

Data Cube Technology

Presentasi Kelompok

Anggota Kelompok

- 5210411132 Ibrahim Al Anshor
- 5200411498 Rabbali Afdhal Yasyfiya
- 5210411106 Ahmad Halim Faizal Zidan

Topik Pembahasan

- Mengenal dan Manfaat Data Cube Technology
- Arsitektur Data Cube : Multidimensional Data Model
- Skema Untuk Multidimensional Data Model
- Hirarki Konsep
- OLAP (Online Analytical Processing)
- Implementasi Nyata Data Cube Technology

Latar Belakang

Data Cube penting dalam dunia teknologi dan bisnis karena memungkinkan analisis data multidimensional yang mendalam, pengambilan keputusan cepat, optimasi operasional, personalisasi produk, penemuan wawasan baru, penanganan Big Data, dan menjaga keamanan data. Data Cube membantu organisasi menjadi lebih efisien, inovatif, dan kompetitif.

MENGENAL DATA CUBE TECHNOLOGY

Apa Itu Data Cube?

Data Cube adalah representasi multidimensional dari data yang memungkinkan penyimpanan, analisis, dan pemahaman yang lebih baik tentang data.

Apa Itu Data Cube

Dalam Data Cube, data diorganisir dalam bentuk tabel tiga dimensi atau lebih, di mana masing-masing dimensi mewakili atribut atau variabel tertentu. Ini memungkinkan pemodelan dan analisis data yang lebih kompleks daripada yang dapat dicapai dengan pendekatan tabel datar.

Apa Itu Data Cube

Data Cube sering digunakan dalam sistem OLAP untuk analisis data interaktif yang efisien. Ini memungkinkan pengguna untuk melakukan operasi seperti slicing, dicing, dan pivoting untuk mendapatkan wawasan dari data yang kompleks.

Cara Kerja Data Cube

Pengumpulan Data,
Pemilihan Dimensi, Mengisi
Data Cube, Pengolahan
dan Agregasi, Analisis
Data, Visualisasi,
Pengambilan Keputusan,
Penyimpanan dan
Pemeliharaan

Konsep dasar dari data multidimensional

Data multidimensional adalah pengorganisasian data dalam bentuk yang mencakup beberapa dimensi atau atribut

Dimensi

Dimensi adalah atribut atau kategori tertentu yang digunakan untuk mengklasifikasikan atau mengelompokkan data.

Misalnya, dalam analisis penjualan, dimensi bisa mencakup waktu (bulan, tahun), produk (jenis produk), dan wilayah geografis.

Fakta

Fakta adalah data aktual yang ingin dianalisis. Ini adalah nilai atau matrix yang diukur atau diamati dalam konteks data multidimensional. Misalnya, fakta dalam analisis penjualan bisa berupa jumlah penjualan atau pendapatan.

Hierarki

Hierarki adalah hubungan hirarkis antara elemen dalam dimensi. Sebagai contoh, dimensi waktu dapat memiliki hierarki yang melibatkan tahun, kuartal, bulan, dan hari.

Operasi OLAP

OLAP (Online Analytical Processing) adalah kumpulan operasi yang digunakan untuk menganalisis data multidimensional. Ini termasuk operasi slicing (memilih satu layer dari Cube), dicing (mengiris Cube dalam dua dimensi), dan pivoting (mengubah orientasi Cube).

MANFAAT DATA CUBE TECHNOLOGY

Bisnis

Data Cube memungkinkan perusahaan untuk menganalisis penjualan berdasarkan berbagai dimensi seperti waktu, produk, dan wilayah, sehingga membantu dalam mengidentifikasi tren penjualan dan peluang pasar.

Analisis Data

Data Cube memungkinkan analis data untuk mengeksplorasi data dengan lebih mendalam, menemukan hubungan yang tersembunyi, dan mengidentifikasi pola yang mungkin tidak terdeteksi dengan data datar.

