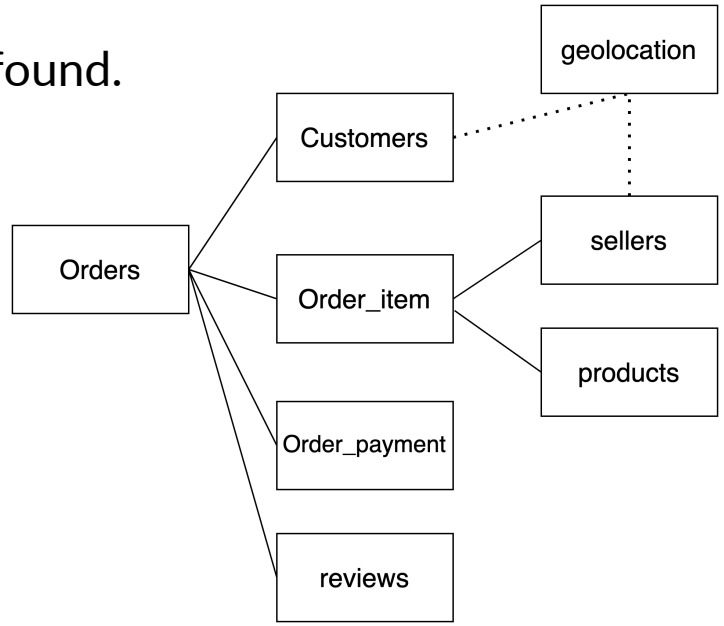
Data description Brazillian ecommerce



4 fact tables include: order_item; order_payment; order_review; orders
No unreasonable missing data found.

4 dim tables include: customers; products; sellers; geolocation
No unreasonable missing data found.

This is the map of how tables connected to each other, not a data model diagram



By number of records, average size of fact tables is ~104k records

Company's "health" or performance is observed with the following metrics.
Metrics built based on:

- Data availability
- Key concerns of ecommerce

1

Customer engagement status

- Number of users who make order, by month
- Number of new users who make order, by month

These metrics demonstrate how is relationship between the company and their customers/users

2

Value

- Total value, by month
- Value/user

These metrics tell the story of company's financial growth and how it keeps the stability of revenue

