

Analyzing eCommerce Business Performance with SQL



Created by:

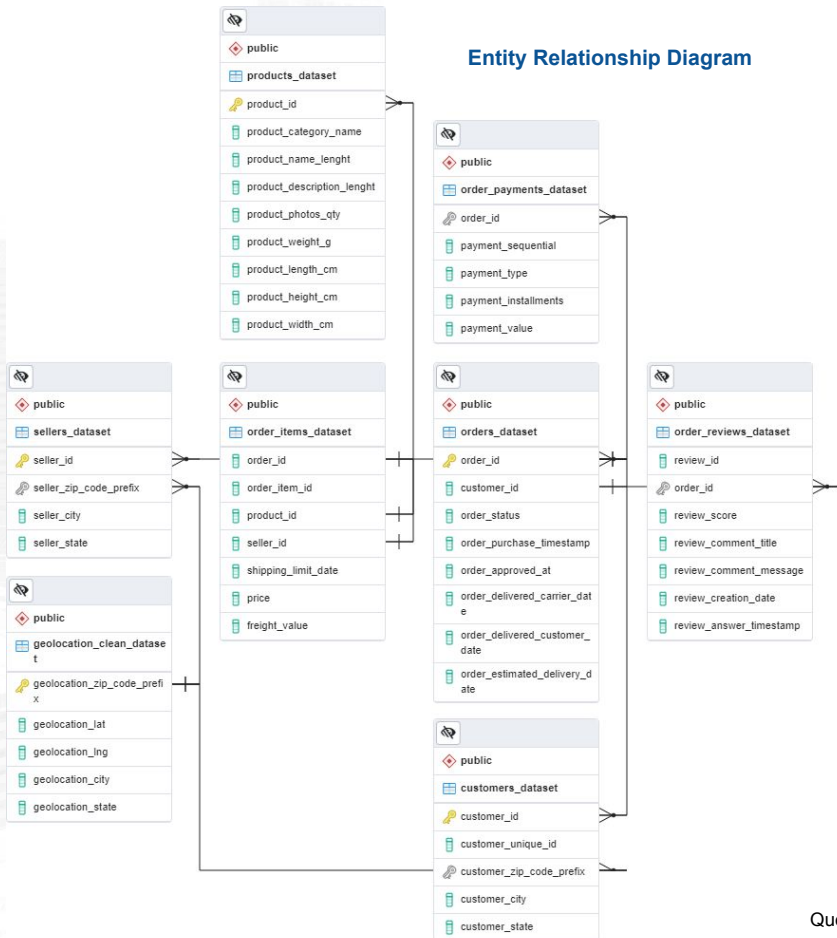
Siti Hajjah Mardiah

sitihamardiah1997@gmail.com

<https://www.linkedin.com/in/sitihajjahmardiah/>

“Graduated from Diponegoro University, Semarang. Having more 3 years of experience working as a Recruiter, Account Manager and Data Analyst. Experienced in recruiting candidates from various fields including Tech and Non-Tech positions. Enthusiastic and open to any opportunity in the field of data, especially as a Data Scientist. Skilled in using PostgreSQL, Python, Git, Github, Looker Studio, Tableau, Power BI, Microsoft Office (Word, Excel, PowerPoint), etc. I am open to any opportunity which relate in data field including Data Scientist, Data Analyst, Business Analyst, Business Intelligence and so on.”

“Dalam suatu perusahaan mengukur performa bisnis sangatlah penting untuk melacak, memantau, dan menilai keberhasilan atau kegagalan dari berbagai proses bisnis. Oleh karena itu, dalam paper ini akan menganalisa performa bisnis untuk sebuah perusahaan eCommerce, dengan memperhitungkan beberapa metrik bisnis yaitu pertumbuhan pelanggan, kualitas produk, dan tipe pembayaran.”



Sebelum memulai pemrosesan data, tahap paling awal yang harus dilakukan adalah melakukan **data preparation** atau mempersiapkan data mentah menjadi data yang terstruktur dan siap diolah.

