



## PART 1 - DATA WAREHOUSE INTRODUCTION

# Data Warehouse



No	Data Warehouse	Data Lake
1	<p><b>Tujuan:</b> Menyimpan data yang terstruktur dan terorganisir untuk analisis dan pelaporan.</p> <p><b>Skema:</b> Menggunakan skema yang telah ditentukan seperti star atau snowflake.</p> <p><b>Jenis Data:</b> Hanya menyimpan data yang terstruktur.</p> <p><b>Proses ETL:</b> Data harus diproses dan diubah (ETL) sebelum dimasukkan ke dalam data warehouse.</p> <p><b>Kinerja:</b> Dioptimalkan untuk query cepat dan analisis kompleks</p>	<p><b>Tujuan:</b> Menyimpan berbagai jenis data dalam bentuk asli untuk analisis lebih lanjut.</p> <p><b>Skema:</b> Skema-on-read; skema ditentukan saat data dibaca.</p> <p><b>Jenis Data:</b> Menyimpan data terstruktur, semi-terstruktur, dan tidak terstruktur.</p> <p><b>Proses ETL:</b> Data dapat disimpan dalam bentuk mentah dan diproses kemudian (ELT).</p> <p><b>Kinerja:</b> Dioptimalkan untuk penyimpanan skala besar dan fleksibilitas dalam analisis.</p>