3

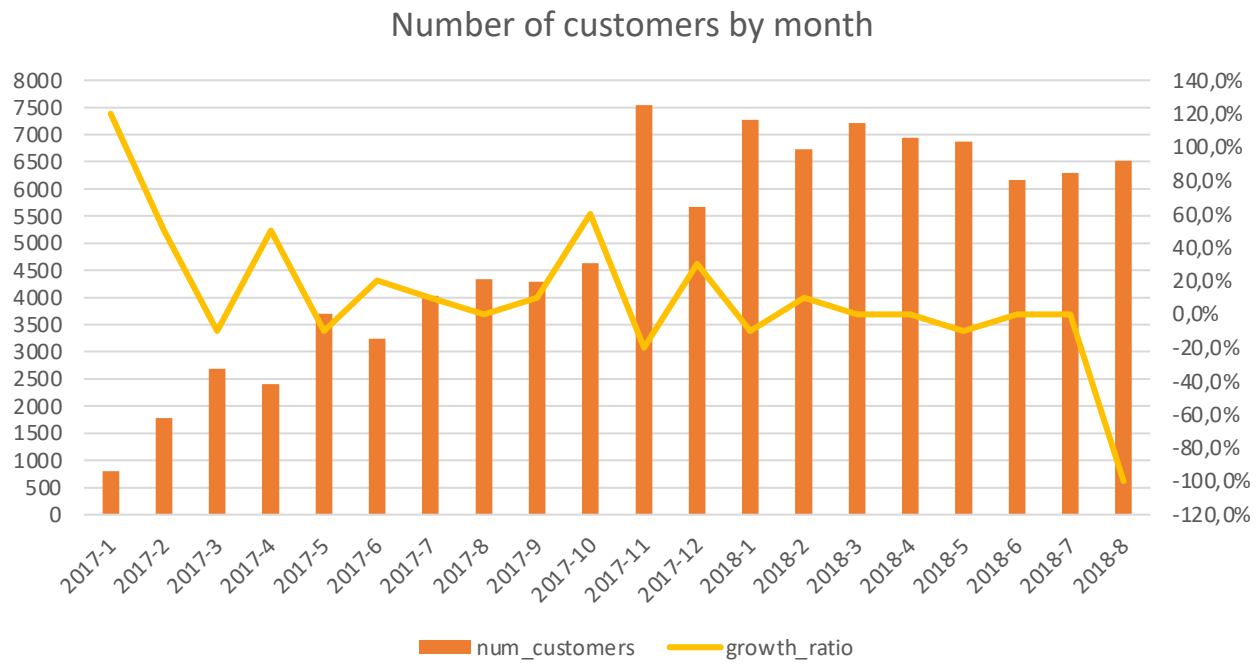
Service quality

- Review score

"How the customers feel about our services and products?" is the question that review score answers

Besides, based on data availability, the company can look at its delivery time to find out if it needs to pay attention to logistics and operation

Customer metrics: signal of the cut-throat in customer engagement

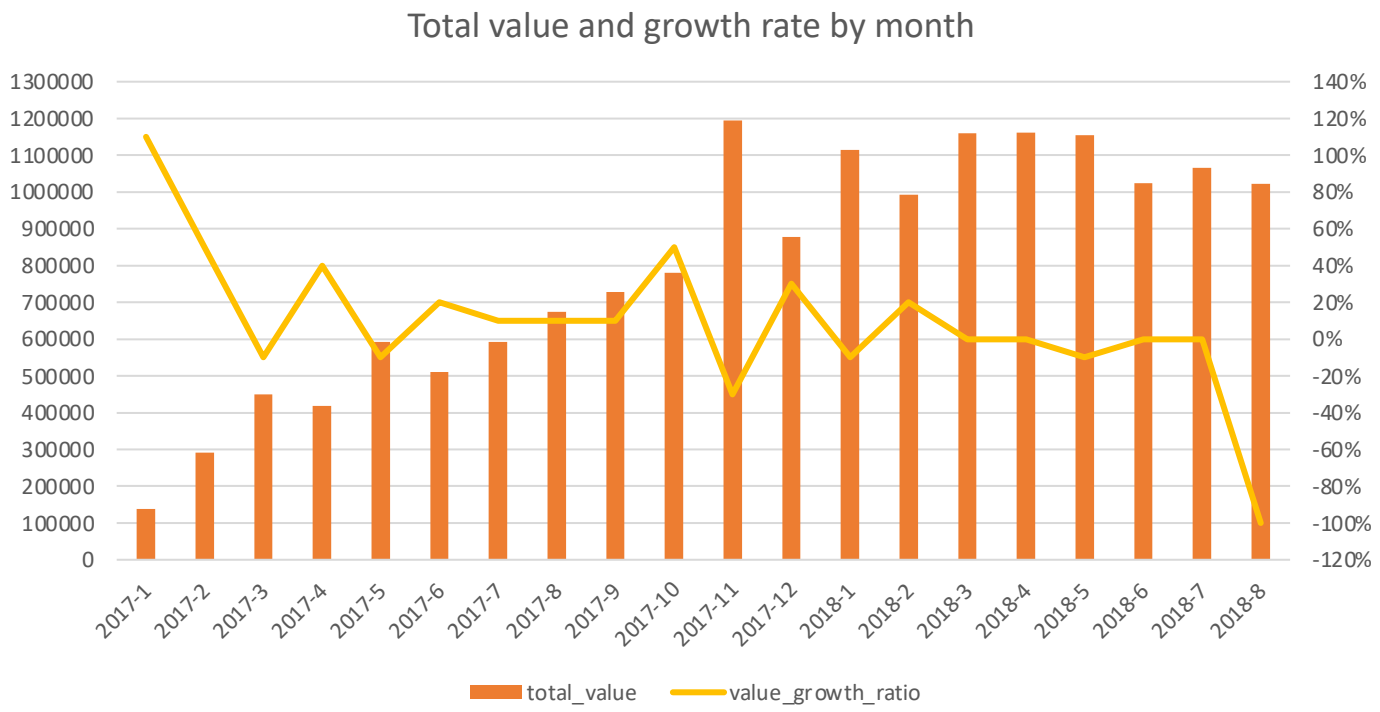


Average number of customer: 4505
Average growth rate in number of customer: 15.8%
Overall, the number of customers each month is fluctuated, showing a signal of cut-throat competition with other platform. However, from Feb 2018 to Jul 2018, the trend has a more harmonized cadence.
It implies that at this moment of business landscape, the company should focus on customer understanding and engagement campaigns.

