

1. Sebutkan perbedaan antara datalake dan datawarehouse

	Data Lake	Data Warehouse
Struktur Data	Raw data	Structured data/ Processed data
Tujuan Data	belum ditentukan	sudah ditentukan/memiliki tujuan tertentu
Pengguna	data engineer, data scientist	business professional
Aksesibilitas	mudah diakses dan dapat diupdate lebih cepat	lebih rumit dan memerlukan biaya untuk merubahnya

2. Apa yang membedakan OLAP dan OLTP

	OLTP (Online Transaction Processing)	OLAP (Online Analytical Processing)
Karakteristik utama	menangani sejumlah transaksi kecil setiap hari secara real time	menangani jumlah data yang besar dalam beberapa database untuk menemukan insight
Sumber data	transaksi	database OLTP dan sumber lainnya
Waktu respon dan tipe query	milidetik dan query sederhana	makin banyak data makin lama waktu responnya dan query rumit
Desain database	normalisasi database untuk efisiensi	denormalisasi database untuk analisa
Pengguna	berorientasi kepada pasar	berorientasi kepada pelanggan

3. Teknologi apa saja yang biasanya dipakai untuk Data Warehouse ?

- AWS Redshift
- Google Big Query
- Clickhouse
- Snowflake
- Databrick
- Apache Doris
- Postgre with citus extension

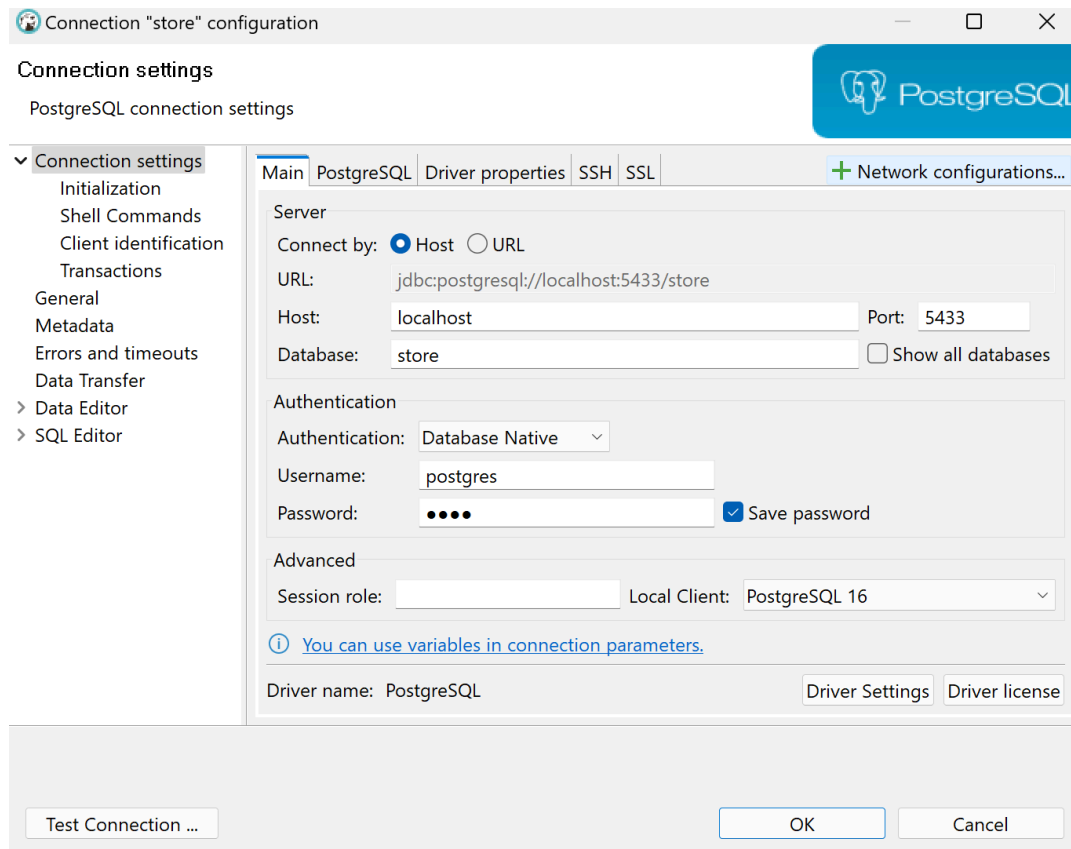
4. Tuliskan setiap perintah dari proses instalasi citus menggunakan docker compose sampai tabel terbentuk

```
Ahmad Fathoni A@Laptoptoni MINGW64 /c/Alterra
$ git clone https://github.com/Immersive-DataEngineer-Resource/citus-demo.git
Cloning into 'citus-demo'...
remote: Enumerating objects: 12, done.
remote: Counting objects: 100% (12/12), done.
remote: Compressing objects: 100% (10/10), done.
remote: Total 12 (delta 0), reused 12 (delta 0), pack-reused 0
Receiving objects: 100% (12/12), 5.02 KiB | 734.00 KiB/s, done.
```

