Case Study: Olist.com

Application of Process Mining to Improve Electronic Commerce

Rene Laub & Bernd Skiera

This version: 2019/ 12/ 03

The case is based on data from the Brazilian E-Commerce company Olist. The data is available on Kaggle (https://www.kaggle.com/olistbr/brazilian-ecommerce).

Department of Marketing, Faculty of Business and Economics, Goethe-University Frankfurt am Main, Theodor-W.-Adorno-Platz 4, 60629 Frankfurt, Germany, phone: +49-69-798-34 561, email: rlaub@wiwi.uni-frankfurt.de(corresponding author), skiera@wiwi.uni-frankfurt.de

1 Company Presentation

Olist is a Brazilian e-commerce company. Olist connects small retailers and merchants with customers. It acts as an intermediary, representing small retailers and merchants under the Olist brand on marketplaces (see Figure 1):

- Small retailers cooperate with Olist to represent them on different marketplaces like Amazon Marketplace, E-bay, Americana, WalMart Marketplace and others
- Olist sells the products of the small retailers on the different marketplaces under the brand "Olist", and is thereby listed as seller on the marketplace (see Figure 2)
- Customers buy the products on the marketplace from Olist
- Olist then handles the order-to-delivery process (including payments) for the small retailers and manages the interaction with the customer
- Customers therefore only deal with one seller, namely Olist
- The small retailers benefit from Olist's reputation and the infrastructure of Olist

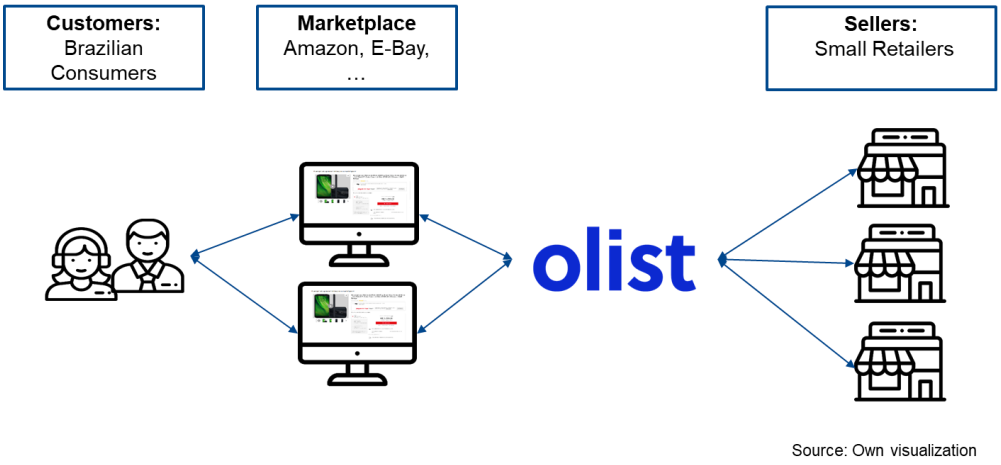

Figure 1: Market Landscape of Olist

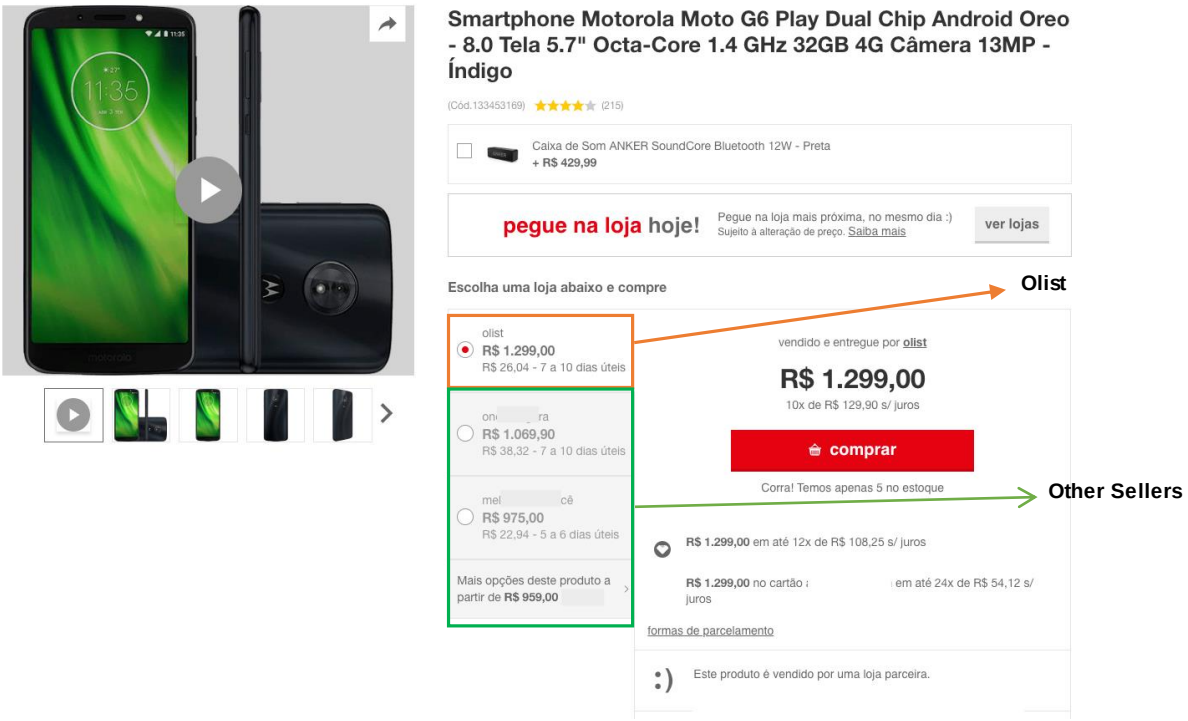

Figure 2: Listing of Olist on a Marketplace

Additional Links:

