Soal & Template Jawaban

Task 5

Nama : Yudhi Ahmadi

Petunjuk

Silahkan merujuk pada Data Source Task 5 yang telah disediakan untuk mengerjakan soal soal di bawah ini

Pada bagian data analytics, terdiri dari 4 soal dengan use case & tabel yang sama. Bayangkan kamu memiliki database erp yang terdiri dari 3 tabel: penjualan, pelanggan, barang. Tabel tersebut akan dibuat menjadi sebuah datamart yang nantinya digunakan untuk visualisasi.

Query

Soal 1 *:

Dari 2 query ini, mana yang bekerja lebih baik? Jelaskan mengapa.

(a) SELECT * FROM pelanggan WHERE SUBSTR(alamat, 1, 3) = Mat;
(b) SELECT * FROM pelanggan WHERE alamat LIKE 'Mat%'

**disclaimer: soal ini tidak terkait dengan data source*

Jawaban : (b) SELECT * FROM pelanggan WHERE alamat LIKE 'Mat%'

Alasan : Menurut literatur yang saya baca penggunaan 'LIKE' akan jauh lebih cepat karena menggunakan pencarian biner pada indeks. Jika menggunakan metode SUBSTR akan menggunakan pemindaian tabel penuh, karena Oracle harus memproses baris demi baris fungsinya.

Query

Soal 2 *:

Anggap kita memiliki tabel pelanggan dengan kolom: id, nama, tanggal_lahir, alamat. Bagaimana cara yang lebih tepat dalam menulis query untuk mendapatkan data pelanggan yang tanggal_lahir nya ada di antara 2000-01-01 sampai 2008-12-31? Pilihlah salah satu jawaban dan berikan alasannya.

(a) SELECT * FROM pelanggan WHERE tanggal_lahir >= '2000-01-01' AND tanggal_lahir <= '2008-12-31'
(b) SELECT * FROM pelanggan WHERE tanggal_lahir BETWEEN '2000-01-01' AND '2008-12-31'

**disclaimer: soal ini tidak terkait dengan data source*

Jawaban : (b) SELECT * FROM pelanggan WHERE tanggal_lahir BETWEEN '2000-01-01' AND '2008-12-31'

Alasan : Kebanyakan orang berpendapat memilih BETWEEN karena membuat programmer menulis dengan jelas untuk memeriksa suatu rentang. Jika RDBMS tidak mendukung BETWEEN maka bisa menggunakan yang manual (‘>= AND <=‘), dan bisa dimodifikasi jika ingin mengubah rentang misal (‘> AND <‘).

Soal 3: Menentukan Primary Key

A. Tugas
   Tentukan primary key dari table penjualan. jelaskan alasannya

B. Jawaban & Penjelasan : id_invoice

Karena kolom id_invoice memenuhi kriteria primary key yaitu tidak memiliki nilai yang sama, tidak kosong (null), tidak melebihi batas kolom (900 byte), tidak bertipe blob, dan berhubungan dengan data (tidak digunakan tabel lain sebagai primary-key).

Soal 4: Design Datamart

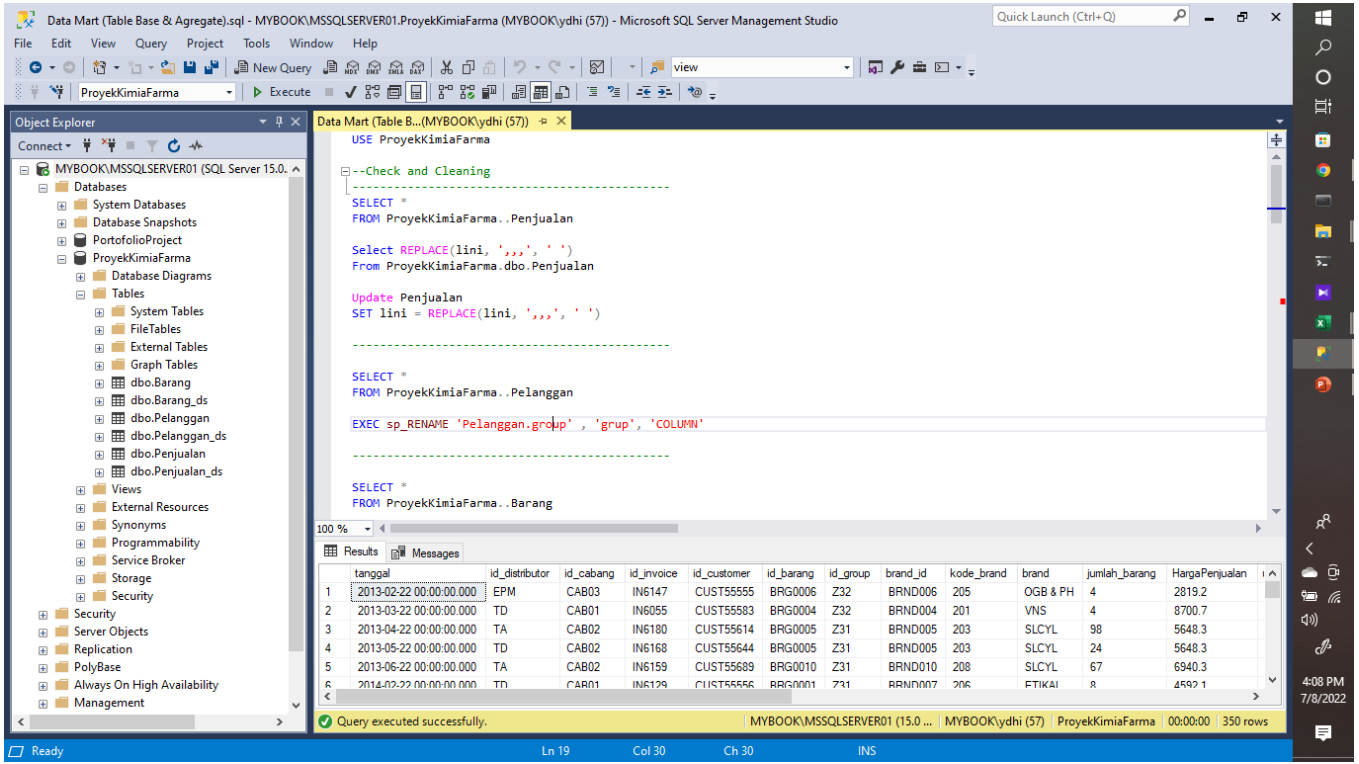
A. Tugas
   Buatlah design datamart (Terdiri dari tabel base, dan tabel aggregate). Upload file query dalam gdrive mu (pastikan dapat diakses public). Lalu masukkan linknya di tabel di bawah, dan cantumkan juga screenshoot query nya (jika lebih dari 1 file, maka masing masing file di-screenshoot)

