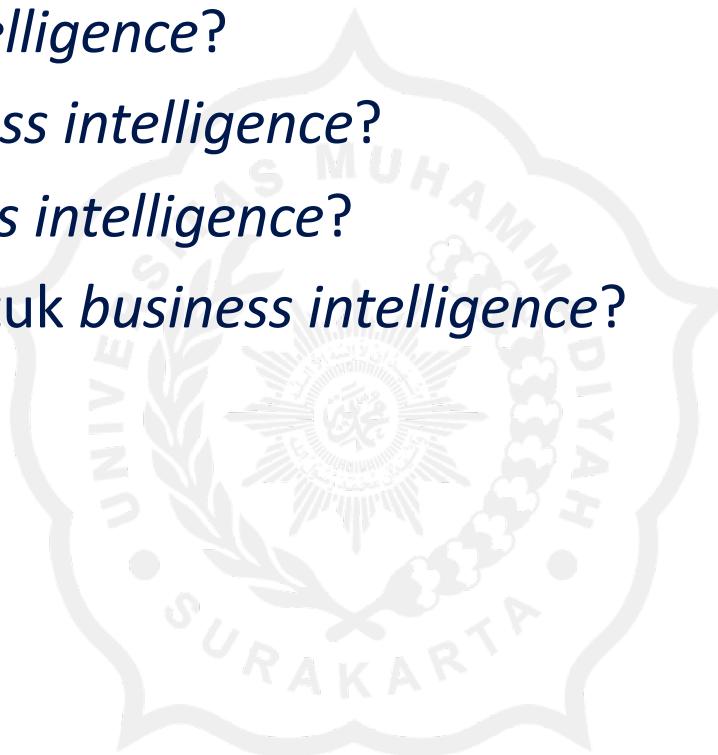


## Bab 2 **Business Intelligence (BI)**

Dedi Gunawan, PhD. & Dr.Eng. Yusuf Sulistyo Nugroho

# OUTLINE

- Apa itu *business intelligence*?
- Kenapa perlu *business intelligence*?
- Proses pada *business intelligence*?
- Desain database untuk *business intelligence*?



# Apa itu business intelligence?

- **Business intelligence** merupakan suatu alat untuk melakukan eksplorasi dan manipulasi data
  1. Data yang merepresentasikan berbagai periode waktu secara historis
  2. Data dengan jumlah dan ukuran yang sangat besar serta selalu bertambah sesuai dengan kebutuhan bisnis.

Business intelligence sangat berkaitan dengan teknologi berikut ini

1. *Decision support system* (Sistem pendukung keputusan)
2. *Online analytical processing* (OLAP)
3. *Data warehouse* (DW)
4. *Data mart*
5. *Data mining*

# BI terpisah dengan DB Operasional

- Data yang digunakan untuk proses business intelligence terpisah dari database operasional
  - Untuk menghindari perubahan data yang merusak transaksi dan histori
  - Untuk menghindari query yang lama dan mengakibatkan error
  - Memiliki tujuan yang berbeda dengan database transaksi dalam hal ruang lingkup, users, metode pemrosesan dan Analisa kebutuhan

Database yang digunakan untuk proses **business intelligence** tidak hanya database operasional saja melainkan ada database lain diantaranya

- *Behavioral data* (data yang mencerminkan perilaku user)
  - Click stream, page visit, bookmark dan perilaku online lainnya
- *External demographic data*
- *Meta data*
  - Informasi mengenai data pada database
  - Informasi physical data, informasi teknis, informasi bisnis proses, rule and constraints, dan struktur data

# Perbedaan BI & Sistem Operasional

	<b>Operational Systems</b>	<b>Business Intelligence</b>
<b>Processing</b>	OLTP, ACID	Load, analyze
<b>Level of detail</b>	Transaction	Aggregate
<b>Timeframe</b>	Current transactions	Historical
<b>Database design</b>	Complex relationships	Star schema

# Kenapa memerlukan business intelligence?

# Manfaat Business Intelligence

- Untuk menghasilkan keputusan yang lebih cepat dan lebih baik
- E-business marketing:
  - Bisa digunakan untuk Analisa karakter invididu dari user dan mensegementasi user menjadi beberapa kelompok
  - Merchandising
- Meningkatkan kualitas bisnis
- Meningkatkan kualitas operasional
- Sebagai bahan pembelajaran untuk strategi bisnis
- Juga bisa digunakan sebagai penghasilan tambahan sekaligus untuk mentargetkan konsumen

## Traditional users:

- Sales management
- Marketing management
- Purchasing
- Operations management
- Human Resources
- Senior management

