



[www.esaunggul.ac.id](http://www.esaunggul.ac.id)

**DATA WAREHOUSE  
PERTEMUAN - 4  
NOVIANDI**

**PRODI MIK | FAKULTAS ILMU-ILMU KESEHATAN**

# KEMAMPUAN AKHIR YANG DIHARAPKAN

Mahasiswa mampu menjelaskan konsep data warehouse dan OLAP

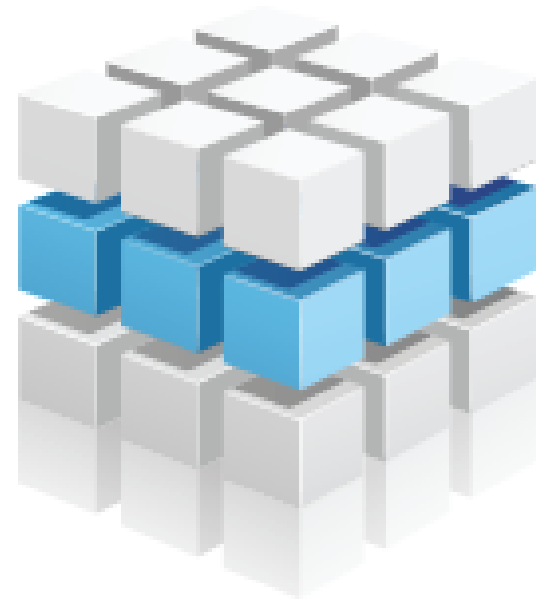


<https://www.bigdata.cam.ac.uk/images/g12.png/@@images/566a1269-d82b-484f-ae92-8c731b0db2ea.png>

# DATA WAREHOUSE

Data warehouse adalah koleksi dari data yang *subject-oriented*, terintegrasi, *time-varian*, dan *nonvolatile*, dalam mendukung proses pembuatan keputusan.

Sering diintegrasikan dengan berbagai sistem aplikasi untuk mendukung pemrosesan dan analisis data dengan menyediakan platform untuk *historical* data.



<http://www.visualmetrics.co.uk/wp-content/uploads/data-warehouse.png>

# KRITERIA DATA WAREHOUSE

## ***SUBJECT ORIENTED***

Diorganisasikan seputar subjek utama seperti customer, produk, sales. Fokus pada pemodelan data dan analisis data untuk pembuatan keputusan

## ***NON-VOLATILE***

Penyimpanan data yang terpisah secara fisik yang ditransformasikan dari lingkungan operasional. Data warehouse tidak memerlukan pemrosesan transaksi

## **TERINTEGRASI**

Mengintegrasikan banyak sumber data yang heterogen seperti: relational database, flatfile, olap. Teknik yang digunakan adalah data *cleaning* dan integrasi data

## **TIME-VARIANT**

Data disimpan untuk menyediakan informasi 5-10 tahun yang lalu. Data mengandung elemen waktu baik eksplisit maupun implisit

# OLAP (*on-line analitical processing*)

- OLAP adalah operasi basis data untuk mendapatkan data dalam bentuk kesimpulan dengan menggunakan agregasi sebagai mekanisme utama.
- OLAP berbeda dengan OLTP. OLTP ( on-line transaction processing) merupakan operasi harian yang dilakukan dalam sistem seperti: pembayaran, registrasi, inventori, dsb

