



Kampus
Merdeka
INDONESIA JAYA

business intelligence

ARSITEKTUR DATA WAREHOUSE

FAST: FOCUS, ADAPTIVE, SINCERE, TRUSTWORTHY

JURUSAN TEKNOLOGI INFORMASI
POLITEKNIK NEGERI MALANG



OUTLINE

1. Review Konsep Dasar Data Warehouse
2. Model-model Arsitektur
3. Struktur Data Warehouse
4. Manajemen Data dalam Data Warehouse



Bagian 1

Review Konsep Dasar Data Warehouse

DATA WAREHOUSE

01 Pengertian Umum
Data warehouse adalah sistem yang mengarsipkan dan menganalisis data historis untuk pengambilan keputusan.

02 Fungsi Utama
Sebagai referensi utama dalam pengambilan keputusan bisnis yang tepat dan efisien.

03 Peran dalam Bisnis
Sebagai referensi utama dalam pengambilan keputusan bisnis yang tepat dan efisien.

04 Pentingnya Data Warehouse
sebagai inti dari infrastruktur informasi bisnis, yang memungkinkan organisasi untuk mengumpulkan, mengintegrasikan, dan menganalisis data dari berbagai sumber secara efisien.

05 Perbedaan dengan Database
Database digunakan untuk operasi sehari-hari data warehouse digunakan untuk analisis bisnis yang mendalam.



Bagian 2

Model – model Arsitektur

Model Arsitektur

Single-tier Architecture

01

02

Two-tier Architecture

Three-tier Architecture

03

04

Integrated Business
Intelligence Architecture

Single-Tier Architecture

- Model ini melibatkan penyimpanan data, proses ETL, dan alat analisis dalam satu sistem tunggal.
- Cocok untuk organisasi kecil dengan kebutuhan yang sederhana dan volume data yang tidak terlalu besar.
- Keuntungan utamanya adalah kesederhanaan dan biaya yang rendah.
- Kelemahannya adalah kurangnya skalabilitas dan keterbatasan dalam mengelola data yang sangat besar atau kompleks

Two-Tier Architecture

- Model ini memisahkan antara penyimpanan data (data warehouse) dan alat analisis (misalnya, OLAP server).
- Data diambil dan dimuat dari sumber eksternal ke dalam data warehouse melalui proses ETL, dan kemudian dianalisis menggunakan alat analisis terpisah.
- Cocok untuk organisasi dengan volume data yang lebih besar dan kebutuhan analisis yang kompleks.
- Keuntungannya termasuk pemisahan antara penyimpanan dan analisis, serta kemampuan untuk meningkatkan kinerja dan skalabilitas.

Three-Tier Architecture

- Model ini membagi arsitektur menjadi tiga lapisan: lapisan sumber data, lapisan penyimpanan data (data warehouse), dan lapisan presentasi (alat analisis dan pelaporan).
- Proses ETL dilakukan antara lapisan sumber data dan lapisan penyimpanan data, sementara analisis dan pelaporan dilakukan pada lapisan presentasi.
- Cocok untuk organisasi dengan kebutuhan yang kompleks dan skala yang besar, yang membutuhkan fleksibilitas dan skalabilitas.
- Keuntungannya meliputi pemisahan tugas, skalabilitas, dan fleksibilitas dalam pengelolaan dan analisis data

Integrated Business Intelligence Architecture

- Model ini mengintegrasikan data warehouse dengan sistem operasional dan aplikasi bisnis lainnya.
- Data diambil secara real-time atau hampir real-time dari sumber operasional ke dalam data warehouse, dan kemudian digunakan untuk analisis dan pelaporan.
- Cocok untuk organisasi yang mengutamakan akses cepat dan informasi yang mutakhir.
- Keuntungannya adalah kemampuan untuk menghasilkan wawasan bisnis yang lebih cepat dan responsif terhadap perubahan pasar.