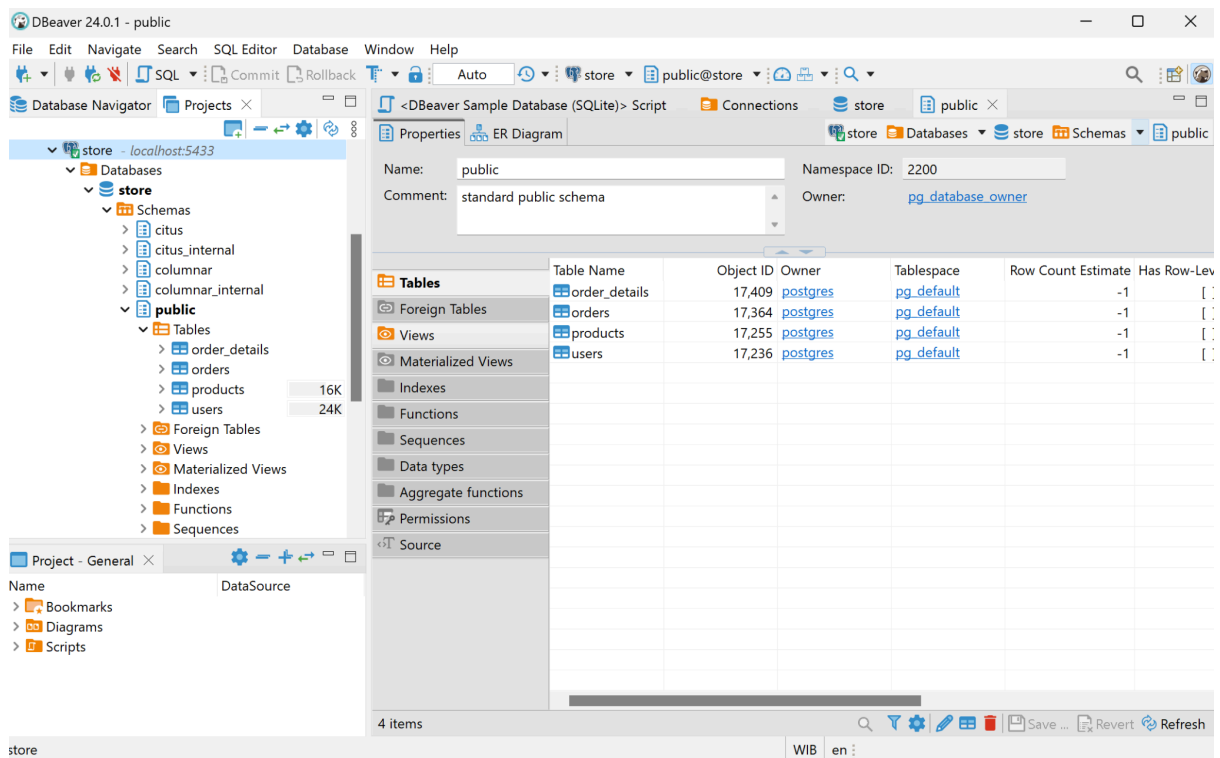
Melakukan git clone terhadap repo citus-demo ke dalam folder Alterra

```
Ahmad Fathoni A@Laptoptoni MINGW64 /c/Alterra/citus-demo (main)
$ docker compose up -d
worker-3 Pulling
worker-1 Pulling
worker-2 Pulling
master Pulling
faef57eae888 Pulling fs layer
a33c10a72186 Pulling fs layer
d662a43776d2 Pulling fs layer
a3ba86413420 Pulling fs layer
a627f37e9916 Pulling fs layer
424bade69494 Pulling fs layer
dd8d4fcd466b Pulling fs layer
a3ba86413420 waiting
a627f37e9916 waiting
424bade69494 waiting
03d0efeea592 Pulling fs layer
dd8d4fcd466b waiting
4f27e1518a67 Pulling fs layer
03d0efeea592 waiting
0c8ac8b8eb90 Pulling fs layer
4f27e1518a67 waiting
c08e79653ad2 Pulling fs layer
d5724e8c22af Pulling fs layer
0c8ac8b8eb90 waiting
```

Melakukan docker compose up terhadap folder citus demo untuk menjalankan aplikasi citus-demo



Mengkoneksikan docker compose citus-demo dengan dbeaver



docker compose citus-demo sudah terkoneksi, dan tabel sudah muncul

5. Jelaskan perbedaan antara access method heap dan columnar pada citus

	Heap	Columnar
Definisi	cara standar untuk menyimpan data dalam PostgreSQL dan banyak sistem basis data relasional lainnya.	pendekatan alternatif untuk menyimpan data di mana nilai-nilai dari setiap kolom disimpan secara terpisah dalam blok penyimpanan.
Penyimpanan	disimpan secara berurutan di blok penyimpanan.	disimpan secara berurutan dalam blok penyimpanan yang terpisah.
Kegunaan	lebih cocok untuk transaksi online yang memerlukan pencarian data berdasarkan beberapa kolom sekaligus, karena semua kolom baris dapat diakses dengan cepat.	digunakan dalam skenario analitis, di mana kueri sering kali memproses sejumlah besar data tetapi hanya memerlukan sebagian kecil dari kolom-kolom tertentu.