Tahap melakukan **data preparation** adalah sebagai berikut:

1. Download file "Dataset.rar", kemudian ekstrak file tersebut.
2. Membuat database baru beserta tabel-tabelnya untuk data yang sudah disiapkan
3. Importing data csv ke dalam database
4. Membuat entity relationship antar tabel (ERD)

Dari ERD di samping, dapat dilihat keterkaitan antar tabel:

- Tabel **order_items_dataset** memiliki keterkaitan antar:
 - Tabel **product_dataset** yang dihubungkan dengan kolom **product_id**.
 - Tabel **sellers_dataset** yang dihubungkan dengan kolom **seller_id**.
 - Tabel **orders_dataset** yang dihubungkan dengan kolom **order_id**.
- Tabel **orders_dataset** memiliki keterkaitan antar:
 - Tabel **order_payments_dataset** yang dihubungkan dengan kolom **order_id**.
 - Tabel **order_reviews_dataset** yang dihubungkan dengan kolom **order_id**.
 - Tabel **order_items_dataset** yang dihubungkan dengan kolom **order_id**.
 - Tabel **customers_dataset** yang dihubungkan dengan kolom **customer_id**.
- Tabel **geolocation_clean_dataset** memiliki keterkaitan antar:
 - Tabel **customers_dataset** yang dihubungkan dengan kolom **zip_code_prefix**.
 - Tabel **sellers_dataset** yang dihubungkan dengan kolom **zip_code_prefix**.

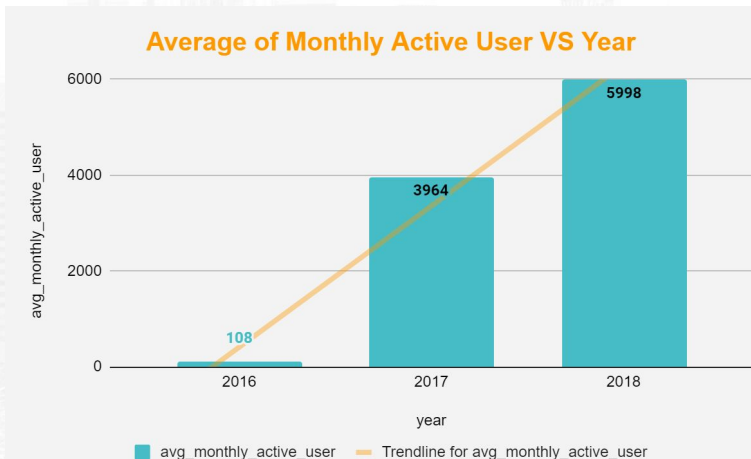
Salah satu metrik yang digunakan untuk mengukur performa bisnis eCommerce adalah aktivitas customer yang berinteraksi di dalam platform eCommerce tersebut. Untuk mengetahui Annual Customer Activity, diperlukan analisis beberapa metrik yang berhubungan dengan aktivitas customer seperti **jumlah customer aktif**, **jumlah customer baru**, **jumlah customer yang melakukan repeat order** dan juga **rata-rata transaksi yang dilakukan customer setiap tahun**.

Tahapan analisis:

1. Menampilkan **rata-rata jumlah customer aktif bulanan (Monthly Active User)** untuk setiap tahun
(Hint: Perhatikan kesesuaian format tanggal)
2. Menampilkan **jumlah customer baru** pada masing-masing tahun
(Hint: Pelanggan baru adalah pelanggan yang melakukan order pertama kali)
3. Menampilkan **jumlah customer yang melakukan pembelian lebih dari satu kali (repeat order)** pada masing-masing tahun
(Hint: Pelanggan yang melakukan repeat order adalah pelanggan yang melakukan order lebih dari 1 kali)
4. Menampilkan **rata-rata jumlah order yang dilakukan customer** untuk masing-masing tahun
(Hint: Hitung frekuensi order (berapa kali order) untuk masing-masing customer terlebih dahulu)
5. Menggabungkan ketiga metrik yang telah berhasil ditampilkan menjadi satu tampilan tabel

