

ANALYZING ECOMMERCE BUSINESS PERFORMANCE WITH SQL



Created by:
Putrini Nur Amalina H.
putrininur@gmail.com
[Linkedin.com/in/putrininur](https://www.linkedin.com/in/putrininur)

A Diponegoro University graduate who experienced working in the Finance and Accounting Department at a Food Distribution company. A data-driven and tech-savvy person who has huge interest in data analytics and currently learning data science at Data Science Bootcamp who is skilled in SQL, Python, and data visualization using Google Data Studio. Highly skilled in Microsoft Excel and able to actively communicate in English fluently.

Supported by :
Rakamin Academy
Career Acceleration School
www.rakamin.com

OVERVIEW

Performance analysis is a tool you can use to check important metrics every month or year and make plans for adjustment and improvement. By regularly evaluating the performance of your organization can help you understand how much progress you're making toward your goals and measure success or failure in each business process.

Business metric is a quantifiable measure businesses use to track, monitor and assess the success or failure of various business processes. In e commerce company there are many metrics that should be regularly checked including customer growth, product quality and payment methods.

In this paper, we will analyze eCommerce Business Performance using SQL.

GOAL

The goal is to measure e-commerce business performance using business metrics related to customers activity, product quality and payment method. By these information, we can find patterns that will lead to business recommendation based on data.

TOOLS



SQL as
Programming language



PostgreSQL as
tools

DATA PREPARATION

CREATE WORKSPACE

- Create database using CREATE DATABASE statement
- Create table based on dataset using CREATE TABLE statement. The tables are customers, geolocation, order_items, order_payments, order_reviews, orders, product, sellers.
- Make sure to set each data type based on its column.

IMPORTING DATA

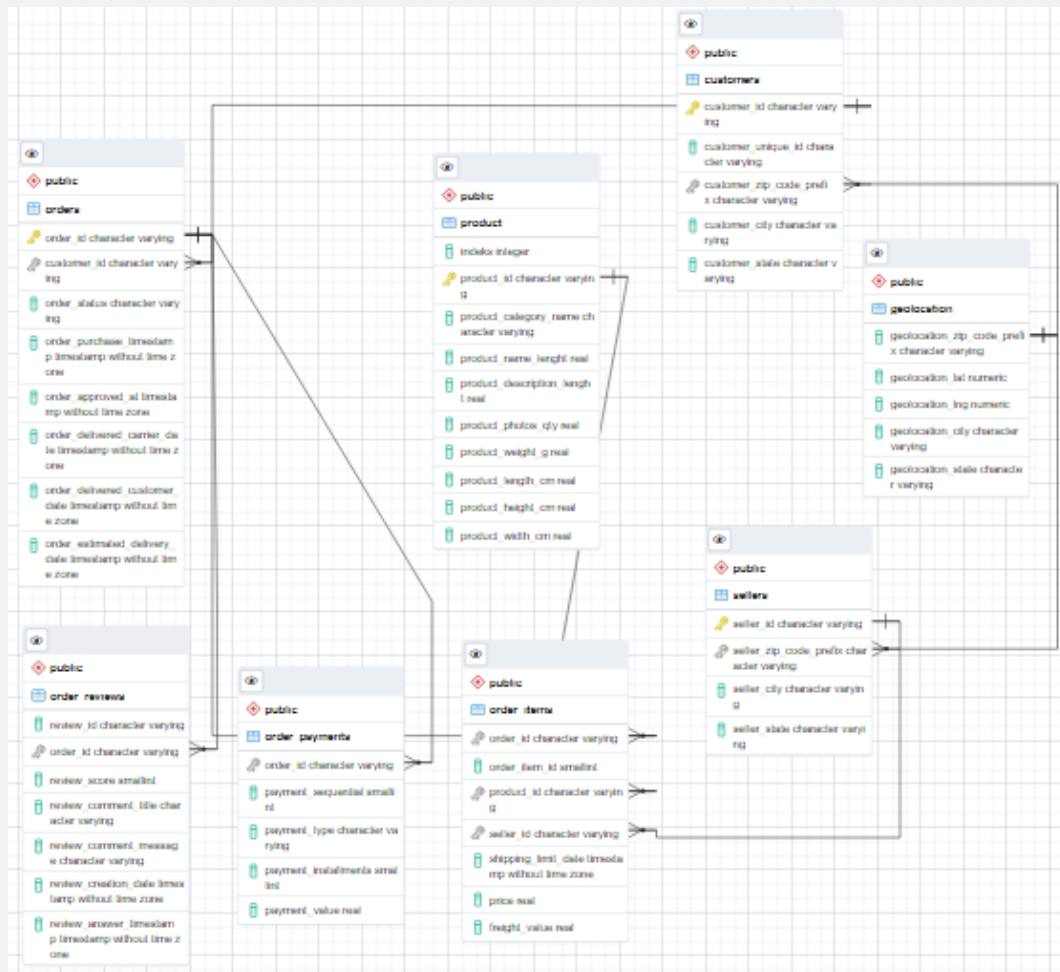
- Import csv dataset using COPY statement to each table.
- Use FROM statement to put directory path of dataset file.