Bagian 3

Struktur Data Warehouse

OLTP vs OLAP

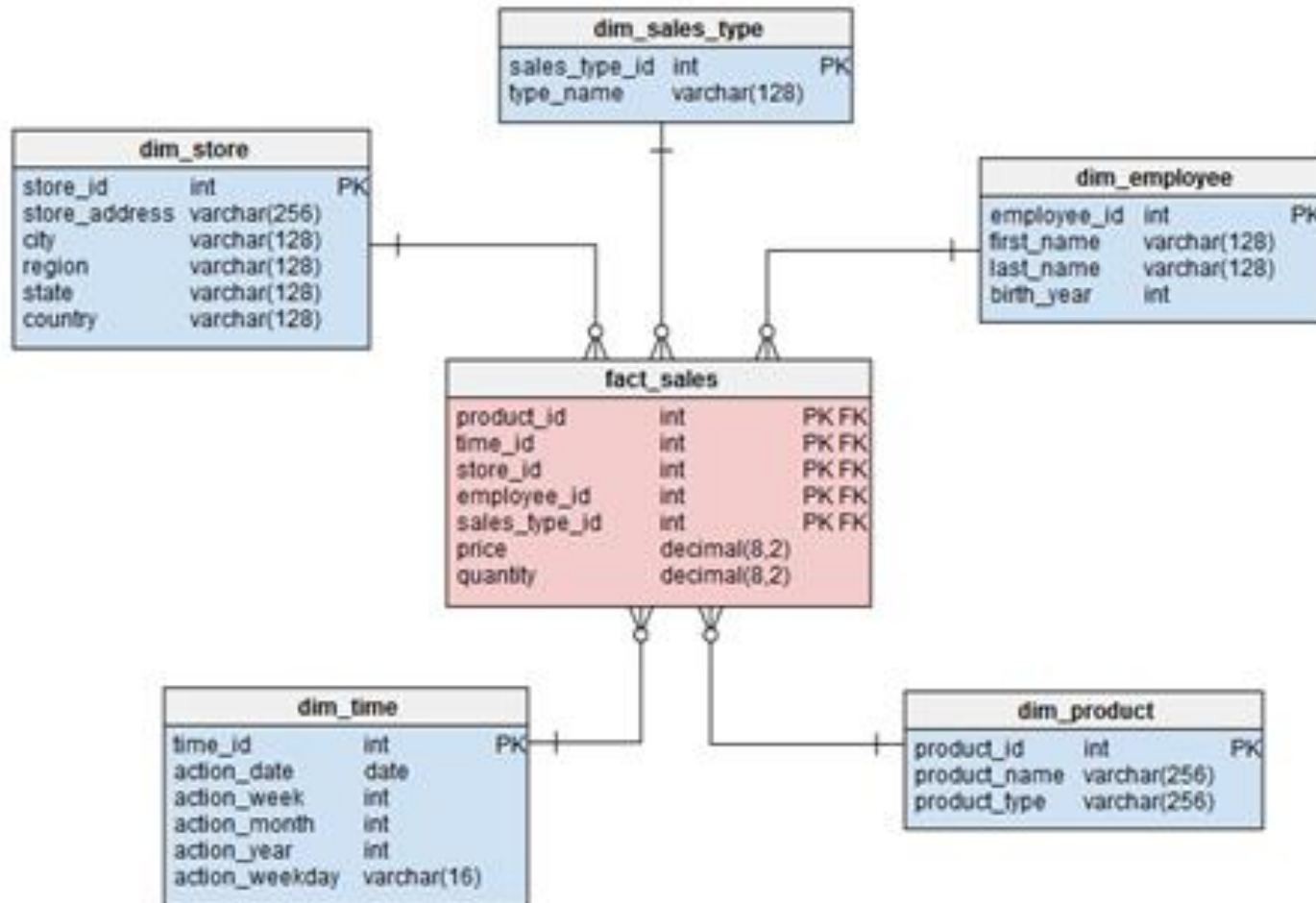
	OLTP	OLAP
Tujuan Utama	transaksi bisnis sehari-hari,	menganalisis data bisnis secara mendalam
Jenis Operasi	operasi transaksi seperti penambahan, penghapusan, dan pembaruan.	Melakukan operasi analisis data kompleks seperti kueri yang kompleks, agregasi, dan pemodelan multidimensional
Volume Data	Mengelola volume data yang besar dalam basis data operasional.	volume data yang besar dari data warehouse.
Struktur Data	Struktur data yang digunakan adalah struktur relasional	Struktur data yang digunakan dapat berupa model dimensional
Pengguna dan Kasus Penggunaan	Biasanya digunakan oleh pengguna operasional	Biasanya digunakan oleh analis bisnis atau manajer
Kecepatan Respon	Memiliki waktu respons yang sangat cepat untuk memproses transaksi secara real-time.	waktu respons mungkin lebih lambat daripada karena melibatkan akses data yang besar dan kompleks.