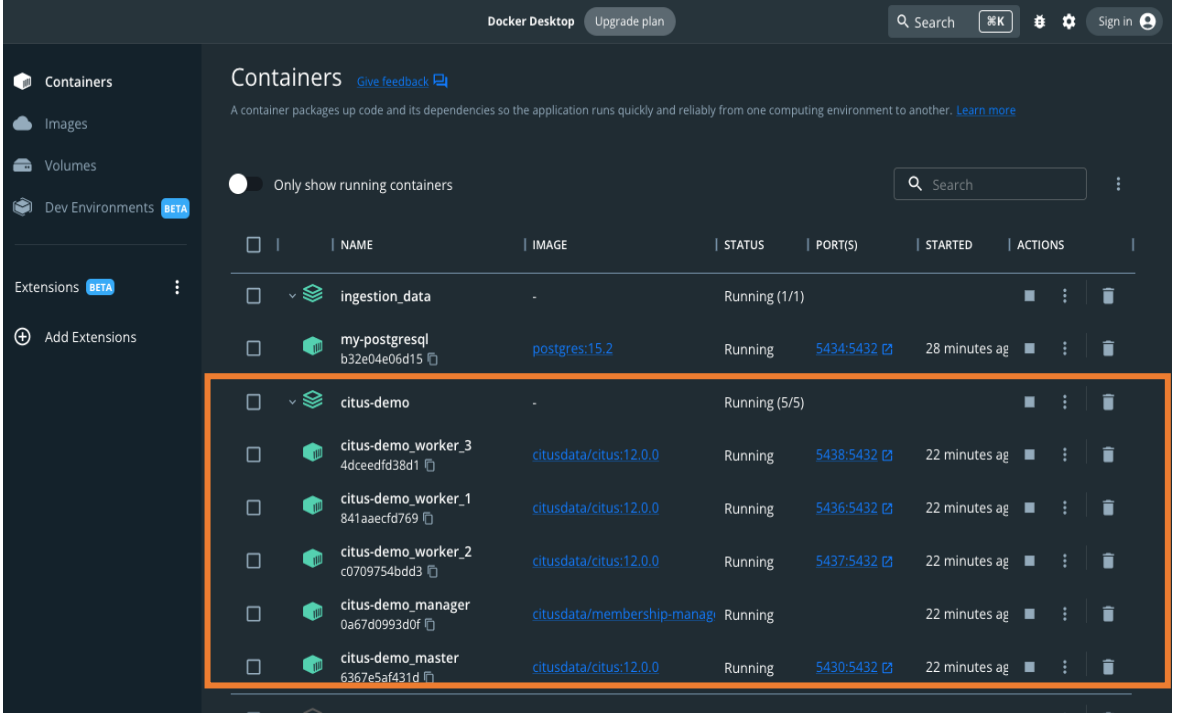
# Data Warehouse



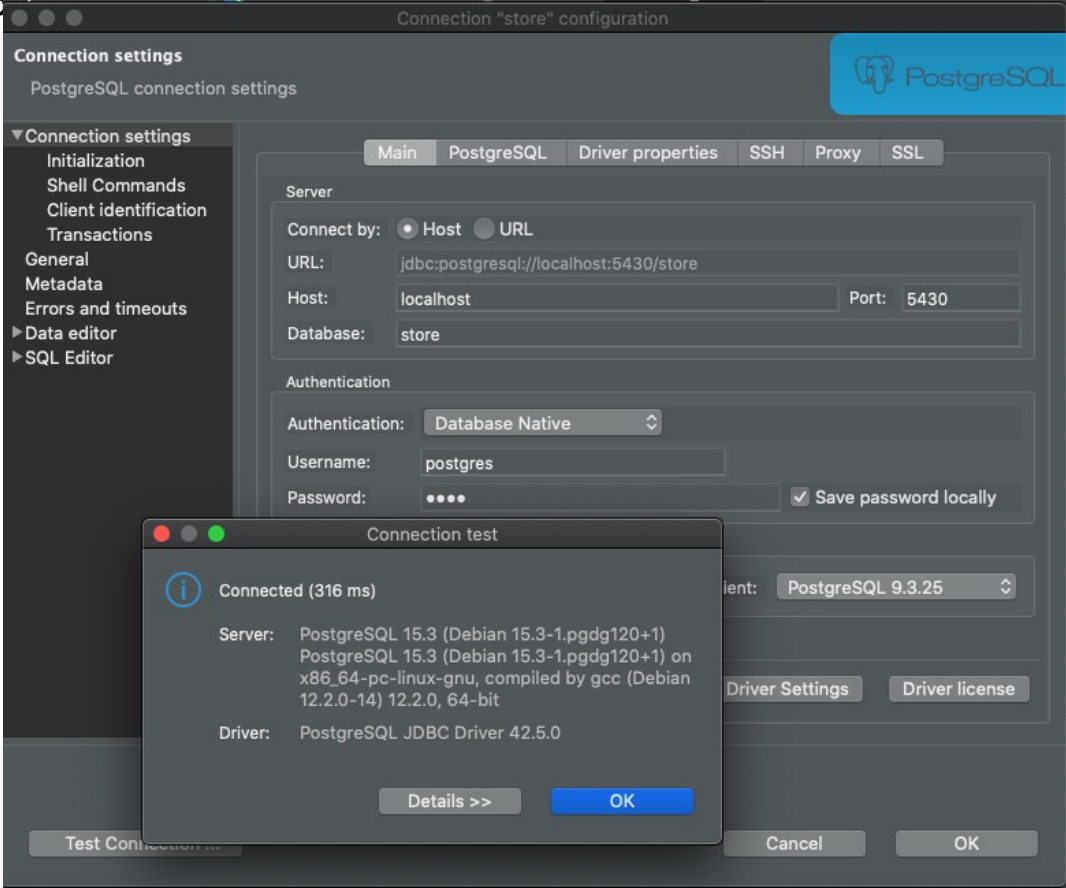
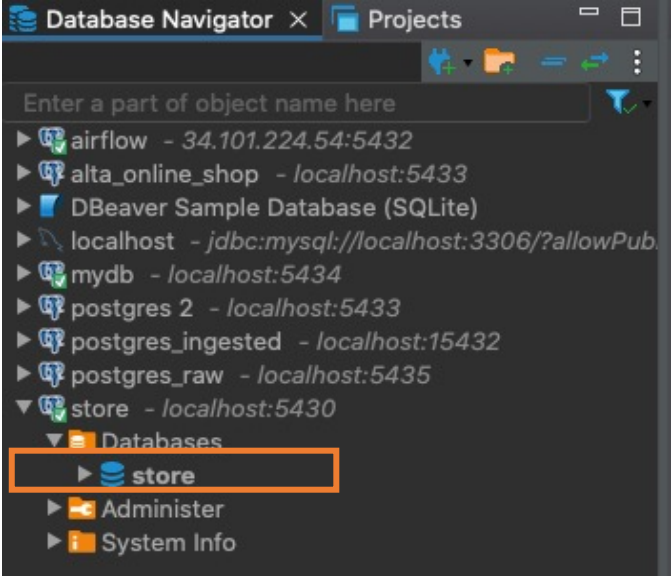
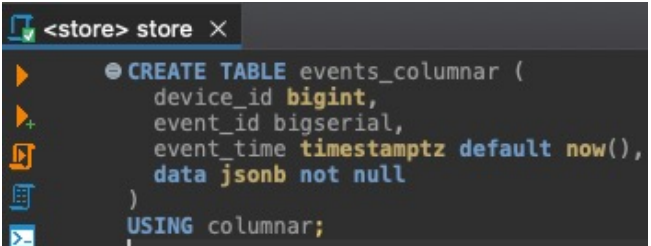
No	Online Analytical Processing (OLAP)	Online Transaction Processing (OLTP)
2	<p><b>Desain:</b> Denormalisasi untuk meningkatkan kinerja query analitis.</p> <p><b>Penggunaan:</b> Digunakan untuk analisis data dan pelaporan.</p> <p><b>Query:</b> Mendukung query yang kompleks dan intensif.</p> <p><b>Data:</b> Data historis dan besar dalam volume.</p> <p><b>Contoh Teknologi:</b> Google BigQuery, Amazon Redshift, Snowflake.</p>	<p><b>Desain:</b> Normalisasi untuk mengurangi redundansi dan meningkatkan integritas data.</p> <p><b>Penggunaan:</b> Digunakan untuk operasi transaksi harian.</p> <p><b>Query:</b> Mendukung query sederhana dan cepat.</p> <p><b>Data:</b> Data transaksi real-time dan kecil dalam volume.</p> <p><b>Contoh Teknologi:</b> MySQL, PostgreSQL, Microsoft SQL Server</p>