CREATE ERD

- Add primary keys and foreign keys to each tables using ALTER statement
- Generate ERD diagram to show the relation of each tables.

[click to show the complete query](#)

ENTITY RELATIONSHIP DIAGRAM



ANNUAL CUSTOMER ACTIVITY GROWTH

A growing customer is a reflection on a growth sales. That's why customer growth is one of the key goals in e commerce company. It is not only important to grow your customer but also to retain them for the long term.

The following is the result of analysis of business metrics related to the growth of annual customer activity including average monthly active user (MAU), number of new and returning customers per year and average customer purchases per year.

year	average_mau	new_customer	returning_customer	average_order
2016	108.67	326	3	1.01
2017	3694.83	43708	1256	1.03
2018	5338.2	52062	1167	1.02

[click to show the complete query](#)

BUSINESS METRICS 1 :

AVG MONTHLY ACTIVE USER PER YEAR

Active user is a metric to track the number of unique visitors who access their website and make a purchasing activity within a 30-day period. Having a high number of active users benefits businesses and developers by increasing their revenue.

Objectives :

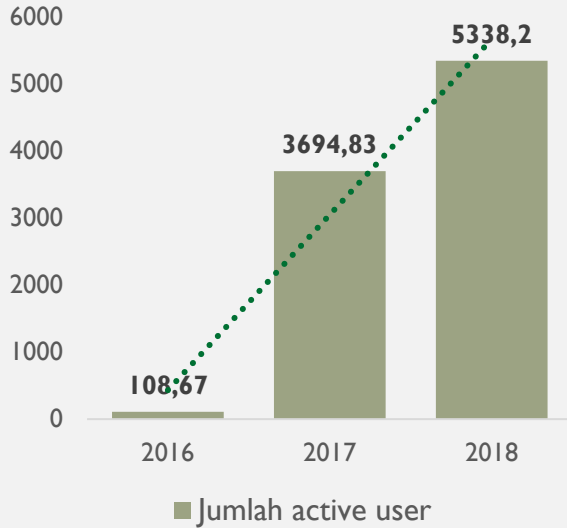
- Identify trends in viewer behaviors, implement features that increase active users and develop marketing strategies to gain additional active users
- Measure the effectiveness of marketing and feature implementation strategies.
- Analyze the customer behavior during key times of the year.

BUSINESS METRICS 1 :

AVG MONTHLY ACTIVE USER PER YEAR

Monthly Active User

Monthly active user increased in last 3 years.



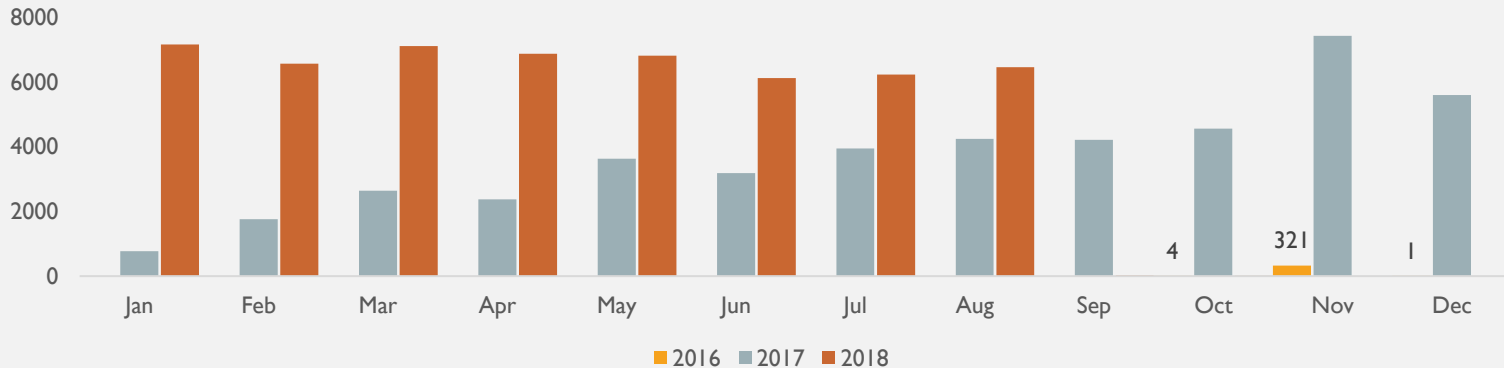
The average of monthly active user is increasing in the last 3 years. This shows that the traffic of customers is increasing that will lead to high revenue.