   Silahkan tambah halaman jika dibutuhkan

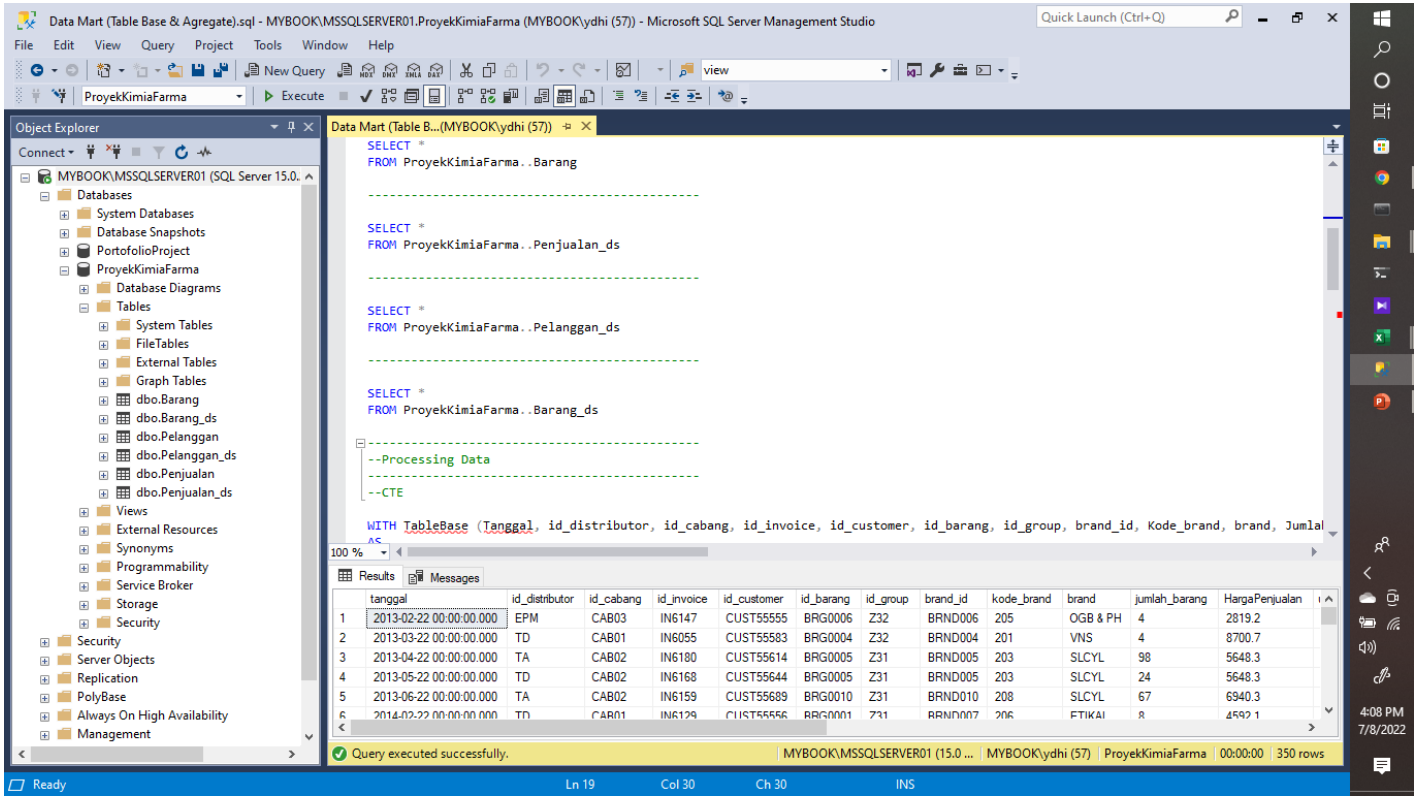
A. Jawaban : Query Dikerjakan Menggunakan SQL SERVER 2019

No	Nama File	Link
1	Data Mart (Table Base & Agregate).sql	https://drive.google.com/drive/folders/1rX4Bz09vzOhvafJgbHg5NZg3dHuh5WMY?usp=sharing

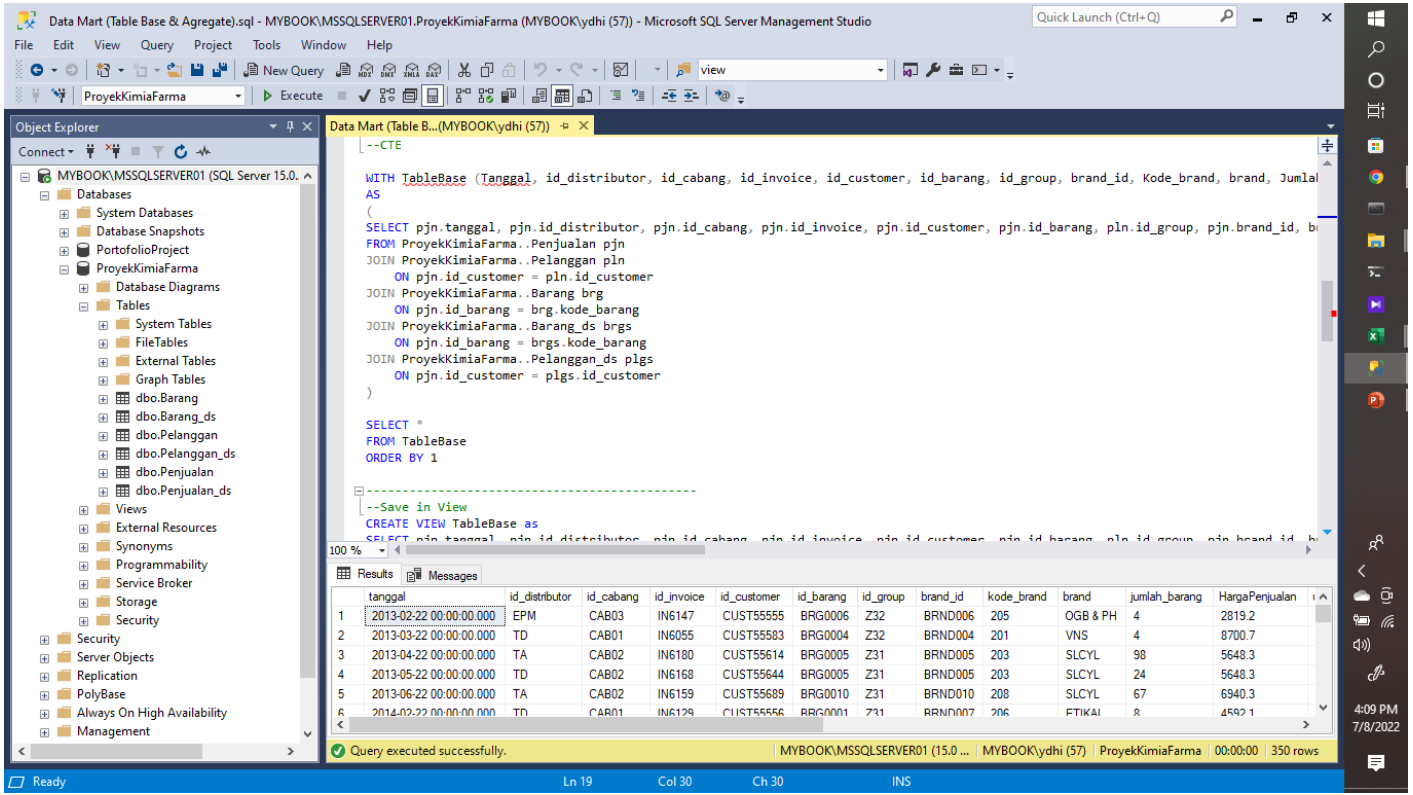
1. Table Base “TableBase”