Tabel Dimensi dan Tabel fakta

- Tabel dimensi adalah tabel yang menyimpan informasi tentang **dimensi atau karakteristik** dari data yang diukur dalam data warehouse.
- Dimensi adalah sudut pandang yang berbeda dari data, seperti waktu, lokasi, produk, atau pelanggan.
- Contoh atribut dalam tabel dimensi waktu dapat mencakup tanggal, bulan, tahun, hari, jam, dll.
- Tabel dimensi biasanya memiliki struktur yang relatif datar dan terdiri dari kunci utama (**primary key**) dan beberapa atribut deskriptif.

- Tabel fakta adalah tabel yang menyimpan data yang **diukur atau diamati** dalam data warehouse, serta informasi numerik yang terkait dengan dimensi tertentu.
- Fakta dalam tabel fakta seringkali adalah **ukuran atau metrik bisnis**, seperti penjualan, pendapatan, atau jumlah barang yang terjual.
- Setiap baris dalam tabel fakta **mewakili sebuah peristiwa atau transaksi** bisnis yang diamati dan direkam.
- Tabel fakta biasanya terdiri dari dua jenis kolom: kunci-kunci dimensi yang berfungsi sebagai kunci luar (**foreign key**) ke tabel dimensi dan kolom-kolom yang menyimpan nilai-nilai fakta.

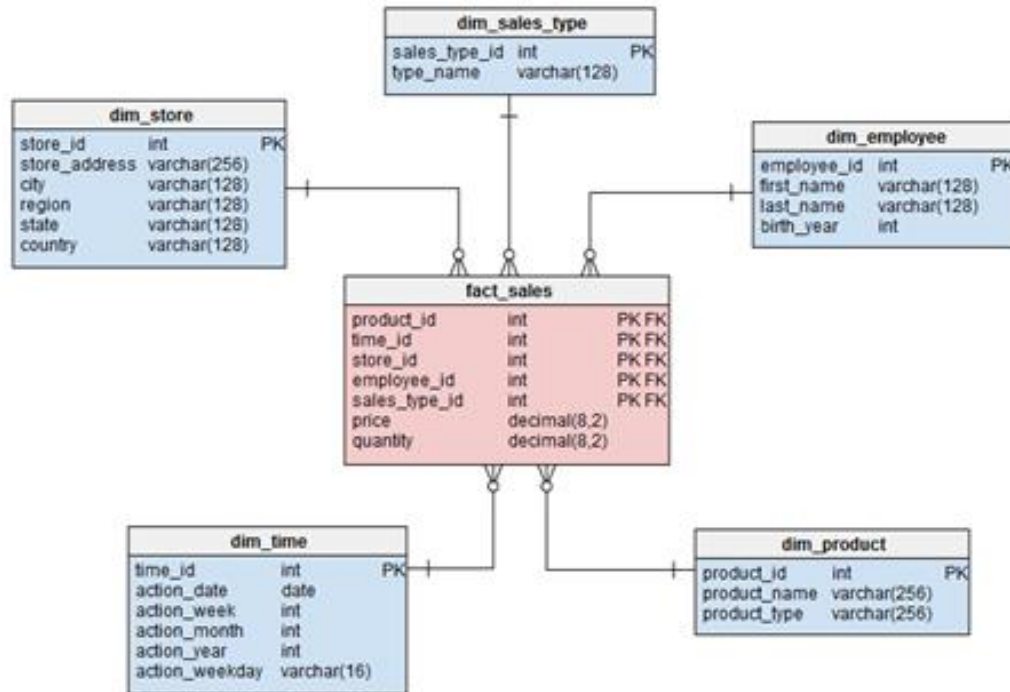
STAR SCHEMA



Sederhana dan Mudah Dipahami:

1. Star schema memiliki struktur yang sederhana dan mudah dipahami, membuatnya menjadi salah satu model yang paling populer dalam desain data warehouse.
2. Struktur yang jelas dan intuitif memfasilitasi pengambilan data dan analisis yang cepat.