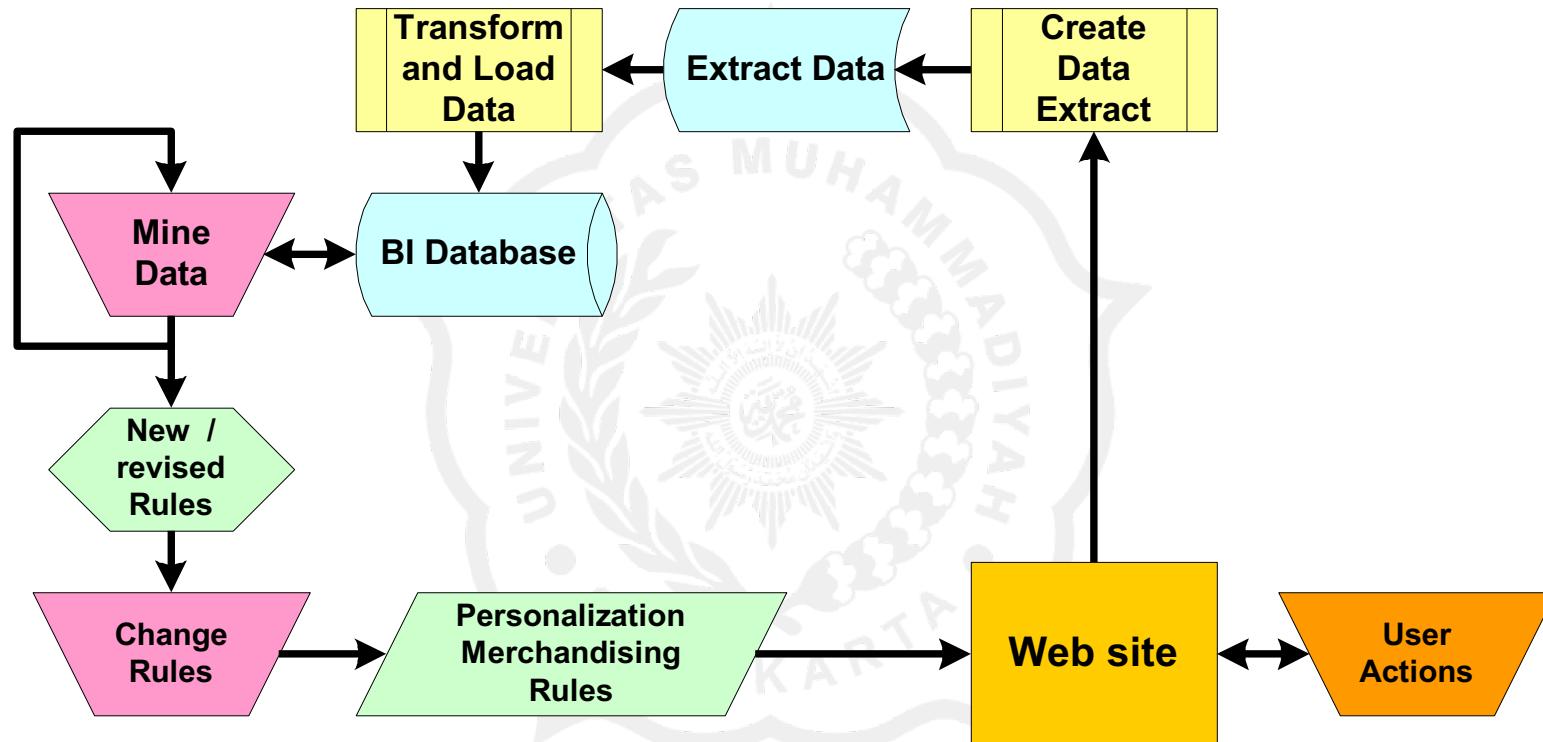
## Expanded users:

- Employees in all parts of the enterprise
- Suppliers
- Channel partners
- Customers

# Proses pada Business Intelligence

- Proses ETL secara periodik dari operasional database dan system
  - **Extract**: Mengambil bagian data yang diperlukan
  - **Transform**: merubah data yang sifatnya individual menjadi agregat atau general (Umum)
  - **Load** : memasukkan data ke dalam system BI
- Proses ETL bervariasi, bisa dilakukan setiap bulan sekali ataupun secara realtime setiap saat dibutuhkan.

