**Part 2**

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

* relational database digunakan jika data terstruktur dengan Skema Tetap. seperti tabel dengan kolom yang ditentukan secara jelas.

**Contoh:** Sistem manajemen sumber daya manusia (HRM), sistem akuntansi, sistem pemesanan tiket.

* nosql database digunakan jika data tidak terstruktur atau Semi-Terstruktur. Data tidak memiliki skema yang tetap atau sering berubah, seperti data dalam format JSON, XML, atau dokumen lainnya.

**Contoh:** penyimpanan log aktivitas pengguna, data sensor IoT.

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

### **Database**

Database adalah kumpulan data yang terstruktur dan dikelola menggunakan sistem manajemen basis data (DBMS).

**Contoh:** MySQL, PostgreSQL, Oracle Database, Microsoft SQL Server.

### **Data Lake**

Data lake adalah repositori penyimpanan yang dapat menyimpan data dalam bentuk mentah, terstruktur, semi-terstruktur, atau tidak terstruktur. Data disimpan dalam format aslinya hingga diperlukan untuk analisis.

**Contoh:** Amazon S3, Azure Data Lake Storage, Google Cloud Storage.

### **Data Warehouse**

Data warehouse adalah sistem penyimpanan data yang dioptimalkan untuk kueri dan analisis. Data dari berbagai sumber diintegrasikan, dibersihkan, dan disusun dalam skema yang konsisten untuk analisis.

**Contoh:** Amazon Redshift, Google BigQuery, Snowflake, Microsoft Azure Synapse.

### **Data Mart**

Data mart adalah subset dari data warehouse yang berfokus pada area bisnis tertentu atau departemen dalam organisasi. Data mart dirancang untuk memenuhi kebutuhan analitik spesifik dari kelompok pengguna tertentu.

**Contoh:** Data mart penjualan yang hanya menyimpan data terkait penjualan untuk analisis performa penjualan.

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

* Apa itu Normalisasi Database?

Normalisasi database adalah proses pengorganisasian atribut dan tabel dalam basis data untuk mengurangi redundansi data dan meningkatkan integritas data. Proses ini melibatkan pembagian tabel besar menjadi tabel-tabel yang lebih kecil dan mendefinisikan hubungan di antara mereka

* Normalisasi Tabel

**Tabel Asal**

| 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 Fransisco | 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 |

**Tabel yang sudah di normalisasi**

* **Tabel Employees**

| 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 Martin | 101 | 210 |

* **Tabel Jobs**

| job\_code | job |
| --- | --- |
| 101 | Software Engineer |
| 102 | Data Analyst |
| 103 | Data Engineer |

* **Tabel Cities**

| city\_code | city\_name | province\_code |
| --- | --- | --- |
| 201 | New York | 301 |
| 202 | Los Angeles | 302 |
| 203 | Chicago | 303 |
| 204 | Houston | 304 |
| 205 | Miami | 305 |
| 206 | Boston | 306 |
| 207 | San Fransisco | 307 |
| 208 | Seattle | 308 |
| 209 | Denver | 309 |
| 210 | Atlanta | 310 |

* **Tabel Provinces**

| 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 |