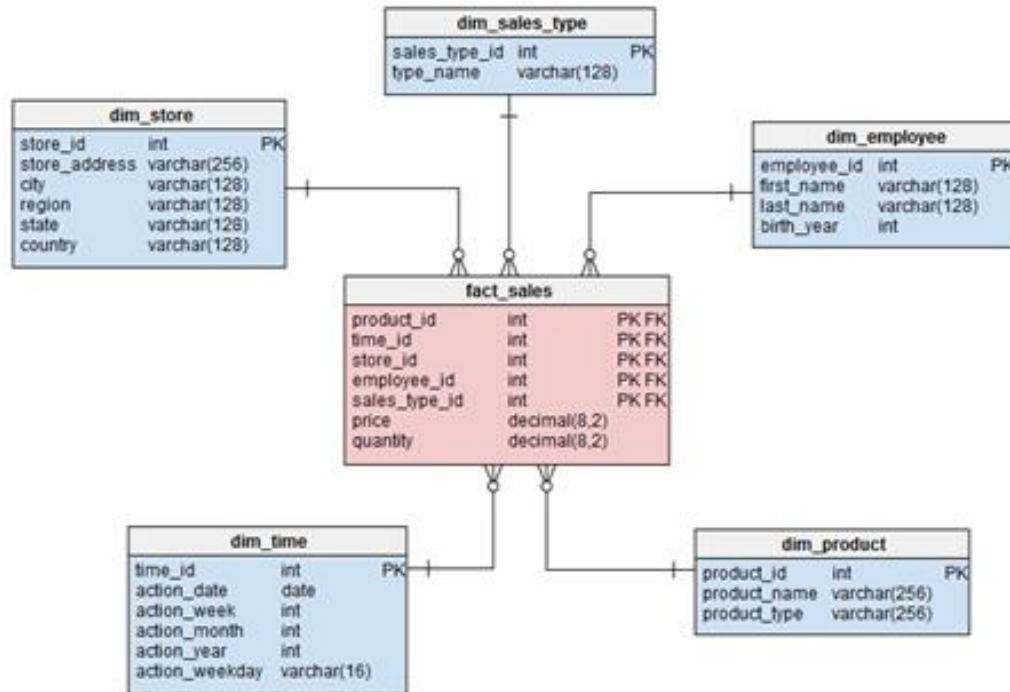
STAR SCHEMA



1. Tabel Fakta:

1. Tabel fakta berfungsi sebagai titik pusat dalam star schema.
2. Tabel fakta menyimpan data yang diukur atau diamati, seperti metrik bisnis atau angka penjualan.
3. Setiap baris dalam tabel fakta mewakili sebuah peristiwa atau transaksi bisnis yang diamati.
4. Tabel fakta biasanya memiliki kunci-kunci dimensi (foreign keys) yang menghubungkannya ke tabel dimensi.