2016 witnessed a dramatical trend in growth of customers. ***The absolute values of Sep, Oct and Dec 2016 were abnormal***. Look at the data below. Hence, this could be the case that there was a flaw in data collection. ***Similar situation is observed with Sep, oct 2018***.
Until we can ***verify the reason of plummet number***, the analysis of customer metrics should get rid of these years with the purpose of finding a more reasonable evaluation on trend.

year-month	num_customers
2016-9	4
2016-10	324
2016-12	1
2018-9	16
2018-10	4

Value metrics: the value undergoing a similar trend to number of customers



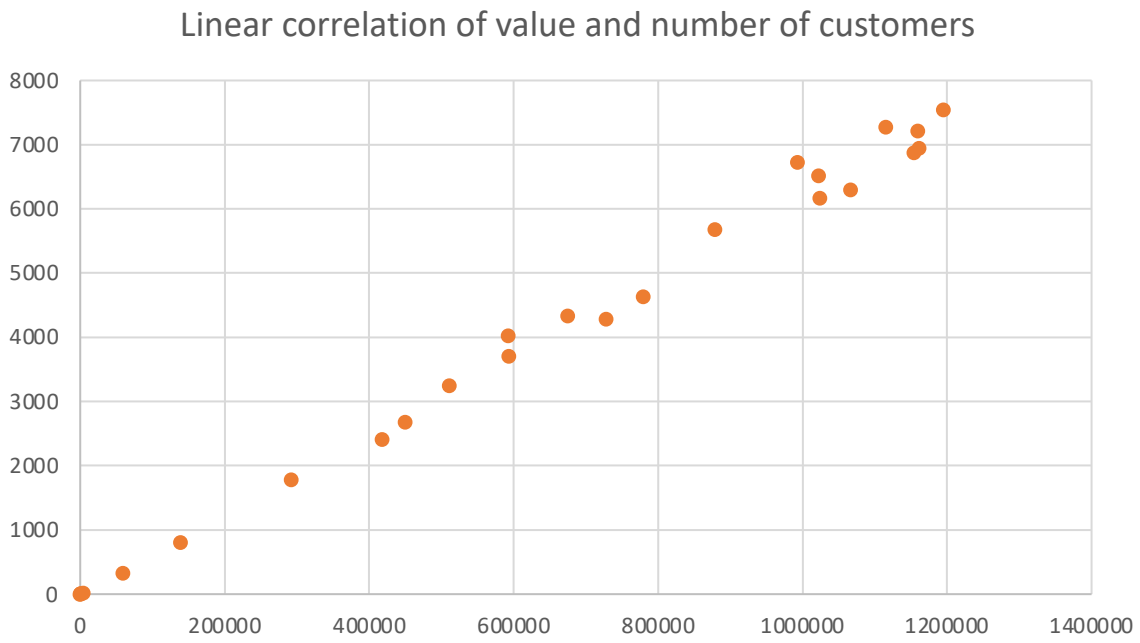
Average total value: 797224
Average growth rate in value: 9%

Also a fierce competition in maintaining the stability, the value growth is fluctuated. From Mar 2018 - Jul 2018, it was more stable.
Value of this platform could be ***strongly correlated to its number of users***.
However, ***Brazillian users seem to enjoy switching among ecommerce platform***, they do not have intention of purchasing most of their time on a particular platform.

Looking at the trend, we find fluctuation. However, to have an insight if the average number and value underperform the market, further data/analysis needs to be acquired.

Survival analysis would be a tool to find out the truth.

