



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

Nama : Satria Rakhmadani

Nomor Urut : 28

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

	Star Schema	Snowflake Schema
Normalisasi	Tidak dinormalisasi, karena tabel dimensi langsung terhubung ke tabel fakta memecahkan sub-dimensi	Ternormalisasi, karena tabel dimensi dipecah menjadi beberapa sub-tabel dimensi
Kompleksitas desain/skema	Sederhana karena hanya ada satu layer dimensi (dimensi langsung terhubung ke fakta).	Lebih kompleks karena dimensi bisa bercabang menjadi sub-dimensi.
Kompleksitas query	Lebih sederhana untuk query karena join langsung antara tabel fakta dan tabel dimensi.	Lebih kompleks karena memerlukan join tambahan pada sub-dimensi.
Performa query	Cepat dalam eksekusi query karena join lebih sedikit.	Cenderung lebih lambat dibanding Star Schema, jika banyak join ke sub-dimensi.
Storage	Biasanya membutuhkan lebih banyak storage karena data dimensi terduplikasi	Lebih hemat storage karena data dimensi dipecah dan dinormalisasi (mengurangi redundansi).
Integritas data	Karena denormalisasi, potensi inkonsistensi data dimensi lebih besar jika tidak dikelola dengan baik.	Integritas data lebih terjaga karena normalisasi, duplikasi data lebih minim.
Maintenance (pengisian data dengan proses ETL dari OLTP)	Lebih mudah dalam proses ETL karena struktur dimensi lebih simpel.	Lebih rumit karena harus mengisi banyak tabel sub-dimensi (butuh logika ETL lebih kompleks).

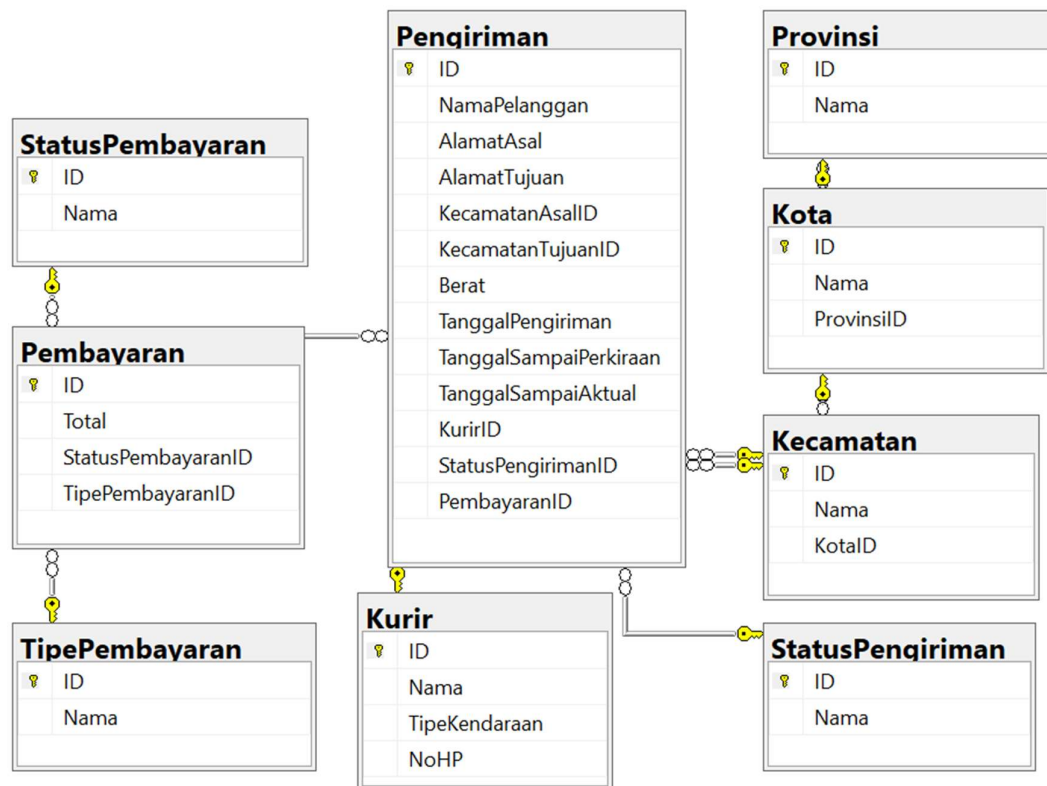
Star Schema:

- Terdiri dari satu tabel fakta di tengah dan beberapa tabel dimensi di sekelilingnya.
- Desain sederhana dan mudah dipahami, performa query lebih cepat karena join yang lebih sedikit.
- Namun, karena denormalisasi, data dimensi bisa jadi lebih besar dan berpotensi duplikasi.

Snowflake Schema:

- Merupakan perluasan Star Schema dengan memecahkan tabel dimensi menjadi beberapa tabel sub-dimensi (normalisasi lebih tinggi).
- Mengurangi duplikasi data (lebih hemat storage) dan meningkatkan integritas data.
- Akan tetapi, query menjadi lebih kompleks (banyak join) dan proses ETL yang lebih rumit.

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.



Jawaban

Tabel Fakta dengan Judul Fact Pengiriman

FactPengirimanID → Primary Key (PK) untuk tabel fakta
TanggalPengirimanID → (FK ke DimDate): Tanggal barang dikirim
TanggalSampaiTujuanID → (FK ke DimDate): Tanggal estimasi sampai
TanggalSampaiAktualID → (FK ke DimDate): Tanggal barang benar-benar sampai
LokasiAsalID → (FK ke DimLokasi): Lokasi asal (kecamatan, kota, provinsi)
LokasiTujuanID → (FK ke DimLokasi): Lokasi tujuan (kecamatan, kota, provinsi)
KurirID → (FK ke DimKurir): Informasi kurir (nama, tipe kendaraan)
StatusPengirimanID → (FK ke DimStatusPengiriman): Status pengiriman
StatusPembayaranID → (FK ke DimStatusPembayaran): Status pembayaran
TipePembayaranID → (FK ke DimTipePembayaran): Metode pembayaran

Untuk Tabel Dimensi lainnya terdapat 6 tabel dimensi.

DimDate

Kolom
DateID (PK)

Tanggal
Bulan
Tahun

DimLokasi

Kolom
LokasiID (PK)
Kecamatan
Kota
Provinsi

DimKurir

Kolom
KurirID (PK)
Nama
TipeKendaraan

DimStatusPengiriman

Kolom
StatusPengirimanID (PK)
Nama

DimStatusPembayaran

Kolom
StatusPembayaranID (PK)
Nama

DimTipePembayaran

Kolom
TipePembayaranID (PK)
Nama