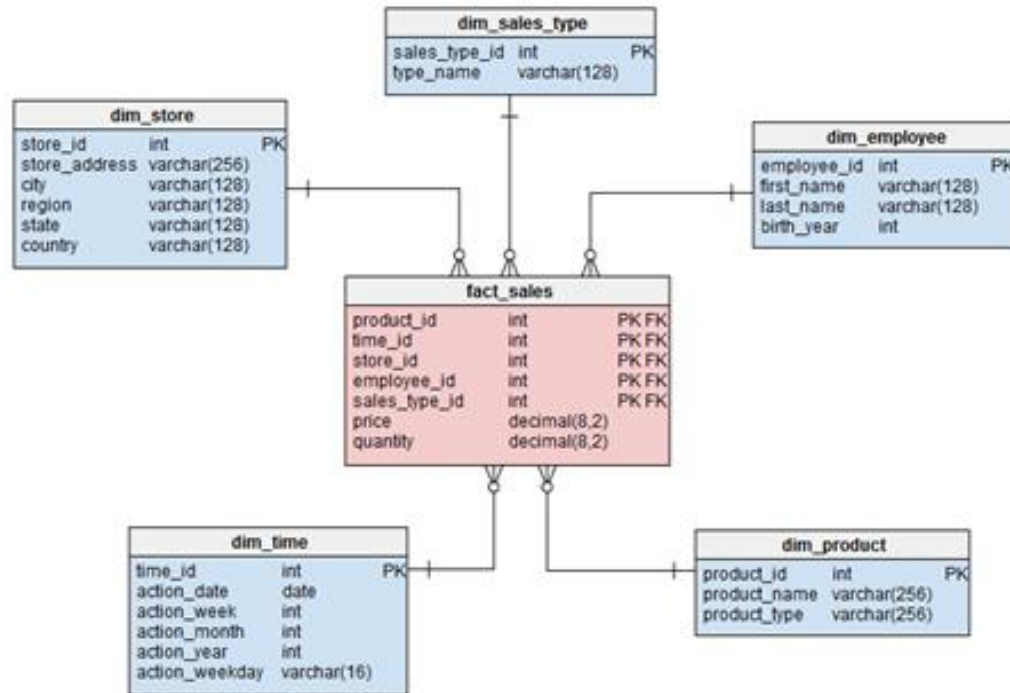
STAR SCHEMA



2. Tabel Dimensi:

1. Tabel dimensi menyimpan informasi tentang dimensi atau karakteristik dari data yang diukur dalam data warehouse.
2. Dimensi bisa berupa waktu, produk, pelanggan, atau dimensi lainnya yang relevan untuk analisis.
3. Setiap tabel dimensi biasanya memiliki satu baris untuk setiap nilai unik dalam dimensi tersebut.
4. Tabel dimensi memiliki hubungan one-to-many dengan tabel fakta, di mana satu baris dalam tabel dimensi dapat berhubungan dengan banyak baris dalam tabel fakta.