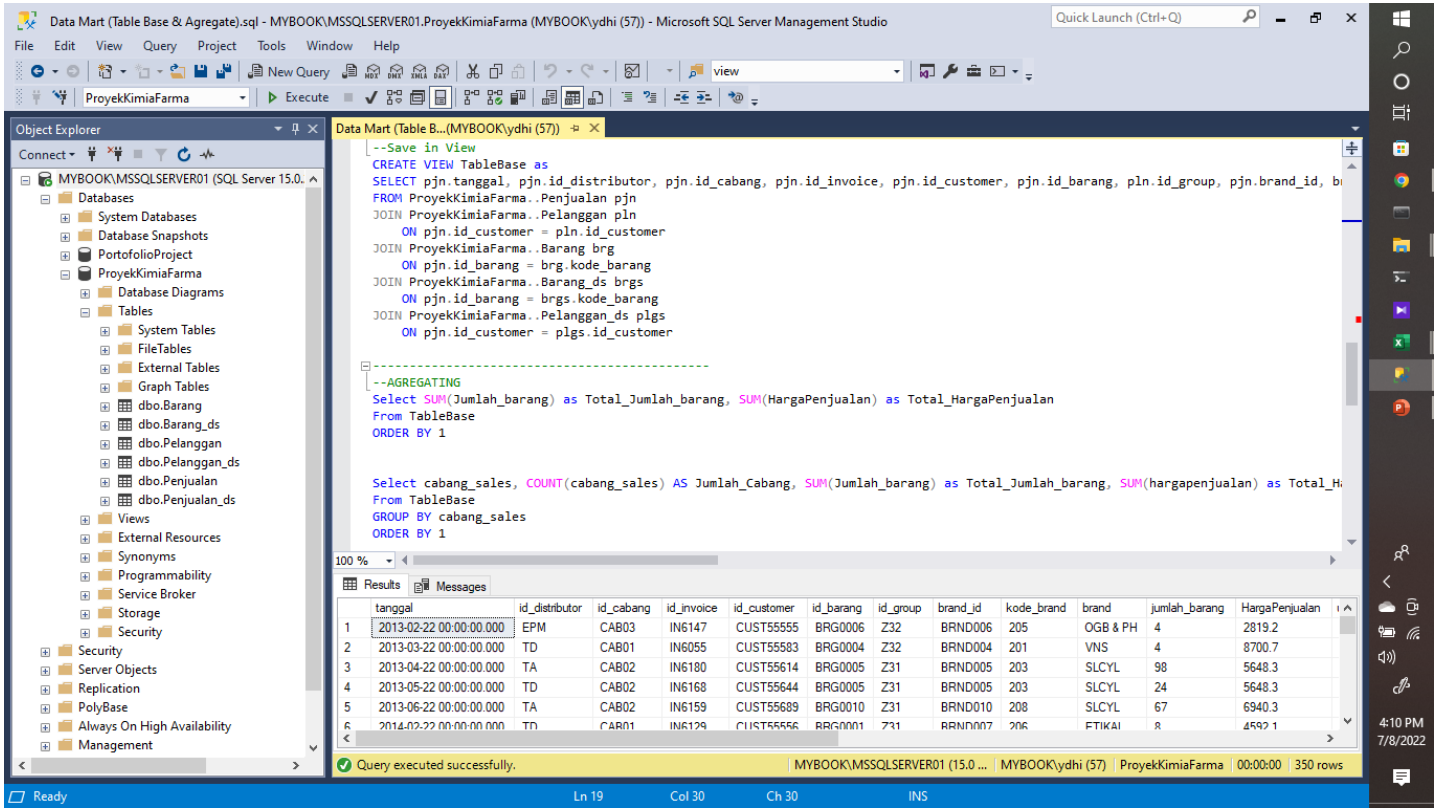
Lanjutan SS Query



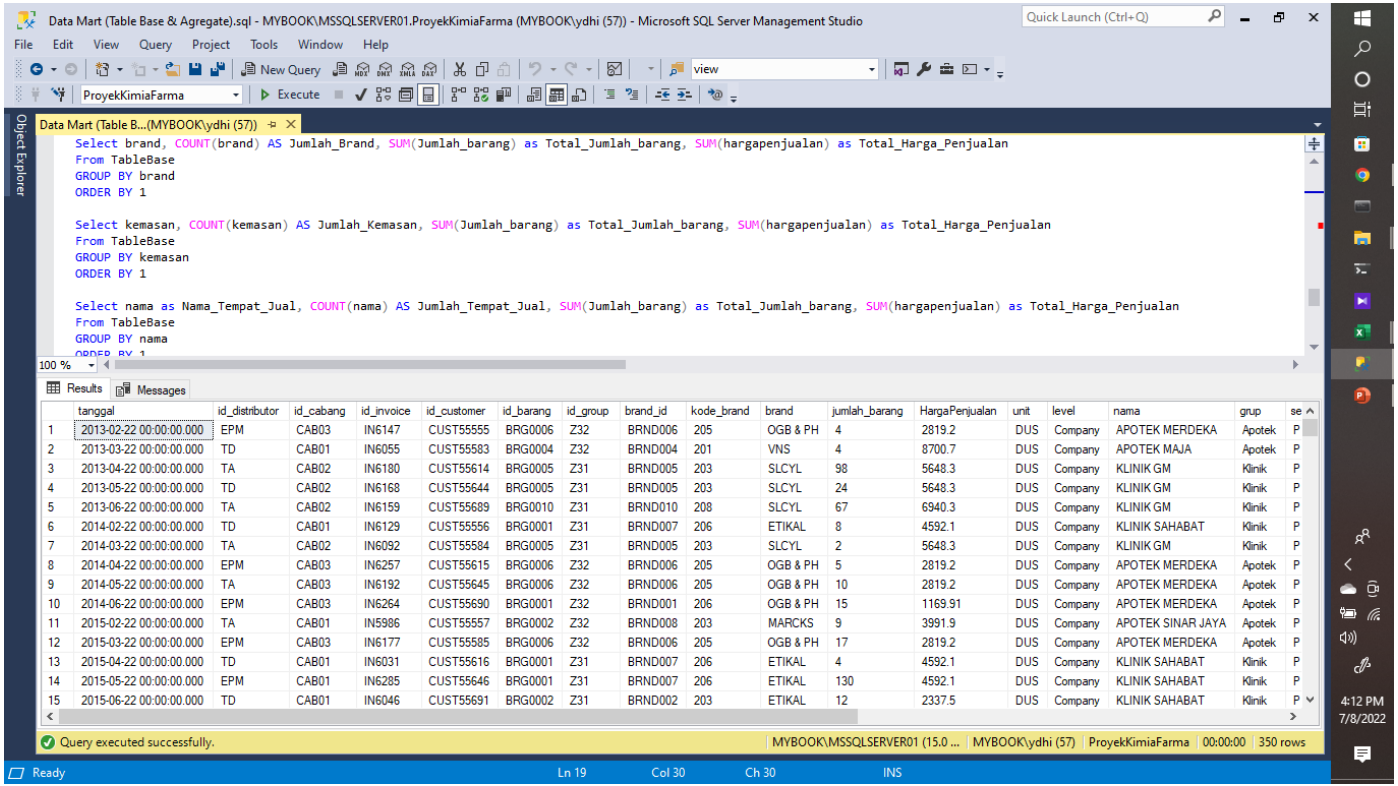
Lanjutan SS Query



