

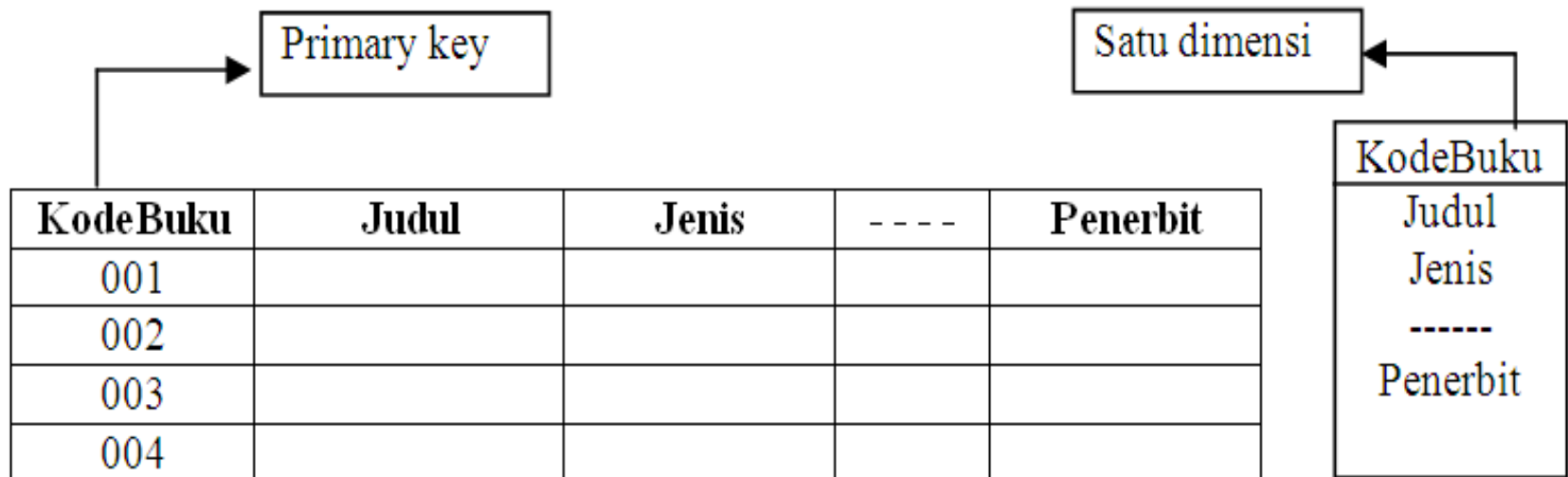
DIMENSIONAL MODELING

Dimension

- Dimension adalah sebuah struktur yang terbentuk dari satu atau lebih hirarki yang mengkategorisasi data
- Dimension terbentuk dari satu atau lebih tabel. Setiap kolomnya merepresentasikan level pada *hierarchy*.
- Model dimensional merupakan rancangan logikal yang bertujuan untuk menampilkan data dalam bentuk standar dan intuitif yang memperbolehkan akses dengan performa yang tinggi.

Tabel Relational

- Tabel Relasional → dibangun oleh baris dan kolom dengan dua sudut pandang , yaitu : baris sebagai sumbu x dan kolom sebagai sumbu y.
- Sebuah tabel hanya mempunyai satu dimensi.



Tabel Relational

Karakteristik :

- setiap record atau baris merepresentasikan data mahasiswa yang berbeda-beda.
- Satu baris dengan baris lainnya diidentifikasi dengan sebuah key yaitu *primary key*.
- Sedangkan bagian kolom seperti, nama, alamat, telepon menyimpan fakta yang sama atau sejenis, dimana setiap fakta tersebut merujuk pada *primary key* yaitu NRP.
- Hal inilah yang menunjukkan bahwa tabel relasional hanya mempunyai satu dimensi.

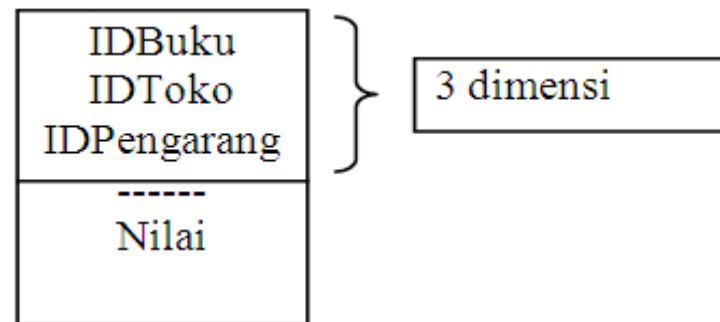
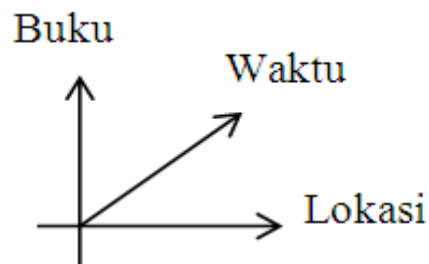
Data Multidimensional

- Data multidimensi adalah “ketika kita dapat melihat sebuah data dari berbagai sudut pandang atau dimensi”
- Sebagai contoh, penjualan buku dapat dilihat dari segi lokasi Buku, waktu, lokasi penjualan / toko dan sebagainya.
- Jika digambarkan , maka akan terdapat tiga koordinat yaitu sumbu x mewakili buku, sumbu y mewakili dimensi waktu dan sumbu z untuk dimensi lokasi.
- Hal inilah yang menjadi perbedaan mendasar antara tabel relasional dan data multidimensi.

Contoh Data Multidimensional

- Spreadsheet adalah salah satu contoh data multidimensional
- Fungsi dapat menyediakan perhitungan sederhana dan dapat diubah menjadi dimensi tunggal

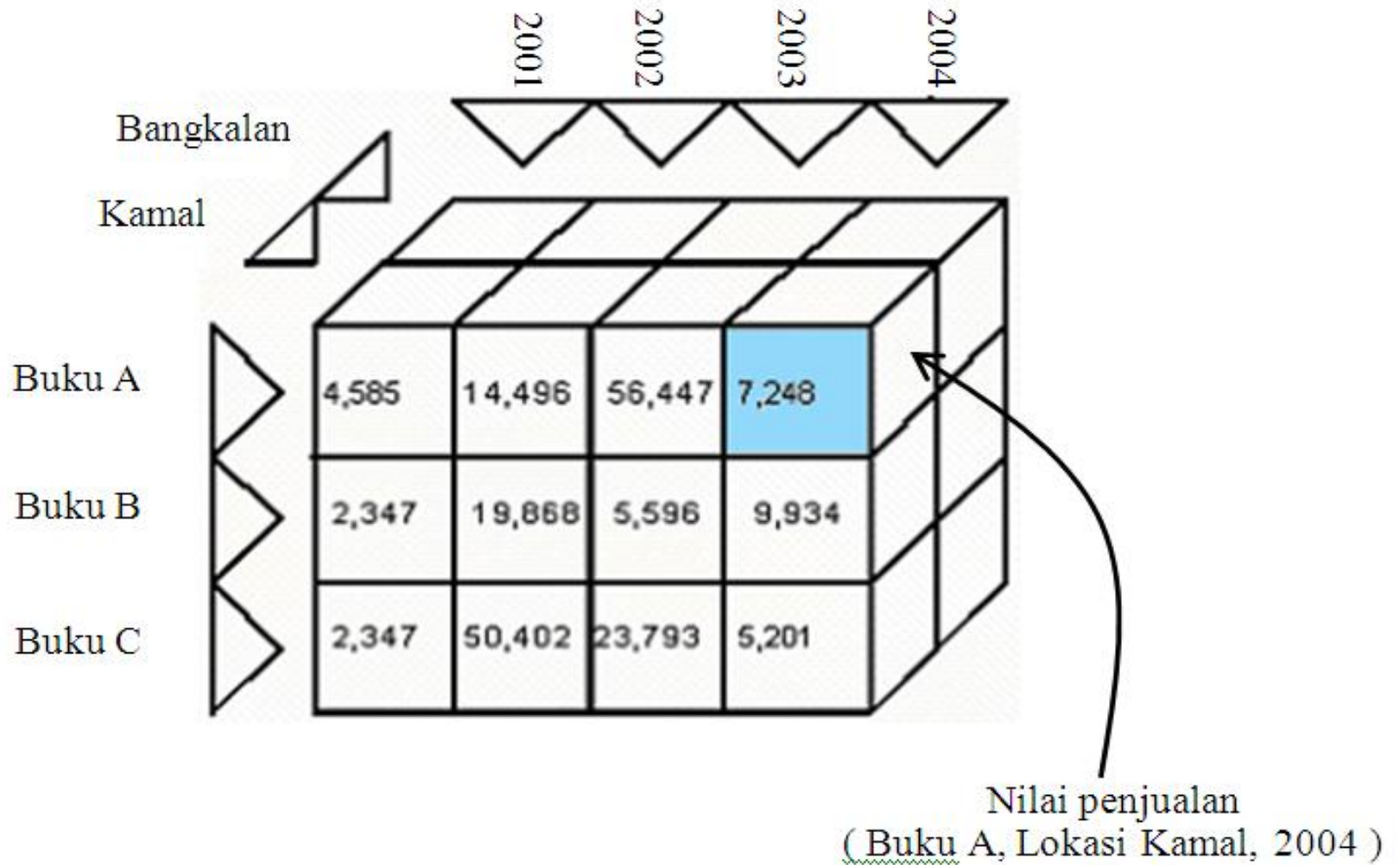
	Kamal	Peb 01	Peb 02	----	Peb 04
Bangkalan	Jan 01	Jan 02	----	Jan 04	
Buku A			----		
Buku B			----		
Buku C			----		