Pengambilan Keputusan

Data Cube memungkinkan pengambilan keputusan yang cepat karena memungkinkan akses instan ke data yang relevan dan kemampuan untuk melihat data dari berbagai sudut pandang.

ARSITEKTUR DATA CUBE : MULTIDIMENSIONAL DATA MODEL

Multidimensional Data Model

Kita biasanya menganggap kubus sebagai struktur geometri 3 Dimensi, dalam pengolahan data, data cube adalah **n-dimensi**.

Untuk memahami data cube dan model data multidimensional dengan lebih baik, mari kita mulai dengan melihat data cube 2 Dimensi sederhana yang sebenarnya adalah tabel atau spreadsheet untuk data penjualan dari *AllElectronics*.

Data 2D

Data penjualan AllElectronics untuk barang yang terjual per kuartal di kota Vancouver.

Dalam representasi 2-D ini, penjualan untuk Vancouver ditampilkan sehubungan dengan dimensi waktu (diorganisir dalam kuartal) dan dimensi item (diorganisir berdasarkan jenis item yang dijual)

2-D view of Sales Data location = "Vancouver"				
time (quarter)	item (type)			
	home entertainment	computer	phone	security
Q1	605	825	14	400
Q2	680	952	31	512
Q3	812	1023	30	501
Q3	927	1038	38	580

Multidimensional Data Model

Sekarang, anggaplah kita ingin melihat data penjualan dengan dimensi ketiga.

Misalnya, kita ingin melihat data berdasarkan waktu dan item, serta lokasi, untuk kota-kota Chicago, New York, Toronto, dan Vancouver.

Data 3D

Kita dapat menampilkan data n-dimensi apa pun sebagai serangkaian "kubus" (n-1)-dimensi.

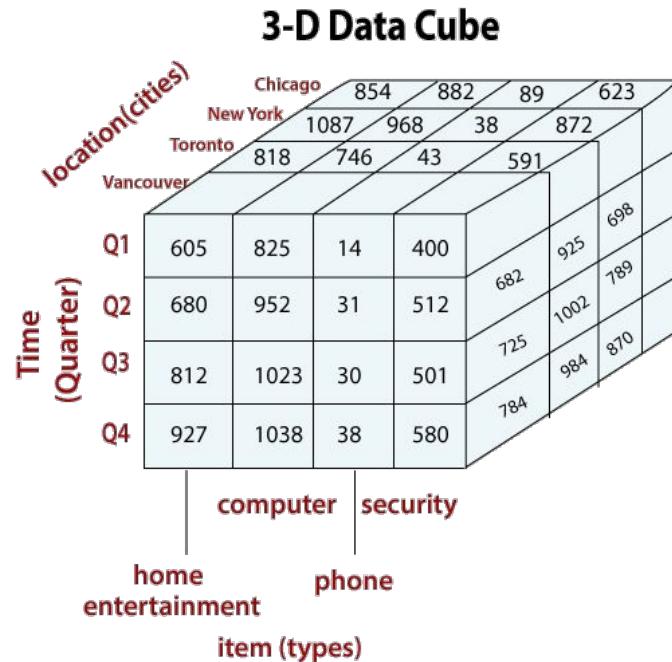
Hal penting yang perlu diingat adalah bahwa data cube adalah n-dimensi dan tidak membatasi data pada 3-D.

3-D view of Sales Data

location ="Chicago"					location ="New York"					location ="Toronto"				
item					item					item				
home					home					home				
time	ent.	comp.	phone	sec.	time	ent.	comp.	phone	sec.	time	ent.	comp.	phone	sec.
Q1	854	882	89	623	1087	968	38	872		818	746	43	591	
Q2	943	890	64	698	1130	1024	41	925		894	769	52	682	
Q3	1032	924	59	789	1034	1048	45	1002		940	795	58	728	
Q4	1129	992	63	870	1142	1091	54	984		978	864	59	784	

