# Analyzing eCommerce Business Performance with SQL

**Created by:**
**Samuel Akwila**
sakwila96@gmail.com
LinkedIn | Github

"I was graduated from Bogor Agricultural University at 2018. I am working as finance and accounting in manufacturing company. I join Rakamin Data Science Bootcamp in January 2022, to enhance my skill in data field and hopefully to switch my career to data field."

Rakamin Academy

"In this project I analyzed business performance of an e-Commerce that involves reviewing customer growth, products quality, and payment methods using PostgreSQL, and the visualisation using Tableau and Ms. Excel.

Datasets that I used in this project was provided by Rakamin Academy, it was from multiple eCommerce in Brazil that covers around 100.000 datasets from 2016-2018. The datasets provides some metrics such: orders, payments, customers, sellers, locations, products, and reviews. These solutions on solving the problem heavily rely on the collection of e-commerce-related data."

# Data Preparation

Before processing the data, first step that should be done is preparing raw data and make it ready to process data. These are the steps on making a new database:

1. Creating the new database with all tables for the prepared data, by paying attention on data types for every columns in tables.
2. Importing the file into the database.
3. Creating the entity relationship diagram (ERD) by paying attention on relationship between every tables.
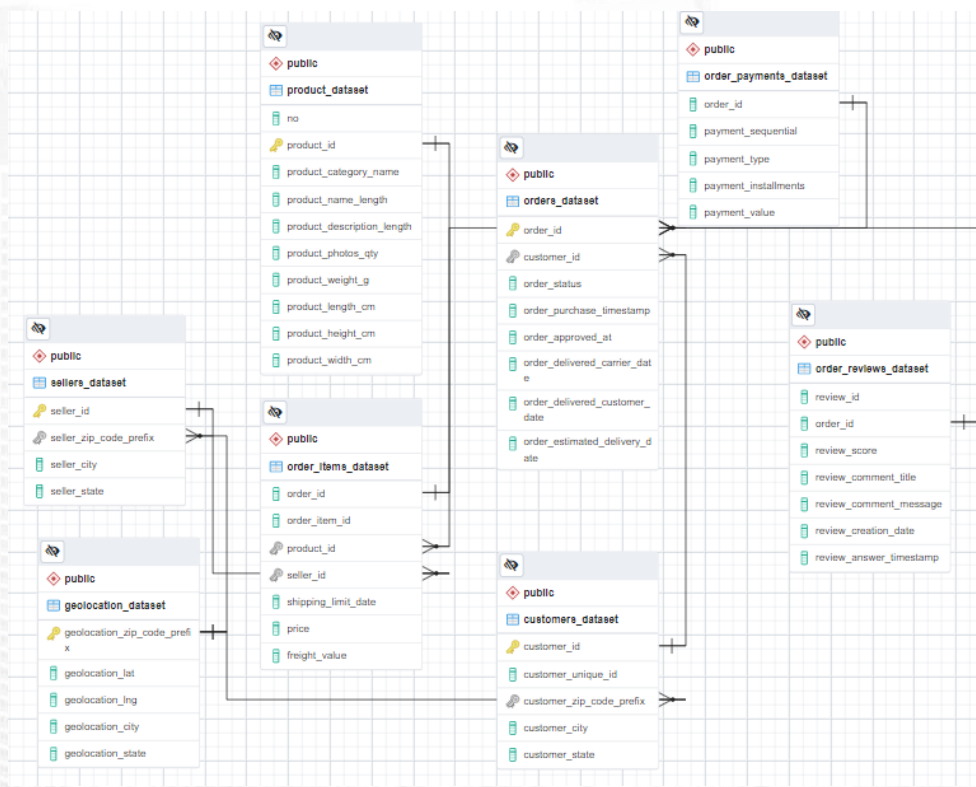


**Figure 1. Entity Relationship Diagram (ERD)**

Open and download ERD here

# Annual Customer Activity Growth Analysis

There are some metrics the needed on analysis customer activity growth, such as monthly active users, new customers, repeat order, and average order frequency per customers (Table 1).

| | year<br>double precision | average_mau<br>numeric | new_customers<br>bigint | repeating_customers<br>bigint | avg_order_freq_per_cust<br>numeric |
|---|---|---|---|---|---|
| 1 | 2016 | 110 | 326 | 3 | 1.01 |
| 2 | 2017 | 3758 | 43708 | 1256 | 1.03 |
| 3 | 2018 | 5401 | 52062 | 1167 | 1.02 |