Contoh Data Multidimensional

- *Cube* merupakan contoh data multidimensi selain spreadsheet. Setiap sumbu *cube* mewakili dimensi-dimensi.
- Terdapat *measure* yaitu nilai quantitative database yang ingin dianalisa. Biasanya *measure* berupa nilai penjualan, biaya, *budget* dan sejenisnya. *Measure* dihitung berdasarkan dimensi-dimensi dari *cube*.
- Misalkan menganalisa penjualan buku, total penjualan (sebagai *measure*) untuk suatu jenis buku tertentu (dimensi) pada lokasi / toko buku yang berbeda-beda (lokasi dapat dipilih sesuai dengan kebutuhan seperti negara, propinsi dll) pada suatu periode waktu tertentu (hari, minggu, bulan, tahun, kuartal).

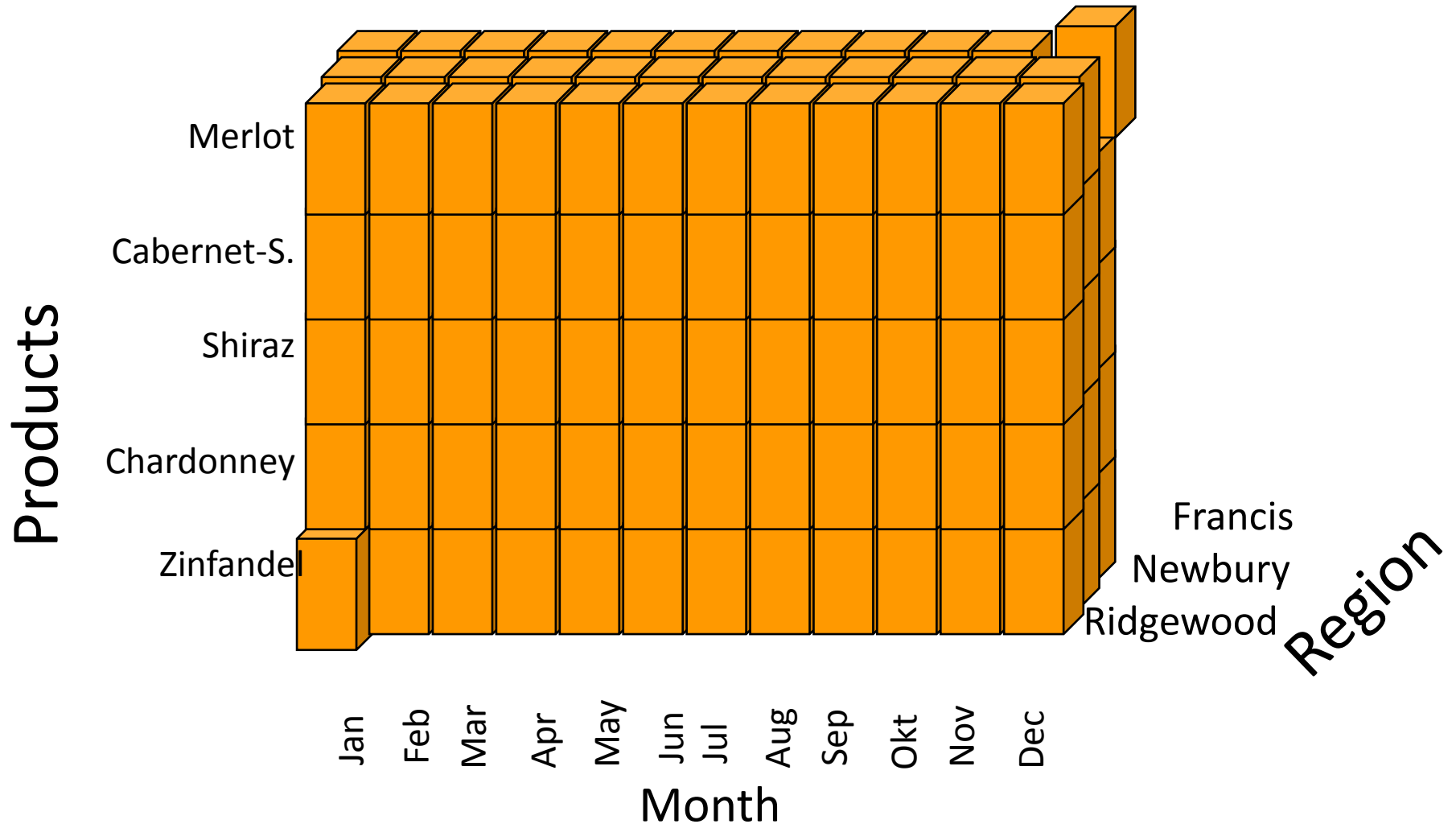
Cube



Cube

- **Representasi kumpulan data** dalam multi-dimensi
- Meskipun dinamakan cube, namun dapat merepresentasikan data dalam **N-dimensi**
- Dengan *cube* data menjadi lebih mudah untuk dimanipulasi.