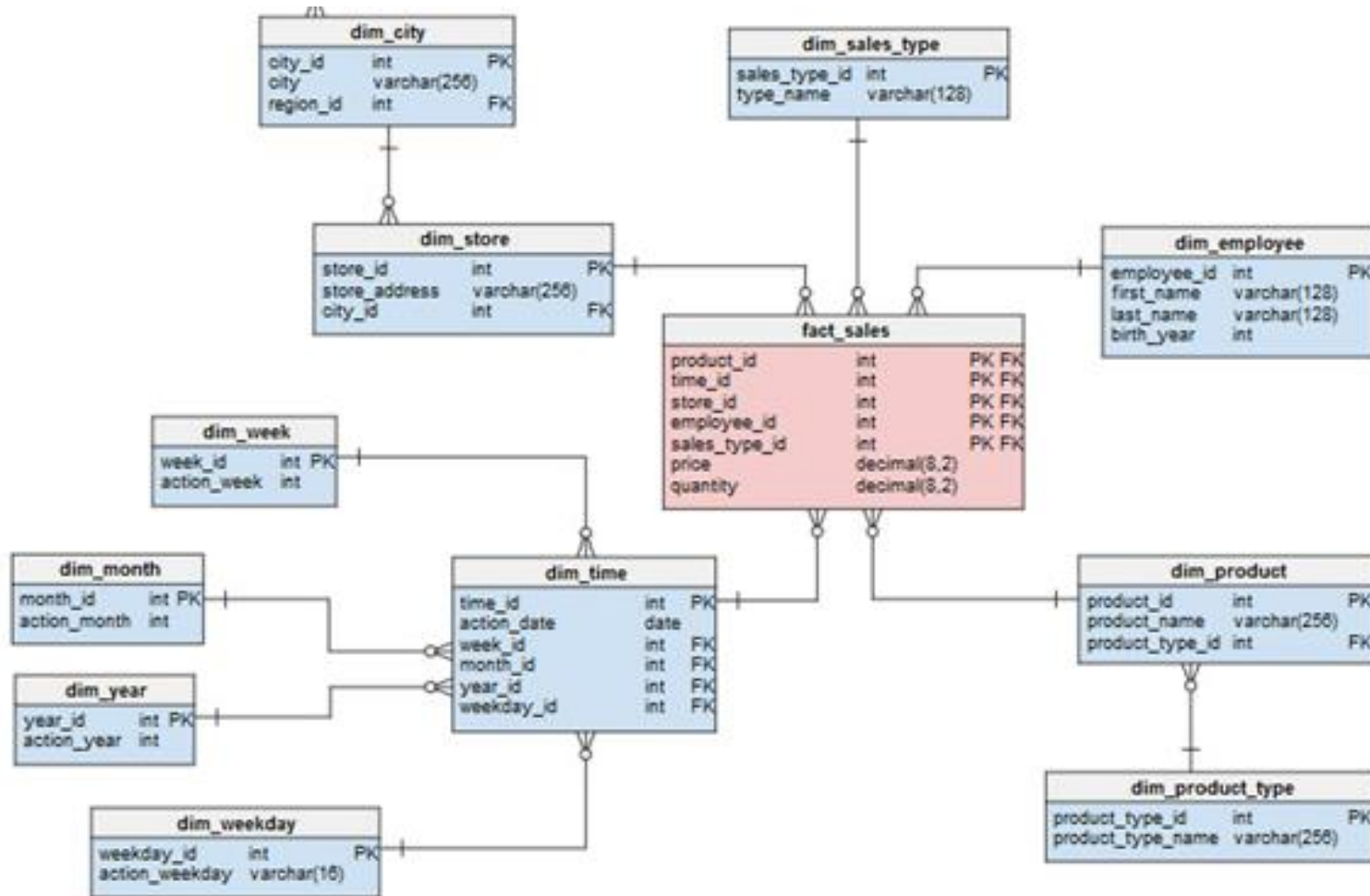
STAR SCHEMA



3. Hubungan Antar Tabel:

1. Hubungan antara tabel fakta dan tabel dimensi bersifat star, di mana tabel fakta berfungsi sebagai pusat dan tabel dimensi berfungsi sebagai "lengan" bintang yang menghubungkannya.
2. Hubungan ini biasanya dilakukan melalui kunci-kunci dimensi, di mana kunci-kunci dimensi dalam tabel fakta adalah kunci luar (foreign keys) yang menghubungkannya ke tabel dimensi.