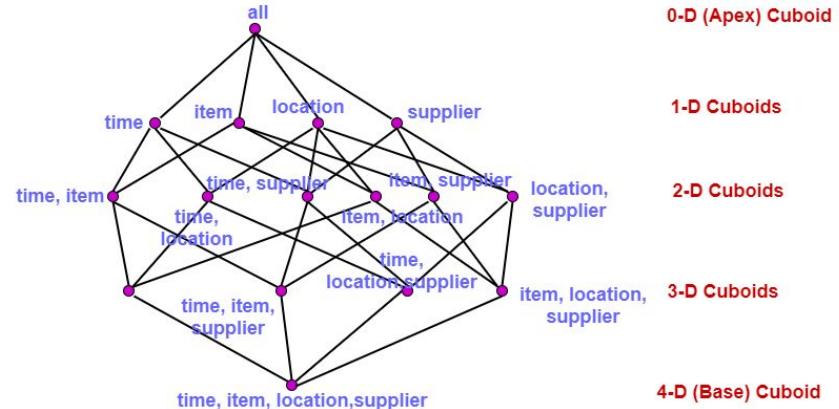
Representasi 3D Data Cube

Secara konseptual, kita juga dapat merepresentasikan data yang sama dalam bentuk data cube 3D.



Lattice Cuboid

Lattice cuboid yang menciptakan data cube 4-D untuk dimensi waktu, item, lokasi, dan pemasok. Setiap cuboid mewakili tingkat ringkasan yang berbeda.



SKEMA UNTUK MULTIDIMENSIONAL DATA MODEL

Skema

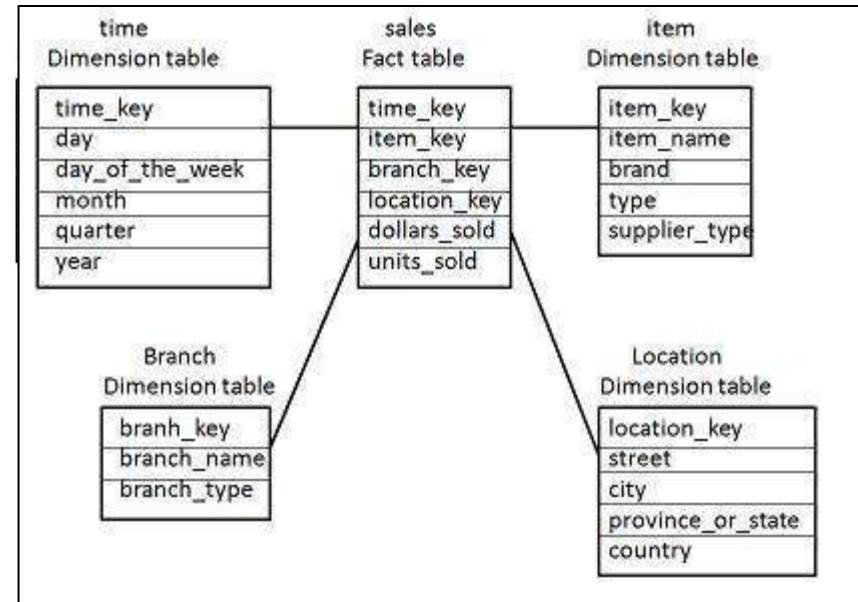
Skema pada data warehouse adalah struktur organisasi data yang menentukan bagaimana data disimpan, diatur, dan dihubungkan. Ini mencakup bagaimana tabel-tabel data dan relasinya dirancang untuk mendukung analisis dan pelaporan bisnis. Skema ini mencakup tabel fakta yang berisi data bisnis inti dan tabel dimensi yang menyediakan konteks terhadap data fakta.

Star Schema

Setiap dimensi direpresentasikan dengan hanya satu tabel dimensi.

Tabel dimensi ini berisi sekumpulan atribut.

Tabel fakta di tengah. Tabel ini berisi kunci untuk masing-masing dari empat dimensi, juga berisi atribut jumlah penjualan dalam dolar dan unit yang terjual.