Cube Representation



Cube dengan N-Dimension

- 0 dimensi:

- Total penjualan sampai sekarang

Total	XXX
--------------	------------

- 1 dimensi:

- Total penjualan untuk waktu tertentu

Waktu 1	AAA
Waktu 2	BBB
Waktu 3	CCC

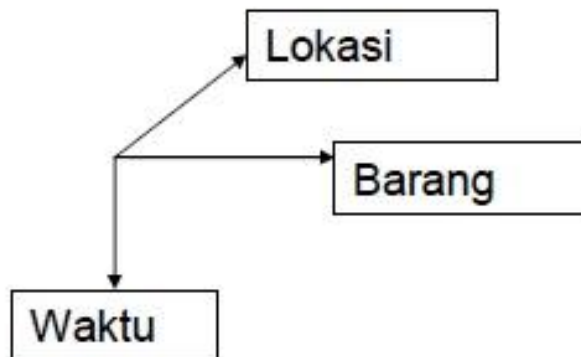
- 2 dimensi:

- Total penjualan pada waktu tertentu untuk barang tertentu

	Barang 1	Barang 2
Waktu 1	AAA	DDD
Waktu 2	BBB	EEE
Waktu 3	CCC	FFF

Cube 3 Dimensi

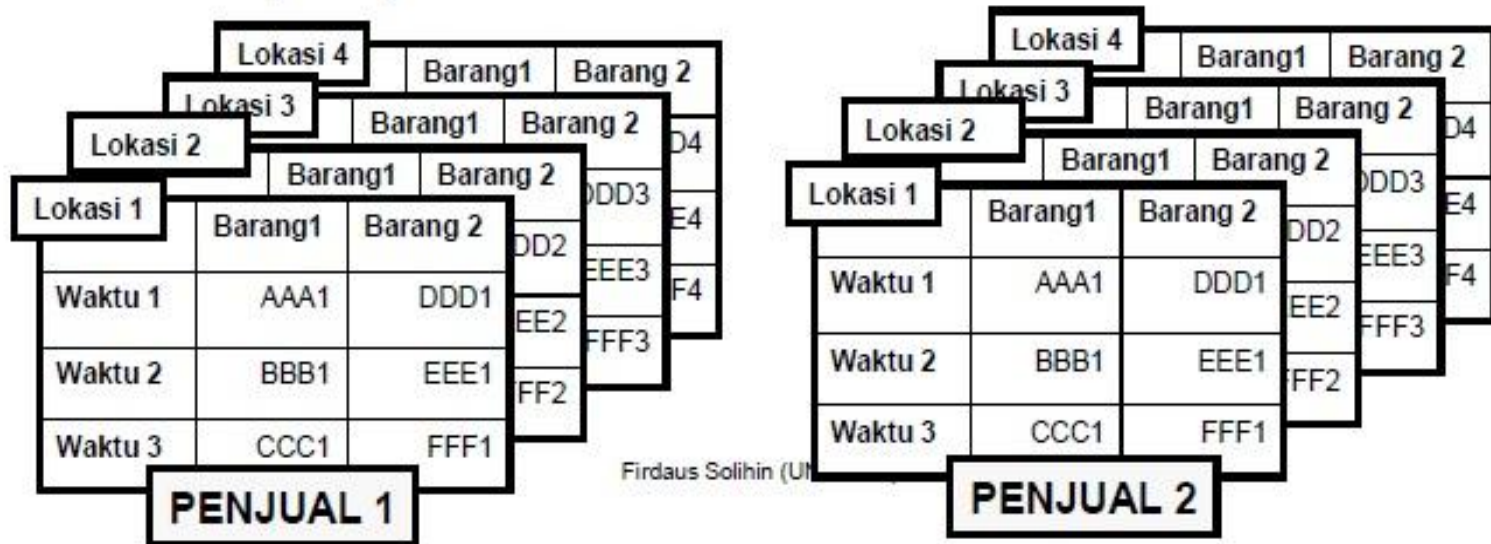
- 3 dimensi:
 - Total penjualan pada **waktu** tertentu untuk **barang** pada **lokasi** tertentu



Lokasi 4			Barang 1	Barang 2
Lokasi 3			Barang 1	Barang 2
Lokasi 2			Barang 1	Barang 2
Lokasi 1			Barang 1	Barang 2
Waktu 1	AAA1	DDD1	DDD2	EEE3
Waktu 2	BBB1	EEE1	FFF2	FFF3
Waktu 3	CCC1	FFF1	FFF2	FFF3

Cube 4 Dimensi

- 4 dimensi:
 - Total penjualan pada waktu tertentu untuk barang pada lokasi tertentu oleh **penjual** tertentu



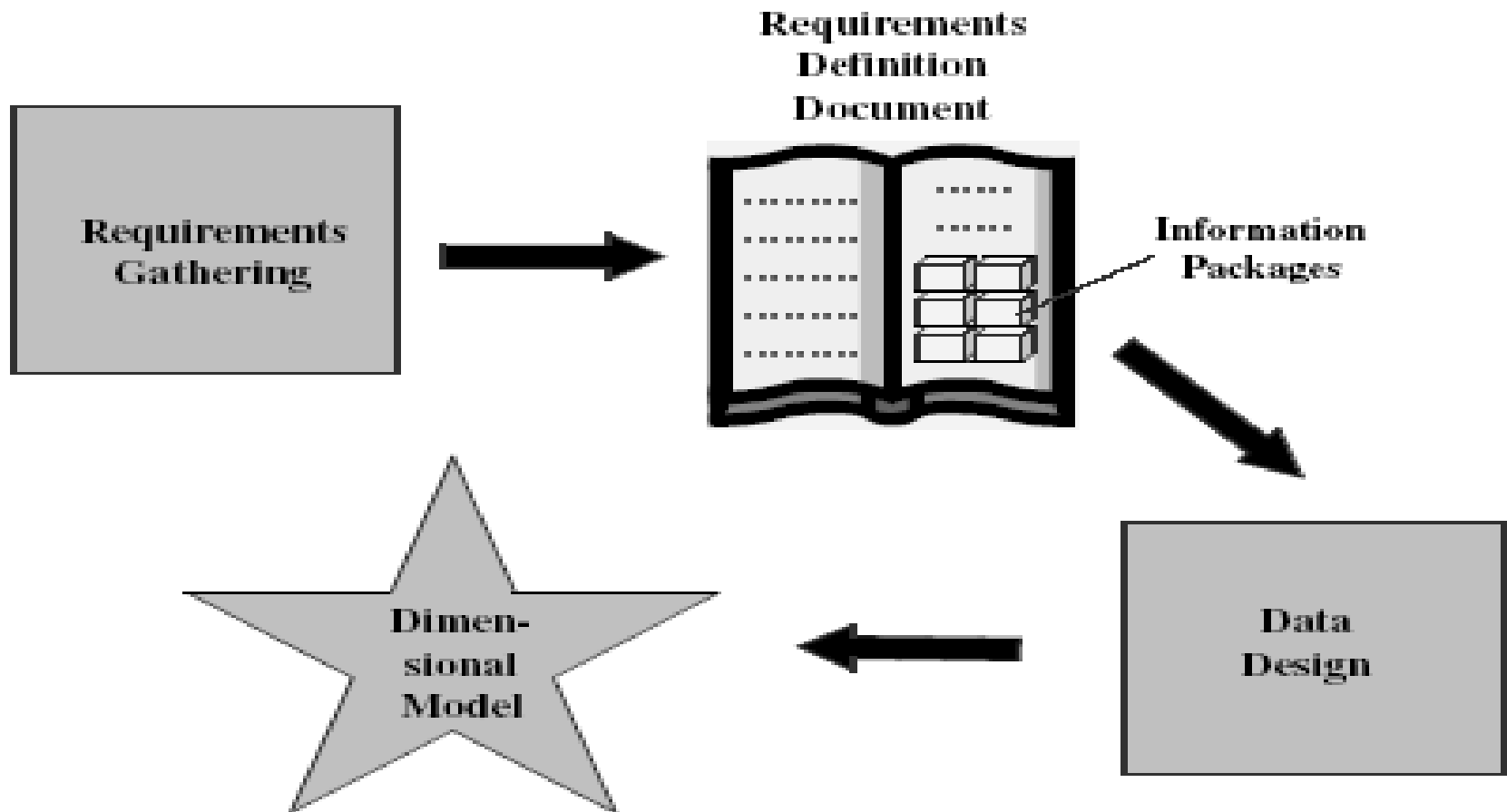
Pemodelan Data Multidimensional

- Menggunakan dua konsep yaitu tabel fakta (fact table) dan tabel dimensi (dimensional table)
- Berbeda dengan konsep normalisasi (3rd normal form)
- Model dimensional menggunakan konsep model hubungan antar entity dengan beberapa batasan yang penting. Setiap model dimensi terdiri dari sebuah tabel dengan sebuah komposit *primary key*, disebut dengan table fakta, dan satu set table yang lebih kecil disebut table dimensi.
- Setiap tabel dimensi memiliki sebuah *simple primary key* yang merespon tepat pada satu komponen *primary key* pada tabel fakta. Dengan kata lain *primary key* pada table fakta terdiri dari dua atau lebih *foreign key*.