# OLTP vs. OLAP

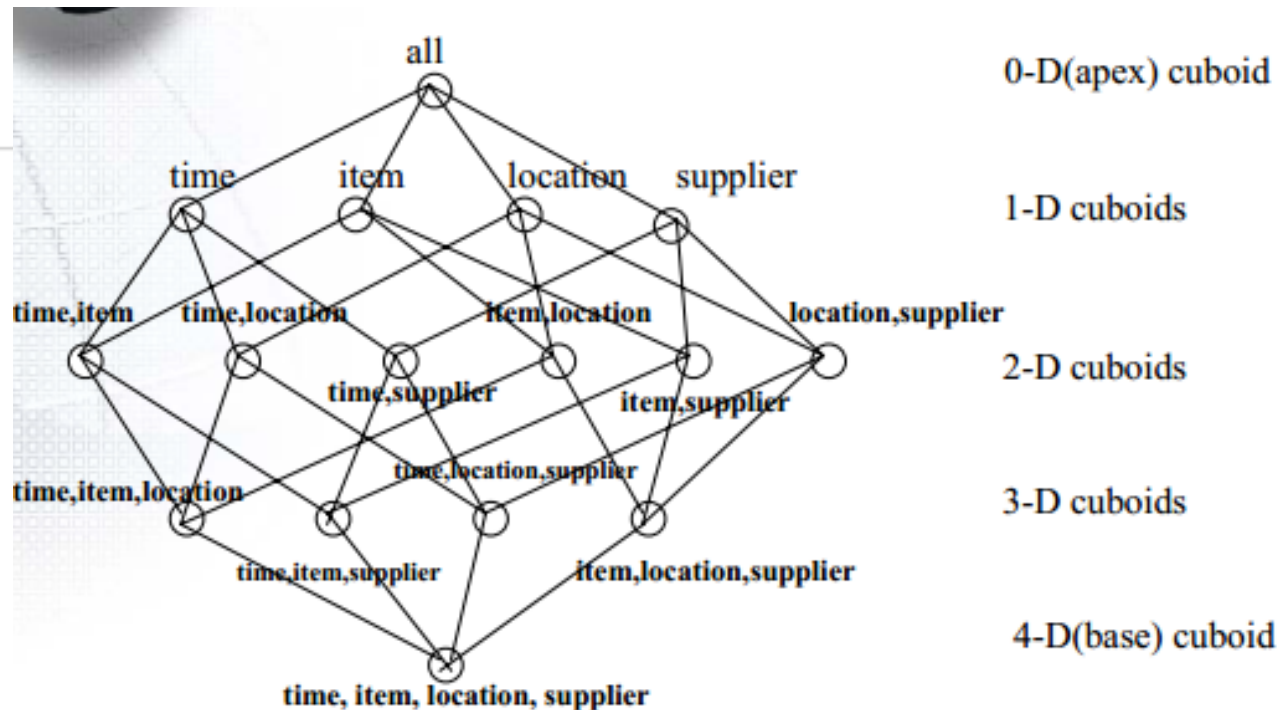
	OLTP	OLAP
<b>users</b>	clerk, IT professional	knowledge worker
<b>function</b>	day to day operations	decision support
<b>DB design</b>	application-oriented	subject-oriented
<b>data</b>	current, up-to-date detailed, flat relational isolated	historical, summarized, multidimensional integrated, consolidated
<b>usage</b>	repetitive	ad-hoc
<b>access</b>	read/write index/hash on prim. key	lots of scans
<b>unit of work</b>	short, simple transaction	complex query
<b># records accessed</b>	tens	millions
<b>#users</b>	thousands	hundreds
<b>DB size</b>	100MB-GB	100GB-TB

# KUBUS DATA

- ❑ Data *warehouse* didasarkan pada model data multidimensional, dimana data dipandang dalam bentuk kubus data.
- ❑ Kubus data, seperti sales, memungkinkan data dipandang dan dimodelkan dalam banyak dimensi
  - Tabel dimensi, seperti *item* (*item\_name*, *brand*, *type*), or *time*(*day*, *week*, *month*, *quarter*, *year*).
  - Tabel fakta mengandung *measures* (seperti *dollars\_sold*) dan merupakan kunci untuk setiap tabel-tabel dimensi terkait.
- ❑ n-D *base cube* dinamakan **base cuboid**. 0-D cuboid merupakan cuboid pada level paling tinggi, yang menampung ringkasan data dalam level paling tinggi, dinamakan **apex cuboid**. Lattice dari *cuboid* membentuk sebuah **data cube**.



