**Nama : Pradana Argo Pangestu**

**Nim : 2311104079/SE-07-02**

**MODUL 10**

**TUGAS PENDAHULUAN**

1. Apa yang dimaksud dengan columnar database ?
2. Apa kelebihan dari columnar database?
3. Berikancontohbagaimana data disimpandalam columnar database.

**1. Apa yang Dimaksud dengan Columnar Database?**

**Columnar database** adalah jenis database yang menyimpan data berdasarkan kolom, bukan baris (seperti pada database tradisional/row-based). Setiap kolom disimpan dalam struktur terpisah, sehingga memungkinkan pembacaan data kolom-spesifik secara efisien. Jenis database ini dirancang untuk **analitik data** dan **query agregasi** (misalnya, menghitung rata-rata, total, atau filter kolom tertentu).

**2. Kelebihan Columnar Database**

* **Efisiensi Query**: Hanya kolom yang dibutuhkan dalam query yang dibaca, mengurangi I/O dan mempercepat pemrosesan**1**.
* **Kompresi Optimal**: Data dalam kolom cenderung homogen (misalnya, angka atau kategori), sehingga teknik kompresi seperti run-length encoding lebih efektif**2**.
* **Analitik Lebih Cepat**: Cocok untuk OLAP (Online Analytical Processing) dan aplikasi Business Intelligence (BI).
* **Skalabilitas**: Mendukung volume data besar (big data) karena desainnya yang optimal untuk operasi kolom.

**3. Contoh Penyimpanan Data dalam Columnar Database**

**Contoh Tabel Data (Format Row-based):**

| **ID** | **Nama** | **Usia** | **Kota** |
| --- | --- | --- | --- |
| 1 | Alice | 25 | New York |
| 2 | Bob | 30 | Los Angeles |
| 3 | Charlie | 28 | San Francisco |

**Penyimpanan Columnar:**

* **Kolom ID**: **[1, 2, 3]**
* **Kolom Nama**: **["Alice", "Bob", "Charlie"]**
* **Kolom Usia**: **[25, 30, 28]**
* **Kolom Kota**: **["New York", "Los Angeles", "San Francisco"]**

**Cara Kerja:**

* Jika query meminta **rata-rata usia**, database hanya membaca kolom **Usia** (tidak perlu akses kolom Nama/Kota).
* Data setiap kolom disimpan dalam blok terpisah (misalnya, dalam file atau memori) dengan metadata kompresi.

**Contoh Teknologi:**

* **Apache Parquet** (format penyimpanan columnar untuk Hadoop).
* **Amazon Redshift**, **Google BigQuery** (database analitik columnar).

**Perbandingan dengan Row-based Database**:

| **Operasi** | **Columnar Database** | **Row-based Database** |
| --- | --- | --- |
| Query agregasi (SUM, AVG) | Sangat cepat | Lambat (baca seluruh baris) |
| Insert/Update data | Lebih lambat | Cepat (optimal untuk OLTP) |