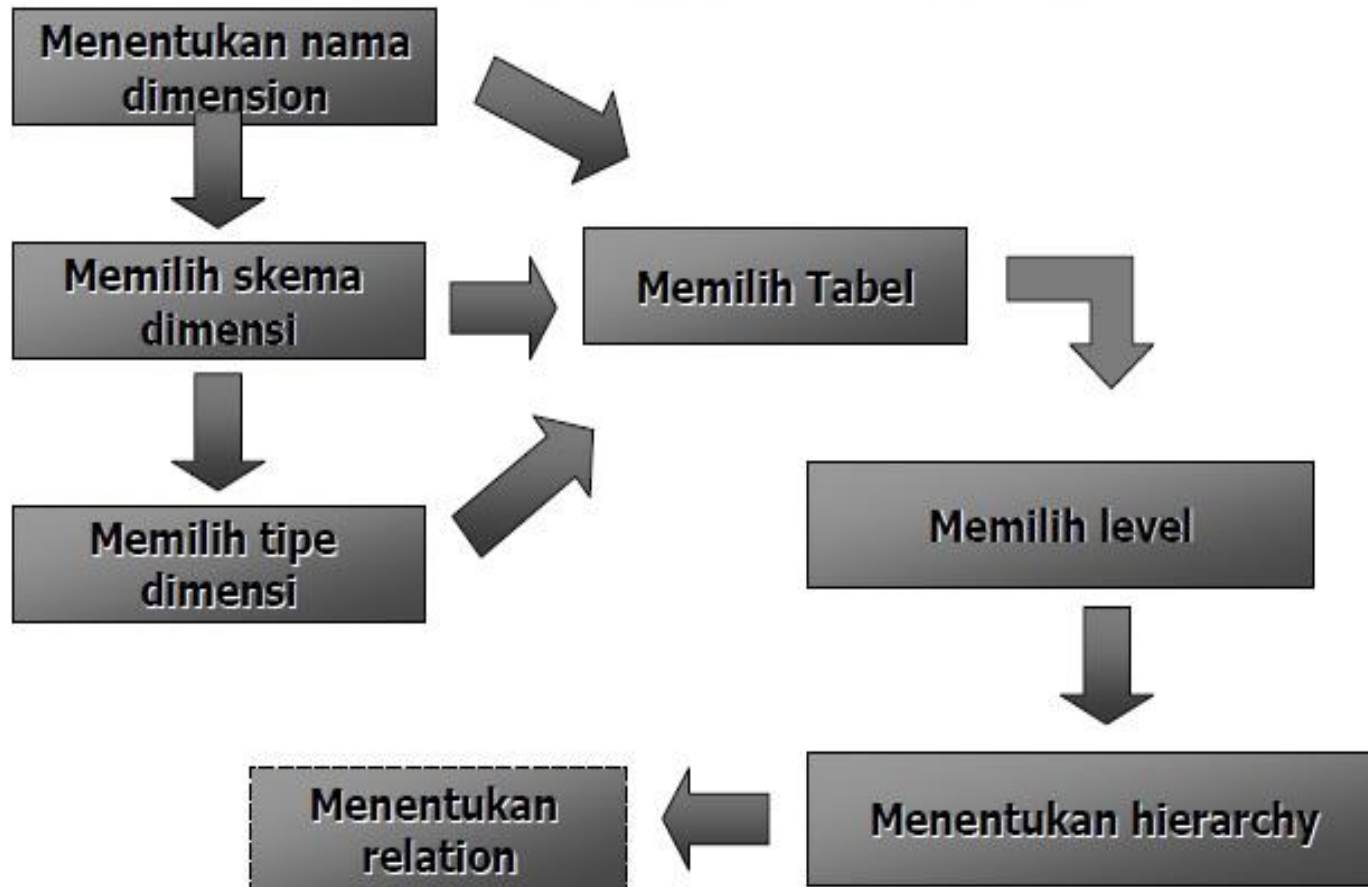
Pemodelan Data Multidimensional

- Fitur terpenting dalam model dimensional ini adalah semua *natural keys* diganti dengan kunci pengganti (*surrogate keys*). Maksudnya yaitu setiap kali join antar tabel fakta dengan tabel dimensi selalu didasari kunci pengganti. Kegunaan dari kunci pengganti adalah memperbolehkan data pada *data warehouse* untuk memiliki beberapa kebebasan dalam penggunaan data, tidak seperti halnya yang diproduksi oleh sistem OLTP.
- Sebuah sistem OLTP memerlukan normalisasi untuk mengurangi redudansi, validasi untuk input data, mendukung volume yang besar dari transaksi yang bergerak sangat cepat. Model OLTP sering terlihat seperti jaring laba-laba yang terdiri atas ratusan bahkan ribuan tabel sehingga sulit untuk dimengerti.

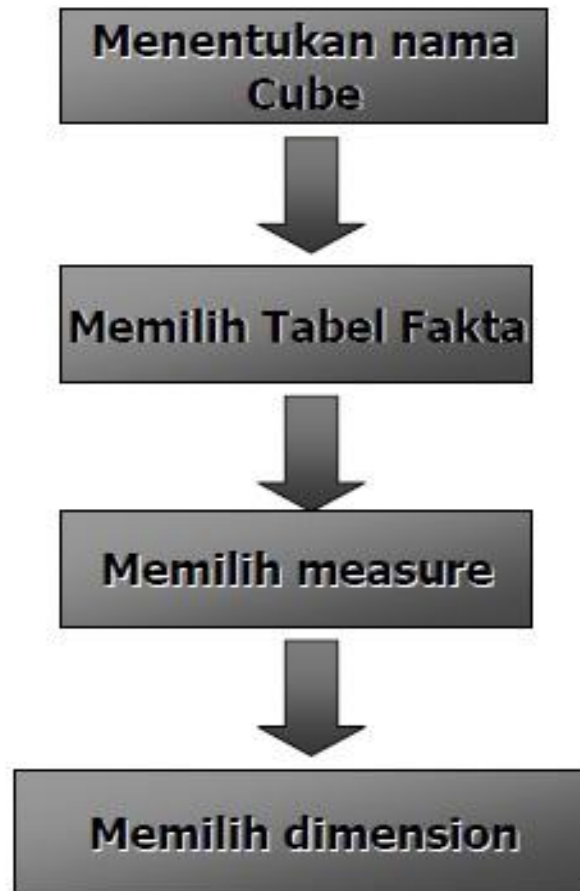
From Requirement (analisa kebutuhan) to Data Design (design data)



Alur Pembuatan Dimension



Alur Pembuatan Cube



Tabel Fakta dan Tabel Dimensi

- Tabel Fakta → berisi measurement atau metric dari proses bisnis dan foreign key dari tabel dimensi.
- Tabel Dimensi → berisi tektual atribut dari measurement yang disimpan pada tabel fakta.
- Tabel dimensi merupakan hierarki, kategori dan logic yang dapat digunakan untuk menganalisa measurement dari sudut pandang tertentu.