# Proses BI



# Vendor Teknologi BI (skala internasional)

- Business Objects
- Cognos
- Information Builders
- MicroStrategy
- SAS Institute
- Blue Martini (*integrated with e-commerce*)

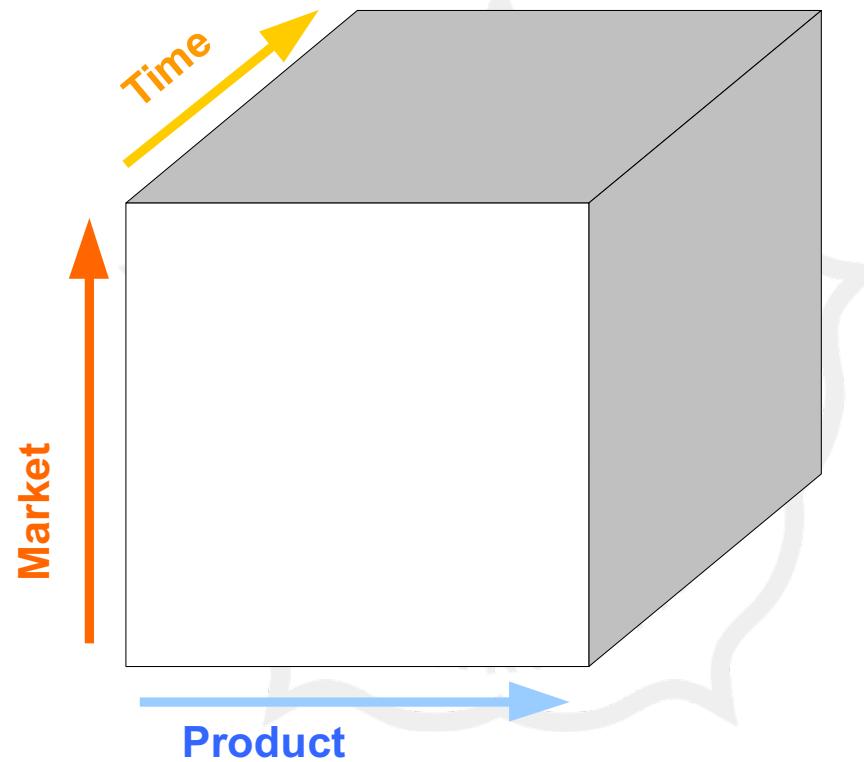


# Desain database untuk Business Intelligence

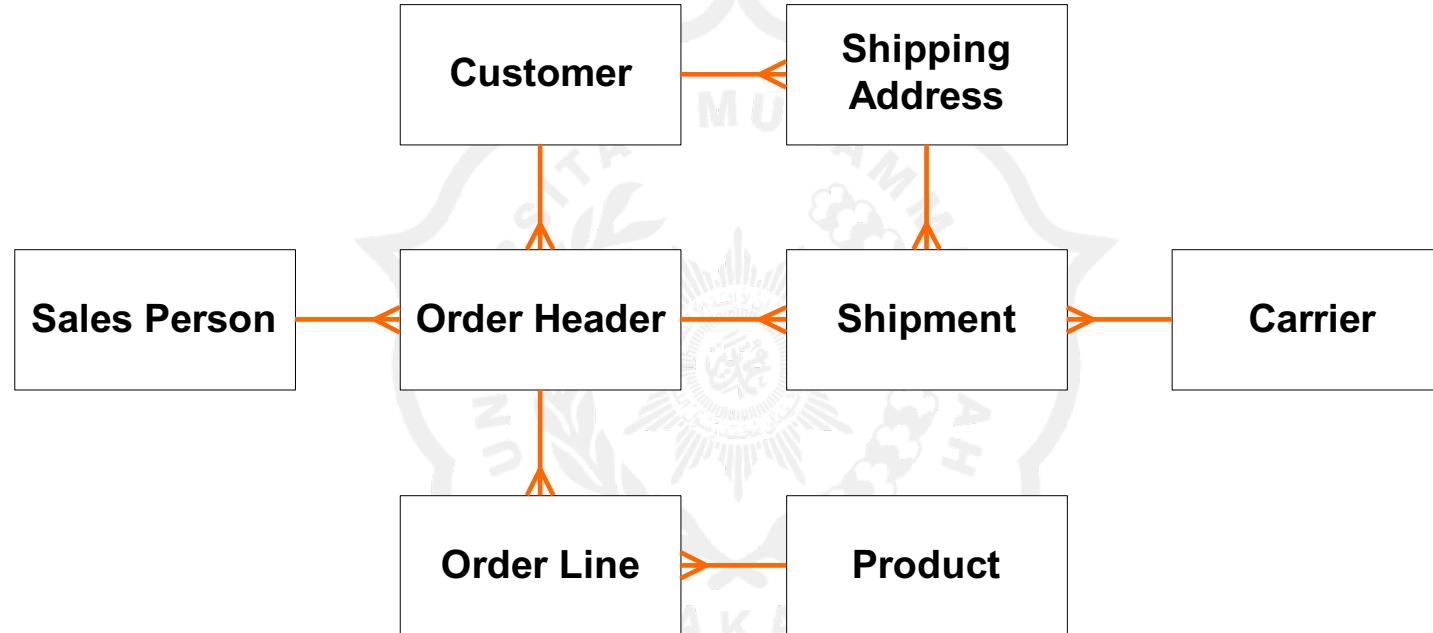
# Desain database untuk proses BI

- Berdasarkan dimensi dari model bisnis
  - Cth : perusahaan menjual **produk** ke berbagai **pasar** dengan melakukan pengukuran performa penjualan perusahaan secara berkala setiap **waktu** tertentu.
- Dimensi merupakan suatu kumpulan jenis data
  - Cth: dimensi produk, dimensi pasar, dimensi waktu
  - Dimensi produk berkaitan dengan jenis, nama, harga, kuantitas dst dari sebuah produk
  - Dimensi pasar berisi data pasar/market yang berupa lokasi, nama market, skala market dst yang menjadi target pemasaran.
- Database umumnya berbentuk dua dimensi (table) namun ada database yang memiliki lebih dari dua dimensi, dinamakan hypercube.

