

Nama : Kharisma Novi Chandramukti

Kelas/Batch : Data Engineering 4

Part 2 – Fundamental DE

1. Kapan kita harus menggunakan relational database atau nosql database ?

JAWAB:

- Relational database : untuk aplikasi yang membutuhkan struktur data yang jelas, transaksi ACID (Atomicity, Consistency, Isolation, Durability) yang kuat, dan konsistensi data yang tinggi
- NoSQL database : untuk aplikasi dengan volume data besar, struktur data yang beragam, dan kebutuhan skalabilitas horizontal yang tinggi

2. Apa perbedaan antara database, data lake, data warehouse, dan data mart ?

JAWAB :

- Database adalah tempat penyimpanan terstruktur untuk data operasional dan transaksional dalam aplikasi bisnis. Contoh :MySQL dan PostgreSQL
- Data lake adalah tempat penyimpanan yang menyimpan data mentah dan tidak terstruktur dari berbagai sumber, seperti data streaming, sensor, atau data historis
- Data warehouse adalah tempat penyimpanan data yang dioptimalkan untuk analisis dan pelaporan bisnis. Contoh : Snowflake dan Google BigQuery
- Data mart adalah subset dari data warehouse yang berfokus pada data tertentu untuk departemen atau fungsi bisnis tertentu

3. Jelaskan apa itu normalisasi database, dan normalisasikan table dibawah !

employee_id	employee_name	job_code	job	city_code	city_name	province_code	province_name
1	John Smith	101	Software Engineer	201	New York	301	New York
2	Alice Johnson	102	Data Analyst	202	Los Angeles	302	California
3	Bob Davis	103	Data Engineer	203	Chicago	303	Illinois
4	Emily Wilson	101	Software Engineer	204	Houston	304	Texas
5	Michael Lee	102	Data Analyst	205	Miami	305	Florida
6	Sarah Brown	103	Data Engineer	206	Boston	306	Massachusetts
7	James Clark	101	Software Engineer	207	San Francisco	307	California
8	Laura Taylor	102	Data Analyst	208	Seattle	308	Washington
9	Daniel White	103	Data Engineer	209	Denver	309	Colorado
10	Olivia Martin	101	Software Engineer	210	Atlanta	310	Georgia

JAWAB:

Normalisasi database adalah proses merancang basis data relasional dengan cara meminimalkan redundansi data memastikan dependensi fungsional antar atribut.

Normalisasi tabel :