- https://www.kaggle.com/olistbr/brazilian-ecommerce
- https://olist.com/d2c/
- https://www.crunchbase.com/organization/olist

2 Your Task at Olist

You recently joined the business analytics unit of Olist. Your new boss, Carla Silva, just came back from a big summit on digital transformation. She is enthusiastically telling you and your colleagues about a new method, called "Process Mining". Process Mining is a method to understand, analyze and improve business processes. Success stories of companies like UBER, BMW or Siemens encouraged her to now introduce the method to Olist. In her weekly E-Mail she outlines her idea:

RECEIVER: Olist Business Analytics Team <BI-Team@olist.com>

SENDER: Silva, Carla <Carla.Silva@olist.com>

RECEIVED: 03/DEC/2019 08:03 am

SUBJECT: Process Mining at Olist!

Dear Team,

Last weekend I learned about a new data science method called "Process Mining" at the South American Digital Transformation Summit. Process Mining is an approach that combines data science with business process modelling and is already applied by many leading companies worldwide. One company reported that they managed to increase their profits by stunning 20% after the introduction of process mining.

How did they do this? The company started out with analyzing their core business processes with Process Mining. Thereby, they were able uncover weaknesses in their processes and then started to dug deeper. After fixing the weaknesses, the company streamlined their operations and interactions with customers and managed to improve the experience of their customers. This improvement lead to a higher customer satisfaction and in return to higher customer retention and ultimately higher revenues, partly even at lower costs. Amazing!

I think it's time, that we apply this method to our processes and examine the weak spots of our processes.

The core of our business is the handling and settlement of orders. As a starting point, I suggest that you start analyzing our order-to-delivery process (including payment).

Please quickly familiarize yourselves with the method. I provide you a list if quick introductions to the methods and the R-Package BupaR at the end of this email.

Please provide me the following insights:

1. **Descriptive Analysis**
    a. How many cases *(-> orders*) does the event log contain?
    b. How many events does the event log contain?
    c. How many unique activities are executed throughout the process?
    d. Provide me with a summary *("summary()"* command in BupaR) of the event log and visualizations of interesting variables.

2. **Control-Flow Perspective/ Structure of the Process**
    a. How many different variants of the process (traces) are in the event log?
    b. What is the most frequent variant of the process?
    c. Visualize the process by using the *"trace_explorer()"* command and the *"process_map()"* command.

d. How many cases (absolute & relative) follow the most frequent process variant?
e. How long are the traces?
f. What are the typical start & end activities?
g. What are the deviations from the most frequent variant of the process (2b)? Hint: *The deviations can be caused either by differences in the presence, the number of occurrences or the ordering of the activities within a trace.*

3. **Performance/ Time Perspective**
   a. What is the minimum, maximum and mean throughput time?
   b. Create and interpret a performance map using *"process_map(eventlog, (performance(mean, "days"))*
   c. Create and interpret a dotted chart using "*dotted_chart(eventlog, x="absolute", y="start"")*

4. **Satisfaction Analysis:** Use the event log to investigate drivers of customer satisfaction.
   Hints:
   - Apply regression analysis with "review score" as dependent variable
   - Add a variable indicating if an order followed the most common process variant
   - Add a variable containing the throughput time
   - Feel free to create additional variables, based on your process analysis

The data scientists at the summit told me that the basis for the analysis is an **event log**. I think, in our case, we should look at an event log that documents the order process.

Unfortunately, we do not have this kind of data structure ready yet. I already discussed the matter with our data engineers. The data engineers will prepare an event log for you, based on the data tables in our different databases.

Please remember, our operations department summarized the **order process** as follows:

After a customer orders a product from Olist (on behalf of a small retailer) on a marketplace, the small retailer gets notified from Olist to fulfill that order. After the customer received the product, or the estimated delivery date was due, Olist invites the customer via email to participate in a satisfaction survey, where the customer should provide a rating to express their customer satisfaction and leave some additional comments.

I am looking forward to your insights! Happy Mining! (…as those process mining people say)

Carla

Data:

Olist_Event_Log_2019.RData (available on Olat)

Links:

- General Introduction: http://processminingbook.com/aboutbook.html
- R-Package:
  - https://www.bupar.net/
  - https://www.datacamp.com/courses/business-process-analytics-in-r
  - https://www.r-bloggers.com/process-mining-part-1-3-introduction-to-bupar-package/
- Python Library: http://pm4py.org/