Value and customers: they are truly correlated



Average value per customer: 162

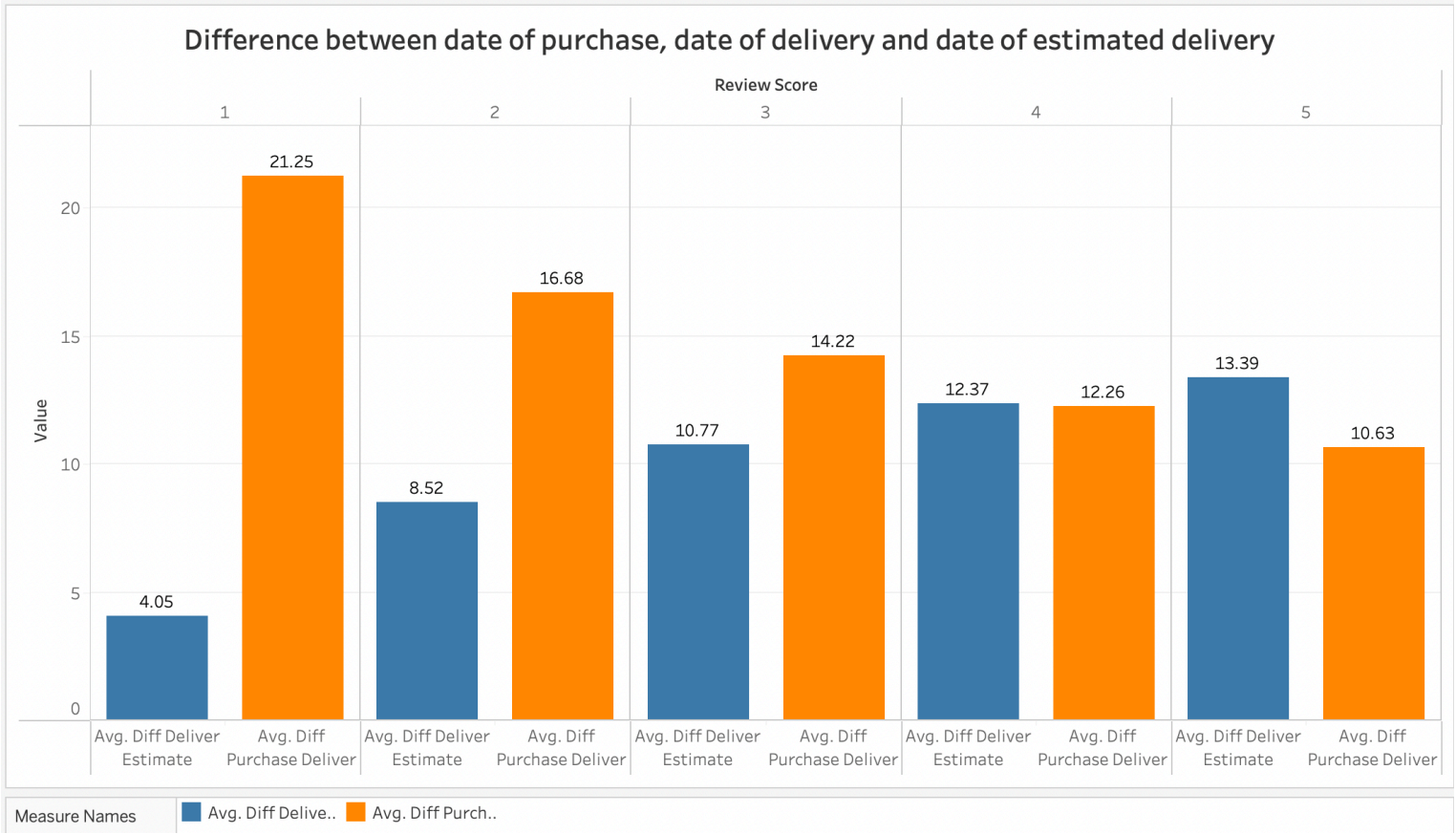
This scatter plot shows the correlation between:

- Our total value each month
- Our number of customers each month

At the beginning of the analysis, we have a sense that these 2 factors could be strongly related. Numbers say it correct.

Hence, the next question is: “is it correc to say that the game of ecommerce to build a platform with short latency, appropriate product category, etc. in order to keep customers stay with us?”