Measurement

- Measurement - data kuantitatif

Misal :

Jika anda mempunyai bisnis penjualan sepeda motor maka measurement dari bisnis anda adalah “***jumlah penjualan motor***” atau “***rata-rata penjualan sepeda motor merk x***”

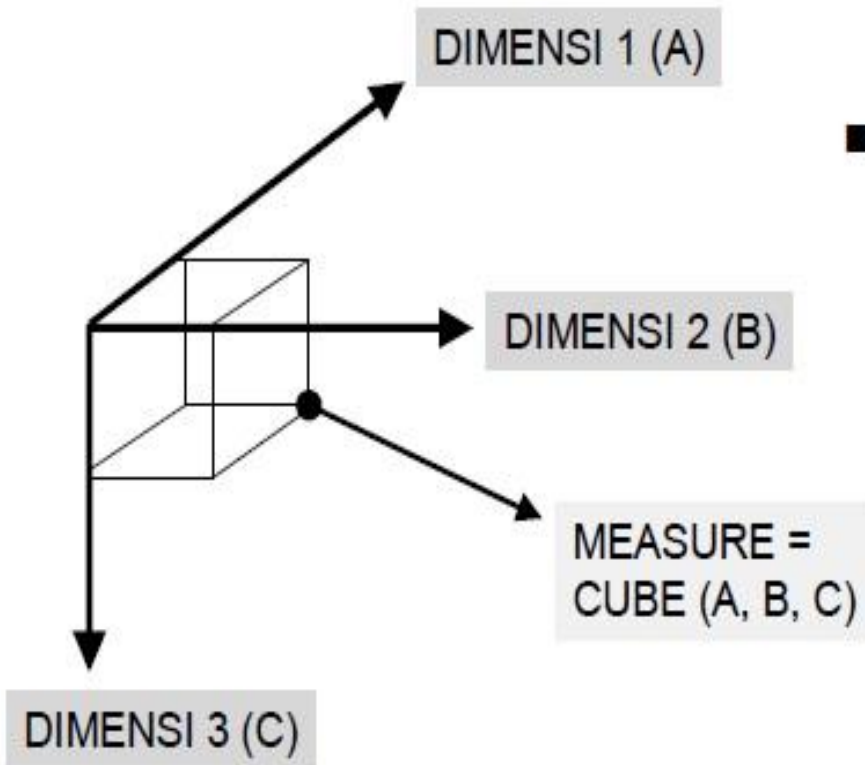
Contoh lain

Jika Proses bisnis anda produksi kertas maka measurementnya adalah “***rata-rata produksi kertas pada suatu mesin***” atau “***produksi kertas setiap bulan***”

Cara Membuat Model Data Multidimensional

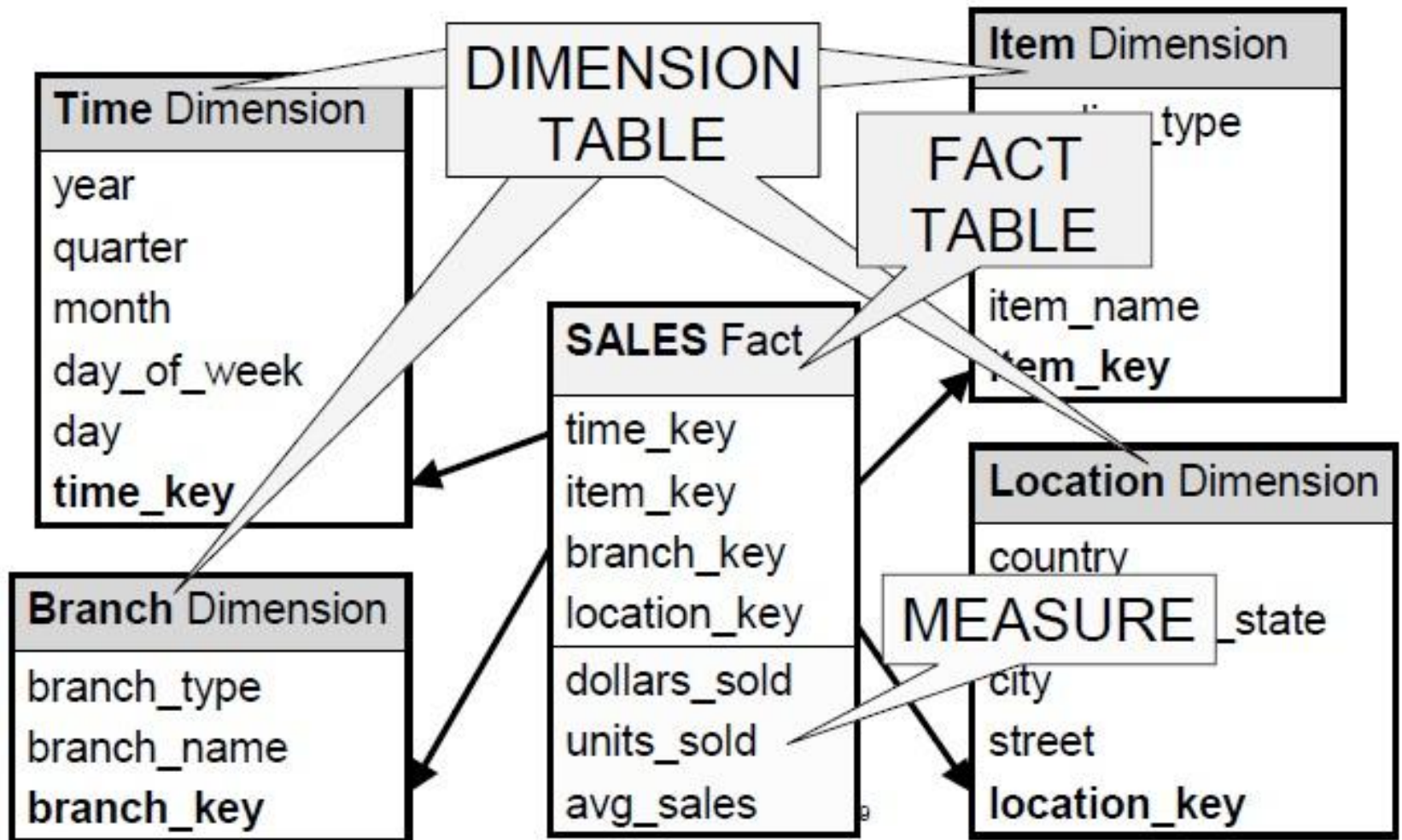
- **Tabel dalam Model Data Multidimensi** dibagi menjadi 2 macam:
 - Fact Table => Measure
 - Dimension Table => Atribut Sumbu
- **Skema Model Data Multidimensi** dibagi menjadi 3:
 - Star Schema
 - Snow Flake Schema
 - Fact Constellation Schema

