

Perancangan Data Warehouse

Data yang disimpan dalam data warehouse adalah data historis berorientasi subjek yang dapat mendukung proses pengambilan keputusan bagi manajemen. Artinya data tersebut harus kita susun sedemikian rupa sehingga dapat dianalisis menjadi berbagai informasi yang dibutuhkan manajemen saat proses pengambilan keputusan.

Oleh karena itu, tahap pertama dari perancangan data warehouse adalah **mendefinisikan informasi-informasi apa saja yang dibutuhkan oleh manajemen**. Supaya kebutuhan ini dapat didefinisikan dengan tepat, maka pemahaman akan peran dan tugas manajemen yang membutuhkan informasi tersebut mutlak harus dilakukan lebih dulu. Jika sudah dipahami, selanjutnya kita hanya tinggal “menjawab” pertanyaan-pertanyaan berikut:

- Siapa yang membutuhkan informasi dari data warehouse?
- Informasi apa saja yang dibutuhkan tersebut?
- Seperti apa layout dan isi informasi-informasi itu?
- Kapan informasi tersebut digunakan?
- Untuk keperluan apa?
- Basis data apa yang menjadi sumber untuk informasi tersebut?

Sebagai contoh, misalkan akan dibuat sebuah data warehouse penjualan (atau data mart penjualan tepatnya) untuk sebuah perusahaan dagang.

- Siapa yang membutuhkan informasi dari data warehouse?
Manager Pemasaran
- Informasi apa saja yang dibutuhkan Manager Pemasaran?
Barang apa yang paling banyak terjual di lokasi tertentu sepanjang tahun
Barang apa yang paling banyak memberikan pendapatan sepanjang tahun
- Seperti apa layout dan isi informasi-informasi itu?
Barang yang paling banyak terjual di lokasi tertentu sepanjang tahun

tahun	kecamatan	kategori	sum(total_penjualan)
2007	ANDIR	KONSUMSI	26
2007	SUMUR BANDUNG	ATK	17
2007	ASTANA ANYAR	ATK	16
2007	BANDUNG WETAN	KONSUMSI	15
2007	SUMUR BANDUNG	KONSUMSI	13
2007	ANDIR	ATK	13
2007	BANDUNG WETAN	ATK	7

Barang yang paling banyak memberikan pendapatan sepanjang tahun

tahun	kategori	subkategori	sum(total_penerimaan)
2007	ATK	ALAT TULIS	4270000
2007	ATK	ELEKTRONIK	511000
2007	KONSUMSI	SEBAKO	116000
2007	KONSUMSI	SNACK	87100

- Untuk keperluan apa informasi tersebut?
Dasar untuk menentukan strategi penjualan barang
- Kapan informasi tersebut digunakan?
Awal periode penjualan
- Basis data apa yang menjadi sumber untuk informasi tersebut?
Basis data penjualan dengan skema sebagai berikut:
 1. Barang (#kode_barang, nama_barang, kelompok, satuan, harga)
 2. Kategori (#kelompok, sub_kategori, kategori)
 3. Pelanggan (#kode_cust, nama_cust, alamat, kota, kode_pos, telepon)
 4. Lokasi (#kode_pos, kelurahan, kecamatan)
 5. Penjualan (#no_faktur, #kode_barang, qty)
 6. Pembayaran (#no_faktur, tanggal, total, discount, #kode_cust)

Tahap berikutnya yang harus dilakukan adalah **menentukan *measure* dan *dimension* untuk semua informasi yang dibutuhkan manajemen**. *Measure* adalah data numerik yang akan dicari jejak nilainya, sedangkan *dimension* adalah parameter atau sudut pandang terhadap *measure* sehingga dapat mendefinisikan suatu transaksi.

Sebagai contoh, untuk informasi “barang yang paling banyak terjual di lokasi tertentu sepanjang tahun”,

- *Measure*: total penjualan
- *Dimension*: barang, tahun (waktu/periode), lokasi

Sedangkan untuk informasi “barang yang paling banyak memberikan pendapatan sepanjang tahun”,

- *Measure*: total pendapatan
- *Dimension*: barang, tahun (waktu/periode)

Dimension mempunyai hirarki. Penentuan hirarki untuk *dimension* ini sepenuhnya tergantung kepada proses *drill down* dan *roll up* yang ingin dilakukan saat melakukan OLAP (OnLine Analytical Processing) nanti.