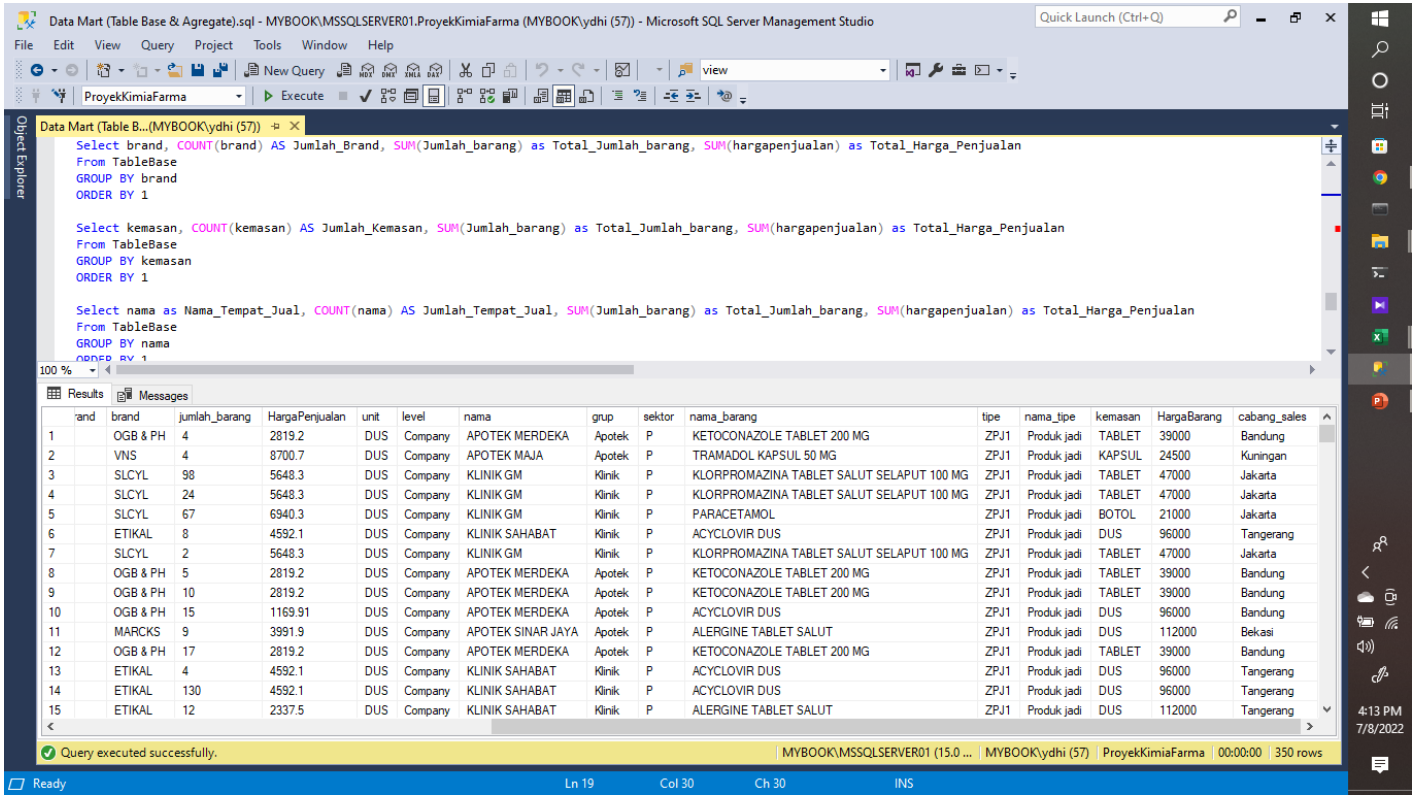
Lanjutan SS Query



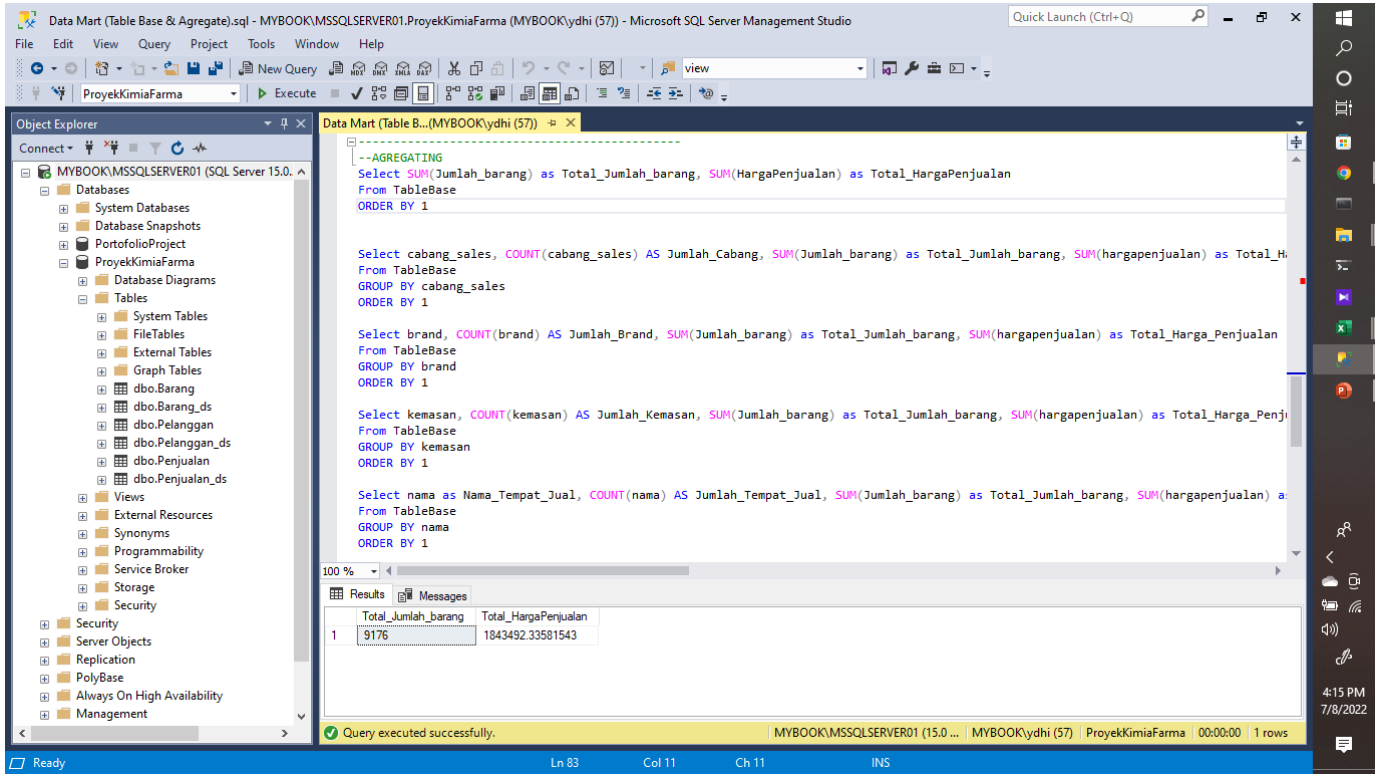
2. Table Base “TableBase”



