

# MATERI : OLAP DAN

## ANALISIS DATA MULTIDIMENSIONAL

---

### OLAP

- Singkatan dari : On line Analytical Processing
- Konsep OLAP pertama diusulkan oleh E.F Codd, bapak dari basis data relasional
- Pada basis data relasional data dimasukkan ke dalam tabel, sementara pada OLAP digunakan representasi array multidimensional



# MEMBUAT SEBUAH ARRAY MULTIDIMENSI

---

2 Langkah dalam mengkonversi data tabular ke dalam array multidimensi

**Pertama** , identifikasi atribut-atribut mana yang merupakan atribut dimensi, dan atribut mana yang merupakan atribut target/tujuan yang nilainya muncul sebagai entri dari array multidimensi :

a. atribut-atribut yang digunakan sebagai dimensi adalah nilai diskrit



## MEMBUAT SEBUAH ARRAY MULTIDIMENSI (lanjutan)

---

- b. Nilai target/tujuan biasanya adalah nilai kontinu, sebagai contoh biaya dari sebuah item
- c. Tidak memiliki nilai target/tujuan sama sekali kecuali jumlah dari objeknya yang memiliki kumpulan yang sama dengan nilai atribut

**Kedua**, cari nilai dari setiap entri dalam array multidimensi dengan menjumlahkan nilai-nilai (dari atribut tujuan) atau dengan menghitung seluruh objek yang memiliki nilai atribut yang berkorespondensi dengan entri tersebut

## CONTOH : DATA BUNGA IRIS

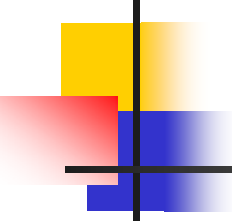


Melalui data-data di bawah ini, akan dibahas bagaimana atribut-atribut, panjang mahkota, lebar mahkota dan spesies dapat dikonversi menjadi array multidimensi

Pertama, kita bedakan antara lebar mahkota dengan panjangnya sehingga didapatkan kategori : *low, medium and high*

Akan diperoleh tabel berikut :

## CONTOH : DATA BUNGA IRIS (lanjutan)

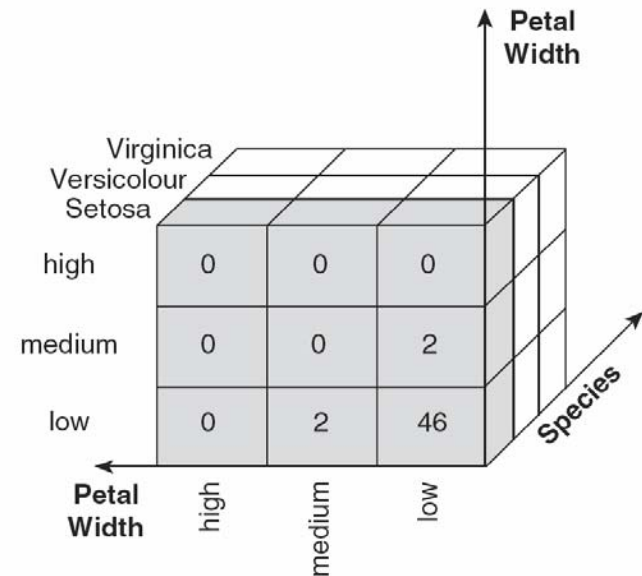


---

Panjang mahkota	Lebar mahkota	Tipe spesies	Jumlah
Low	low	Setosa	46
Low	medium	Setosa	2
Medium	low	Setosa	2
Medium	medium	Versicolour	43
Medium	medium	versicolour	3
Medium	high	virginica	3
High	medium	versicolour	2
High	medium	virginica	3
High	high	versicolour	2
high	high	virginica	44

## CONTOH : DATA BUNGA IRIS (lanjutan)

- Setiap tuple unik dari lebar mahkota, panjang mahkota dan tipe spesies menunjukkan satu elemen dari array
- Elemen ini menunjukkan korespondensi nilai jumlah
- Gambar di samping menjelaskan hal tsb
- Semua tuple yang tidak dispesifikasikan adalah nol



Irisan dari array multidimensi ditunjukkan oleh tabulasi silang berikut ini

Apa sebenarnya yang ingin ditunjukkan oleh tabel berikut ?

