

Analyzing eCommerce Business Performance with SQL



Created by:

Irham Maulana Ahmadi

irhammahmadi@gmail.com

www.linkedin.com/in/irhamahmadi

www.linktr.ee/IrhamMAhmadio7

“I am a recent graduate of the Faculty of Petroleum Engineering UPN “Veteran” Yogyakarta, seeking for opportunity to apply skills in data analyst and data science. Skilled in operating Python for data processing, visualization, machine learning. Tableau for data reporting, and SQL. highly enthusiastic individual, eager to learn new things, and can work in a team or individually, “

“Measuring business performance is very important in a company to track, monitor, and evaluate the success or failure of various business processes. Therefore, in this paper, I will analyze business performance in an eCommerce company by considering several business metrics such as customer growth, product quality, and payment type.”



Fig. 1 : ERD (Entity Relationship Diagram)

Source : Personal Documentation

- All data processing was carried out using PostgreSQL
- At the Data Preparation stage, several processes are carried out, including creating a new database, importing data, and creating relationship between table.
- The process of providing relationships between tables is carried out by inputting them via the Properties menu (Properties > Constraints) on each table.
- In the geolocation_dataset table, duplicate data was found in the geolocation_zip_code_prefix column which was to be used as the primary key, so it was necessary to delete the duplicates first.
- The final result of providing relationships between tables can be seen in the ERD diagram

View or Download ERD [here](#)
The complete query can be seen [here](#)

Annual Customer Activity Growth Analysis

Table 1 : Annual Customer Summary

Year	Avg. Monthly Customer	New Customer	Repeat Order	Avg. Order
2016	103.00	305	36	1.21
2017	3,653.33	42885	5360	1.19
2018	5,907.78	52230	6146	1.17

From Table 1, it can be seen that there is an increase in the average monthly customers and new customers every year. This is also reinforced by Figure 2, which shows a tendency to increase the number of customers every month. The highest value was reached in November 2017, but after that, a downward trend began to occur, although it was not very significant.

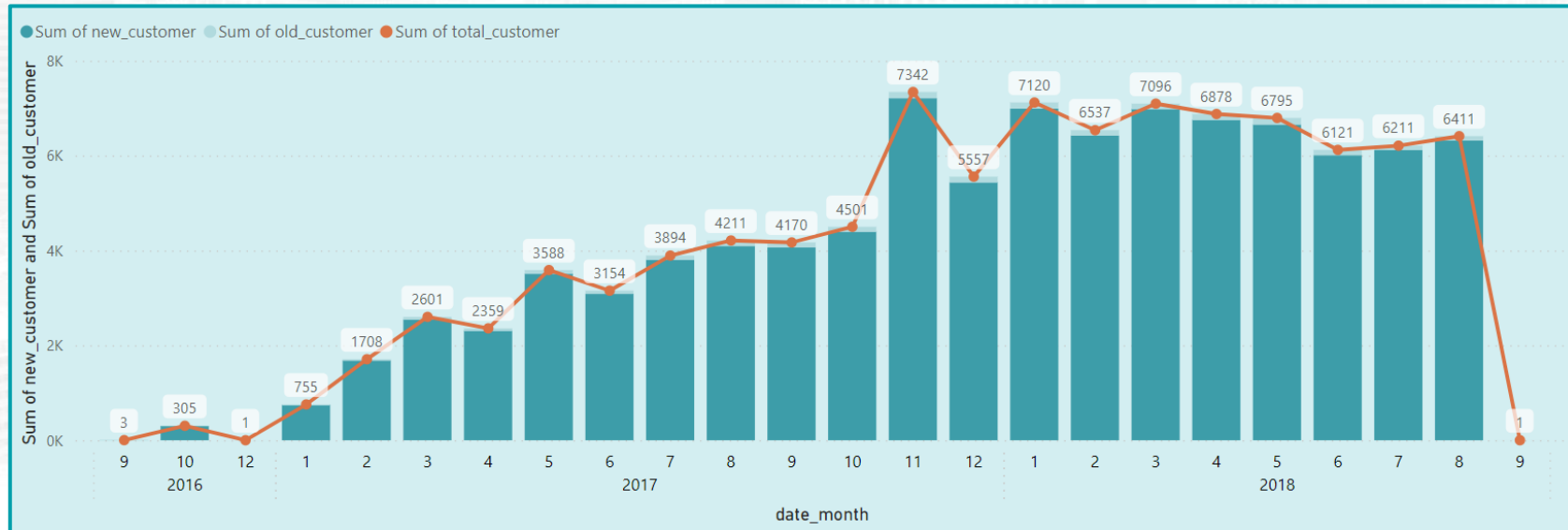


Fig. 2 : Monthly Customer Line Chart

Source : Personal Documentation

The complete query can be seen [here](#)