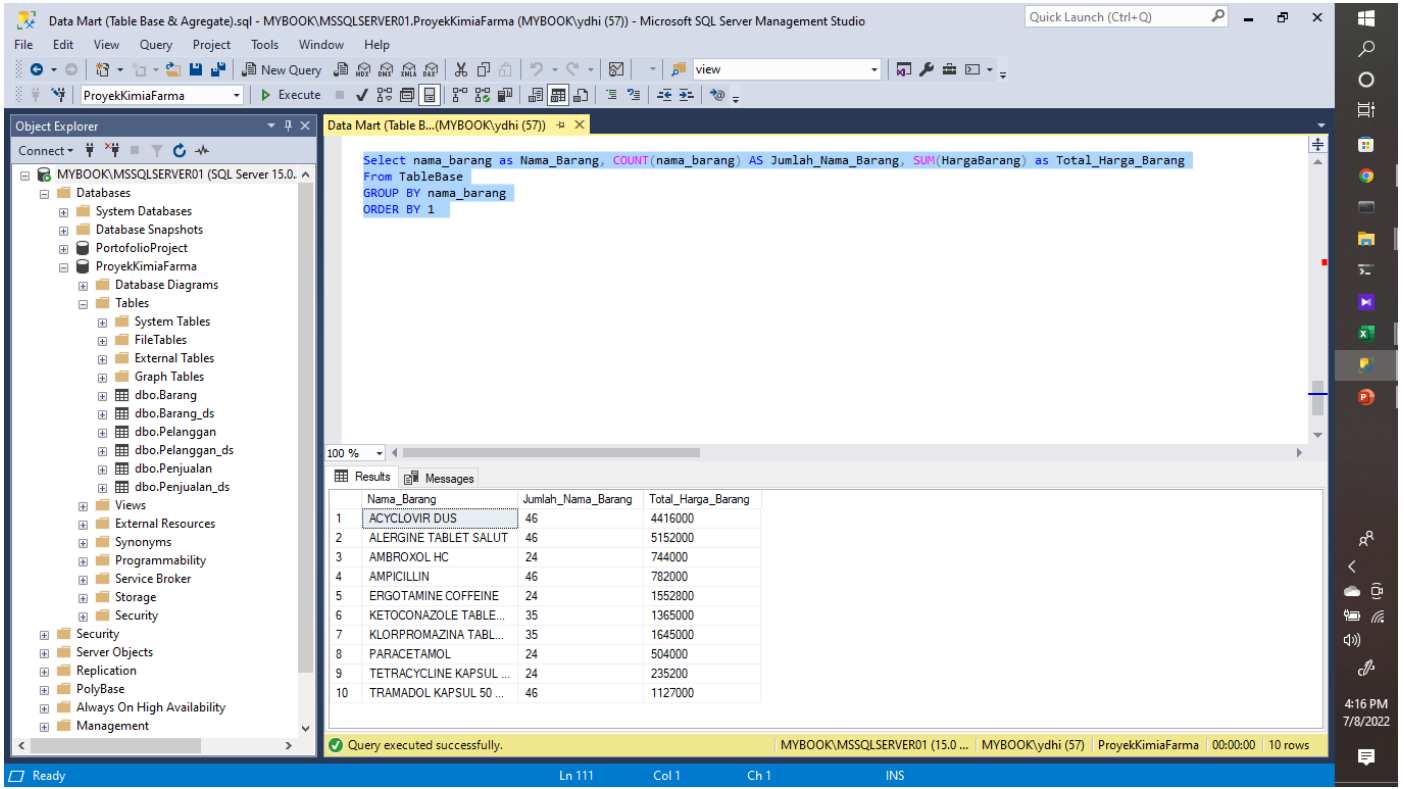
Lanjutan TableBase



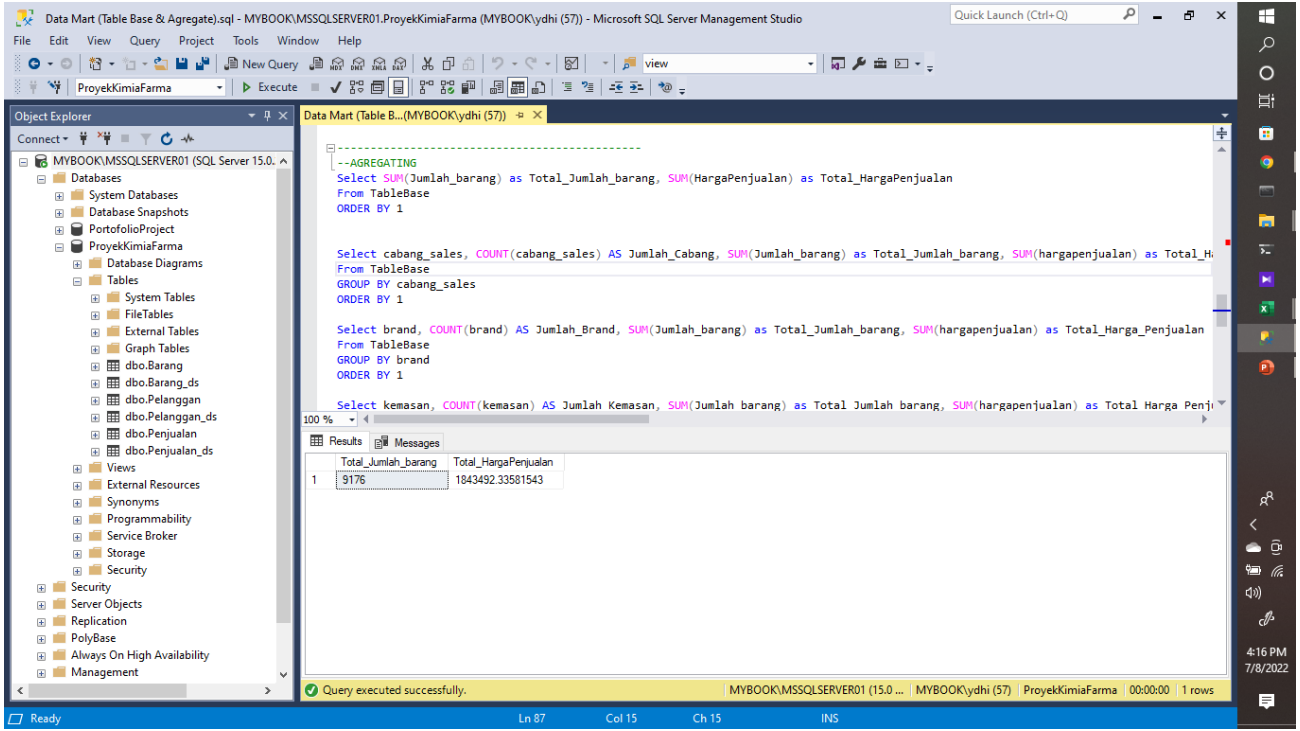
3. Table Aggregate “TableBase”