# Menjalankan Citus via Docker Compose Sampai Table Terbentuk



No	Deskripsi	Output
4	<pre>● wartadi@wartadis-MacBook-Pro citus-demo % docker compose up -d [+] Running 5/5  # Container citus-demo_master   Started    1.0s  # Container citus-demo_manager  Started    2.8s  # Container citus-demo_worker_1 Started    6.8s  # Container citus-demo_worker_2 Started    6.5s  # Container citus-demo_worker_3 Started    5.9s</pre> <p>Dalam tugas nomor 1 perintah untuk menjalankan Citus di computer kita dengan docker compse</p> <p>Perintah yang diberikan :</p> <p><b>docker compose up -d</b></p> <p>dalam hal ini port yang digunakan disesuaikan agar citus yang kita running tidak ada error. Port yang digunakan antara lain</p> <p>Master : 5430 Worker 1 : 5436 Worker 2 : 5437 Worker 3 : 5438</p>	 <p>Dalam citus-demo disini ada 5 container yang sedang running antara lain citus-demo-worker1, citus-demo-worker2, citus-demo-worker3, citus-demo-manager dan citus-demo-master</p>

# Menjalankan Citus via Docker Compose Sampai Table Terbentuk

No	Perintah	Output
4	<p>Dalam membuat database perlu membuat connection database postgresQL terlebih dahulu. Input yang digunakan seperti nama <b>Database</b>, <b>Port :5430</b> (<b>master</b>) , <b>Username</b> dan <b>password</b> disesuaikan dengan yang ada di <b>docker compose</b></p> 	<p>Pada Database Navigator sudah terdapat connection database dengan nama <b>store</b> dengan <b>localhost:5430</b> dan ada database dengan nama <b>store</b></p>  <p>pembuatan table melalui script</p> 

# Data Warehouse



No	Soal	Jawaban
3	Teknologi yang biasa digunakan untuk data warehouse	<ul style="list-style-type: none"><li>• Amazon Redshift</li><li>• Google BigQuery</li><li>• Snowflake</li><li>• Microsoft Azure Synapse Analytics</li><li>• Apache Hive</li><li>• Apache Airflow</li></ul>
5	Perbedaan antara Access Method Heap dan Columnar pada Citus	<p><b>Heap:</b> <b>Penyimpanan:</b> Data disimpan dalam urutan baris. <b>Kinerja:</b> Baik untuk operasi penulisan dan query yang mengambil seluruh baris. <b>Penggunaan:</b> Umum digunakan untuk OLTP.</p> <p><b>Columnar:</b> <b>Penyimpanan:</b> Data disimpan dalam urutan kolom. <b>Kinerja:</b> Baik untuk query analitis yang mengakses beberapa kolom. <b>Penggunaan:</b> Umum digunakan untuk OLAP.</p>

THANK YOU 😊