Annual Customer Activity Growth Analysis

Table 2 : Annual Transaction Stats

Year ▲	Avg. Total Transaction	Highest Transaction by User	Avg. Distance Between Transaction, Day	Avg. Unique Items Each Transaction	User With 1 Transaction, Percentage (%)	Avg. Rating
2016	1.21	8	40.10	1.20	88.35	3.57
2017	1.19	24	18.38	1.15	87.60	4.05
2018	1.17	24	15.80	1.15	88.31	4.02

- In Table 2, it can be seen that the average annual transactions still show fluctuating figures. Still, although the average number of transactions per user in 2018 decreased, the number of transactions increased quite significantly (Table 1).
- There is an increase in the number of transactions per user in one year, with the average daily distance per user to reorder getting smaller. This shows positive potential in the future.
- Only around 12% of the total users who made transactions made repeat orders in the same year
- There was an increase in customer satisfaction regarding goods and sellers. Although it showed a decline in 2018, the difference was insignificant.

Annual Customer Activity Growth Analysis

Table 3: Customer Origin Summary

Year	Total State	Total Transaction, State	Highest Customer, State	Comparison with Overall Transaction (%), State	Total City	Highest Transaction, City	Total Transaction, City	Comparison with Overall Transaction (%), City
2016	21	126	SP	33.96	173	sao paulo	41	11.05
2017	27	20210	SP	39.95	3148	sao paulo	7360	14.67
2018	27	27384	SP	43.64	3389	sao paulo	10532	16.64

- With the increasing number of new customers every year, it can also be seen that there is an increase in customer coverage, which can be seen from the increasingly varied states and cities of customer origin.
- Most customers come from the state of Sao Paulo, which covers around 34% (in 2016) and continues to increase every year.

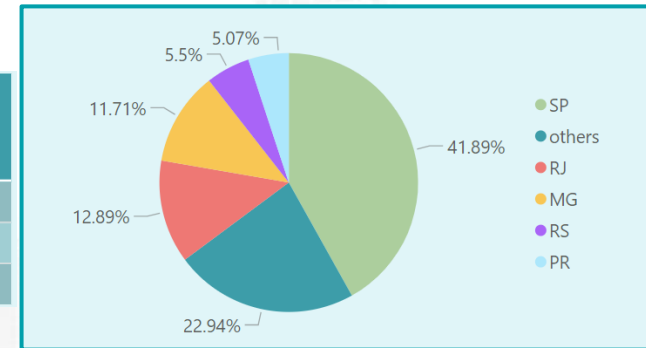


Fig. 3: Customer Origin Percentage (State)

Source : Personal Documentation

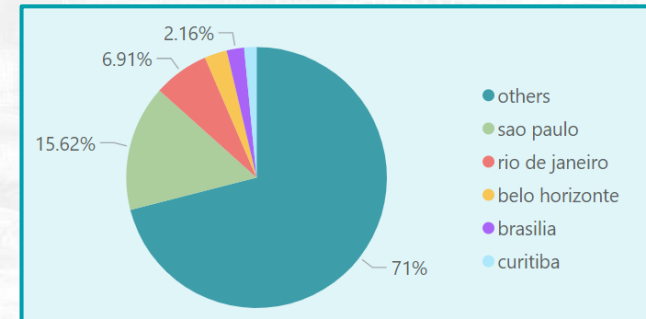


Fig. 4: Customer Origin Percentage (City)

Source : Personal Documentation

Annual Product Category Quality Analysis

Table 4: Product Revenue Summary

Year	Revenue	Top Product	Most Canceled	Total Cancel
2016	45.89K	furniture_decor	toys	15
2017	6,190.34K	bed_bath_table	sports_leisure	247
2018	9,183.54K	health_beauty	health_beauty	280

Table 5: Top Product Summary

Year	First Top Category, Transaction	First Top Category, Revenue	First Top Category, Reorder	First Top Product id	First Top Product Category, Reorder
2016					
1	furniture_decor	furniture_decor	furniture_decor	bdcec8e22b04a8f4241d87bcfe9ea877	furniture_decor
2	health_beauty	perfumery	health_beauty	d9894482fba41f536a273ba2276d951f	market_place
3	perfumery	toys	market_place	cf43c5e600d45ac029357406b05dd82	air_conditioning
2017					
1	bed_bath_table	bed_bath_table	furniture_decor	422879e10f46682990de24d770e7f83d	garden_tools
2	furniture_decor	health_beauty	bed_bath_table	d5991653e037ccb7af6ed7d94246b249	computers_accessories
3	sports_leisure	sports_leisure	sports_leisure	389d119b48cf3043d311335e499d9c6b	garden_tools
2018					
1	health_beauty	health_beauty	bed_bath_table	422879e10f46682990de24d770e7f83d	garden_tools
2	bed_bath_table	watches_gifts	furniture_decor	aca2eb7d00ea1a7b8ebd4e68314663af	furniture_decor
3	computers_accessories	bed_bath_table	computers_accessories	368c6c730842d78016ad823897a372db	garden_tools

- Total revenue continues to show an increase every year.
- Trends in purchasing goods constantly change every year, as seen from the top product categories, which are not always the same. Even so, the furniture_decor category is always the leading product line often reordered by customers.
- The number of transactions experiencing cancellation always increases every year. This is directly proportional to the total number of transactions, which increases yearly.