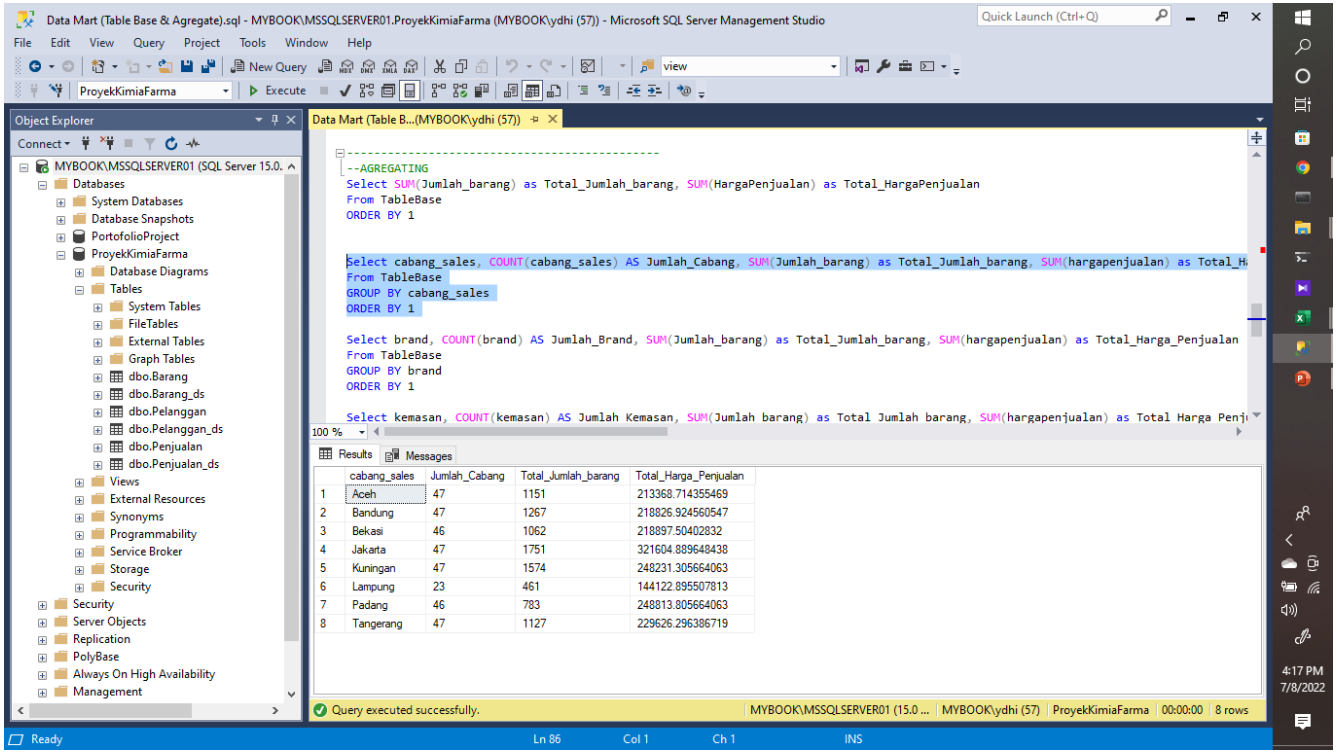
Lanjutan Query Table Aggregate “TableBase”



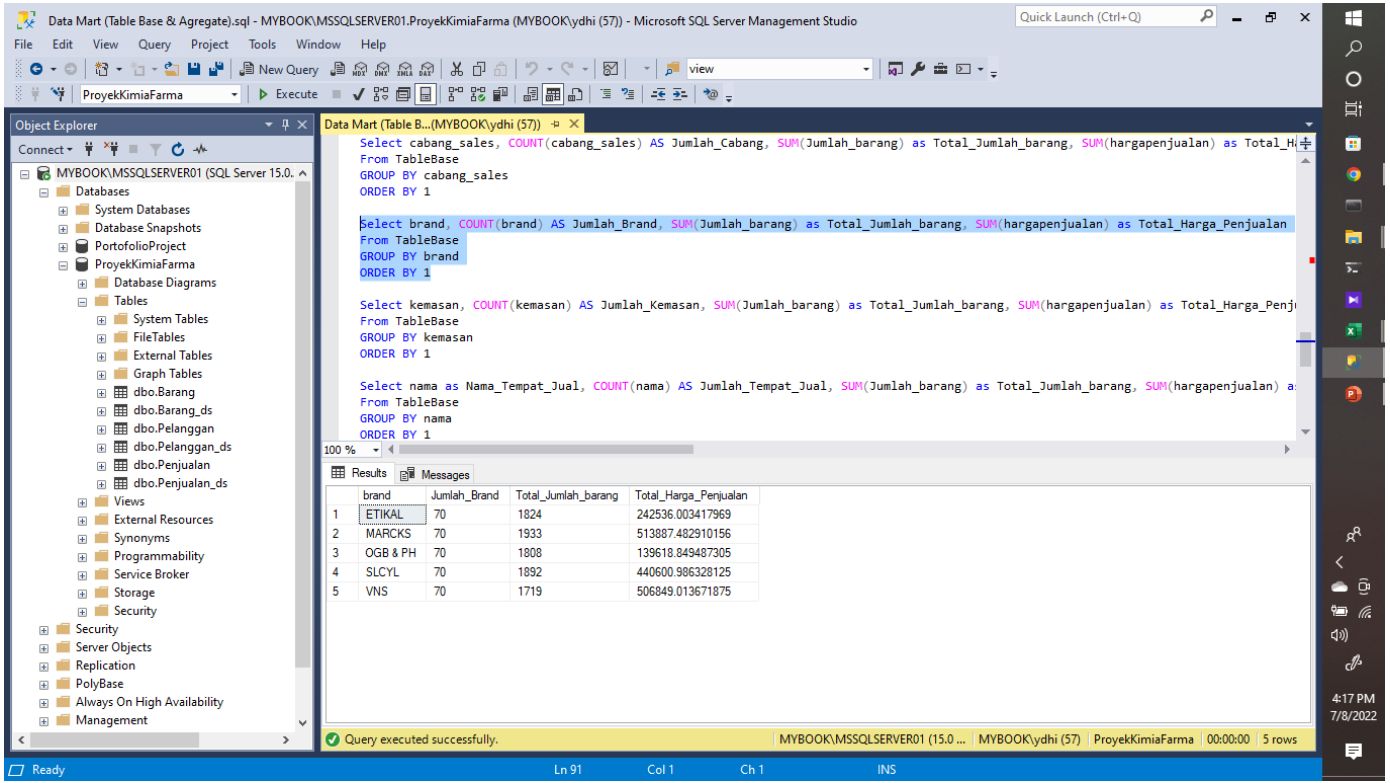
4. Table Aggregate “TableBase”



