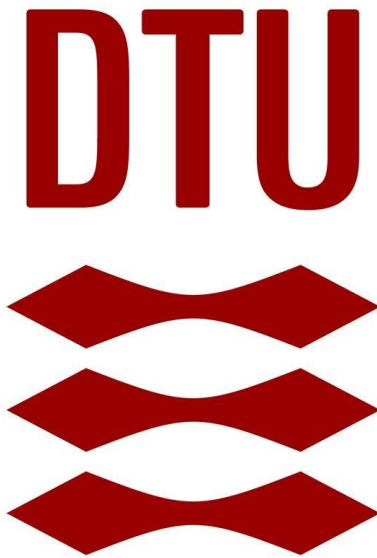


DECEMBER 6, 2019

# TECHNICAL REPORT

FINAL PROJECT - OLIST

42578 - ADVANCED BUSINESS ANALYTICS



TECHNICAL UNIVERSITY OF DENMARK

---

Andreas Kaae, s154395

This technical report is constructed on the basis of the work presented in the attached Jupyter notebook.

# 1 Introduction

Olist is a Brazilian e-commerce company which sells a variety of different products from different product categories. As any other e-commerce company they receive customer ratings and comments reviewing the company's performance of the orders. These reviews can help attract more customers as online shoppers often will check the overall ratings. It is therefore very important that Olist keeps a high rating to ensure that new customers choose Olist for their purchase.

This summary shows the results of working with data provided by Olist on 100.000 orders from 2016 to 2018. The summary focuses on the product review scores and the product review comments to identify which product categories are more prone to bad reviews and what causes a bad review.

Currently Olist has a mean review score of 4.1, the score is distributed across the 5 possible ratings as seen in figure 1. Thereby 15% of the reviews are of review score 1 or 2 which is categorised as a bad review score.

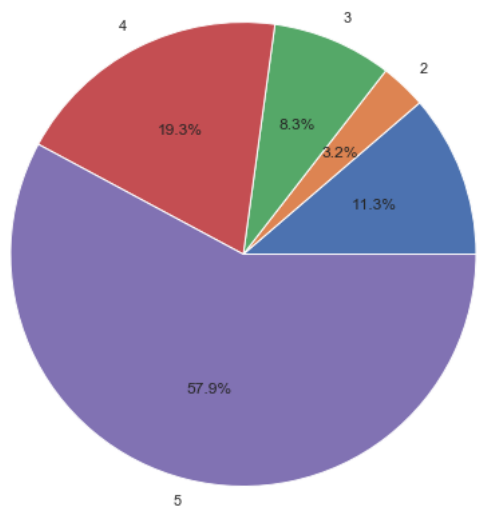


Figure 1: Distribution of the review scores for all orders.

By examining returning customers it was found that Olist only has 3% of these. The mean review score for these customers is 4.2, and thereby only marginally higher than the overall mean score. This indicates that increasing customers satisfaction level will not yield in more returning customers which should be taken into account.

## 2 Identifying critical categories

Olist in total has 73 different product categories of which 15 are responsible for 80% of the reviews. To identify which product categories were most critical with regards to bad reviews 3 things were considered. 1) The size of the category - out of all the orders how many orders is the category responsible for. 2) The amount of bad reviews - out of all

the bad reviews how many is the category responsible for. And 3) what was the overall mean of the category. By analysing all the product categories the 4 categories *Bed, bath and table*, *Computers accessories*, *Furniture decor* and *Office furniture* were identified. In figure 2 the % of bad reviews within the category, % of bad reviews across all categories and % of orders out of all orders is illustrated for the 4 categories.



Figure 2: Measure for the 4 chosen product categories.

All 4 categories are chosen as they proportional are responsible for a larger share of the bad reviews than the orders. The 3 categories *Bed, bath and table*, *Computer accessories* and *Furniture decor* are also chosen as they further are among the 15 categories responsible for 80% of the orders, where *Bed, bath and table* is the category responsible for most orders. *Office furniture* is further chosen as this category was among the categories which had the worst mean ratings.

### 3 Causes for bad review scores

By analysing the data it was found that 8.1% of all orders were delivered late and that the more orders a category would have the more late orders is would also cause. The mean score for all late orders were 2.5 and there by 39% worse than the review score for all the reviews. The distribution of the scores are showed in figure 3 which illustrates that 54.6% of the late orders have a bad review. It can further be be derived that of all the bad reviews 33.7% were also late thus indicating that the amount of bad reviews potentially could have been reduced with 1/3 by eliminating late orders.

By analysing the review comments it was found that the biggest problems were either with regards to the delivery, late deliveries, problems with the post service or with regards to defect products, product quality and missing products. By examining at the review scores over time it was further found that the mean review score is stable over time. However, during Black Friday 2017 and the period of march 2018 to april 2018 the mean review score is below the average. Within these periods customers were complaining more about

late deliveries than normally, especially during Black Friday while in the period of march 2018 to april 2018 the company Lannister is mentioned.

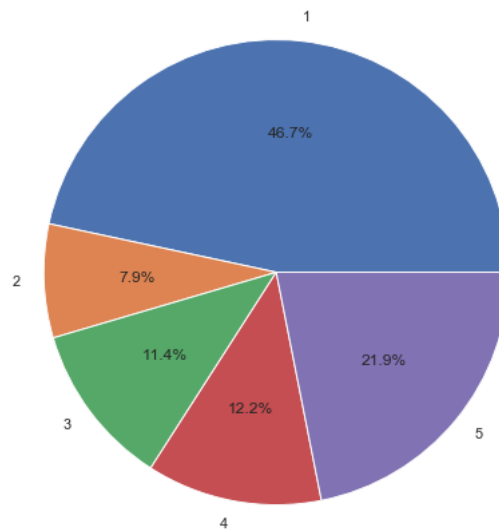


Figure 3: Distribution of the review scores for late orders.

For the category *Bed, bath and table* it was found that problems were with regard to the delivery and product quality, specifically bed covers were mentioned. For *Computer Accessories* the biggest problems were with defect products were especially printer cartridges are mentioned. For the category *Furniture decor* the problems were mainly about the delivery. Lastly for the category *Office furniture* customers were often complaining about chairs and missing or wrong parts.

## 4 Conclusion

In conclusion it has been established that returning customers generally do not give better reviews than customers who only place one order. Improving the overall review score will therefor not necessarily yield in more returning customers. 4 specific product categories have been identified: *Bed, bath and table*, *Computers accessories*, *Furniture decor* and *Office furniture*, as these all had a significant amount of bad reviews and orders. It was further found that during Black Friday 2017 and in the period of march 2018 to april 2018 the review score deviated from the mean of 4. The reason for bad reviews were overall found to be due to problems with late deliveries or product quality and defects. During the two periods it was found that delivery especially was a problem. Olist main focus should therefor be to decrease the amount of late deliveries as this potentially can prevent 1/3 of future bad ratings.