Data Cube dan Skema Multidimensi

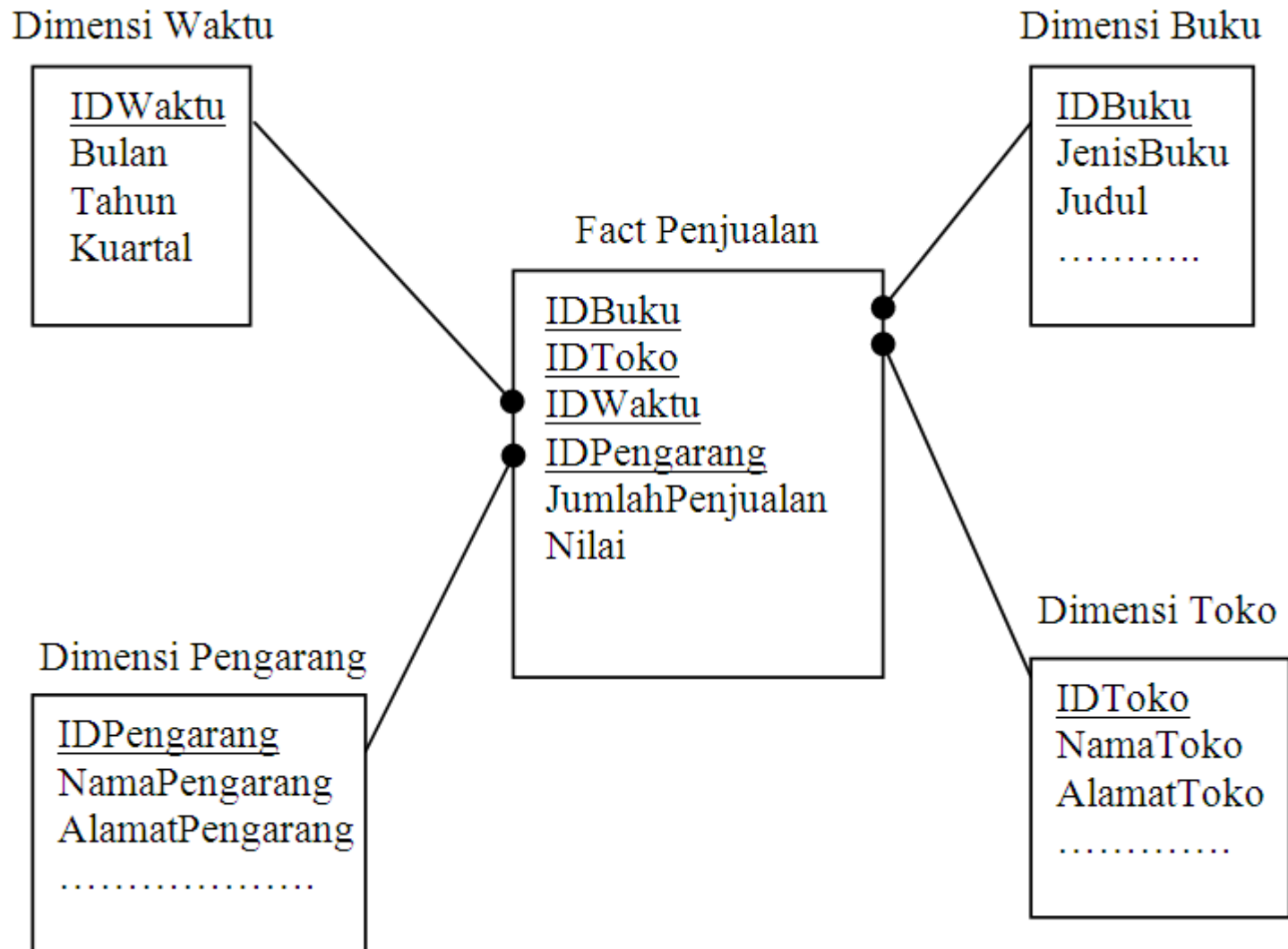


- **Fact Table**
menyimpan Measure
- **Dimension Table**
menyimpan Atribut
Sumbu Cube /
Dimensi
 - Dimensi 1 (A)
 - Dimensi 2 (B)
 - Dimensi 3 (C)

Star Schema



Star Schema



Karakteristik Star Schema

- Pusat dari star disebut fact table
- Fact table mempunyai sebuah nilai aggregate dari data-data yang berasal dari tabel dimensi
- Setiap tabel dimensi berelasi langsung dengan fact table
- Tabel dimensi beisikan data tentang informasi atau waktu
- Relasi antara fact table dengan dimensi-dimensinya adalah 1 – N (one to many)
- Primary key pada tabel dimensi akan menjadi key pada fact table atau dapat dikatakan bahwa fact table memiliki kombinasi key dari tabel dimensi tersebut

Keuntungan dan Kerugian Star Schema

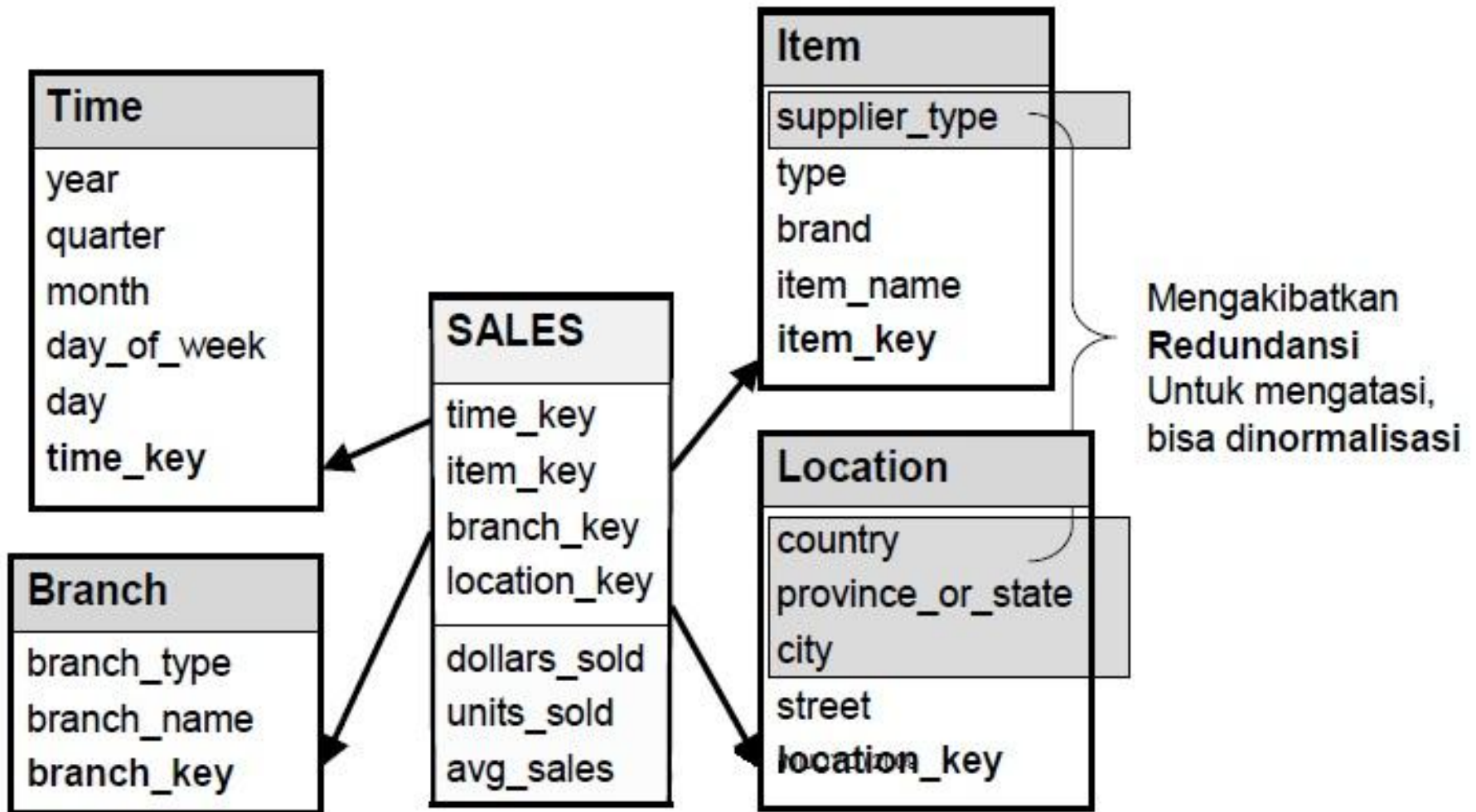
Keuntungan :

- Lebih simple
- Mudah dipahami.
- Hasil dari proses *query* juga relatif lebih cepat.