# Model dimensi data pada BI



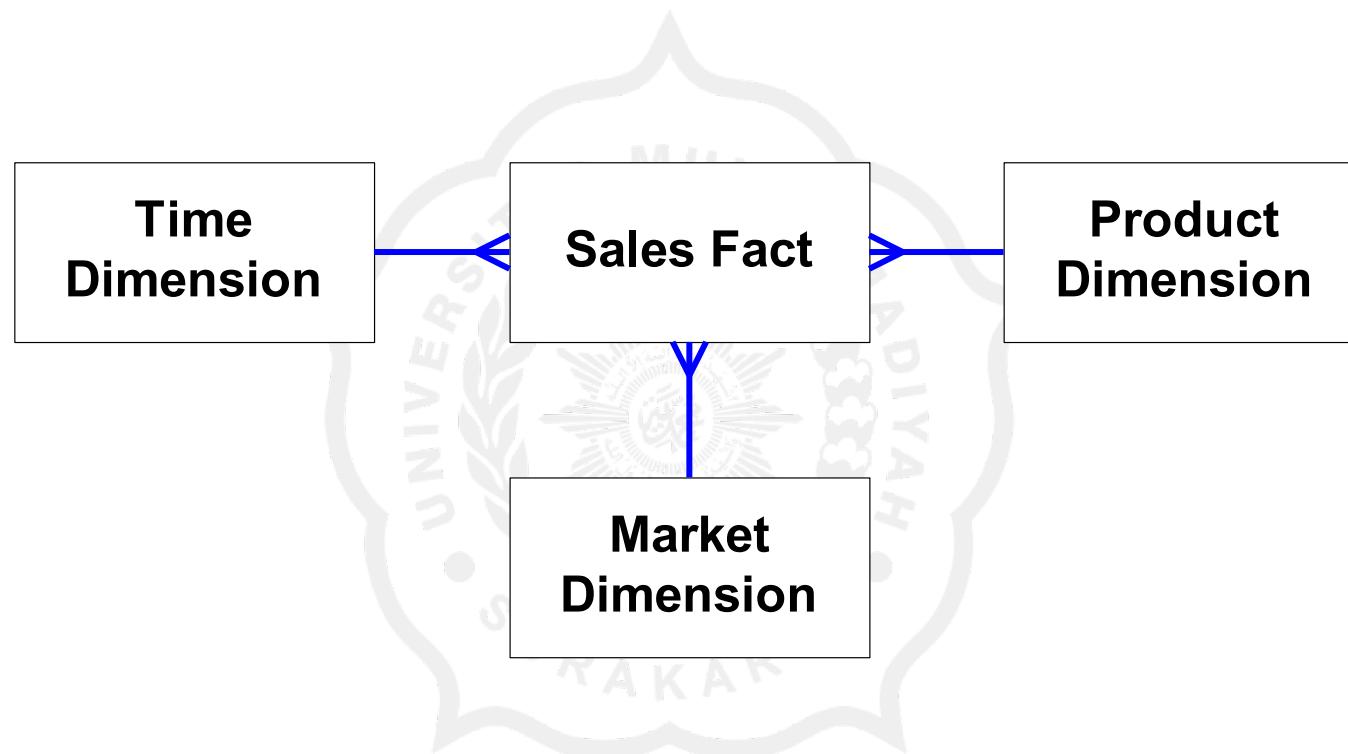
# Contoh Database Operasional



# Desain database untuk proses BI

- Dimensi menetukan desain dari database yang akan dibuat, yang nantinya akan menjadi tabel-tabel dimensi
- *Fact table* atau table fakta adalah tabel yang menjadi fokus utama Ketika mendesain database
- Tabel dimensi mempunyai relasi dengan tabel fakta
  - *Star join schema (skema star)*
- Setiap join (gabungan) memiliki makna untuk menjelaskan hubungan antara item yang ada pada proses bisnis

# Contoh desain database untuk BI



- Sales fact (table fakta penjualan)
  - Time key
  - Product key
  - Market key
  - Units sold
  - Dollars sold
  - Dollars cost
- Market Dimension
  - Market key
  - Region, territory, etc.
- Time Dimension
  - Time key
  - Day of week
  - Month
  - Quarter
  - Year
- Product Dimension
  - Product key
  - Category
  - Description



# Desain BI Database: Fact table / tabel fakta

- Mengukur keterkaitan antar dimensi
  - Cth: penjualan suatu produk di suatu pasar pada waktu tertentu