Untuk contoh diatas, hirarki masing-masing *dimension* adalah:

- Barang: nama barang, sub-kategori, kategori
- Periode: minggu, bulan, tahun
- Lokasi: kelurahan, kecamatan, kota

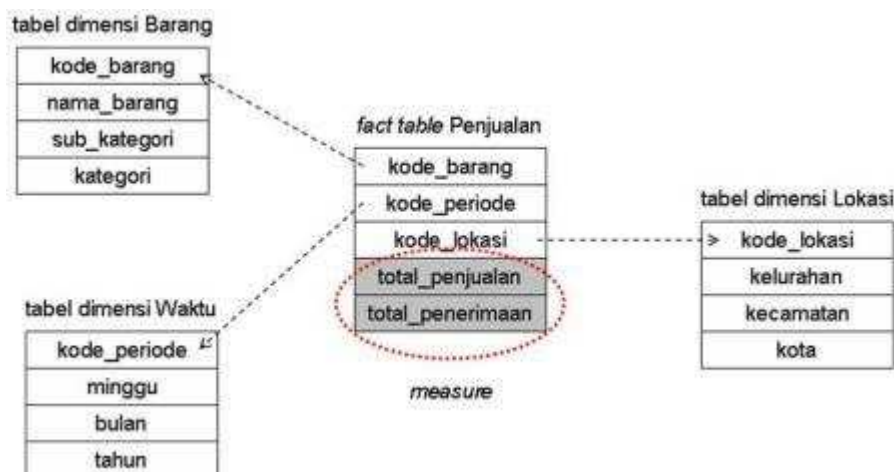
Sedangkan *layout* dan isi informasinya dapat ditunjukkan oleh gambar berikut ini:

BARANG				PERIODE			LOKASI			TOTAL	TOTAL	
KODE	NAMA	SUB KATEGORI	KATEGORI	MINGGU	BULAN	TAHUN	KEKURAHAN	KEDAMATAN	KOTA	PEJUALAN	PENERIMAAN	
B-001	SUKRO	SNACK	KONSUMSI	44	11	2007	ARJUNA	ANDR	BANDUNG	5	12.500	
				45	11	2007	ARJUNA	ANDR	BANDUNG	1	2.500	
				45	11	2007	SUKALUYU	BANDUNG WETAN	BANDUNG	1	2.500	
				45	11	2007	BRAGA	SUMUR BANDUNG	BANDUNG	2	5.000	
B-002	BENGBENG	SNACK	KONSUMSI	44	11	2007	ARJUNA	ANDR	BANDUNG	1	1.250	
				45	11	2007	BRAGA	SUMUR BANDUNG	BANDUNG	5	6.000	
				45	11	2007	SUKALUYU	BANDUNG WETAN	BANDUNG	2	2.400	
B-003	POP MIE	SEBAKAD	KONSUMSI	45	11	2007	ARJUNA	ANDR	BANDUNG	10	15.000	
				45	11	2007	SUKALUYU	BANDUNG WETAN	BANDUNG	3	4.500	
				45	12	2007	ARJUNA	ANDR	BANDUNG	2	3.000	
TOTAL												54.600

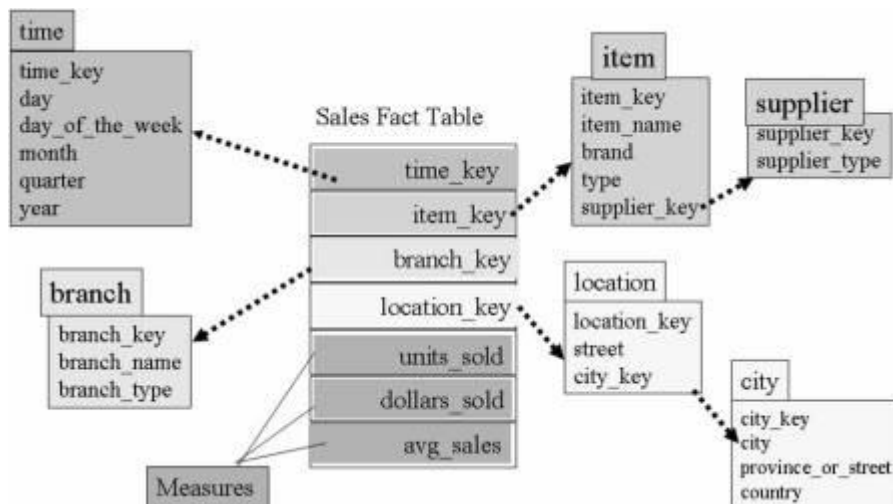
Perancangan model konseptual data warehouse adalah tahap berikutnya yang harus dilaksanakan setelah tahap penentuan *measure* dan *dimension*. Pada tahap ini dibuat suatu model yang dapat menggambarkan data atau tabel apa saja yang akan disimpan dalam data warehouse, berikut keterhubungan diantaranya.