Annual Product Category Quality Analysis

Table 6: Seller Origin Summary

Year ▲	Total State	Total Transaction, State	Highest Transaction, State	Comparison with Overall Transaction (%), State	Total City	Highest Transaction, City	Total Transaction, City	Comparison with Overall Transaction (%), City
2016	12	212	SP	56.60	80	sao paulo	65	21.02
2017	20	36283	SP	70.94	420	sao paulo	10666	23.93
2018	21	44339	SP	71.69	530	sao paulo	13857	25.42

- As with customers, the domicile of active sellers also varies more and more every year.
- Sao Paulo, which is the state of residence of the highest active sellers, covers at least 56% of all transactions per year, and this figure continues to increase every year until it reaches 72% in 2018
- By combining the 5 states with the highest number of sales transactions, they cover 90% of all transactions.
- The highest distribution of sellers' domicile cities comes from Sao Paulo, while for other cities, it can be said to be relatively even.

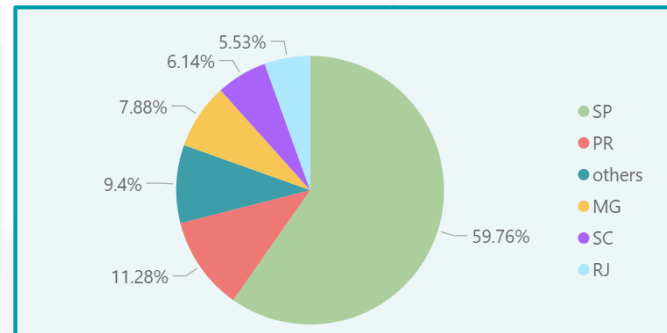


Fig. 5: Seller Origin Percentage (State)

Source : Personal Documentation

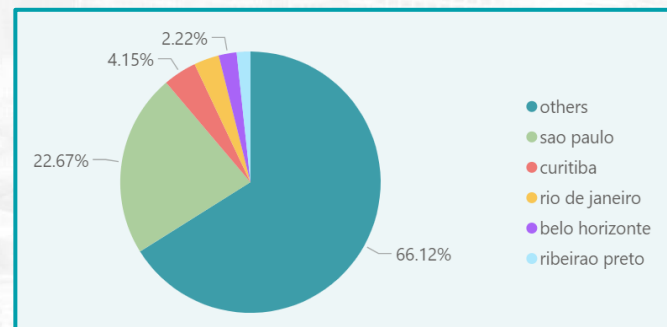


Fig. 6: Seller Origin Percentage (City)

Source : Personal Documentation

Analysis of Annual Payment Type Usage

Table 7: Payment Type Summary

Payment Type	Total Transaction	2016	2017	2018
credit_card	76795	256	30880	45659
boleto	19784	62	8529	11193
voucher	5775	23	2766	2986
debit_card	1529	2	373	1154

Customers prefer to use the credit card payment method because payments can be made in installments, and there is more potential for discounts.

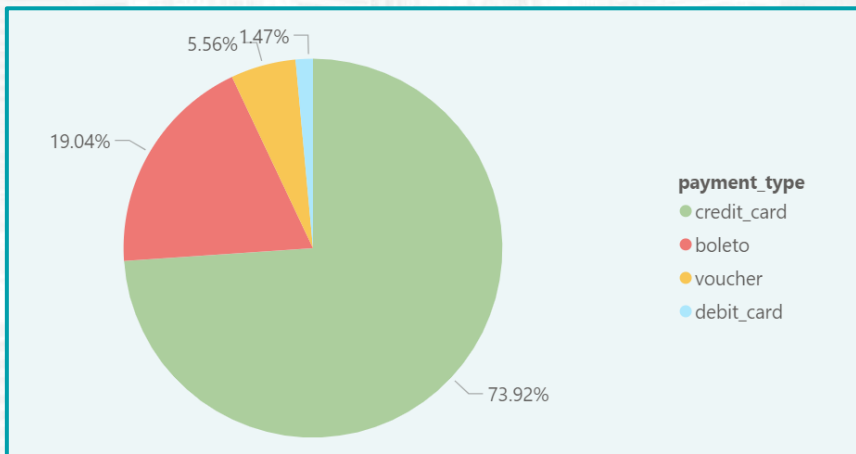


Fig. 7: Payment Type Percentage

Source : Personal Documentation