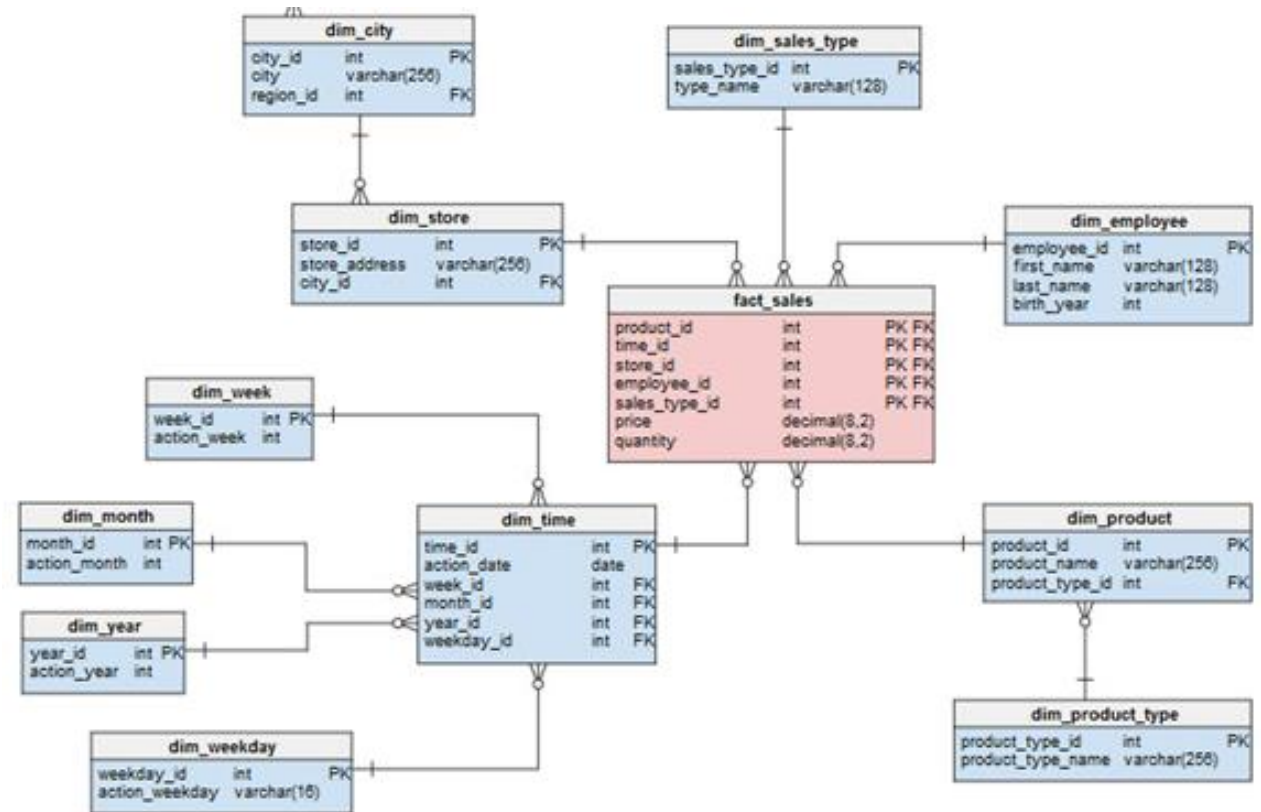
SNOWFLAKE SCHEMA



SNOWFLAKE SCHEMA

1. Tabel Fakta:

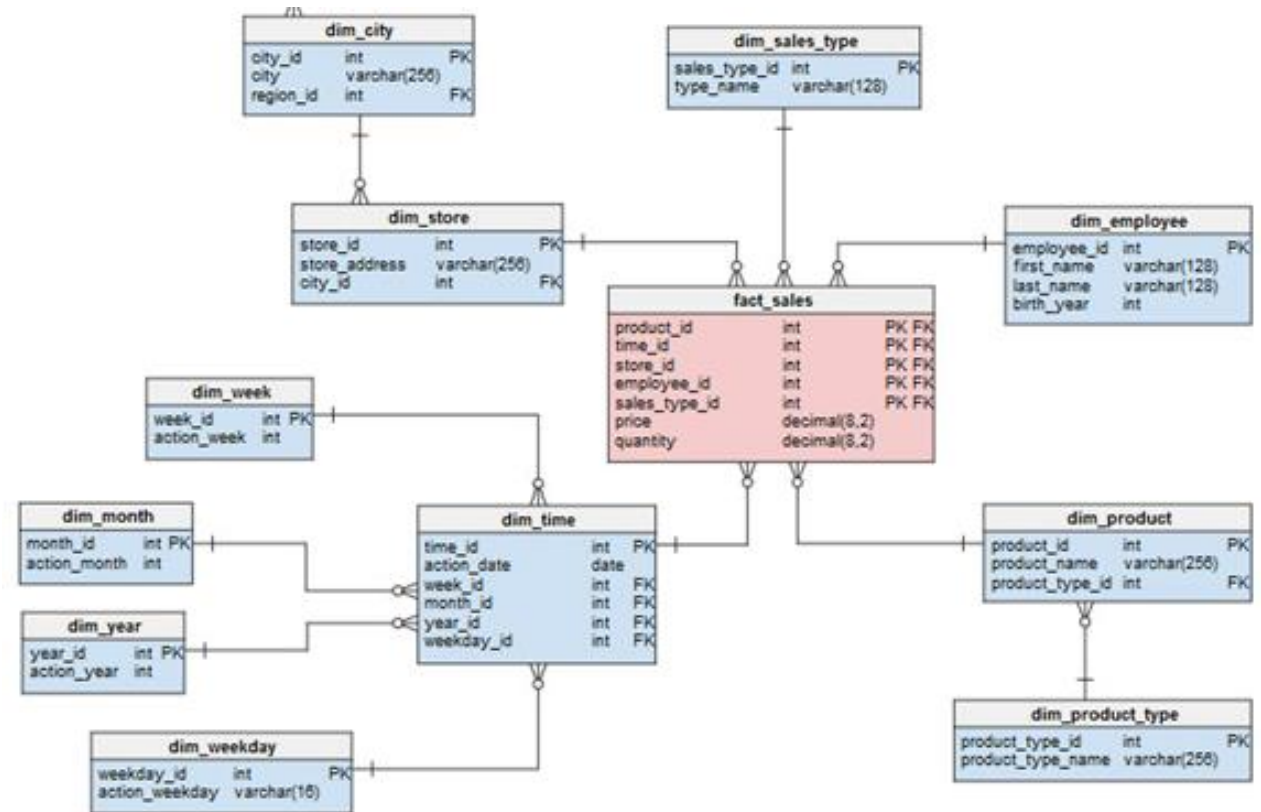
1. Tabel fakta adalah pusat dari snowflake schema, yang menyimpan data yang diukur atau diamati, serta kunci-kunci dimensi yang terkait.
2. Setiap baris dalam tabel fakta mewakili sebuah peristiwa atau transaksi bisnis yang diamati dan direkam.
3. Kolom-kolom dalam tabel fakta mencakup kunci-kunci dimensi yang menghubungkannya ke tabel dimensi, serta atribut-atribut fakta seperti metrik bisnis atau ukuran.