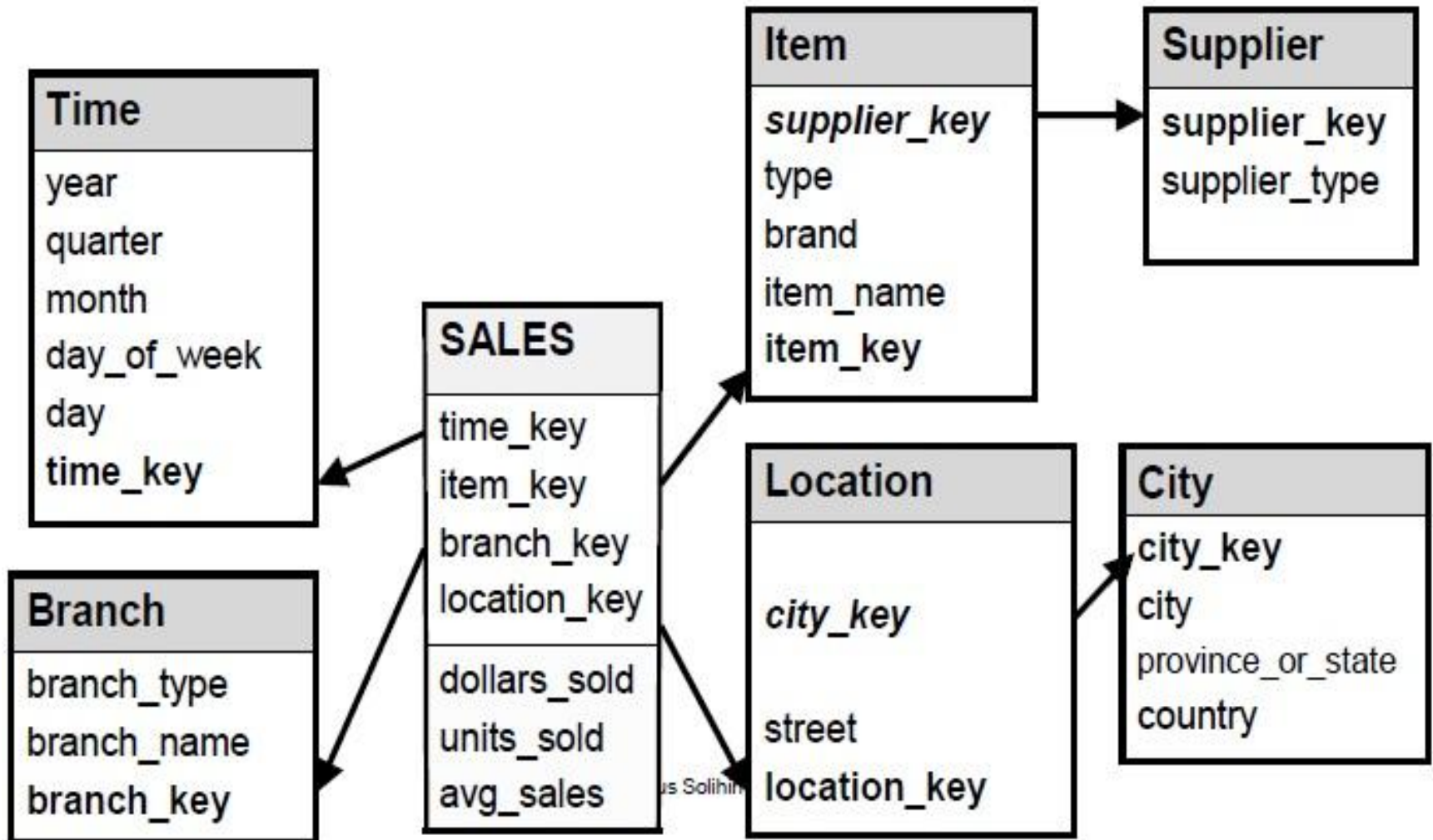
Kerugian :

- boros dalam space.

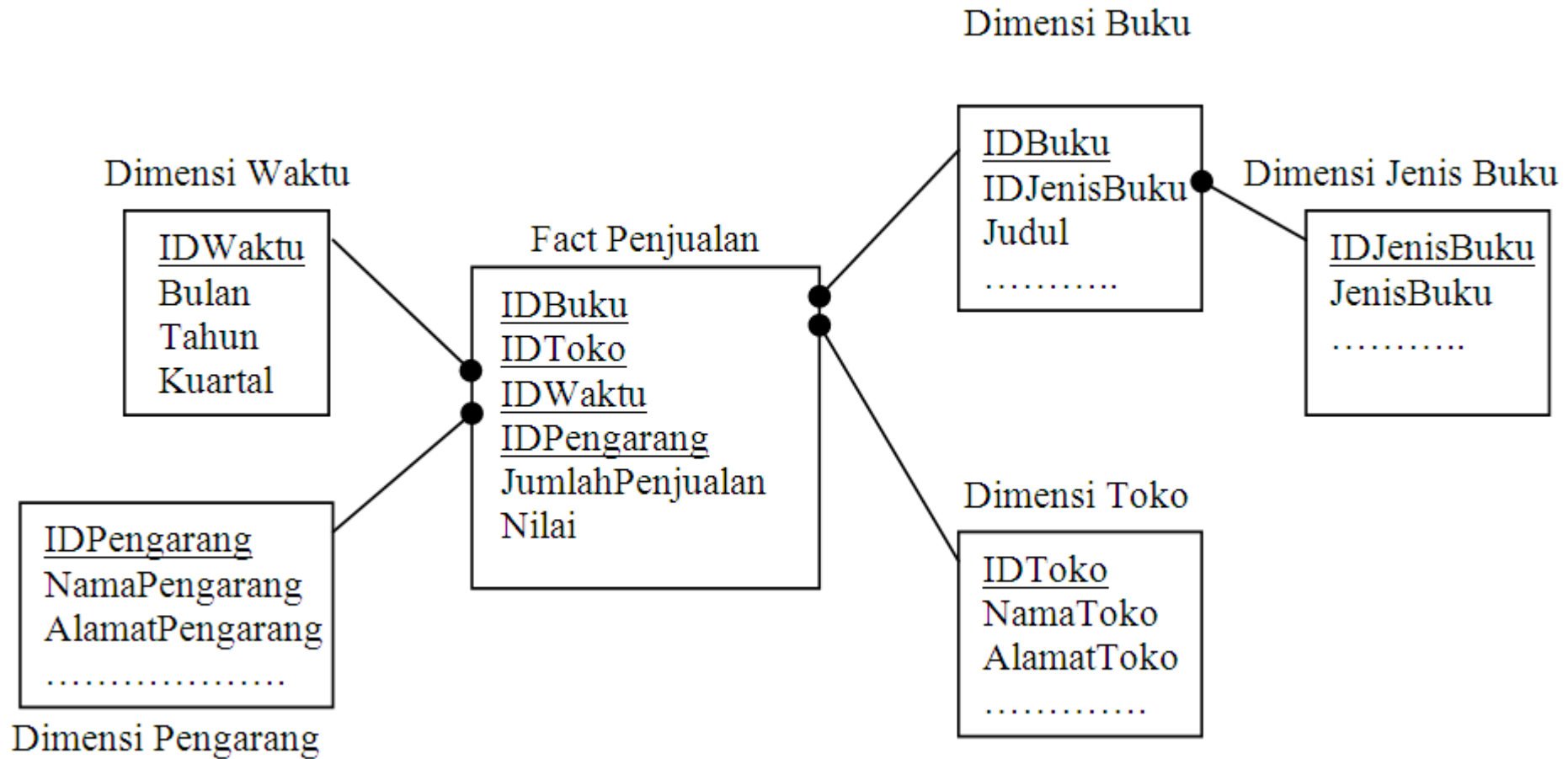
Star Schema



Snowflake Schema



Snowflake Schema



Snowflake Schema

- Model snowflake merupakan perluasan dari star dimana ia juga mempunyai satu atau lebih dimensi. Hanya saja pada snowflake, tabel yang berelasi pada fact table hanya tabel dimensi utama, sedangkan tabel yang lain dihubungkan pada tabel dimensi utama.
- Model snowflake ini hampir sama seperti teknik normalisasi