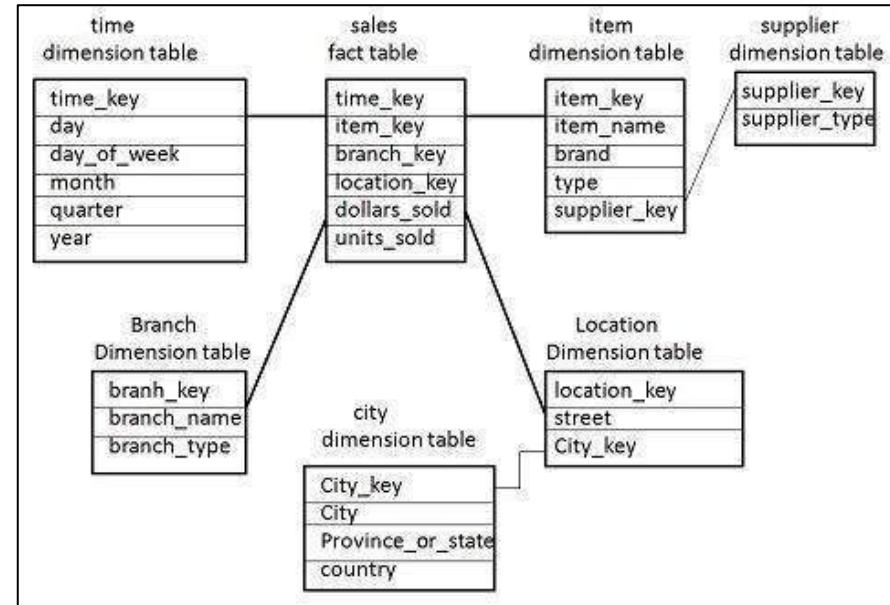


Snowflake Schema

Tabel dimensi dalam bentuk yang dinormalisasi.

Sekarang tabel dimensi item berisi atribut-atribut item_key, item_name, tipe, merek, dan supplier-key.

Supplier key terhubung ke tabel dimensi supplier. Tabel dimensi supplier berisi atribut-atribut supplier_key dan supplier_type.

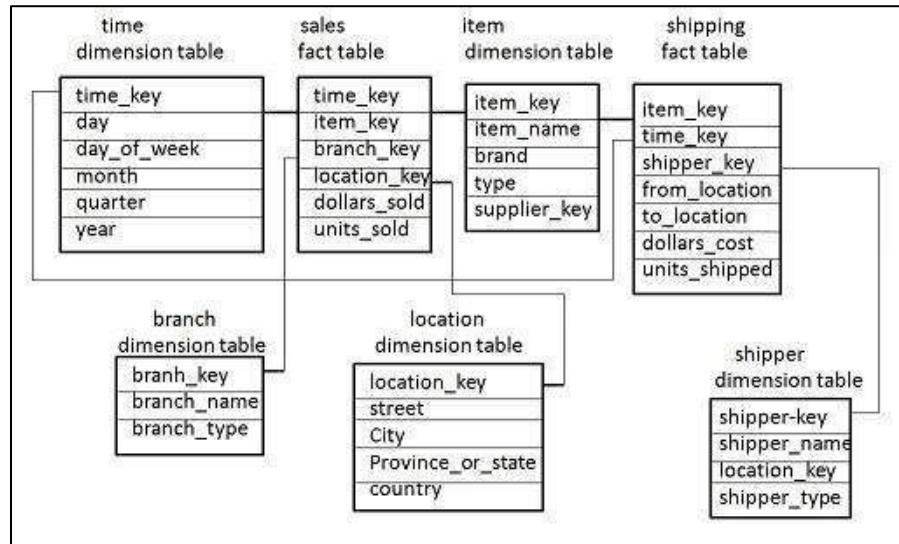


Fact Constellation Schema

Memiliki beberapa tabel fakta, juga dikenal sebagai *galaxy schema*.

