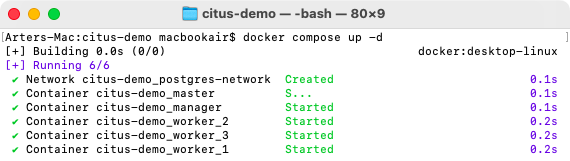
1. Data lake adalah sekumpulan data bervolume sangat besar yang belum diproses dan tujuan dari penggunaan datanya masih belum jelas. Sedangkan data warehouse merupakan gudang data yang terstruktur dan terfilter dengan data yang telah melalui proses untuk mencapai tujuan tertentu. Data lake dan data warehouse sama-sama digunakan untuk menstore data dengan volume yang besar. Data lake umumnya digunakan oleh technical users misalnya data scientist karena akses yang mudah dan proses pembaruan yang cepat (flexibel). Sedangkan untuk data warehouse umumnya digunakan oleh non-technical users misalnya bisnis profesional karena cenderung lebih rumit dengan biaya yang cukup mahal untuk melakukan perubahan (secure).
2. Database data warehouse (OLAP) dirancang untuk analisis dan pemrosesan data secara lebih mendalam untuk mendukung proses pengambilan keputusan dan analisis bisnis, memiliki struktur yang berbeda dengan menggunakan skema bintang atau salju yang memungkinkan penyimpanan data agregat dan terstruktur secara hierarki untuk analisis yang lebih cepat, volume datanya jauh lebih besar, jenis query komplex, kinerja berfokus pada query analitis, menggunakan indeks yang lebih kompleks, waktu respon cenderung lebih lambat. Sedangkan database konvensional (OLTP) digunakan untuk operasi transaksi sehari-hari dalam suatu organisasi, memiliki banyak tabel yang dioptimalkan untuk operasi penulisan, pembacaan, dan pembaruan data transaksi, volume data relatif lebih kecil, umumnya query berfokus pada operasi pencarian, pembaruan dan penghapusan data, kinerja berfokus pada pemrosesan transaksi secepat mungkin, memiliki indeks yang lebih sederhana dan sedikit untuk meningkatkan kinerja, waktu responnya cepat dan konsisten.
3. Data warehouse umumnya menggunakan sistem manajemen database (DBMS) yang berfungsi untuk menyimpan dan mengelola data, misalnya RDBMS (Relational Database Management System), columnar database management systems, atau NoSQL database. Terdapat proses ETL (Extract, Transform, Load), menggunakan sistem operasi untuk menjalankan server dan software data warehouse, menggunakan penyimpanan data, menggunakan teknologi kolom dan kompresi, online analytical processing server, software query dan pelaporan, data mining dan analisis, security dan manajemen akses.
4. Menjalankan docker-compose

|  |
| --- |
| Docker compose up -d |

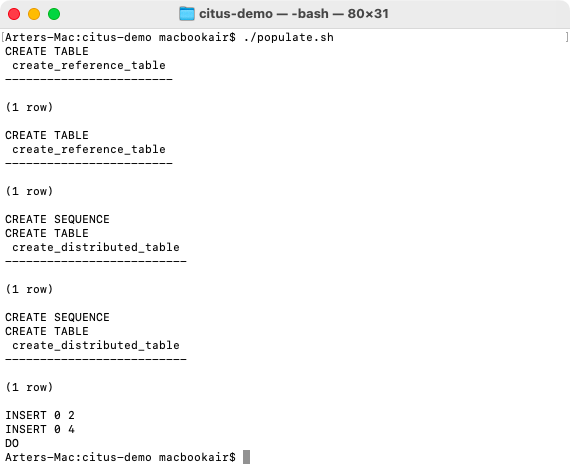
Output



Menjalankan perintah untuk create tabel

|  |
| --- |
| ./populate.sh |

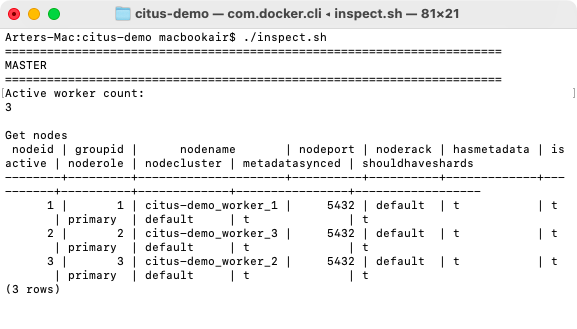
Output



Menampilkan tabel dan node

|  |
| --- |
| ./inspect.sh |

Output



1. Akses method heat menggunakan penyimpanan yang berurutan dalam format baris, cocok dengan query yang hanya mengambil beberaoa kolom dari tabel, cocok untuk transaksi online dimana operasi write and read data singkat dan sering dilakukan, akan tetapi method ini kurang efisien untuk digunakan pada query analitis atau yang memerlukan agregasi data dalam skala besar. Sedangkan access method columnar adalah penyimpanan yang menggunakan format kolom, cocok dengan query yang memerlukan read data kolom secara efisien, method ini cocok untuk query analitis yang memerlukan agregasi data karena memungkinkan kompresi kolom yang efisien, tetapi method ini tidak efisien untuk transaksi online yang membutuhkan read dan write data yang singkat dan sering.