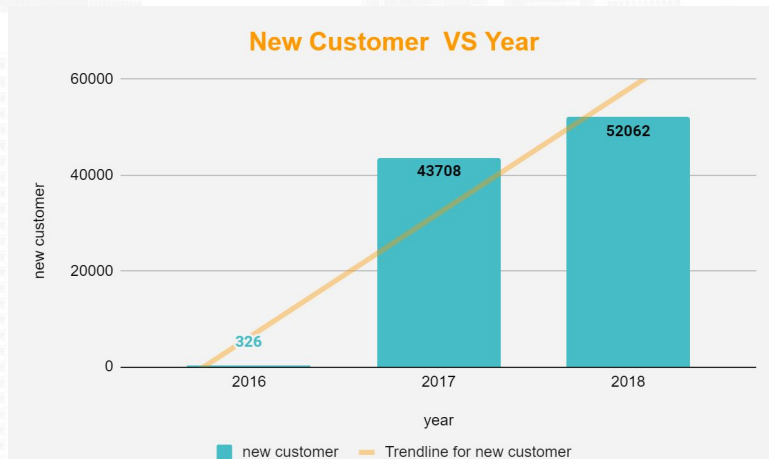
Annual Customer Activity Growth Analysis

1. Rata-rata Jumlah Customer Aktif Bulanan Setiap Tahun



Berdasarkan bar chart di atas, ditemukan bahwa **Rata-rata (Average) Monthly User Active (MAU)** terbanyak terjadi di tahun 2018 dengan jumlah customer sebanyak **5.998 active users**. Berdasarkan trend, setiap tahunnya jumlah customer active selalu meningkat.

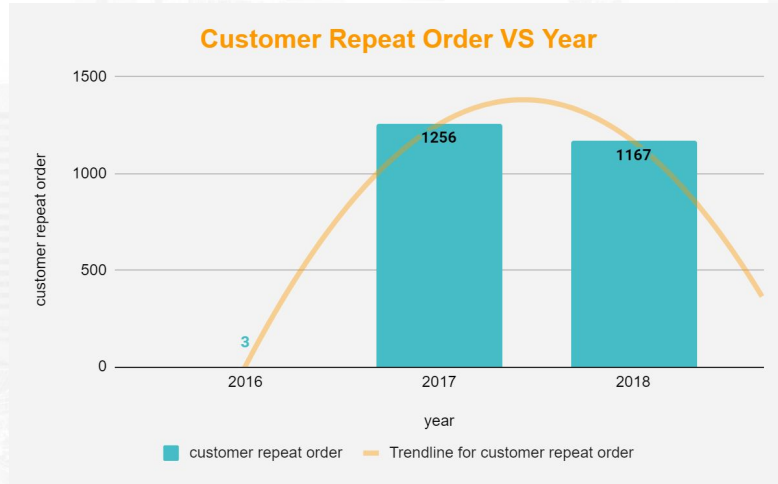
2. Jumlah Customer Baru Setiap Tahun



Berdasarkan bar chart di atas, ditemukan bahwa **Jumlah Customer Baru** terbanyak berada di tahun 2018 sebanyak **52.062 users**. Berdasarkan trend, setiap tahunnya jumlah customer baru selalu meningkat.

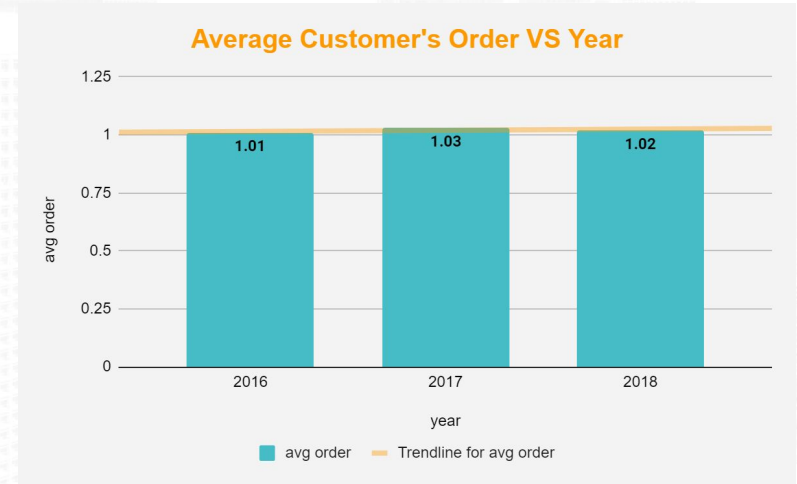
Annual Customer Activity Growth Analysis

3. Jumlah Customer Repeat Order



Berdasarkan bar chart di atas, ditemukan bahwa **Jumlah Customer Repeat Order terbanyak berada di tahun 2017 sebanyak 1.256 users**. Berdasarkan trend, dari tahun 2016 menuju tahun 2017 terjadi kenaikan drastis jumlah customer repeat order sebanyak 1.256 users di tahun 2017, dan kemudian di tahun 2018 mengalami penurunan dengan jumlah customer sebanyak 1.167 users.

