

olist
empowering commerce

Introduction

- Seller side to increase sales
- Customer needs and satisfaction
- Overall Website Performance to attract more customers

Dataset

- Was collected from sellers that filled-in requests of contact to sell their products on Olist.
- Consists of **8,000 leads** that requested to join the Olist e-commerce platform between 2017 and 2018.
- They were randomly selected from the total leads.
- Allow views of a sales process from multiple dimensions, such as business type, average stock, lead category, catalog size, behavior profile and etc.

Methods

- Used **Tableau** to do the Exploratory Data Analysis to visually summarize the main characteristics of the datasets. We viewed orders from multiple dimensions: from sales performance, payment method, and product attributes to customer reviews and marketing channels.
- For **supervised machine learning**, we first made hypotheses on the factors that would affect review scores (Y). Then used forward and backward methods to prove their significance. Finally, used Random Forest and XGboosting to build model with significant variables. The model would help the sellers predict their review scores and make proper business decisions.
- For **unsupervised machine learning**, we used text analysis to gain valuable service insights from comments written by the customer. The top topics returned from the results would reflect the problems that the sellers have.

Results

Overall sales performance in 2017

Month of Order Date	Number of Orders
January 2017	341
February 2017	1,591
March 2017	2,801
April 2017	2,079
May 2017	4,289
June 2017	3,741
July 2017	4,074
August 2017	5,113
September 2017	4,618
October 2017	5,369
November 2017	5,541
December 2017	8,425

- ### Payment in month
-
- | Payment Ty.. | Order | Delivered | Carrier | Date |
|--------------|-------|-----------|---------|--------|
| Boleto | 975 | 6,196 | 324 | 40,737 |
| Credit | 844 | 10,496 | 40,694 | 81,718 |
| debit_card | 90 | 444 | 1,216 | 5,043 |
| voucher | 567 | 929 | 2,174 | 3,885 |

Origin	2018.1	2018.2	2018.3	2018.4	2018.5
Paid Search	152	259	448	392	116
Organic Search	383	273	218	362	135
Unknown	834	144	140	102	96
Social	188	56	38	130	39
Direct Traffic	67	12	27	81	33
Other	6			93	
Referral	16	2	1	30	29
E-mail	1	16	5	2	
Display		6	1		
Total	1647	768	878	1192	448
N of New Sellers	71	72	64	93	45
Conversion Rate	0.04	0.09	0.07	0.08	0.10

[illegible]

- 20% of the products brought nearly 80% of the total revenue, meaning that the sales volume is **seriously polarized**.
- The sales volume of many product categories is close to 0, and the sales volume of the top categories is far ahead.
- A limited selection of product types for customers may lead to insufficient products to meet customer needs.

Discussion

- Data contain NAs and not up to date, so the results may be partial with low reliability.
- We only have completed data of 2017 and there is no other full year's data to compare and to support our findings.
- In text analysis, two topics are too closely related because there are only 36,000 reviews out of 100,000 orders, may lead to biased results.

produto

entregue entrega dentro ser pedido ótima apenas super compra recebi previsto excelente comprar o tudo loja agora dias nada certo parabéns sempre 2 porém antes > site dia rápida produtos nao é boa bom ótimo ainda lammister pois recomendo comprei bem embalagem rápido gostei nota conforme qualidade perfeito chegou prazo

- | | IncNodePurity | rownames(importance) |
|---|---------------|----------------------|
| 1 | 1365.29563 | delivered_days |
| 5 | 888.79392 | payment_value |
| 2 | 755.50874 | freight_value |
| 4 | 727.27485 | price |
| 3 | 638.52315 | product_weight_g |
| 6 | 469.62263 | product_width_cm |
| 7 | 314.89008 | payment_installments |
| 8 | 57.12157 | payment_sequential |

	Train MSE	Test MSE
Linear Regression	1.42	1.51
Random Forest	0.21	1.12
XG Boosting	1.19	1.42

- Linear Regression
- RandomForest
- XGboosting

- Delivered_days
- Payment_value
- Freight_value
- Price

Any improvement on these variables can potentially increase the review score and customer satisfaction.

Conclusions

- **To improve Olist Sellers service**, Olist sellers should set up social media care and live chat to increase the efficiency of customer support and to help with more urgent cases to gain direct feedback from customers and increase review score.
- **To increase Olist customer satisfaction**, new coming Olist sellers should provide Boleto as one of the payment methods, as many find this convenient. More convenient method options can help to increase the speed of payment, simplify online shopping process, and shorten online shopping delivery time.
- **To increase Olist Platform overall quality**, Olist can recruit more diversified sellers on its platform to increase its diversity and to attract more potential buyers or set up a reward system that can award customers for their purchase.

References

Olist. (2018, November 29). Brazilian E-Commerce Public Dataset by Olist. Retrieved from https://www.kaggle.com/olistbr/brazilian-ecommerce#olist_order_payments_dataset.csv