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

**Perbedaan Database, Data Lake, Data Warehouse, dan Data Mart**

**Database**

* **Definisi**: Sistem untuk mengelola dan menyimpan data yang dapat diakses dan diubah oleh aplikasi.
* **Struktur**: Biasanya relasional dengan tabel berstruktur tetap. NoSQL lebih fleksibel.
* **Tujuan**: Menyimpan data operasional harian (transaksi, data pelanggan).
* **Contoh**: MySQL, PostgreSQL, Oracle, MongoDB.
* **Karakteristik**: Mendukung transaksi ACID, operasi CRUD cepat.

**Data Lake**

* **Definisi**: Sistem penyimpanan untuk berbagai format data (mentah, terstruktur, semi-terstruktur, tidak terstruktur).
* **Struktur**: Data disimpan dalam bentuk asli tanpa pengolahan.
* **Tujuan**: Penyimpanan skala besar untuk analitik big data.
* **Contoh**: Hadoop HDFS, Amazon S3, Azure Data Lake.
* **Karakteristik**: Menyimpan berbagai jenis data dalam skala besar, memerlukan alat analitik untuk pemrosesan.

**Data Warehouse**

* **Definisi**: Sistem penyimpanan untuk pelaporan dan analisis data, mengintegrasikan data dari berbagai sumber.
* **Struktur**: Skema terstruktur (skema bintang, skema salju).
* **Tujuan**: Optimalkan query dan analisis data historis untuk keputusan bisnis.
* **Contoh**: Amazon Redshift, Google BigQuery, Snowflake.
* **Karakteristik**: Data telah diproses dan dibersihkan, mendukung query analitik kompleks.

**Data Mart**

* **Definisi**: Subset dari data warehouse untuk analisis dan pelaporan spesifik departemen atau unit bisnis.
* **Struktur**: Fokus pada data relevan untuk area bisnis tertentu.
* **Tujuan**: Menyediakan data terfokus dan mudah diakses untuk analisis departemen.
* **Contoh**: Data mart untuk penjualan, pemasaran, atau keuangan.
* **Karakteristik**: Lebih kecil dari data warehouse, dioptimalkan untuk kebutuhan spesifik.

**Ringkasan**

* **Database**: Data operasional harian, struktur tetap atau fleksibel.
* **Data Lake**: Data mentah berbagai format untuk analitik besar.
* **Data Warehouse**: Data terproses untuk analisis bisnis.
* **Data Mart**: Subset data warehouse untuk analisis departemen spesifik.