

# Analyzing eCommerce Business Performance with SQL



**Created by:**

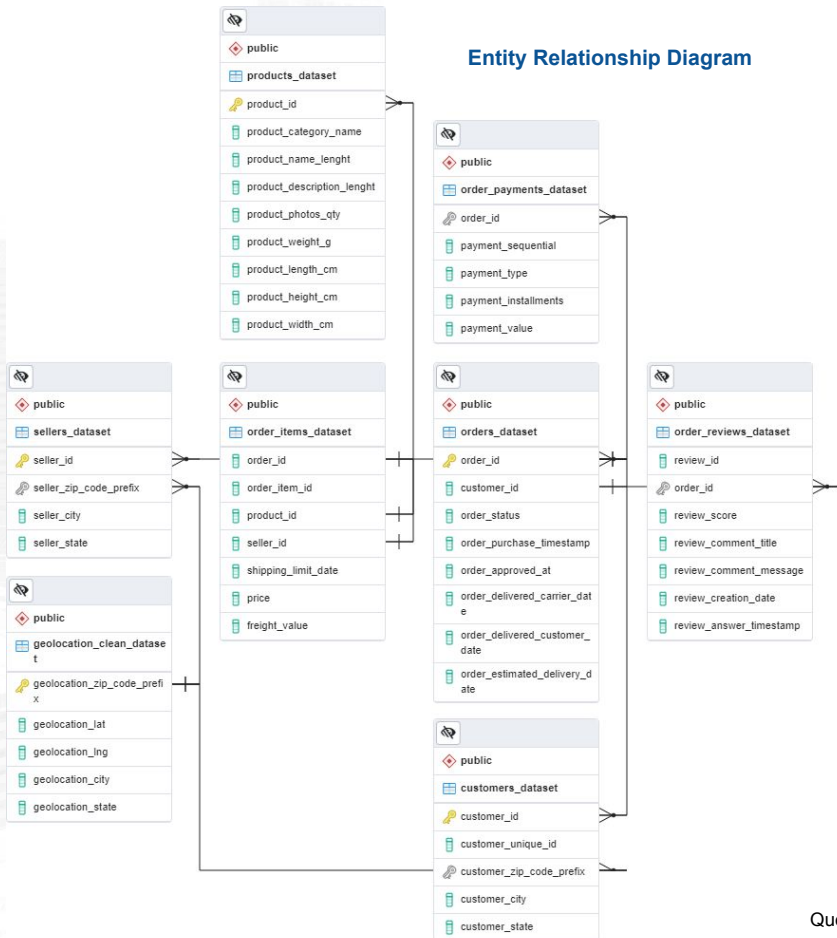
**Siti Hajjah Mardiah**

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**.