# A Lattice of Cuboids

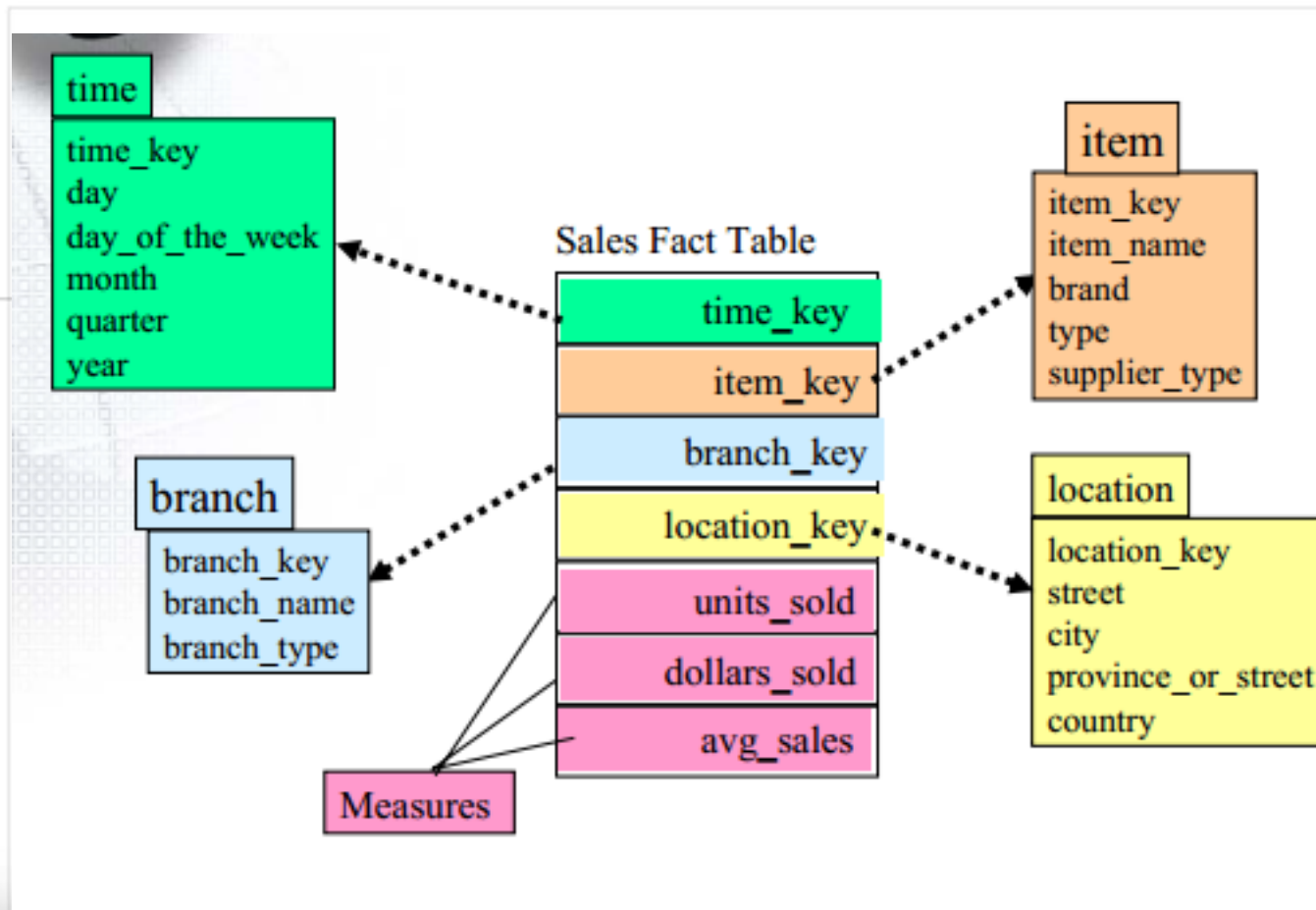




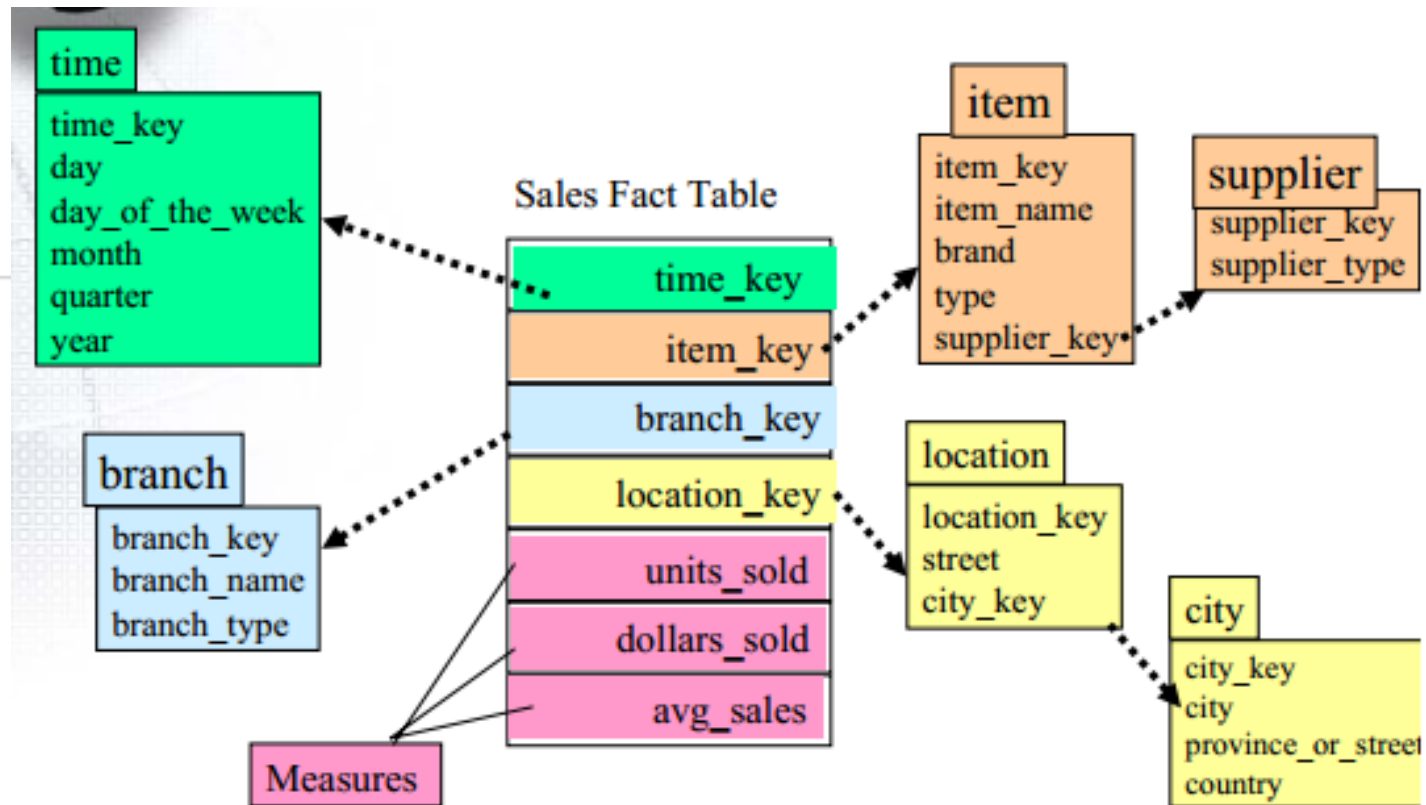
# Pemodelan Konseptual *Data Warehouse*

- **Star schema**: Sebuah tabel fakta di tengah-tengah dihubungkan dengan sekumpulan tabel-tabel dimensi
- **Snowflake schema**: perbaikan dari skema star ketika hirarki dimensional dinormalisasi ke dalam sekumpulan tabel-tabel dimensi yang lebih kecil
- **Fact constellations**: Beberapa tabel fakta dihubungkan ke tabel-tabel dimensi yang sama, dipandang sebagai kumpulan dari skema star, sehingga dinamakan skema galaksi atau *fact constellation*.

