Analyzing eCommerce Business Performance with SQL

Kevin Haposan Aripratama
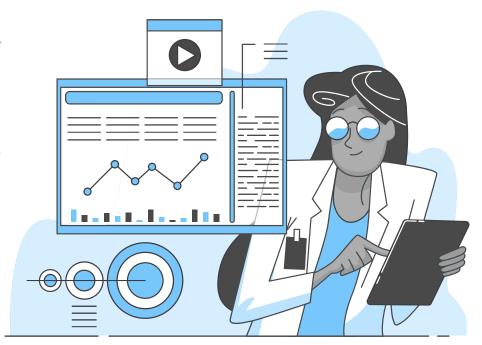
Table of Contents
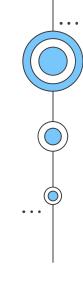
01
Data Preparation

Overview

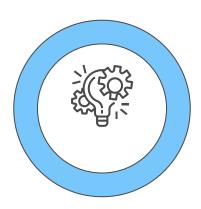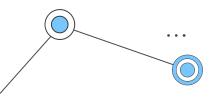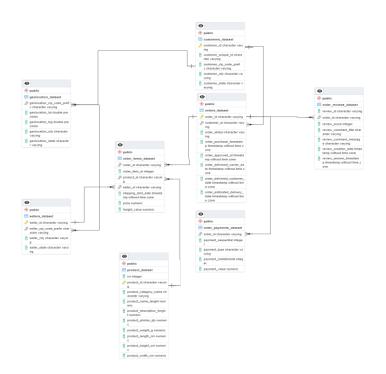In a company, measuring business performance is crucial for tracking, monitoring, and assessing the success or failure of various business processes. Therefore, this paper will analyze the business performance for an eCommerce company, taking into account several business metrics, namely customer growth, product quality, and payment methods
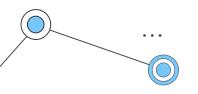
Data Preparation

Entity-Relationship Diagram
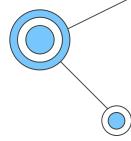
A. Creating queries to generate tables.

B. Importing datasets into the database.

C. Setting up all tables to have a primary key
The columns in geolocation_dataset (geolocation_zip_code_prefix column), order_items_dataset (seller_id column), and seller_dataset (seller_id column) cannot be set primary keys due to the possibility of duplicate data.

D. Checking whether the dataset contains duplicate data
The Geolocation_dataset has duplicate data, necessitating the cleaning of duplicate entries. A check was conducted on the order_items_dataset and sellers_dataset. It was found that both datasets contain unique rows (no duplicates).

Click ERD to see the result
Click here to see the SQL query

Data Preparation
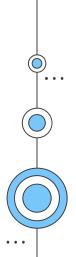
E. Removing duplicated data

F. Checking the dataset column-wise
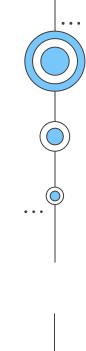   When checking the columns that should be primary keys in the geolocation_dataset, order_items_dataset, and sellers_dataset, it was found that these columns contain duplicates. Therefore, the creation of the ERD is performed using a menu (without query).

G. Creating an ERD using queries and menu in PgAdmin.

Summary:
• Out of the 8 datasets, there are datasets that are unique (can have primary keys assigned) and non-unique (contain duplicate data or non-unique primary keys).
• Duplicate dataset: geolocation_dataset
• Datasets with non-unique columns: geolocation_dataset (geolocation_zip_code_prefix column), order_items_dataset (seller_id column), and seller_dataset (seller_id column)
• Unique datasets have undergone ETR creation using queries, while non-unique datasets were created using the menu in pgAdmin.

Click ERD to see the result
Click here to see the SQL query

02

Annual Customer Activity Growth Analysis

Annual Customer Activity Growth Analysis
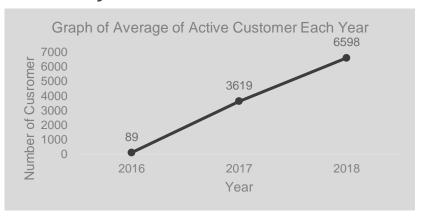
Master Table

	year double precision 🔒	average_active_customer numeric 🔒	number_of_new_customer bigint 🔒	number_of_customer_repeat_order bigint 🔒	average_of_customer_order numeric 🔒
1	2016	89	264	3	1
2	2017	3619	42059	1170	1
3	2018	6598	51444	1086	1

Click here to see the SQL query

Annual Customer Activity Growth Analysis

Average of Active Customer Each Year
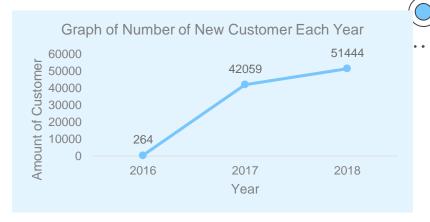
The increase in average active customers may indicate effective strategies such as promotional activities, membership programs, or offering trendy products, which have successfully attracted more customers to engage with the business over time.

Number of New Customer Each Year

The growth of new customers from year to year may be due to the development of the store, which might follow trends and attract many customers to shop at this store. Additionally, the store may frequently hold promotions for new customer.

