



Jurusan Teknologi Informasi Politeknik Negeri Malang
Mata Kuliah Data Warehouse
Kuis 1

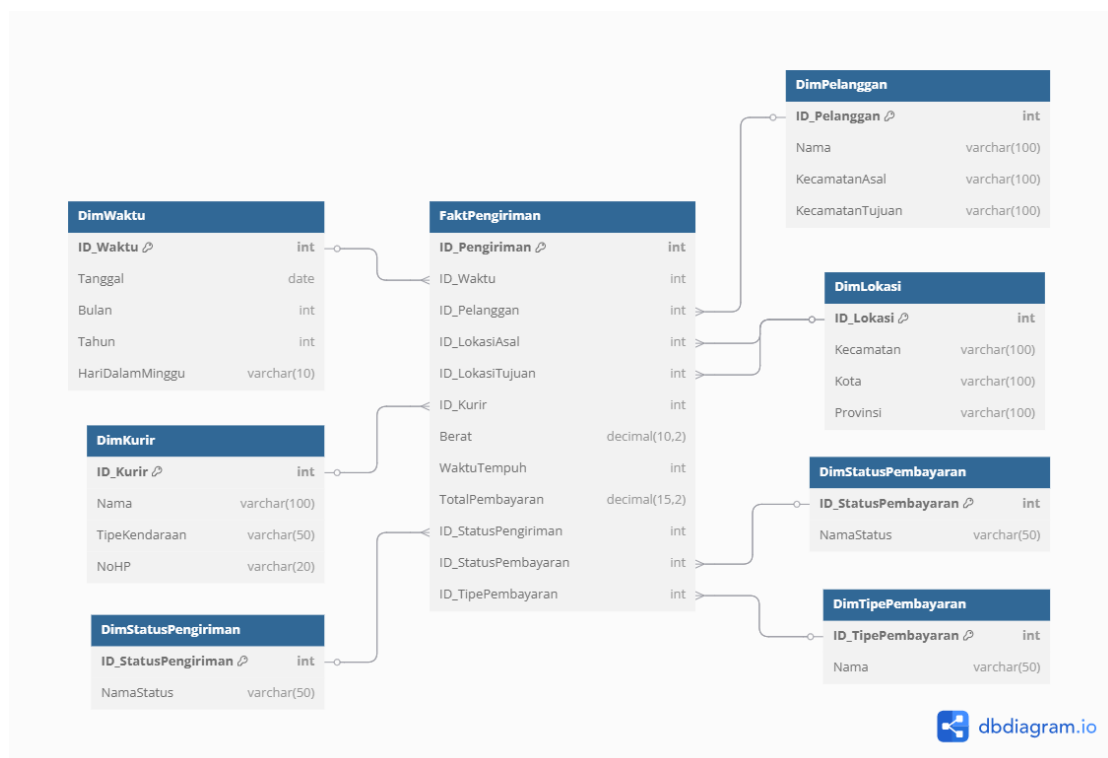
Nama : Dhevina Agustina

Nomor Urut : 06 (2341760065)

1. Tuliskan perbandingan star schema dan snowflake schema pada tabel berikut:

	Star Schema	Snowflake Schema
Normalisasi	Denormalisasi (tabel dimensi tidak dipecah)	Normalisasi (tabel dimensi dipecah menjadi sub-tabel)
Kompleksitas desain/skema	Sederhana, karena tabel dimensi langsung terhubung ke tabel fakta	Lebih kompleks, karena tabel dimensi dipecah ke dalam beberapa tabel kecil
Kompleksitas query	Lebih sederhana, karena jumlah join lebih sedikit	Lebih kompleks, karena membutuhkan lebih banyak join antar tabel
Performa query	Lebih cepat, karena jumlah join lebih sedikit	Lebih lambat, karena query membutuhkan lebih banyak join
Storage	Membutuhkan lebih banyak ruang karena data redundan	Lebih hemat storage karena data tidak redundan
Integritas data	Lebih rentan terhadap inkonsistensi karena tidak ada normalisasi	Lebih baik dalam menjaga integritas data karena adanya normalisasi
Maintenance (pengisian data dengan proses ETL dari OLTP)	Lebih mudah karena skema lebih sederhana	Lebih kompleks karena tabel lebih banyak dan perlu transformasi tambahan

2. Gambar berikut menunjukkan skema OLTP database dari sebuah sistem informasi ekspedisi. Buatlah data warehouse dalam star schema yang digunakan sebagai dasar analisis performa ekspedisi.



Analisis Performa:

- Query lebih cepat
Query untuk analisis dan laporan menjadi lebih cepat karena jumlah join lebih sedikit dan bisa langsung mengambil data dari tabel dimensi tanpa banyak join ke tabel lain.
- Struktur sederhana
Karena data hanya terdiri dari beberapa tabel utama seperti DimWaktu, DimKurir, dan DimLokasi, struktur ini lebih mudah digunakan oleh pengguna non-teknis.
- Mengurangi beban query pada OLAP
Karena data sudah denormalisasi, pengolahan query di sistem OLAP (Online Analytical Processing) lebih efisien, terutama dalam analisis ketepatan waktu pengiriman, performa kurir, dan tren pembayaran.
- Redundansi Data
Tabel DimLokasi telah menggabungkan kecamatan, kota, dan provinsi, menyebabkan pengulangan data yang dapat meningkatkan ukuran penyimpanan. Misalnya, jika ada ribuan pengiriman dari kecamatan yang sama, maka data kota dan provinsi akan diulang dalam tabel.
- Konsumsi Storage Lebih Tinggi
Karena tidak ada normalisasi pada tabel dimensi, penyimpanan data bisa lebih besar dibandingkan Snowflake Schema, terutama pada tabel pelanggan dan lokasi.
- Update Lebih Lambat
Jika ada perubahan pada atribut dimensi (misalnya perubahan nama kecamatan, tipe kendaraan kurir, atau metode pembayaran), maka harus diperbarui di banyak baris, yang bisa menyebabkan latensi pada update data.