Service quality metrics: date difference, a signal of logistics services quality



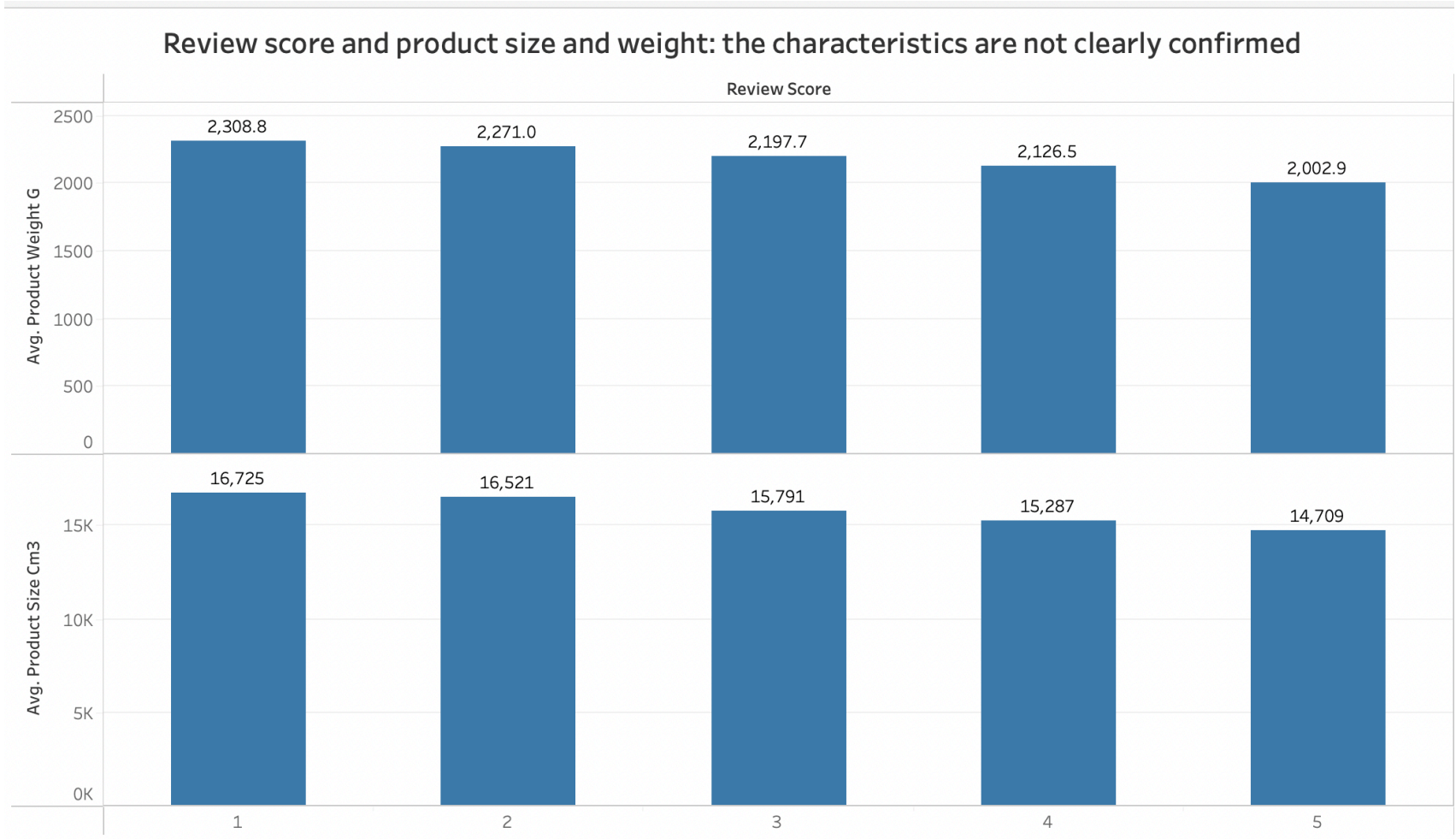
*Avg diff deliver estimate is **the date of estimated delivery minus the date of delivery**.*
*Avg diff purchase deliver is calculated by **the date of delivery minus the date of purchase**.*

1 stands for most unsatisfied.
5 stands for most satisfied.

Ecommerce users do care about logistics services. Date difference stands for waiting time and how customers expect the time they receive packages.