Data atau tabel dalam data warehouse tersebut dapat dimodelkan dengan menggunakan alat bantu pemodelan seperti E-R diagram, [star schema](#), [snowflake schema](#), atau FCO-IM (*Fully Communication Oriented Information Modelling*). Tetapi pada umumnya alat bantu yang digunakan adalah *star schema* atau *snowflake schema*.

Star schema akan menggambarkan *fact table*, yaitu tabel yang merepresentasikan *measure*, sebagai “pusat data”. Tabel ini nantinya akan terkoneksi dengan tabel-tabel yang mendeskripsikan dimensi untuk *measure* tersebut (*dimension table*). Sebagai contoh, *star schema* untuk data warehouse penjualan adalah:



Snowflake schema merupakan perbaikan dari *star schema*, sehingga cara penggambarannya pun mirip. Bedanya, satu atau beberapa hirarki yang ada pada *dimension table* dinormalisasi (dekomposisi) menjadi beberapa tabel yang lebih kecil. Sebagai contoh,



Tahap terakhir dari perancangan data warehouse adalah **membuat rancangan skema data warehouse**, yaitu kumpulan objek-objek basis data seperti tabel, *view*, indeks, dan objek lainnya yang mendeskripsikan suatu data warehouse.

Sebagai contoh, rancangan skema untuk data warehouse Penjualan yang akan diimplementasi dalam MySQL adalah:

Tabel SalesFact

Field	Type	Null	Key	Default	Extra
kode_barang	varchar(5)	YES		NULL	
kode_periode	varchar(8)	YES		NULL	
kode_lokasi	varchar(5)	YES		NULL	
total_penjualan	double	YES		NULL	
total_penerimaan	double	YES		NULL	

Tabel Barang

Field	Type	Null	Key	Default	Extra
kode_barang	varchar(5)	YES		NULL	
nama_brg	varchar(30)	YES		NULL	
subkategori	varchar(20)	YES		NULL	
kategori	varchar(20)	YES		NULL	

Tabel Periode

Field	Type	Null	Key	Default	Extra
kode_periode	varchar(8)	YES		NULL	
minggu	varchar(2)	YES		NULL	
bulan	varchar(2)	YES		NULL	
tahun	varchar(4)	YES		NULL	

Tabel Lokasi

Field	Type	Null	Key	Default	Extra
kode_lokasi	varchar(5)	YES		NULL	
kelurahan	varchar(30)	YES		NULL	
kecamatan	varchar(30)	YES		NULL	
kota	varchar(30)	YES		NULL	

View Datacube (untuk keperluan OLAP)

Field	Type	Null	Key	Default	Extra
kode_barang	varchar(5)	YES		NULL	
nama_brg	varchar(30)	YES		NULL	
subkategori	varchar(20)	YES		NULL	
kategori	varchar(20)	YES		NULL	
kode_periode	varchar(8)	YES		NULL	
minggu	varchar(2)	YES		NULL	
bulan	varchar(2)	YES		NULL	
tahun	varchar(4)	YES		NULL	
kode_lokasi	varchar(5)	YES		NULL	
kelurahan	varchar(30)	YES		NULL	
kecamatan	varchar(30)	YES		NULL	
kota	varchar(30)	YES		NULL	
total_penjualan	double	YES		NULL	
total_penerimaan	double	YES		NULL	

Perlu diketahui, rancangan skema yang dihasilkan pada tahap ini pada umumnya sudah mengikuti format DBMS yang akan digunakan. Pada beberapa DBMS, seperti Oracle atau SQL Server misalnya, disediakan fasilitas maupun *wizzard* untuk membantu merancang dan mengimplementasi skema data warehouse ini.