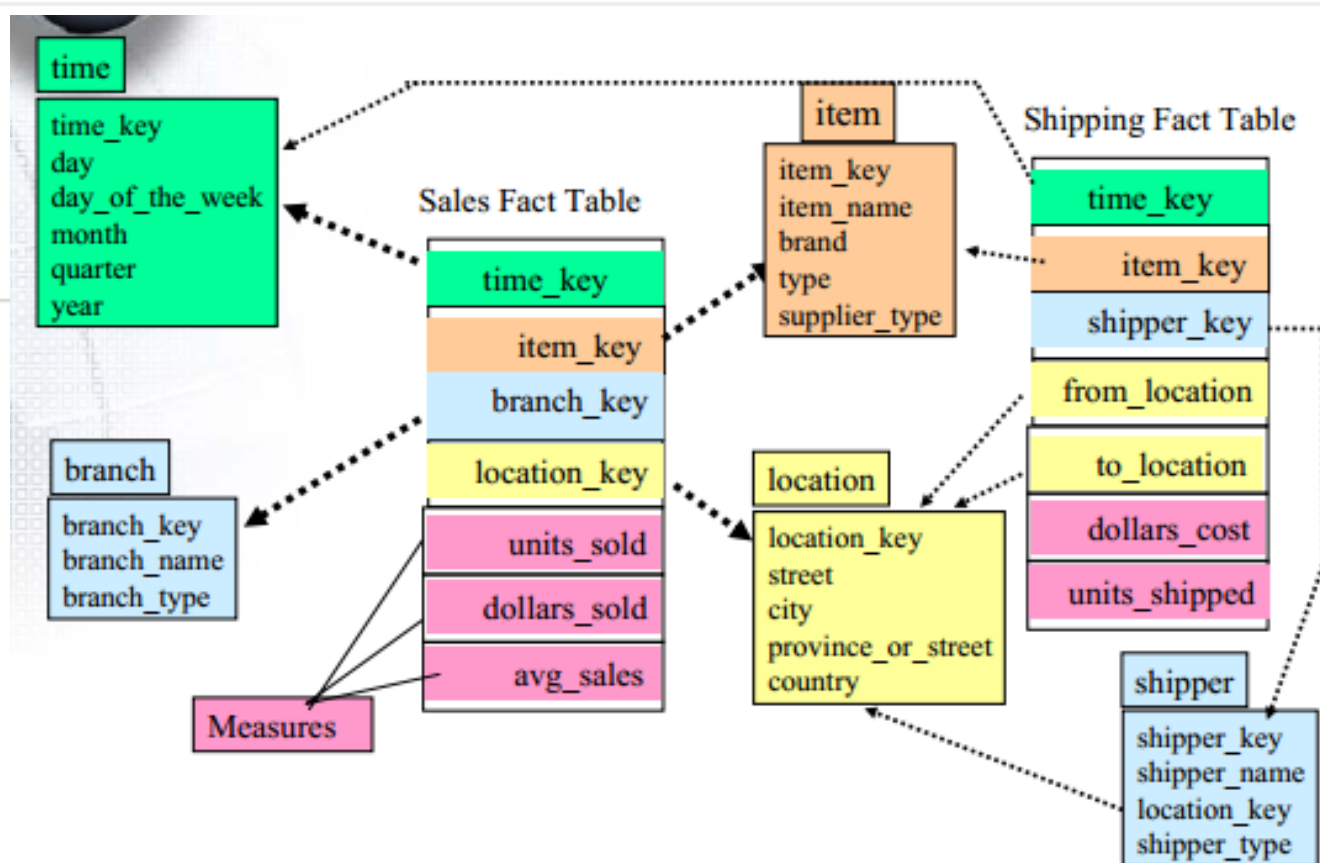
# CONTOH SKEMA STAR



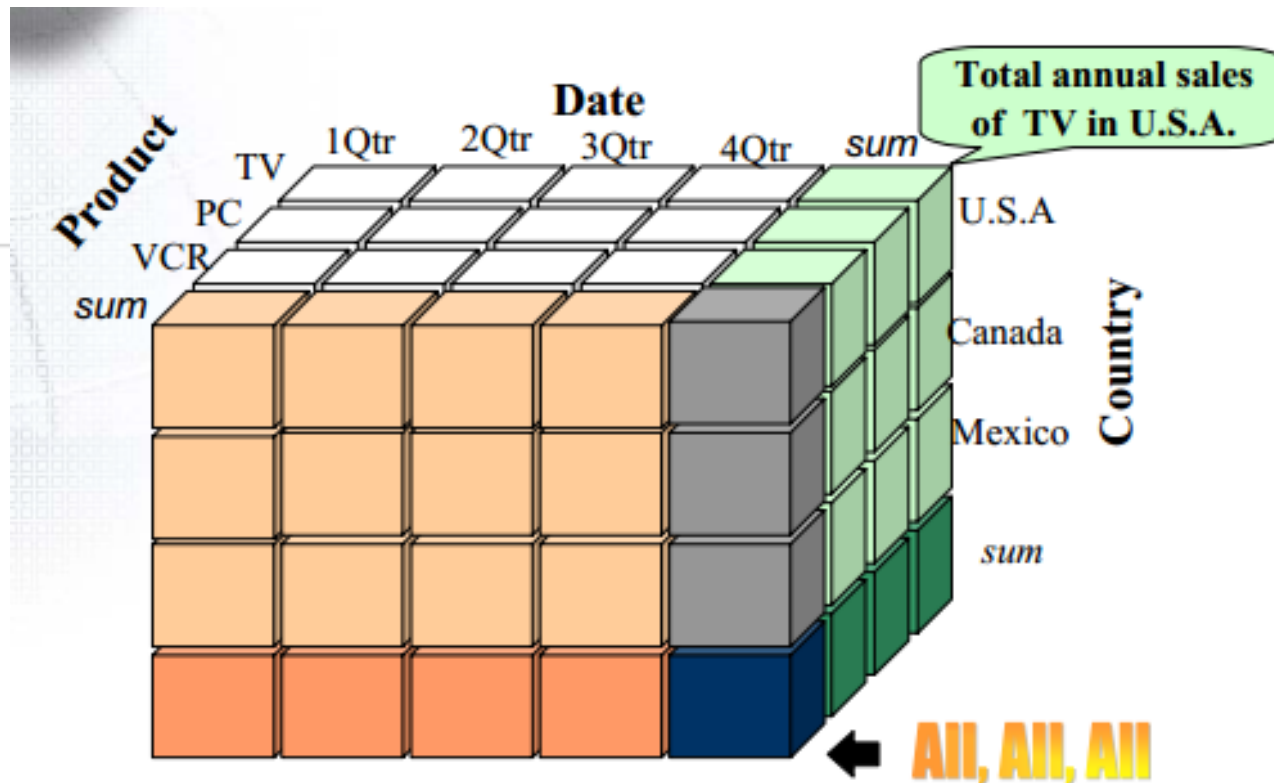
# CONTOH SKEMA SNOWFLAKE



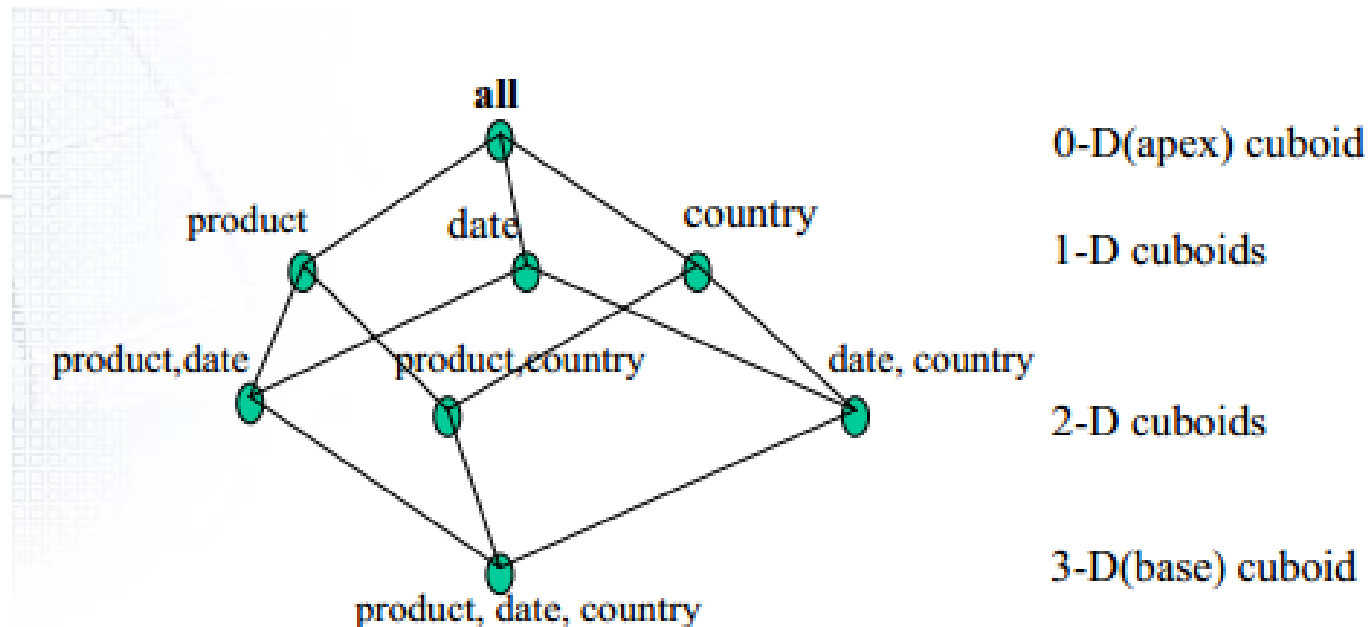