BUSINESS METRICS 1 :

AVG MONTHLY ACTIVE USER PER YEAR

Monthly Active User

Active user keeps on hitting its peak on November as the holiday season starts.



In 2016, the amount of active user is barely seen as it was the beginning of the e commerce activity as a company. But there is a slight increase at November then decreased again on the next month. In 2017, there are significant grows on the active user amount. The active user keeps on increasing and hits the peak at November. In 2018, the active user average amount is above 6000 users.

Based on data, the active user is always hits its peak on November. We assumed it because the black Friday event which is the busiest shopping day of the year, because it kicks off the holiday season.

BUSINESS METRICS 2 :

NEW CUSTOMER VS RETURNING CUSTOMER

New customer are users who made the first purchase and returning customer are users who have made a purchase before or having a purchasing history for more than one order. For a business to be successful, both new and returning users are needed. They have to fill the pipeline with new users that become returning users who are more likely to convert and spend more money.

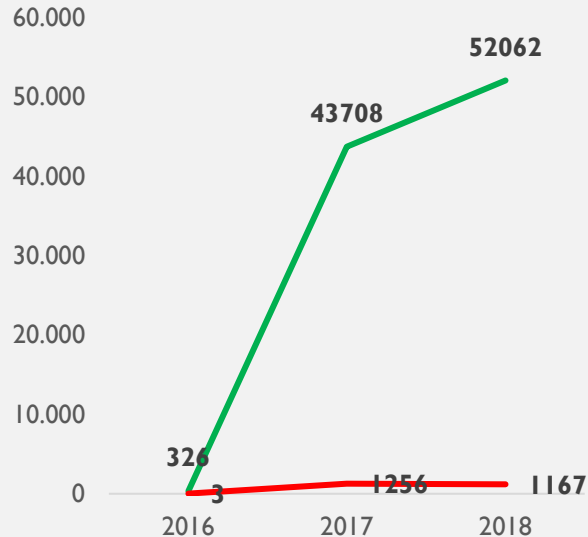
Objectives :

- Measure the effectiveness of your inbound digital marketing techniques across the web by analyzing the ratio between new visitors and returning visitors
- Gain valuable insights about your website content, structure, and design, and it can reveal new opportunities for providing value to your audience by looking at behavioral data for these two groups provides.

BUSINESS METRICS 2 : NEW CUSTOMER VS RETURNING CUSTOMER

New Customer VS Returning Customer

New customers keeps on growing but only few of them make repeat order



The number of new customer keeps on growing in the last 3 years. This measured the increasing traffic of new customer in the ecommerce.

On the other side, the returning customer is decreasing since the last year. This shows that the number of user who do repeat order is very low compared to the amount of user who made first order.

Based on data, we found that the average amount of order each year is only 1 order per customer. Further research needed to analyze the factors of low returning customer.

ANNUAL PRODUCT CATEGORY QUALITY ANALYSIS

The performance of e commerce is relating to the products they provide. By analyzing the quality of product in e commerce, we can gather insights on how to develop the business better. Product quality is important because it affects the success of the company and helps establish its reputation in customer markets. When companies can create high-quality products that continue to meet customer demands, it can lead to higher investment returns and increases in revenue.

The following are the results of business metrics analysis related to product quality which is total revenue per year, number of canceled orders per year, top product categories that generate the largest revenue per year and product categories that generate the most canceled orders per year.

year	total_revenue	cancelled_orders	highest_revenue_product	most_cancelled_product
2016	39827	26	furniture_decor	toys
2017	3388337	265	bed_bath_table	sports_leisure
2018	4374294	334	health_beauty	health_beauty

[click to show the complete query](#)

BUSINESS METRICS 1 :

TOTAL REVENUE

Revenue analysis is a way to measure and track the amount of their sales and profits over a given period. In this dataset, revenue is generated by adding the product price and shipping cost.