Table 1. Customer growth metrics

Query selengkapnya dapat dilihat disini
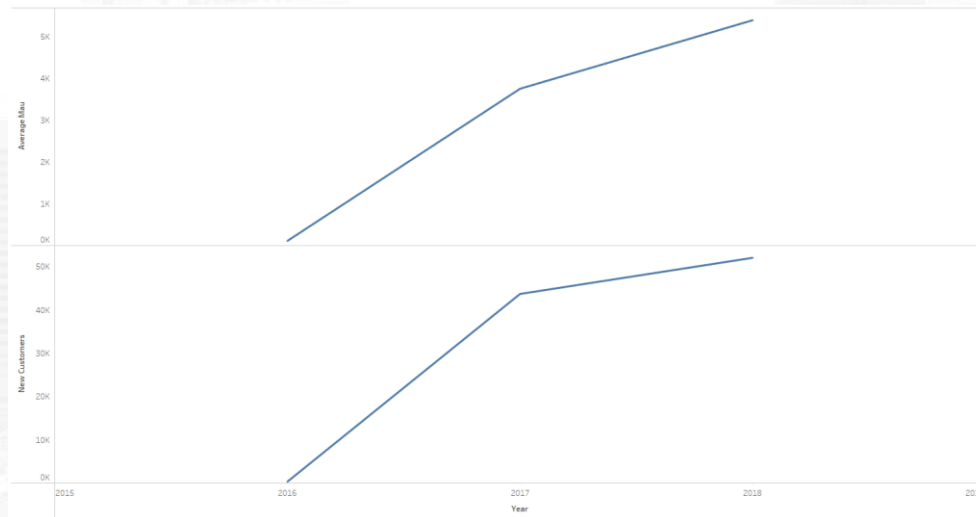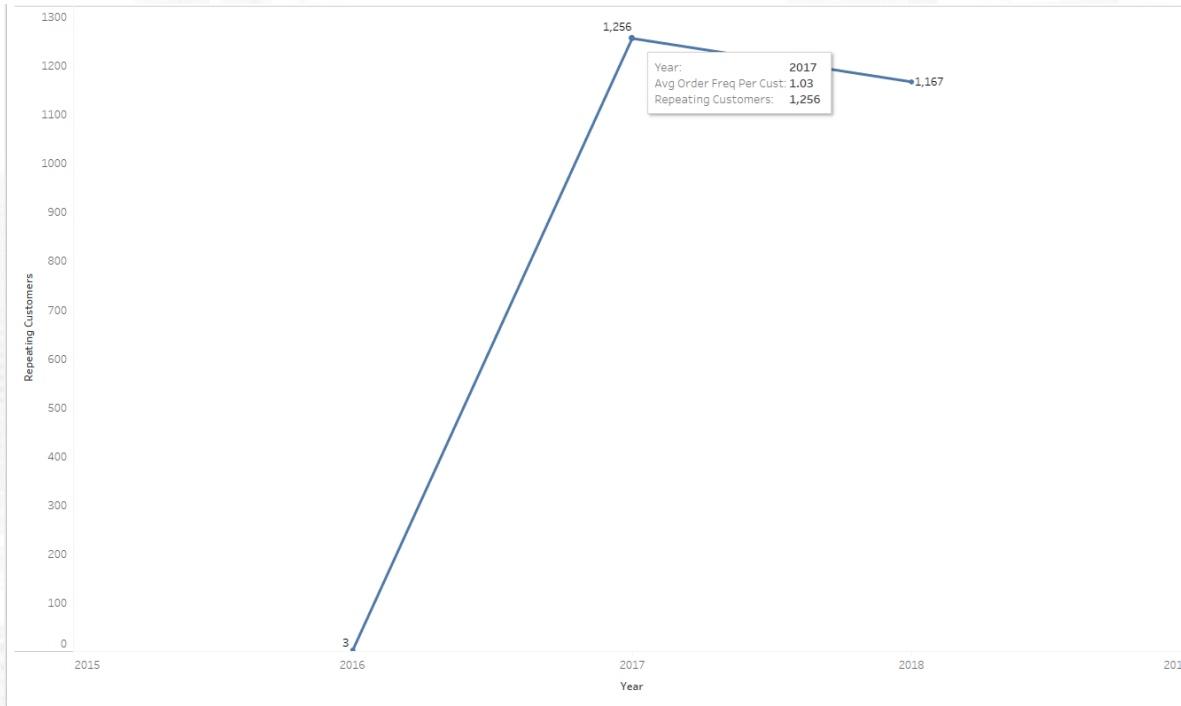
# Annual Customer Activity Growth Analysis