4. Rata-rata Order Customer Setiap Tahun



Berdasarkan bar chart di atas, ditemukan bahwa **Rata-rata Jumlah Order yang dilakukan customer setiap tahunnya adalah 1**.

5. Gabungan Metrik Annual Customer Activity Growth

Year	Average of Monthly Active User	New Customer	Customer Repeat Order	Average of Order per Customer
2016	108	326	3	1,01
2017	3.964	43.708	1.256	1,03
2018	5.998	52.062	1.167	1,02

Tabel diatas adalah beberapa metrik yang berhubungan dengan aktivitas customer yang terdiri dari Rata-rata (Average) Monthly User Active, Jumlah Customer Baru, Jumlah Customer yang melakukan Repeat Order dan Rata-rata Transaksi yang dilakukan customer setiap tahun. Berikut adalah detailnya:

- Metrik **Rata-rata (Average) Monthly User Active (MAU) terbanyak terjadi di tahun 2018** dengan jumlah customer sebanyak **5.998 active users**.
- Metrik **Jumlah Customer Baru** terbanyak berada di **tahun 2018** sebanyak **52.062 users**.
- Metrik **Jumlah Customer Repeat Order** terbanyak berada di **tahun 2017** sebanyak **1.256 users**.
- Metrik **Rata-rata Jumlah Order** yang dilakukan customer **setiap tahunnya** adalah **1**.

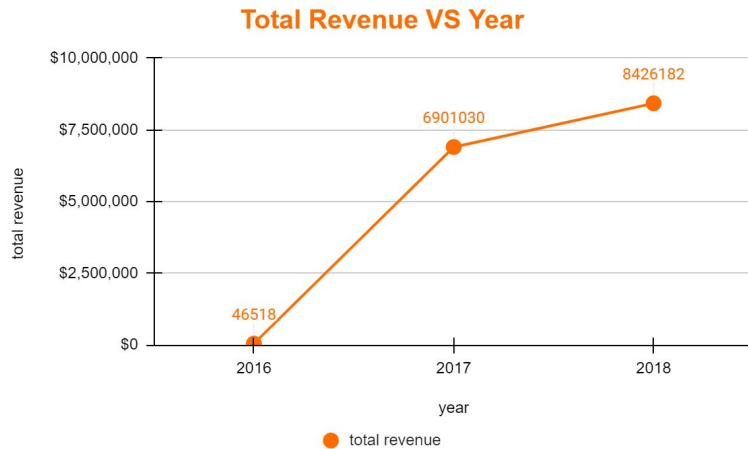
Performa bisnis eCommerce tentunya sangat berkaitan erat dengan produk-produk yang tersedia di dalamnya. Menganalisis kualitas dari produk dalam eCommerce dapat memberikan keputusan untuk mengembangkan bisnis dengan lebih baik. Langkah selanjutnya adalah menganalisis performa dari masing-masing kategori produk yang ada dan bagaimana kaitannya dengan pendapatan perusahaan. Berikut adalah tahap analisisnya:

Tahapan analisis:

1. Membuat tabel yang berisi informasi **pendapatan/revenue** perusahaan total untuk masing-masing tahun
2. Membuat tabel yang berisi informasi **jumlah cancel order** total untuk masing-masing tahun
3. Membuat tabel yang berisi nama **kategori produk** yang memberikan **pendapatan total tertinggi** untuk masing-masing tahun
4. Membuat tabel yang berisi nama **kategori produk** yang memiliki **jumlah cancel order terbanyak** untuk masing-masing tahun
5. Menggabungkan informasi-informasi yang telah didapatkan ke dalam satu tampilan tabel