Tabel fakta penjualan sama dengan yang ada dalam skema bintang (star schema).

Tabel fakta pengiriman memiliki lima dimensi, yaitu item_key, time_key, shipper_key, dari_lokasi, ke_lokasi.



HIRARKI KONSEP

Hirarki Konsep

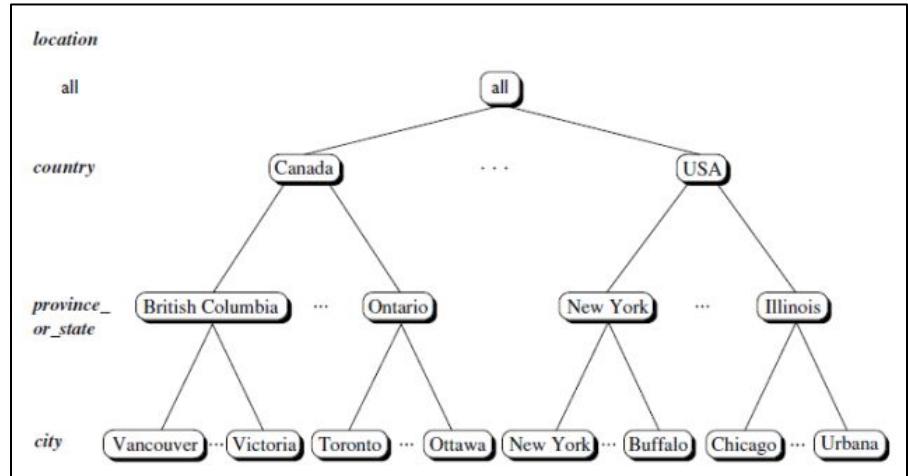
Hirarki Konsep mendefinisikan urutan pemetaan dari sekumpulan konsep tingkat rendah ke konsep yang lebih tinggi dan lebih umum.

Hirarki Konsep

Misal hirarki konsep untuk dimensi lokasi. Nilai kota untuk lokasi meliputi Vancouver, Toronto, New York, dan Chicago.

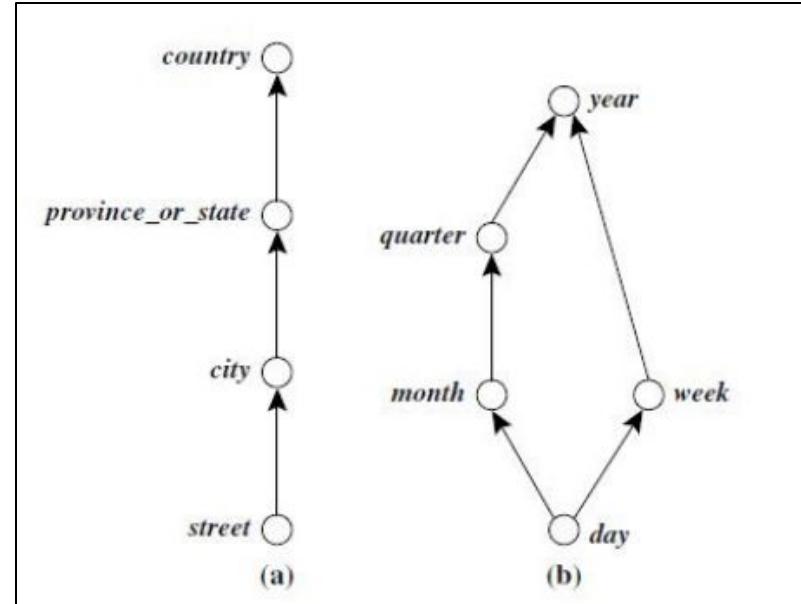
Namun, setiap kota dapat dipetakan ke provinsi atau negara bagian tempatnya berada.

Sebagai contoh, Vancouver dapat dipetakan ke British Columbia, dan Chicago ke Illinois.