Figure 2. Monthly active users (MAU) vs new customers graphic in 2016-2018

1. For both metrics (MAU and new customers), was increasing from 2016-2018
2. Both metrics significantly increase in 2017, it happened because of the data in 2016 only provided from September-December

Figure 3. Repeat orders graphic in 2016-2018

- Repeat orders in 2017 has significantly increase because data in 2016 only from September - December
- In 2018 repeat orders was decreasing around 7%, it would be happen because the data is only until September 2018

# Annual Product Category Quality Analysis

| year double precision | category_most_revenue character varying | most_revenue double precision | total_revenue_per_year double precision | category_most_cancel character varying | num_most_cancel bigint | total_cancel bigint |
|---|---|---|---|---|---|---|
| 1 | 2016 | furniture_decor | 6899.35 | 46653.74000000001 | toys | 3 | 26 |
| 2 | 2017 | bed_bath_table | 580949.2000000012 | 6921535.239999682 | sports_leisure | 25 | 265 |
| 3 | 2018 | health_beauty | 866810.3399999978 | 8451584.769999918 | health_beauty | 27 | 334 |

Table 2. Product category analysis

<u>2016 Recap</u>
- Product category furniture_decor got highest revenue than other product, which is 6,899.35
- Total revenue is 46,653.74
- This means, furniture_decor contribute around 14.8% of total revenue in this year

- The most cancel category is toys with 3 cancel orders. Total cancel in this year is 26 orders
- This means toys contribute around 11.5% of total cancel orders in 2016

Query selengkapnya dapat dilihat disini

## 2017 Recap

- Product category bed_bath_table got highest revenue than other product, which is 580,949.20
- Total revenue is 6,921,535.24
- This means, bed_bath_table contribute around 8.4% of total revenue in this year

- The most cancel category is sports_leisure with 25 cancel orders. Total cancel in this year is 265 orders
- This means, sports_leisure contribute around 9.4% of total cancel in 2017

## 2018 Recap

- Product category health_beauty got highest revenue than other product, which is 866,810.34
- Total revenue is 8,451,584.77
- This means, health_beauty contribute around 10.25% of total revenue in this year

- The most cancel category in this year also health_beauty with 27 from 334 total cancel in this year
- This means, health_beauty contribute around 8% of total cancel in 2018
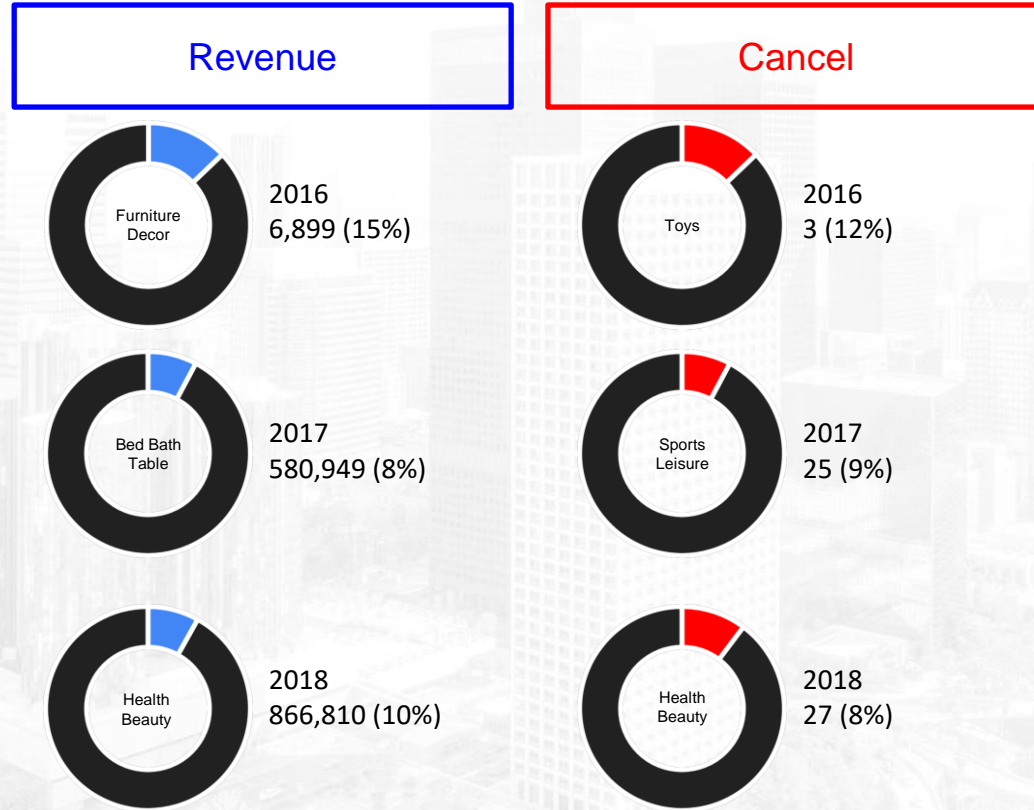
# Annual Product Category Quality Analysis