SNOWFLAKE SCHEMA

2. Tabel Dimensi:

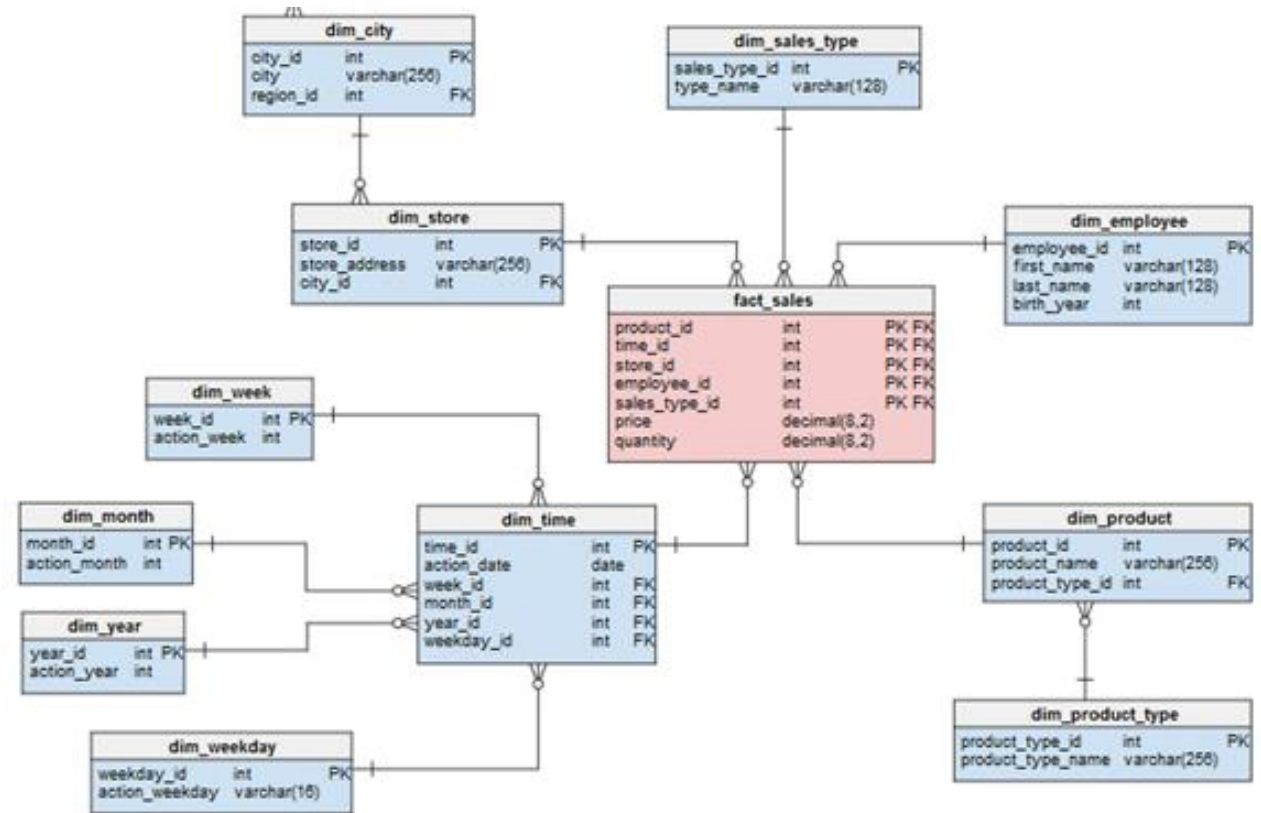
1. Tabel dimensi adalah tabel yang menyimpan informasi tentang dimensi atau karakteristik dari data yang diukur dalam data warehouse.
2. Dimensi tersebut bisa mencakup informasi seperti waktu, produk, pelanggan, dan lokasi.
3. Dalam snowflake schema, tabel dimensi bisa terbagi menjadi beberapa tabel yang lebih kecil atau ter-normalisasi, membentuk "cabang" atau "ranting" yang mengarah ke dimensi yang lebih detail atau spesifik.



SNOWFLAKE SCHEMA

3. Hubungan Antar Tabel:

1. Hubungan antar tabel dalam snowflake schema menggambarkan bagaimana tabel fakta terhubung ke tabel dimensi dan bagaimana tabel dimensi mungkin terhubung satu sama lain.
2. Hubungan antar tabel biasanya terjadi melalui kunci asing (foreign keys) di tabel fakta yang merujuk ke kunci utama (primary keys) di tabel dimensi.
3. Selain itu, dalam snowflake schema, tabel dimensi bisa saja terhubung satu sama lain melalui kunci asing yang membentuk struktur normalisasi.





Bagian 4

Manajemen Data Warehouse

Manajemen Data Warehouse

Tabel Fakta

- Pemeliharaan Data
- Performa Query
- Pembaharuan Data

Tabel Dimensi

- Teknik Pengelolaan Dimensi
- Dimensi Waktu
- Dimensi Hierarki

Kualitas Data

- Teknik Pembersihan Data
- Pengendalian Kualitas

Kemanan Data

- Resiko Keamanan Data
- Strategi Proteksi Data
- Regulasi



Terima kasih