Hirarki Konsep

Struktur Hierarki dan Struktur Jaringan Atribut dalam Dimensi Gudang: (a) hirarki untuk lokasi dan (b) jaringan untuk waktu.



OLAP (Online Analytical Processing)

OLAP

OLAP (Online Analytical Processing) adalah metode komputasi yang memungkinkan pengguna untuk dengan mudah dan selektif mengekstrak dan mengajukan data guna menganalisisnya dari berbagai sudut pandang. Kueri bisnis dan analisis OLAP sering digunakan dalam analisis tren, laporan keuangan, peramalan penjualan, perencanaan anggaran, dan tujuan perencanaan lainnya.

Operasi OLAP

Macam-macam Operasi OLAP

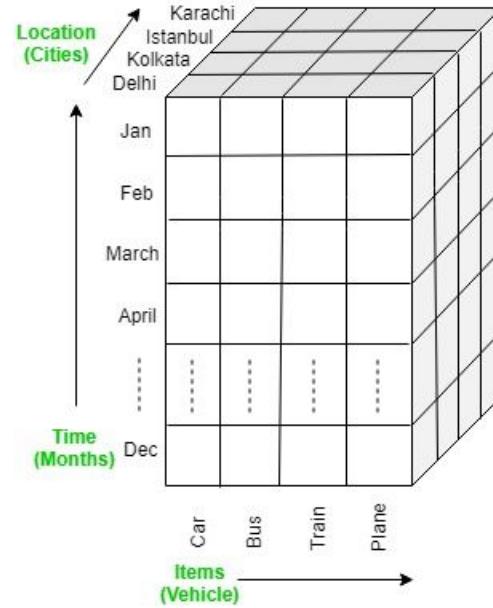
- Drill-down
 - Roll-up
 - Slice
 - Dice
 - Pivot
-

Drill Down

Data yang kurang terperinci dikonversi menjadi data yang sangat terperinci. Ini dapat dilakukan dengan:

1. Pergi ke bawah dalam hierarki konsep
2. Menambah dimensi baru

Contoh, operasi drill-down dilakukan dengan memindah ke bawah dalam hierarki konsep dimensi Waktu (Kuartal -> Bulan).

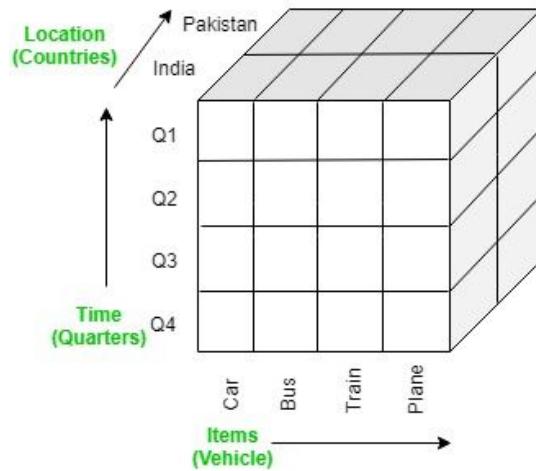


Roll Up

Kebalikan dari operasi drill-down. Ini melakukan agregasi pada kubus OLAP. Ini dapat dilakukan dengan:

1. Naik ke atas dalam hierarki konsep
2. Mengurangi dimensi

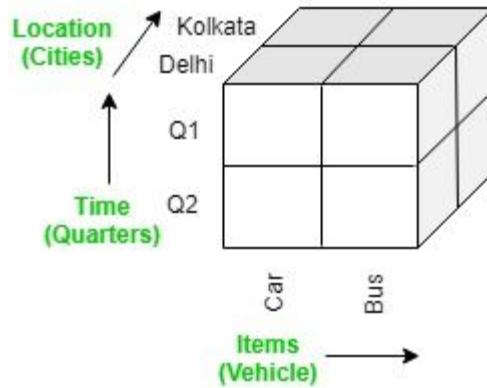
Contoh, operasi roll-up dilakukan dengan naik ke atas dalam hierarki konsep dimensi Lokasi (Kota -> Negara).



