

PART 2- FUNDAMENTAL DE

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

Jawaban :

- Relational Database digunakan jika data tersebut memiliki struktur yang terdefinisi dengan baik dan realasi antar data penting. Relational Database mungkin lebih cocok untuk aplikasi yang memerlukan transaksi kompleks dan konsistensi data yang tinggi.
- Nosql Database digunakan jika data memiliki struktur yang tidak jelas, atau jika kita memiliki kebutuhan untuk menyimpan data semi-struktural atau tidak terstruktur. Database NoSQL biasanya lebih baik dalam menangani skala yang sangat besar atau data dengan pertumbuhan yang cepat karena mereka dirancang untuk menyebarkan beban kerja secara horizontal.

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

Jawab :

- Database adalah kumpulan data yang terstruktur dan tersimpan secara sistematis di dalam sebuah sistem komputer. Data disimpan dalam bentuk tabel dengan relasi antar tabel yang terdefinisi.
- Data lake adalah penyimpanan besar yang menyimpan data mentah (raw) dalam berbagai format, termasuk struktur, semi-struktur, dan tidak terstruktur. Data lake tidak memerlukan skema sebelumnya dan memungkinkan penyimpanan data dalam jumlah besar dengan biaya yang relatif rendah.
- Data warehouse adalah kumpulan data yang terstruktur dan terintegrasi yang diambil dari berbagai sumber data operasional. Data warehouse didesain untuk analisis dan pelaporan bisnis dengan menyediakan data yang bersih, terpadu, dan terorganisir.
- Data mart adalah subset dari data warehouse yang berisi data yang diorganisir untuk kebutuhan spesifik dari departemen atau tim tertentu dalam organisasi.

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

Jawab :

Normalisasi database adalah proses desain database yang bertujuan untuk mengurangi redundansi data dan meningkatkan integritas data. Tujuannya adalah untuk mengatur struktur data dalam database sehingga data disimpan dalam tabel dengan cara yang paling efisien dan efektif.

employee_id	city_code	province_code	employee_name	city_name	province_name	job_code *
1	201	301	John Smith	New York	New York	101
2	202	302	Alice Johnson	Los Angeles	California	102
3	203	303	Bob Davis	Chicago	Illinois	103
4	204	304	Emily Wilson	Houston	Texas	101
5	205	305	Michael Lee	Miami	Florida	102
6	206	306	Sarah Brown	Boston	Massachusetts	103
7	207	307	James Clark	San Fransisco	California	101
8	208	308	Laura Taylor	Seattle	Washington	102

job_code	job
101	Software Engineer
102	Data Analyst
103	Data Engineer

Keterangan : Tabel yang diberikan mempunyai 3 kolom unik yang dapat dijadikan primary key, saya memilih menggunakan 3 kunci gabungan pada table tersebut, berdasarkan Langkah-langkah yang dilakukan pada 1NF, 2NF, dan 3NF didapatkan normalisasi seperti diatas.