# CONTOH SKEMA GALAKSI



# CONTOH KUBUS DATA

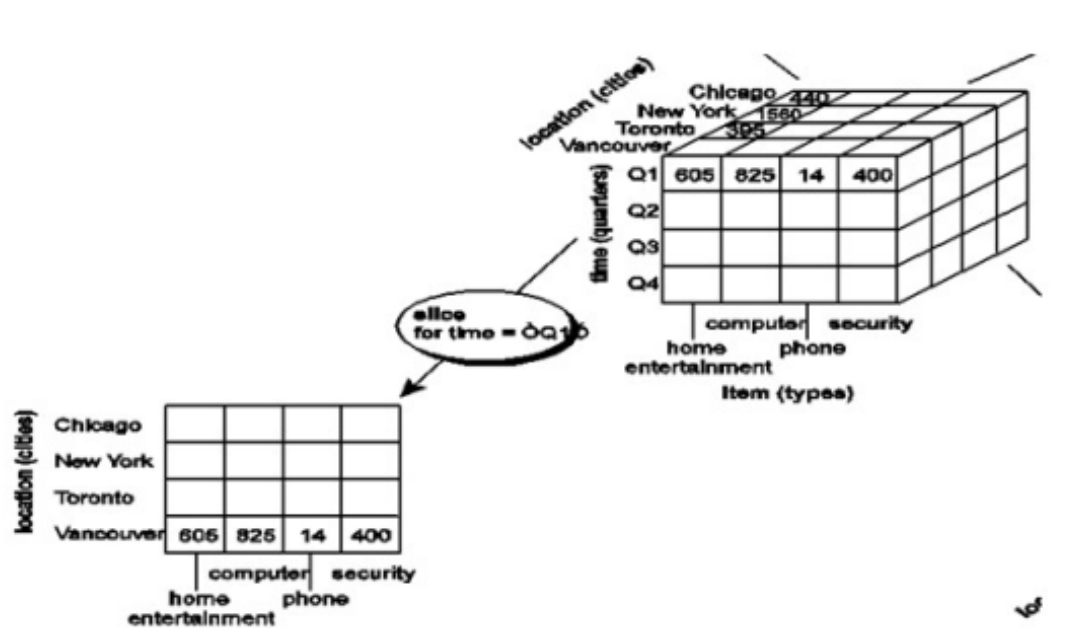


# Cuboid yang terkait dengan kubus



# OPERASI- OPERASI OLAP

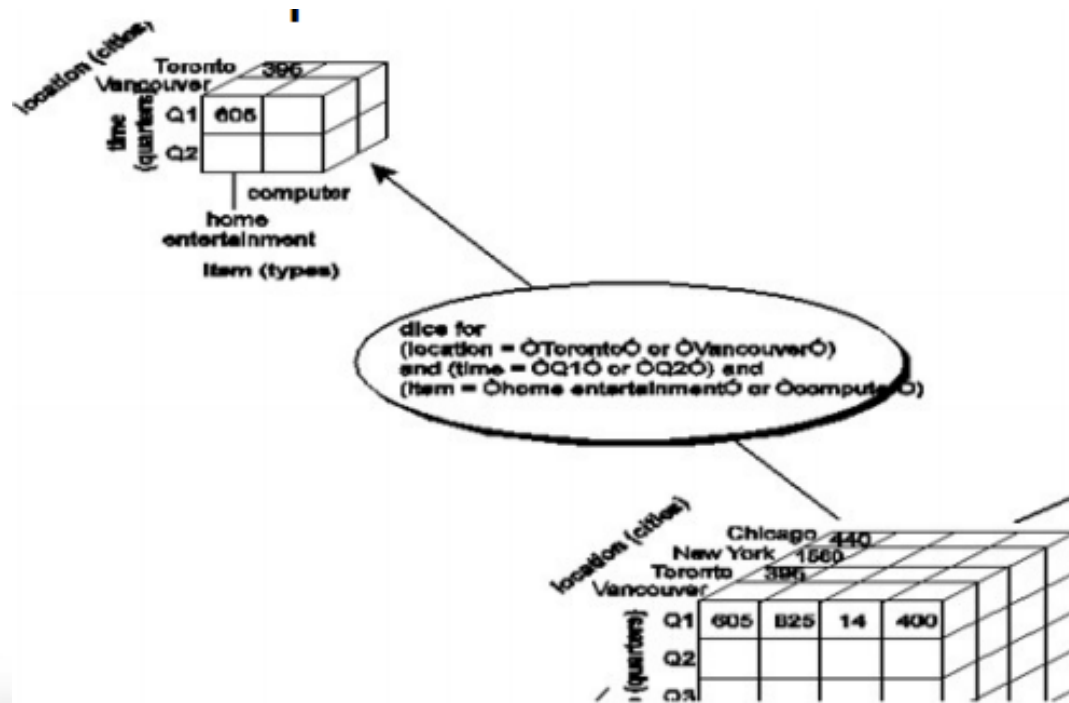
- ❖ **Slicing** adalah seleksi sekumpulan sel dari seluruh array multidimensi dengan menetapkan nilai spesifik untuk satu atau lebih dimensi