Keuntungan dan Kerugian Snowflake Schema

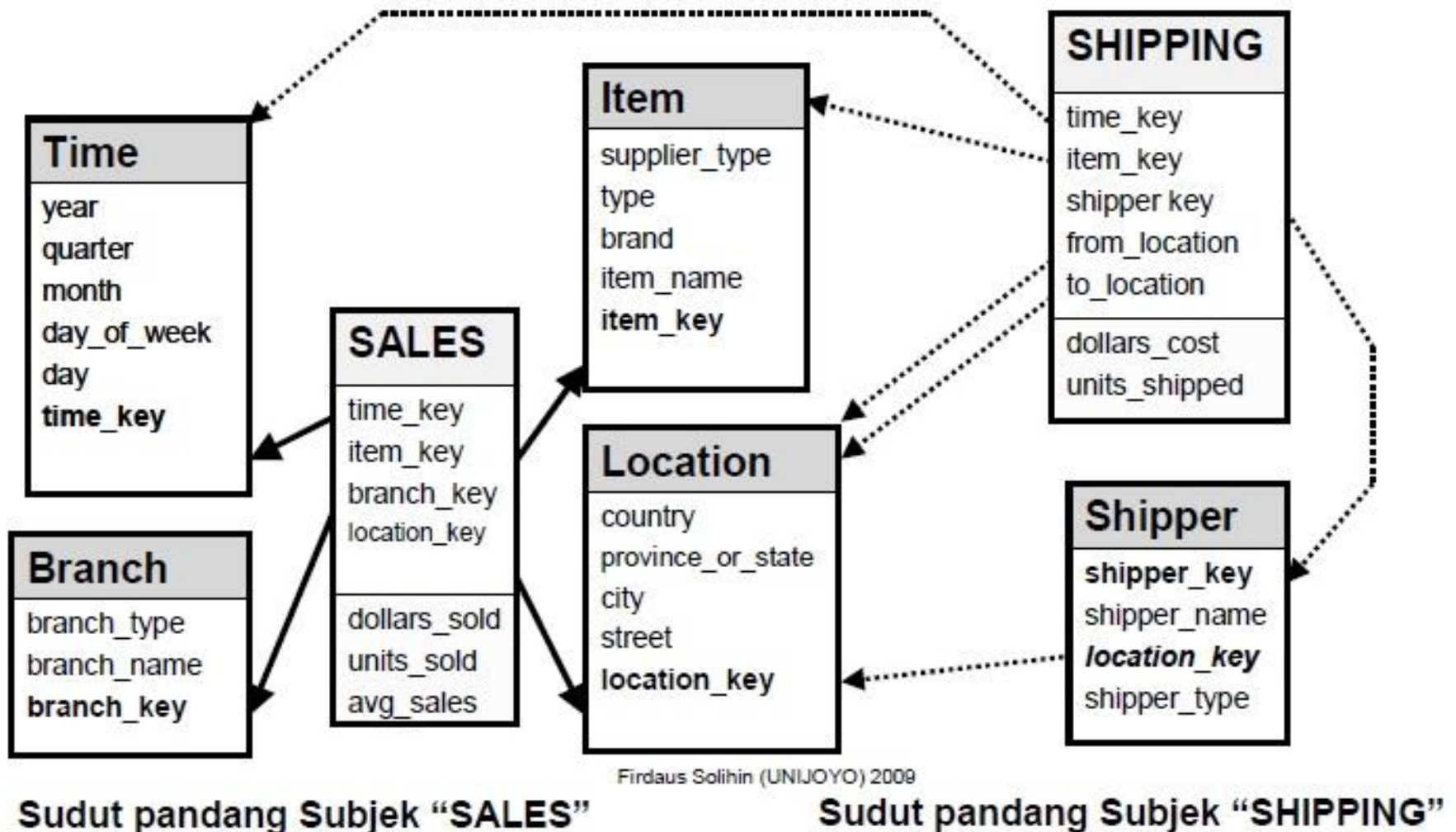
Keuntungan menggunakan model Snowflake

- Pemakaian *space* yang lebih sedikit
- *Update* dan *maintenance* yang lebih mudah

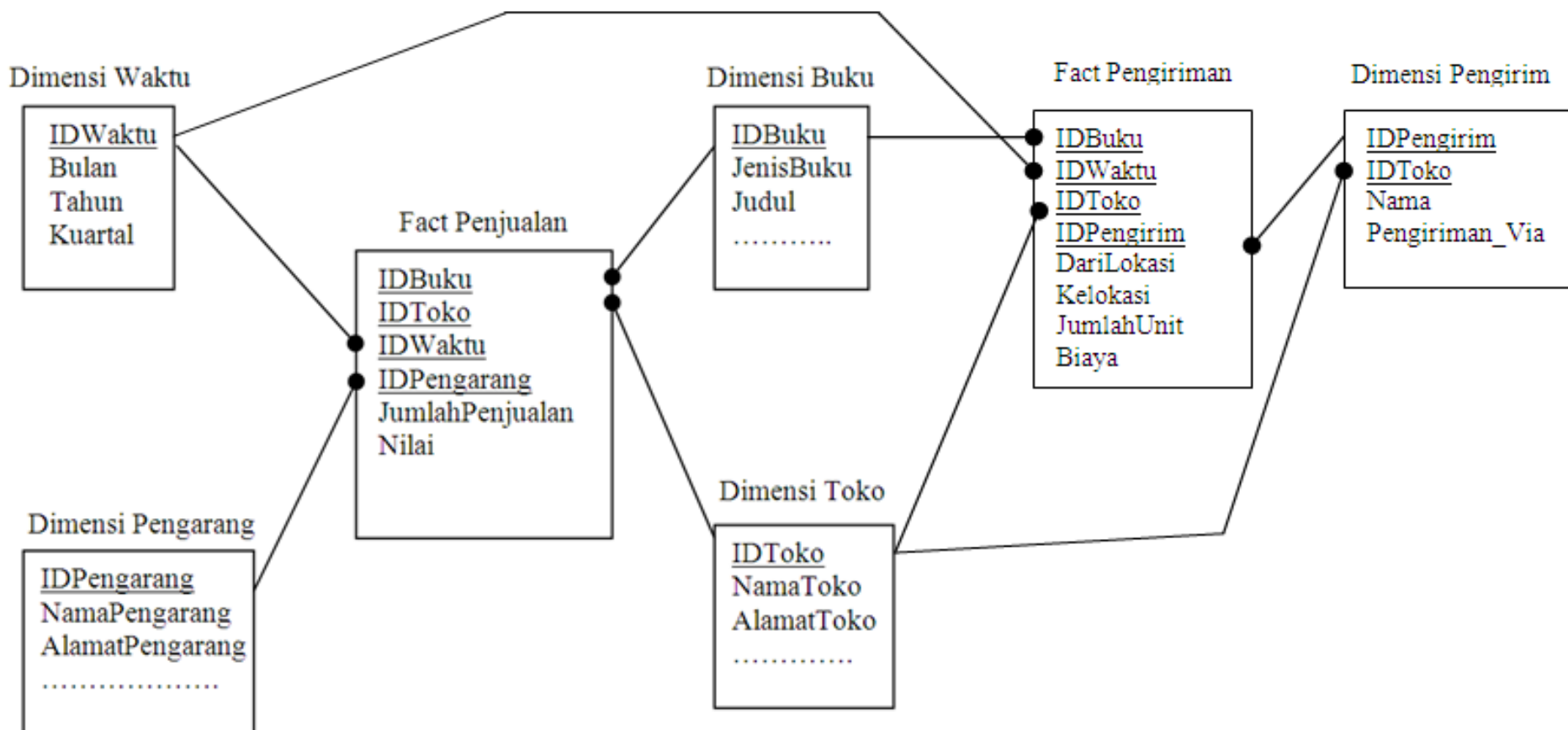
Kerugian menggunakan model ini yaitu :

- Model lebih komplek dan rumit
- Proses *query* lebih lambat
- *Performance* yang kurang bagus

Fact Constellation



Fact Constellation



Fact Constellation Schema

- Berbagi tabel dimensi
- Dapat dilihat sebagai kumpulan beberapa star schema
- Untuk Enterprise-wide
- Dikenal juga dengan istilah galaxy schema