Objective :

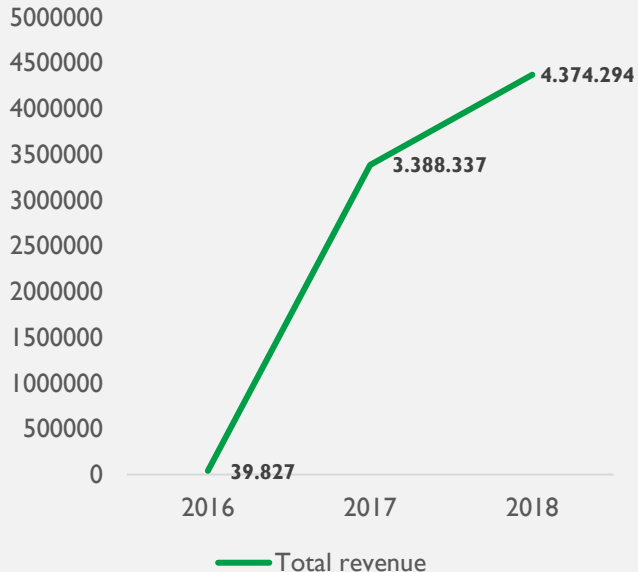
- Identify trends and help make decisions about pricing, marketing, and product development.
- Identify the trend of business market each year
- Recognize which product are most profitable and which product need to be improved or discontinued

BUSINESS METRICS 1 :

TOTAL REVENUE

The annual revenue of the company




The revenue keeps increasing each year



The amount of revenue keeps on growing on the last 3 years. This indicates the company have made a success on their marketing strategy that generates the high amount of sales and revenue.

BUSINESS METRICS 2 : HIGHEST REVENUE PRODUCT

By recognizing the highest sales product, we can determine the customer profile which is a critical section of a company's business plan or marketing plan.

Top product of the year	2016	2017	2018
			
Revenue on each category	furniture_decor \$ 7190	bed_bath_table \$ 590,673	health_beauty \$ 885,183

After further analysis, we found that there are several product categories that generates the highest revenue each year. There are changes on product trend in market as the product that hit the highest revenue are changing every year.

In 2016 and 2017, the product dominated by furniture and appliances category. Based on study, the majority customer of these product are male adult and mostly are business people as these products having a high average of price. Meanwhile in 2018, the product dominated by health and beauty category. Based on study, women are more likely to purchase beauty products than men.

BUSINESS METRICS 3 : TOTAL CANCELLED ORDER

Cancelled order is purchased order cancelled by the consumer.

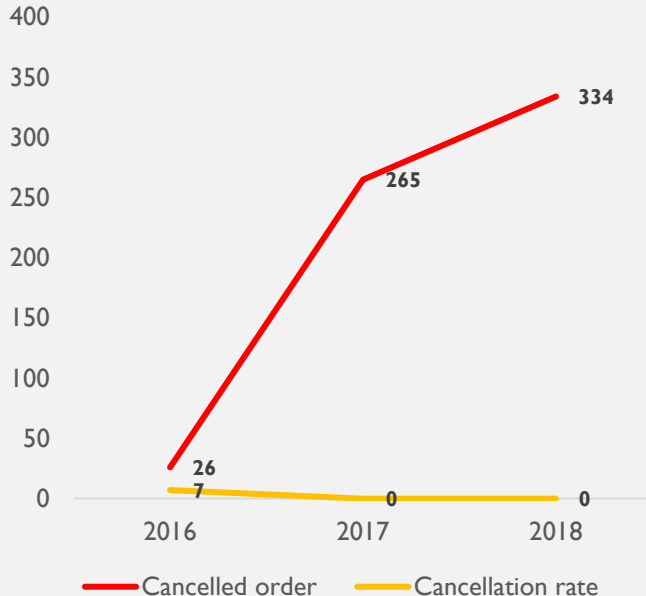
When an order cancelled, it will inflate the operational cost and don't generate any return on investment. Then affecting profit margins, reducing retention rates, and making it difficult for businesses to scale.

Customers who cancel orders also often leave negative reviews, hampering businesses' reputations and leading to customer churn.

BUSINESS METRICS 3 : TOTAL CANCELLED ORDER

The annual cancelled order

Despite the low cancellation rate, the number of cancelled order keeps increasing



Based on data, the amount of cancelled order is increasing each year. But when compared to the total order each year, the cancellation rate is decreasing.