## Revenue

**Furniture Decor**
2016
6,899 (15%)

**Bed Bath Table**
2017
580,949 (8%)

**Health Beauty**
2018
866,810 (10%)

## Cancel

**Toys**
2016
3 (12%)

**Sports Leisure**
2017
25 (9%)

**Health Beauty**
2018
27 (8%)

Figure 4. Product category performance 2016-2018

# Analysis of Annual Payment Type Usage

| # | payment_type<br>character varying | year_2016<br>numeric | year_2017<br>numeric | year_2018<br>numeric | growth_pct<br>numeric |
|---|---|---|---|---|---|
| 1 | not_defined | 0 | 0 | 3 | [null] |
| 2 | debit_card | 2 | 422 | 1105 | 1.62 |
| 3 | credit_card | 258 | 34568 | 41969 | 0.21 |
| 4 | boleto | 63 | 9508 | 10213 | 0.07 |
| 5 | voucher | 23 | 3027 | 2725 | -0.1 |

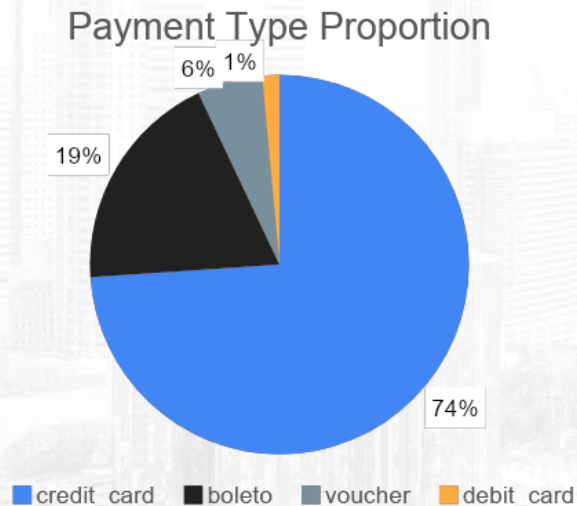Table 3. Annual payment type usage

## Payment Type Proportion



Figure 5. Payment type proportion

- Credit_card was all time most favorite payment method
- There are 76795 payments using credit_card (74%)
- Second was boleto with 19784 payments (19%)
- The least is debit_card besides the not_defined payment_type (6%)
- Yearly, credit_card still the most favorite based on total usage
- The order is still the same as all time

# Analysis of Annual Payment Type Usage



Figure 6. Payment type performance
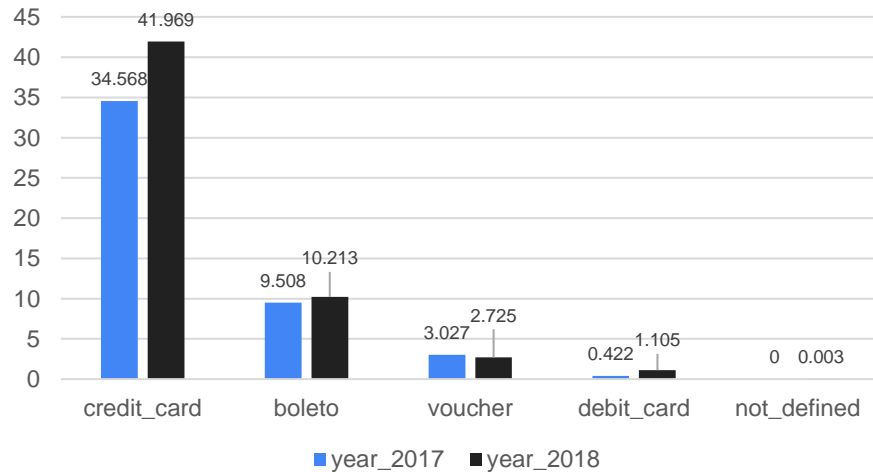
- From 2017 to 2018, debit_card payment_type has the biggest growth than the others
- Voucher has a negative growth which means the usage in 2018 was decreasing from 2017