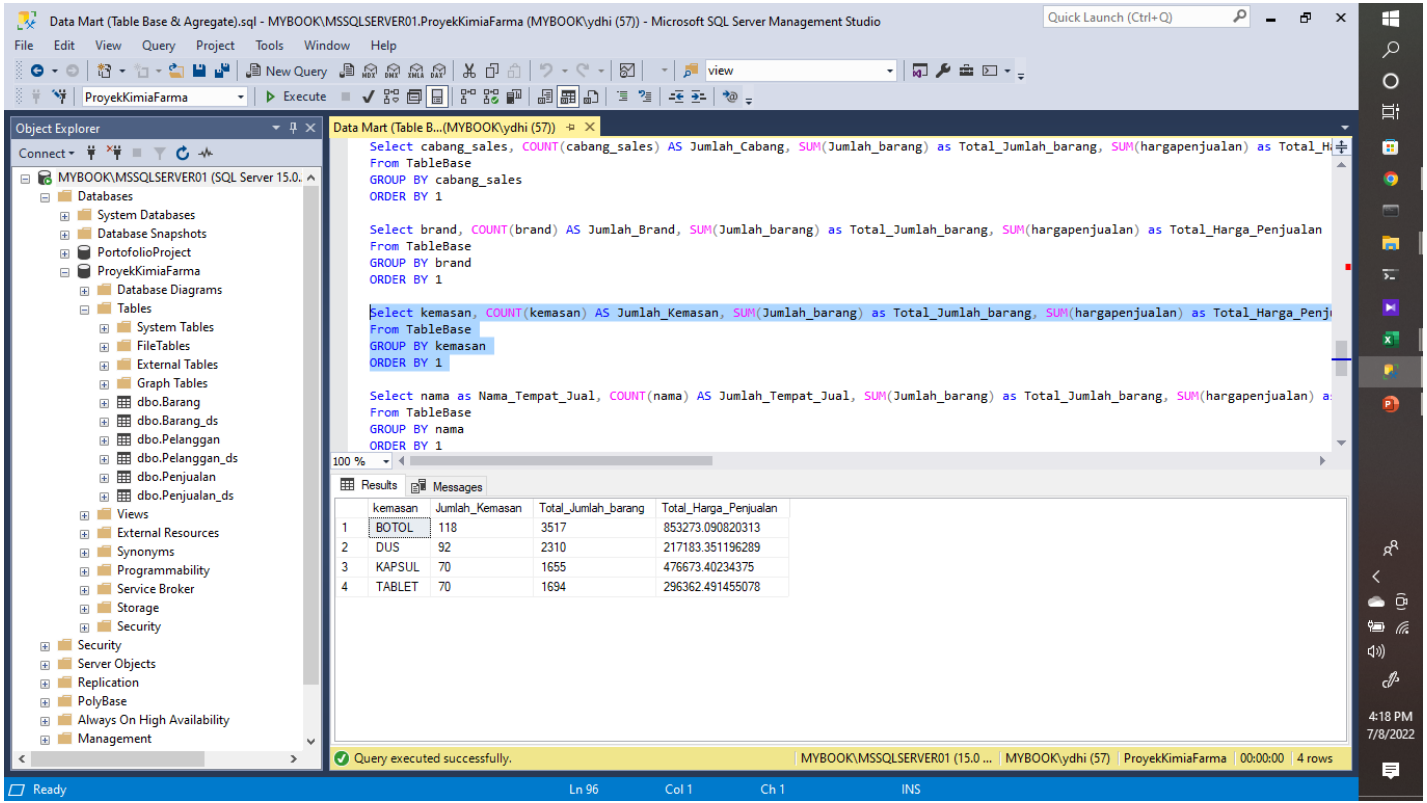
Lanjutan Table Aggregate “TableBase”



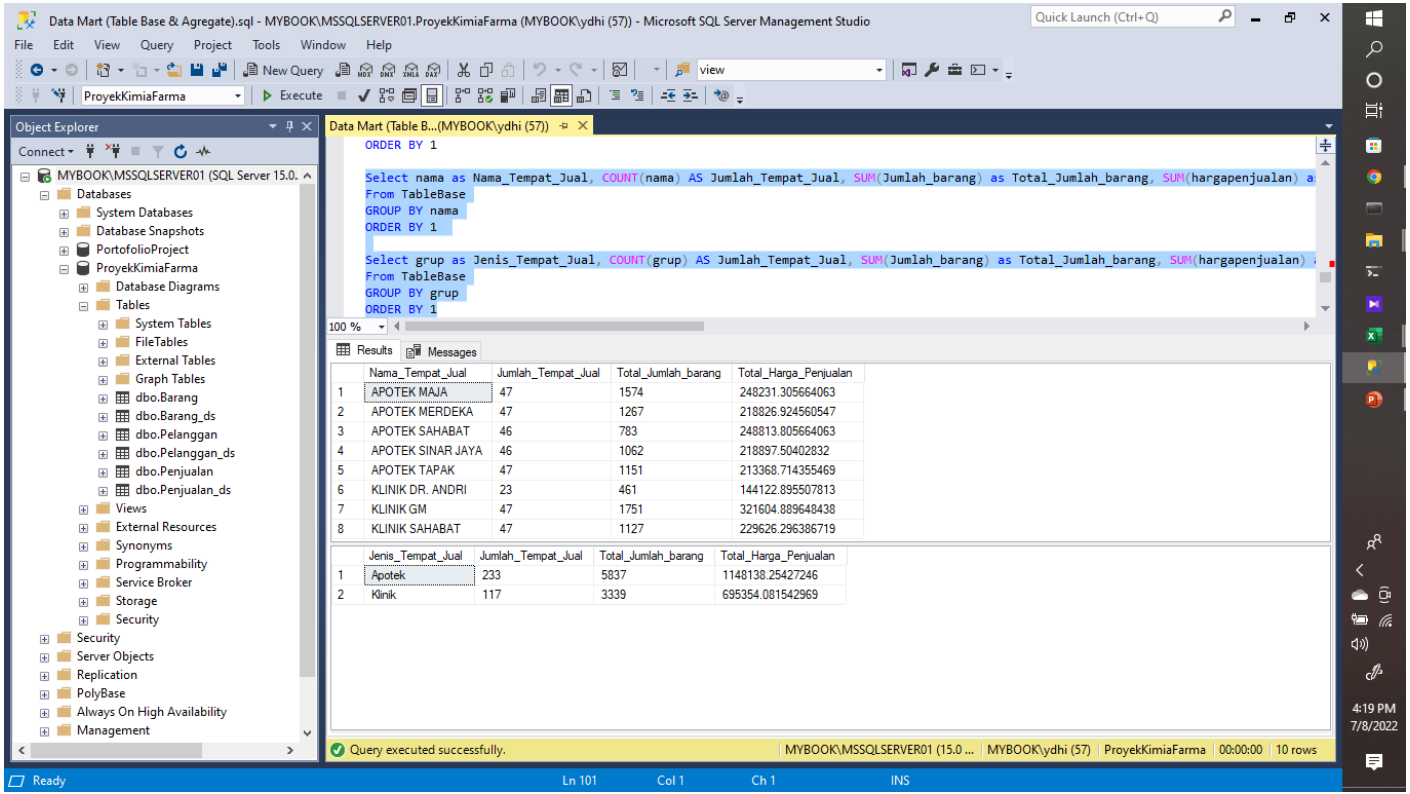
Lanjutan Table Aggregate “TableBase”