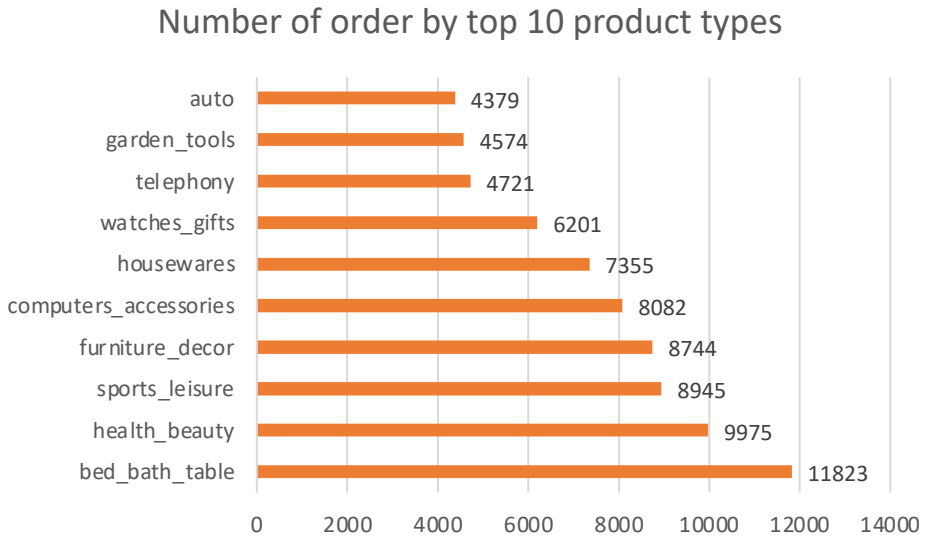
- The more similar between 2 metrics of date difference, the higher the score of review. Be consistent in what you promise to the customer is important.
- The shorter the difference between date of purchase and date of delivery ***(the shorter the waiting time)***, the higher the score of review.
- Those orders whose review around 4-5 have long difference between estimated and real delivery date.
- Case 1: It tells us that if the customer orders a product that could take a long time to be handed to them, the status of receiving it earlier could maitain the feeling of happiness. However, everyone has their threshold.
- Case 2: Right after the order is made, they see the info of estimated delivery, hence, if their package takes much longer than the estimate to arrive, they are unhappy.

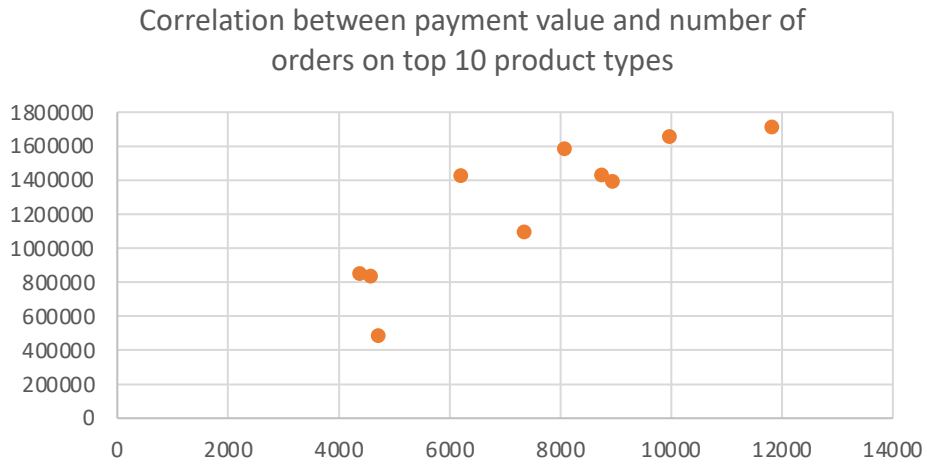
Service quality metrics: do product characteristics relate to score?



The trend is not strong. Not enough evidence to confirm.

Bonus: preference of product





Total payment value that a product type brings to the company does not necessarily correlated to its number of orders.

The company can put next questions of:

- Does the number of orders by product_type correlate to number of customers?
- Does product type make customers stay longer?

