1. Untuk menggunakan relational database atau Nosql database itu tergantung pada kebutuhan spesifik proyek seperti relational database bisa digunakan ketika :

-Data memiliki skema yang jelas dan tetap dengan hubungan yang terdefinisi antara tabel, misalnya melalui foreign keys.

-Memerlukan transaksi ACID (Atomicity, Consistency, Isolation, Durability) untuk memastikan konsistensi dan integritas data.

-Membutuhkan kemampuan untuk menjalankan query yang kompleks menggunakan JOIN, subquery, dan fungsi agregat.

-Memerlukan penggunaan SQL standar untuk query, yang sudah dikenal luas dan didukung oleh banyak tool.

-Volume data yang disimpan dan diakses masih dalam batasan yang dapat dikelola oleh RDBMS.

Contoh penggunaannya seperti managemen keuangan, sistem inventaris, aplikasi CRM(customer relationship management).

Sedangkan Nosql digunakan ketika :

-Data tidak memiliki skema tetap atau sering berubah.

-Memerlukan kemampuan untuk dengan mudah menambah lebih banyak server untuk mengakomodasi peningkatan beban kerja.

-Memerlukan kinerja tinggi dan latensi rendah, terutama untuk operasi baca/tulis yang cepat.

-Berurusan dengan volume data yang sangat besar yang sulit ditangani oleh RDBMS tradisional.

-Memerlukan model data tertentu seperti dokumen (misalnya, JSON), graf, key-value, atau kolom lebar yang lebih cocok untuk kebutuhan aplikasi.

-Memerlukan fleksibilitas untuk menyimpan data dengan struktur yang bervariasi tanpa perlu mengubah skema database.

Contohnya seperti aplikasi media sosial, penyimpanan data log aktivitas pengguna, sistem rekomendasi, penyimpanan data loT(internet of things).

1. A. Database adalah sistem terstruktur untuk menyimpan, mengelola, dan mengakses data. Data dalam database biasanya disimpan dalam tabel dengan skema yang jelas dan tetap. Contohnya, MySQL, PostgreSQL, Oracle Database, Microsoft SQL Server.

B. Data lake adalah repositori besar yang menyimpan data dalam format asli atau mentahnya, termasuk data terstruktur, semi-terstruktur, dan tidak terstruktur. Data lake dirancang untuk menyimpan data dalam volume besar dan beragam jenis. Contohnya, Apache Hadoop, Amazon S3, Azure Data Lake.

C. Data warehouse adalah sistem terstruktur untuk mengumpulkan dan mengelola data dari berbagai sumber untuk analisis dan pelaporan. Data dalam data warehouse biasanya disusun dalam skema yang dioptimalkan untuk query kompleks dan analisis. Contohnya, Amazon Redshift, Google BigQuery, Snowflake, Microsoft Azure SQL Data Warehouse.

D. Data mart adalah subset dari data warehouse yang difokuskan pada area atau departemen tertentu dalam organisasi dan data mart juga dirancang untuk memenuhi kebutuhan analitik spesifik dari kelompok pengguna tertentu. Contohnya, Data mart untuk analisis penjualan yang menarik data dari data warehouse penjualan perusahaan.

3. Normalisasi database adalah proses pengorganisasian tabel dan kolom dalam sebuah database untuk mengurangi redundansi dan meningkatkan integritas data. Tujuannya adalah untuk meminimalkan duplikasi data dan memastikan data tergantung pada kunci utama.

Employee Table

|  |  |  |  |
| --- | --- | --- | --- |
| **Employee\_id** | **Employee\_name** | **Job\_code** | **City\_code** |
| 1 | John Smith | 101 | 201 |
| 2 | Alice Johnson | 102 | 202 |
| 3 | Bob Davis | 103 | 203 |
| 4 | Emily Wilson | 101 | 204 |
| 5 | Michael Lee | 102 | 205 |
| 6 | Sarah Brown | 103 | 206 |
| 7 | James Clark | 101 | 207 |
| 8 | Laura Taylor | 102 | 208 |
| 9 | Daniel White | 103 | 209 |
| 10 | Olivia Martin1 | 101 | 210 |

Jobs Table

|  |  |
| --- | --- |
| **Job\_code** | **Job** |
| 101 | Software Engineer |
| 102 | Data Analyst |
| 103 | Data Engineer |

Cities Table

|  |  |  |
| --- | --- | --- |
| **City\_code** | **City\_name** | **Province\_code** |
| 201 | New Yotk | 301 |
| 202 | Los Angeles | 302 |
| 203 | Chicago | 303 |
| 204 | Houston | 304 |
| 205 | Miami | 305 |
| 206 | Bostom | 306 |
| 207 | San Francisco | 307 |
| 208 | Seattle | 308 |
| 209 | Denver | 309 |
| 210 | Atlanta | 310 |

Provinces Table

|  |  |
| --- | --- |
| **Province\_code** | **Province\_name** |
| 301 | New York |
| 302 | California |
| 303 | Illinois |
| 304 | Texas |
| 305 | Florida |
| 306 | Massachusetts |
| 307 | California |
| 308 | Washington |
| 309 | Colorado |
| 310 | Georgia |