		Width		
		low	medium	high
Length	low	46	2	0
	medium	2	0	0
	high	0	0	0

		Width		
		low	medium	high
Length	low	0	0	0
	medium	0	43	3
	high	0	2	2

		Width		
		low	medium	high
Length	low	0	0	0
	medium	0	0	3
	high	0	3	44

## OPERASI DARI OLAP : KUBUS DATA



Operasi kunci dari OLAP adalah formasi dari kubus data

---

Kubus data adalah representasi multidimensi dari data, bersama dengan seluruh jumlah kemungkinan

Dengan seluruh jumlah kemungkinan, yaitu sekumpulan kemungkinan yang dihasilkan dari proses pemilihan subset dari dimensi-dimensi dan menjumlahkan seluruh dimensi yang tersisa

Sebagai contoh, jika kita memilih dimensi tipe spesies dari data Iris dan menjumlahkan seluruh dimensi yang lain, hasilnya adalah masukan/entry satu dimensi dengan tiga masukan/entry, dimana masing-masing akan terdapat jumlah bunga dari masing-masing jenis

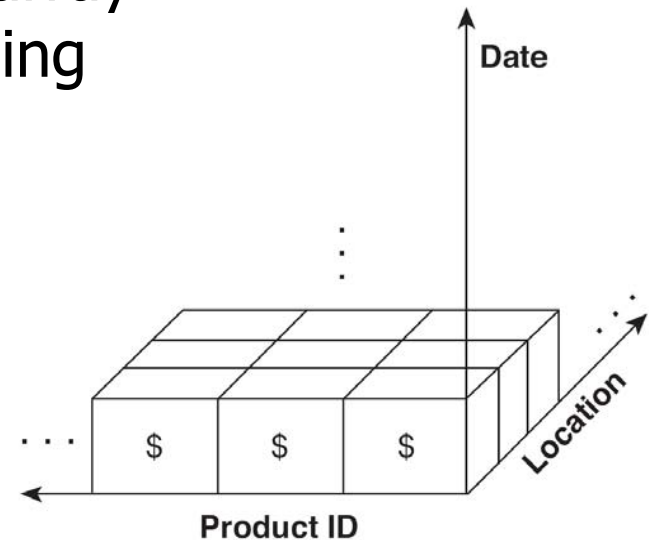


Diumpamakan sekumpulan data yang berisi penjualan produk dari beberapa perusahaan eceran/toko dalam beberapa hari

Data tersebut dapat dinyatakan dalam array 3 dimensi seperti pada gambar di samping

Terdapat 3 agregat two-dimensional (3 choose 2)

3 agregat one-dimensional dan 1 agregat zero-dimensional (total seluruhnya)



## CONTOH KUBUS DATA (lanjutan)

Tabel berikut ini menunjukkan satu dari agregat two-dimensional, bersama dengan agregat one-dimensional dan total seluruhnya

product ID	date				total
	Jan 1, 2004	Jan 2, 2004	...	Dec 31, 2004	
1	\$1,001	\$987	...	\$891	\$370,000
⋮	⋮			⋮	⋮
27	\$10,265	\$10,225	...	\$9,325	\$3,800,020
⋮	⋮			⋮	⋮
total	\$527,362	\$532,953	...	\$631,221	\$227,352,127

OPERASI OLAP :

SLICING DAN DICING

---

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

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

Hal ini ekuivalen dengan mendefinisikan subarray dari array lengkap

Singkatnya, kedua operasi dapat juga disertai dengan pengumpulan beberapa dimensi

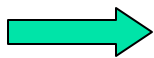
# OPERASI OLAP :

## ROLL -UP DAN DRILL-DOWN

---

Nilai atribut kadang-kadang memiliki struktur hirarki

- Setiap data berasosiasi dengan tahun, bulan dan minggu
- Sebuah lokasi berasosiasi dengan benua, negara, negara bagian (propinsi, dst) dan kota
- Produk dapat dipisahkan dalam beberapa kategori seperti
- pakaian, elektronik dan furniture

Ingat bahwa kategori ini kadang membentuk pohon atau lattice  satu tahun terdiri dari 12 bulan dimana 1 bulan terdiri dari 28, 29, 30 / 31 hari

## OPERASI OLAP :

### ROLL-UP DAN DRILL-DOWN (lanjutan)

---

Struktur hirarki ini akan meningkatkan operasi roll up dan drill down

Untuk data penjualan, kita dapat mengumpulkan data (roll up) penjualan dari seluruh tanggal (harian) dalam satu bulan

Sebaliknya, untuk data yang dimensi waktunya dalam bulanan, kita dapat membagi total data penjualan bulanan (drill-down) ke dalam total penjualan harian

Dengan cara yang sama kita dapat melakukan drill down dan roll up untuk atribut lokasi atau ID produk