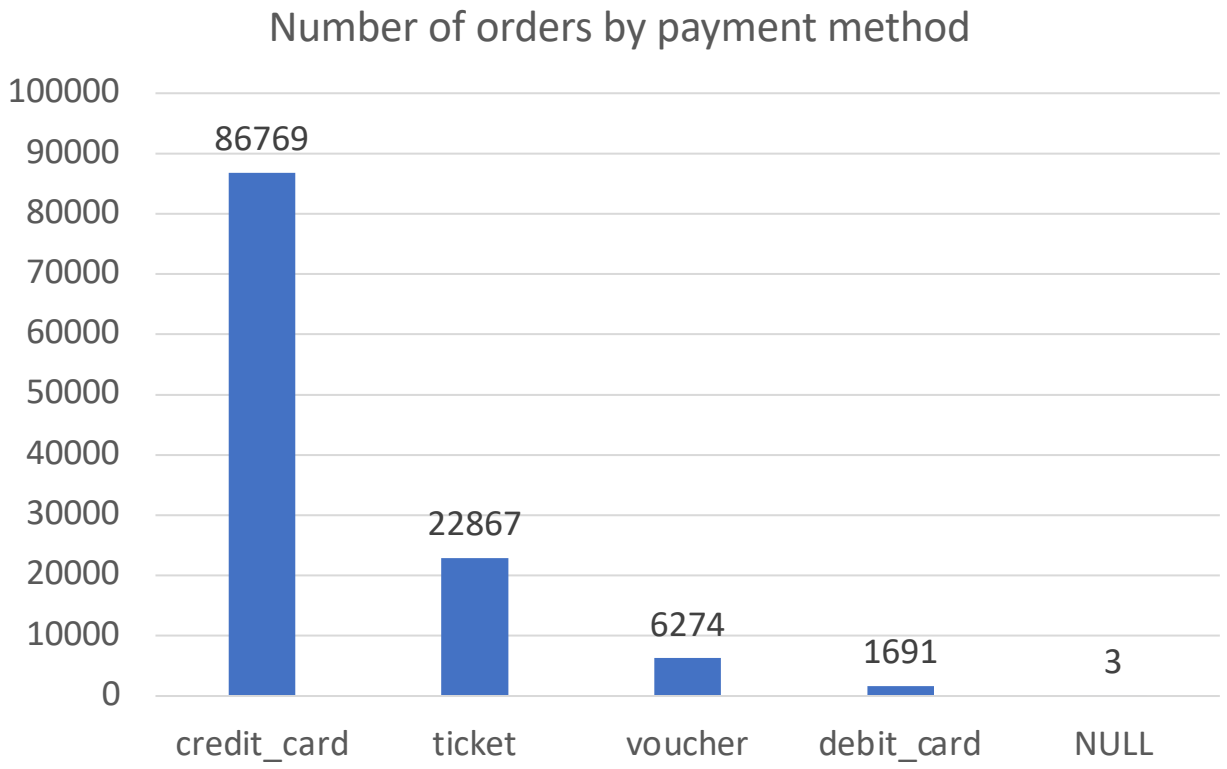
These questions intend to help build campaigns to increase customer relationship.

Or

- Should we promote high value product_type in order to increase our revenue?

These are two choices at this moment, to foster customer base or to focus on revenue?

Bonus: preference of payment



In a short sentence, people love purchasing on debts!

Final thoughts and further data collection

1. Customer metrics:
Each customer has only one time of order. the company should extend the timeframe of data collection or re-design their way of data collecting and data model to see if this customer base has any behavior of reorder. This action will give the company a better look on customer engagement.
2. Value:
Does the value strongly correlated to Brazillian users preference of switching among platform?
$\Rightarrow$ To answer, do more survey on users, collect data about promotion to see how this factor triggers our users.
Work on survival analysis to see when customers leave the platform. Is it too soon?
3. Correlation between users and value:
Collect data of latency, traffic when users browsing the app/web => the latency is the factor that strongly affects customer experience. Furthermore, do analysis on which customer segments cause traffic jam, but no increase in value.
4. Customer understanding:
For customer understanding, do customer segmentation, using clustering, and naming the customer segments (choose the most outstanding characteristics of the customer clusters to do naming).

=> These are suggestion to initiate customized offers to users. To do better offerings, we should build ML model for the purpose of predicting their intention of purchasing (kind of product matching)