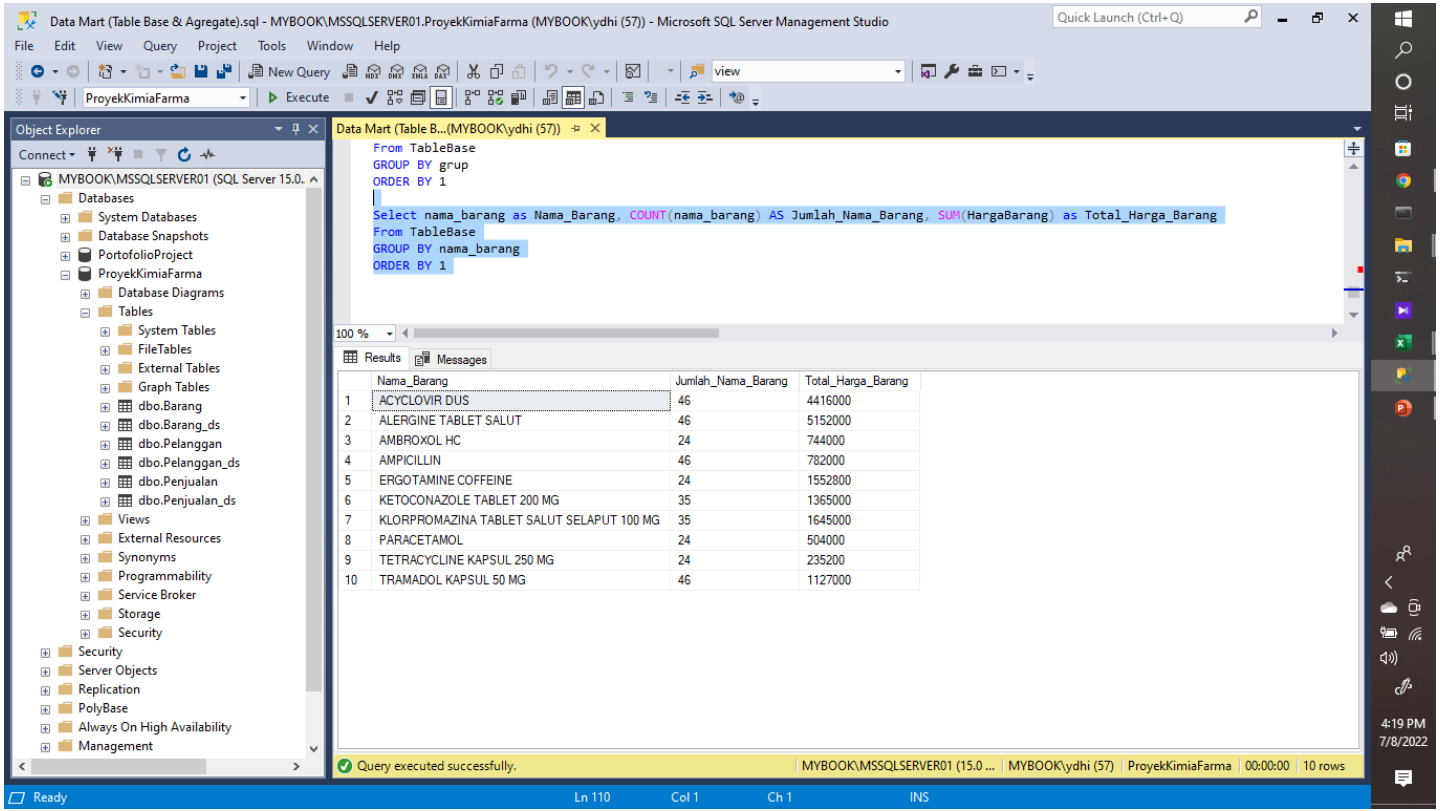
Lanjutan Table Aggregate “TableBase”



Lanjutan Table Aggregate “TableBase”



Lanjutan Table Aggregate “TableBase”



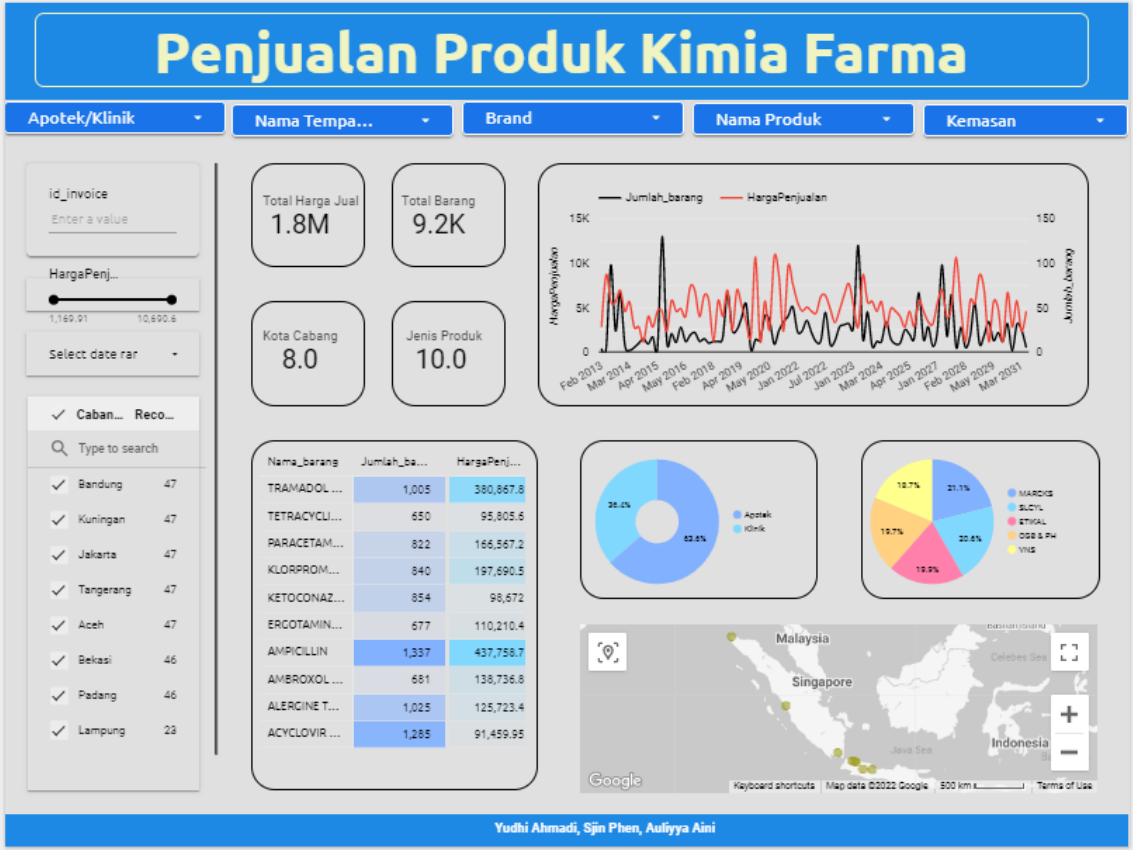
Soal 5 : Data Visualization

A. Tugas
buatlah data visualiasasi nya, dan cantumkan linknya di bawah (pastikan bisa diakses publik). Lalu cantumkan juga screenshot visualisasinya

Silahkan tambah halaman jika dibutuhkan

A. Jawaban :

Link visualisasi (ex link Google Data Studio) : https://datastudio.google.com/reporting/0a488c1f-56c7-4b03-97b8-13d72f0006ac



4:22 PM
7/8/2022

Soal 6 : Additional Complementary Data

A. Tugas :
   Dari data yang tersedia, menurut kamu untuk melengkapi analisis nya apakah diperlukan data lain juga? jika iya, sebutkan data apa yang kamu maksud dan mengapa memerlukan data tersebut

A. Jawaban : Diperlukan Data
- Produk yang Terjual Per Daerah
- Produk yang Tersisa (Tidak Terjual) Per Daerah
- Data Kategori Konsumen (Pembeli Produk)

Karena sebagai data analytics saya mencoba untuk menemukan bagaimana pola pelanggan dalam membeli produk. Pola yang dimaksud ialah barang yang dibeli secara bersamaan. Hal tersebut mungkin dapat digunakan sebagai pengambilan kebijakan untuk penawaran produk dan tata letak produk kepada pelanggan dengan kecerendungan membeli produk yang sama. Contoh : Pembeli produk obat sakit tenggorokan mungkin cenderung untuk membeli produk vitamin C jika kita menawarkannya.