However, in order to prevent loss in revenue and churn customers, the best decision is to keep the number of cancelled order to be minimum. Therefore, further research are needed to determine the factors of cancelled order.

BUSINESS METRICS 4 : MOST CANCELLED ORDER PRODUCT

Based on studies, customer preference-based (style, size, fit) returns take 72% of why customer cancel order and non preference based (broken, faulty items) take up to 10%. Another studies said that the various reason are extended delivery period, high shipping expenses, expensive service charge, and other reasons.

Customer preference-based reason is having high relation to the quality of the product. Therefore we need to analyze which product category have the most cancelled order.

BUSINESS METRICS 4 : MOST CANCELLED ORDER PRODUCT

Most cancelled product	2016	2017	2018
			
	toys	sport_leisure	health_beauty
The amount of loss each year	\$ 3391	\$ 52,120	\$ 50,398

In last 3 year, the trend of most cancelled product category is always changing. Despite being the highest revenue product in 2018, health and beauty product also being the most cancelled product in the same year. We assumed it because the number of order in this category is high so the probability of the product being cancelled is also high. Further research needed to analyze why these categories having most cancelled order.

ANNUAL PAYMENT TYPE USAGE ANALYSIS

Payment method is one of the aspects to consider in revenue analysis by comparing the performance of different payment methods. Based on study, 50% of costumers will abandon their purchase if they don't see their preferred payment method available. So it would be better to provide alternative payment method to check out. Also if one of your payment systems causes problems, having further options available is a good way to sustain conversions and avoid a drop in buyer numbers due to technical faults.

The transaction in e commerce usually provide open payment system so the customers able to chose various type of payment methods. Analyzing performance of each payment type will generate insights to open strategic partnership with better payment provider company. The following are the results of business metrics analysis related to payment type which is most usage payment method and the amount of each payment type per year.

payment_type	amount
credit_card	76795
boleto	19784
voucher	5775
debit_card	1529
not_defined	3

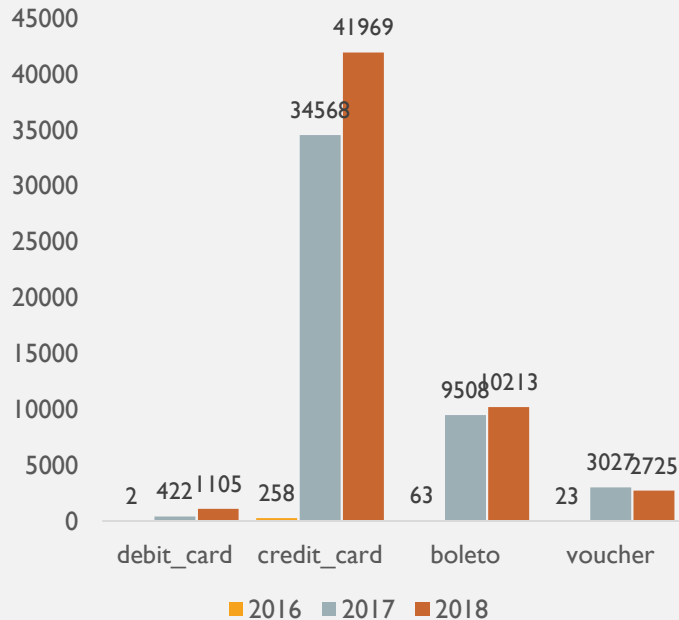
payment_type	year_2016	year_2017	year_2018	pct_change_2017_2018
not_defined	0	0	3	NULL
debit_card	2	422	1105	161.85
credit_card	258	34568	41969	21.41
boleto	63	9508	10213	7.41
voucher	23	3027	2725	-9.98

[click to show the complete query](#)

BUSINESS METRICS 1 : TOP PAYMENT METHOD PER YEAR

Analysis on annual payment method users

Credit card is the most popular payment method of all years



Based on data, the most popular payment method on the last 3 years is credit card, then followed by boleto, voucher and debit card. This supported by study, that 80% of transactions made by millennials are made by card. Payment method using credit card offers many benefits including rewards and promo.

Payment method by debit card have significant increase from 2017 to 2018 with percentage more than 100%. On the other side, payment method using voucher is decreasing from 2017 to 2018.

Based on these findings, we can dig deep to each payment method by determining each customer behavior including most product category checked out by each payment method, payment tenor, etc. The trend of payment method is also gives insights for decision maker to make business plan either to open better strategic partnership with payment provider or create marketing campaign.