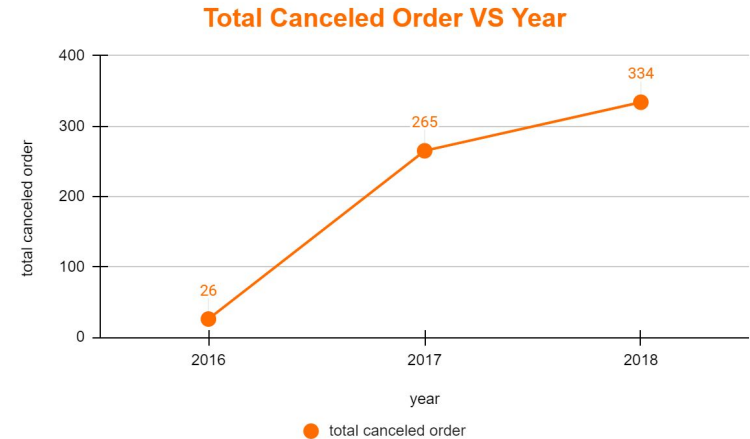
Annual Product Category Quality Analysis

1. Total Revenue per Year



Berdasarkan grafik di atas, dapat disimpulkan bahwa **total revenue perusahaan selalu meningkat** dari periode tahun 2016, 2017 dan 2018. **Total revenue tertinggi** berada di **tahun 2018** dengan total revenue berjumlah **\$8.426.182**.

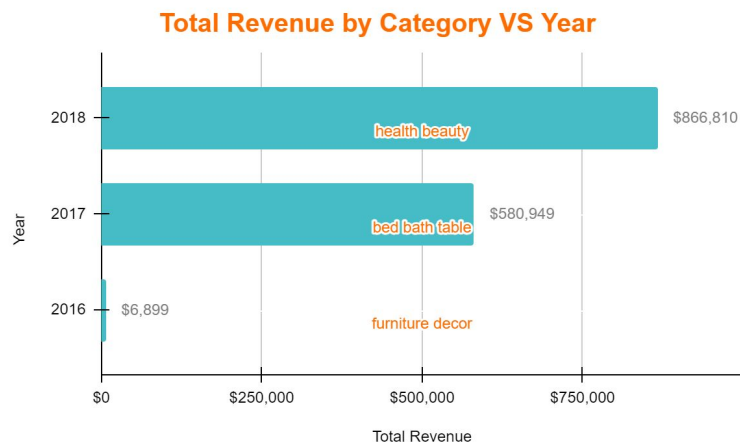
2. Total Canceled Order per Year



Berdasarkan grafik di atas, dapat disimpulkan bahwa total canceled order meningkat setiap tahunnya, dan **total canceled order tertinggi** terjadi pada **tahun 2018** sebanyak **334**.

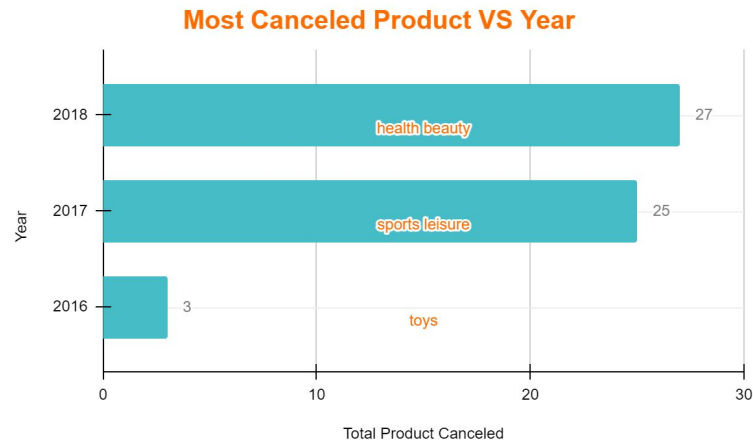
Annual Product Category Quality Analysis

3. Highest Revenue by Category Product per Year



Grafik di atas menunjukkan Highest Revenue berdasarkan kategori produk setiap tahunnya. Di **tahun 2016**, kategori produk terbanyak terjual adalah **furniture decor** dengan **revenue \$6.899**. Di **tahun 2017**, kategori produk terbanyak terjual adalah **bed bath table** dengan **revenue \$580.949**. Dan di **tahun 2018**, kategori produk terbanyak terjual adalah **bed bath table** dengan **revenue \$866.810**.

4. Most Canceled Product per Year



Grafik di atas menunjukkan kategori produk terbanyak yang di canceled oleh customer. Di **tahun 2016**, kategori produk terbanyak di canceled adalah toys dengan **total 3 produk**. Di **tahun 2017**, kategori produk terbanyak di canceled adalah sport leisure dengan **total 25 produk**. Di **tahun 2018**, kategori produk terbanyak di canceled adalah health beauty dengan **total 27 produk**.