Dice

Operasi Ini memilih sub-kubus dari kubus OLAP dengan memilih dua atau lebih dimensi. Contoh, sebuah sub-kubus dipilih dengan memilih dimensi-dimensi berikut dengan kriteria:

- Lokasi = "Delhi" atau "Kolkata"
- Waktu = "Q1" atau "Q2"
- Item = "Mobil" atau "Bus"



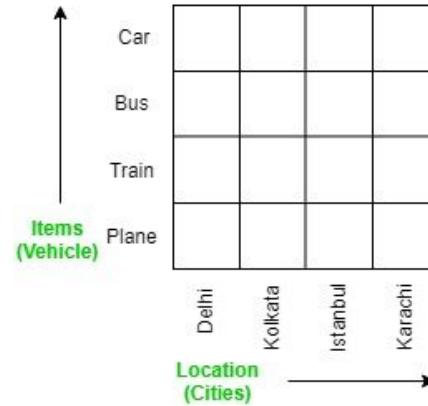
Slice

Operasi Ini memilih satu dimensi dari kubus OLAP yang menghasilkan pembuatan sub-kubus baru. Contoh, Slice dilakukan pada dimensi Waktu = "Q1".



Pivot

Dikenal sebagai operasi rotasi karena memutar tampilan saat ini untuk mendapatkan tampilan representasi yang baru. Dalam sub-kubus yang diperoleh setelah operasi Slice, melakukan operasi pivot memberikan tampilan yang baru darinya.



IMPLEMENTASI NYATA DATA CUBE TECHNOLOGY

Implementasi : Sumber

Dikutip dari Dr. Omid Shabestari, Director of Health Analytics dan Shea Jessee, Team Lead, dari Carilion Clinic.

Implementasi Data Cube dilakukan di Carilion Clinic, Virginia, Amerika Serikat.

Implementasi : Health Data Cube

Implementasi "Health Data Cube" di Carilion Clinic, dengan fokus pada penggunaan data cube dalam analitik kesehatan. Data cube digunakan untuk menjawab pertanyaan penting, seperti rata-rata jumlah pesanan obat per kunjungan pasien dan peringkat kepuasan pasien. Implementasi data cube ini memungkinkan akses lebih cepat dan efisien ke data yang sebelumnya hanya tersedia melalui alat pelaporan.

Implementasi : Cara Kerja

Data cube ini dibangun di atas data dari berbagai sumber, termasuk data rekam medis elektronik (EMR), survei kepuasan pasien, dan data lainnya. Memungkinkan perhitungan agregat dan perbandingan data dengan cepat. Data cube diproses setiap malam sehingga data yang tersedia selalu up-to-date dalam waktu 24 jam. Dapat digunakan dengan berbagai alat analisis, seperti lembar kerja dan alat visualisasi data.

Implementasi : Kesimpulan

Pengembangan data warehouse dan data cube memerlukan perubahan organisasi, termasuk pendirian dewan pengaturan data dan kelompok pengurus data. Cube ini bertujuan untuk menjadi sumber data tunggal dalam organisasi dan mendukung pengambilan keputusan yang lebih baik.

KESIMPULAN

Kesimpulan

Data Cube adalah konsep yang mewakili representasi multidimensional dari data, memungkinkan penyimpanan dan analisis data dari berbagai dimensi seperti waktu, lokasi, dan atribut lainnya.

Kesimpulan

Data Cube memungkinkan pengguna untuk melihat data dari berbagai sudut pandang dan menggabungkan konsep hirarki dalam analisis. Berhubungan erat dengan OLAP, yang memungkinkan pengguna menjelajahi dan menganalisis data dengan lebih mendalam, membuat keputusan yang lebih baik, dan merespons perubahan dengan lebih cepat dalam berbagai konteks bisnis dan ilmiah.

Terima Kasih