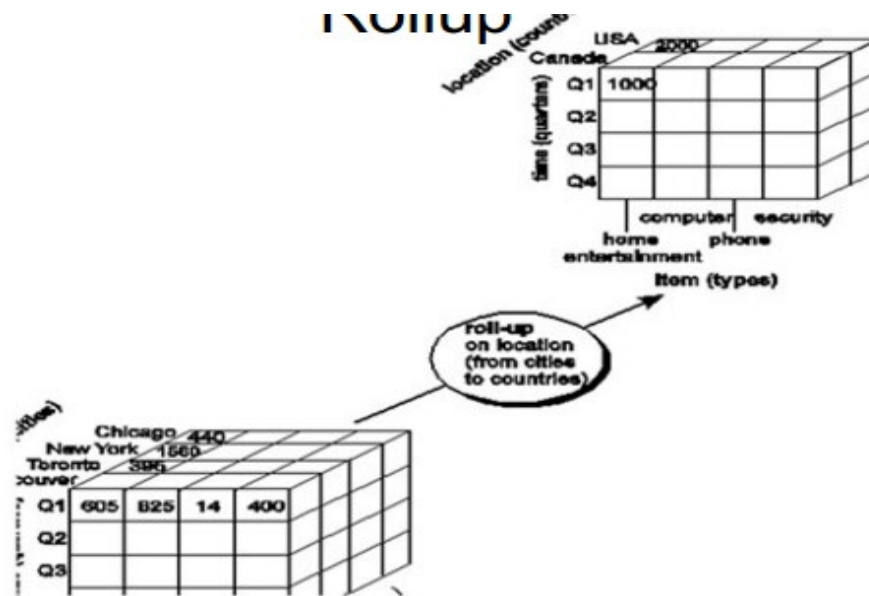
# OPERASI- OPERASI OLAP

- ❖ **Dicing** melibatkan proses seleksi dari subset sel dengan menetapkan interval nilai atribut



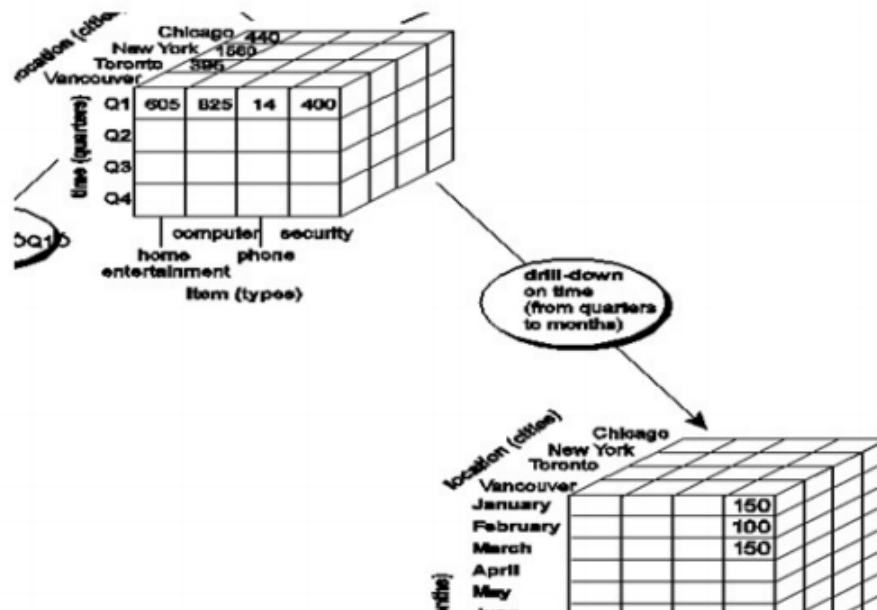
# OPERASI- OPERASI OLAP

- ❖ **Role up** adalah mengumpulkan data-data yang cakupannya kecil hingga menjadi besar. Contoh: Untuk data penjualan, kita dapat mengumpulkan data penjualan dari seluruh tanggal (harian) dalam satu bulan.



# OPERASI- OPERASI OLAP

- ❖ **Drill down** adalah membagi data yang cakupannya besar menjadi lebih kecil. Contoh: Untuk data penjualan, kita dapat membagi total data penjualan bulanan (drill-down) ke dalam total penjualan harian.



**TERIMA KASIH 😊**