Click here to see the SQL query

Annual Customer Activity Growth Analysis

Number of Customer Who Make Repeat Order Each Year

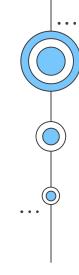Graph of Number of Customer Who Make Repeat Order Each Year

The increase in the number of customers making repeat orders from 2016 to 2017 indicates the effectiveness of marketing strategies or loyalty programs implemented. Despite a slight decline in 2018, the number of customers making repeat orders remains relatively high, indicating success in retaining customer base. Recommendations include strengthening loyalty programs, continuous evaluation of promotions, and improving service quality to prolong customer retention.

Average of Customer Order Each Year

Graph of Average of Customer Order Each Year

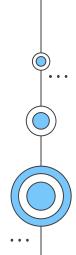The average of one order per customer each year from 2016 to 2018 indicates potential to enhance customer engagement and increase purchase frequency. To boost the number of customer orders, more aggressive promotional strategies should be introduced, such as seasonal discounts, bundle offers, or loyalty incentives.

Click here to see the SQL query

03

Annual Product Category Quality Analysis

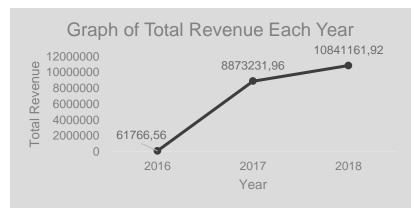# Annual Product Category Quality Analysis

Master Table

	year double precision 🔒	total_revenue numeric 🔒	total_canceled_order bigint 🔒	highest_revenue_product text 🔒	highest_cancel_product text 🔒
1	2016	61766.56	28	furniture_decor - 10699.61	toys - 3
2	2017	8873231.96	333	bed_bath_table - 786557.77	sports_leisure - 22
3	2018	10841161.92	384	health_beauty - 1008401.09	health_beauty - 27

Click here to see the SQL query

Annual Customer Activity Growth Analysis

Total Revenue Each Year

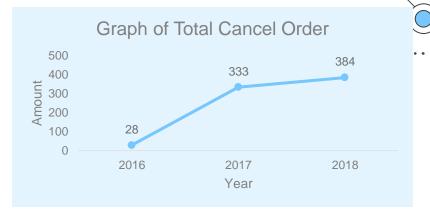### Graph of Total Revenue Each Year

The revenue is taken from orders with the status "delivered" as there's a possibility of other statuses like cancellations, which could impact the accuracy of the calculations. The total revenue showed a significant increase from 2016 to 2018, indicating robust business growth. To sustain this upward trend, implementing more aggressive marketing strategies, enhancing customer service, and innovating products in line with current market trends are recommended.

Total Cancel Order Each Year

Graph of Total Cancel Order

The total number of canceled orders saw an increase over the years, signaling potential issues in customer satisfaction or operational efficiency that may impact revenue. To maintain revenue growth, it's crucial to focus on improving processes, addressing customer concerns promptly, and optimizing product quality to minimize cancellations and enhance overall customer experience.
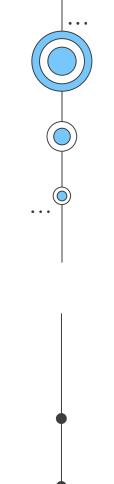
Click here to see the SQL query

Annual Customer Activity Growth Analysis

	2016	2017	2018
Highest Revenue Product			
	Furniture Décor	Bed bath table	Beauty Health
	$10,699.61	$ 786,557.77	$1,008,401.09
Highest Cancel Product			
	Toys	Sports Leisure	Beauth Health
	3 orders	22 orders	27 orders

Products with the highest revenues that fluctuate every year indicate dynamic consumer trends and market demands that need to be monitored carefully by the business world. Businesses must monitor consumer trends, implement flexible inventory management, prioritize high-performing categories, adjust marketing strategies, and foster supplier partnerships for sustainable revenue growth.

In 2018, the total revenue and the highest number of canceled items were found in the health beauty category. This may indicate that there were many sales in the health beauty category that year.
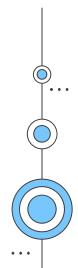
The consistently high cancellation rates for different product categories each year highlight potential areas of improvement in product quality or customer satisfaction. To mitigate cancellations and enhance business performance, companies should focus on quality assurance measures, gather customer feedback for product enhancements, and implement proactive customer service initiatives to address concerns promptly.
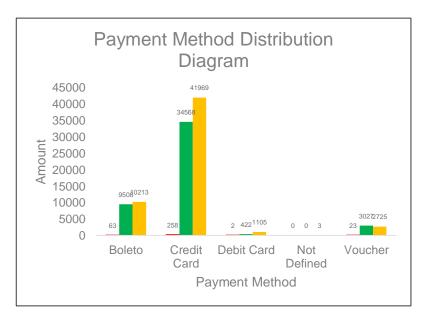
Click here to see the SQL query

04

Analysis of Annual Payment Type Usage

… Analysis of Annual Payment Type Usage

Payment Method Distribution Diagram

The use of credit cards becomes the most preferred payment method each year. This could be due to the option for installment payments when purchasing goods. To enhance this payment… method, an analysis can be conducted regarding the number of installments taken by customers.

Debit card payments increased by more than 150% from 2017 to 2018. This increase could be attributed to partnerships with companies.

Conversely, payments with vouchers experienced a decline. This decline might have occurred due to a reduction in voucher quotas, thus diminishing customer interest in purchasing.

Recommendations that can be given include analyzing purchases made with debit cards and vouchers to increase customer interest in buying products."

Click here to see the SQL query

Thank You