- Idealnya bersifat adaptif dan nilai yang ada di dalam table terus berubah
- Pertimbangkan dengan data-data yang sifatnya semi-additive dan no-additive
  - Cth: nilai rata-rata penjualan, jumlah pelanggan, dll.
- Tabel fakta biasanya bersifat sparse (beragam dan memiliki konten yang berbeda serta tersebar)
  - Cth: penjualan yang hanya 10% dari total produk

# Desain BI Database: dimension / dimensi

- Dimensi menjelaskan tentang anggota atau isi dari table dimensi
- Idealnya memiliki karakter berikut
  - Text (bukan numerik data)
  - Digunakan sebagai constraint (Batasan) untuk menjawab suatu pertanyaan spesifik
  - Bersifat konstan/tetap seiring berjalan waktu
- Tidak bersifat sparse.

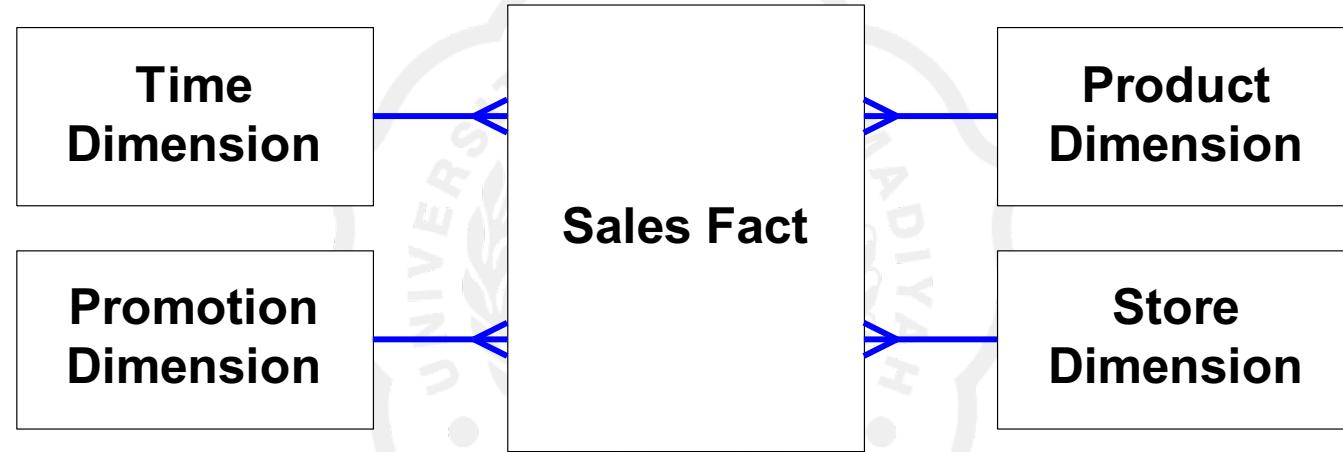
# Desain Database untuk BI

- Pertimbangkan proses dan identitas dari table-table fakta
- Detail dari setiap table fakta sangat penting
  - Detail berdasarkan waktu
- Dimensi setiap tabel fakta perlu dijelaskan dan rancang sesuai keperluan
- Atribut data pada table fakta
- atribut dimensi dan sumber data
- Rentang waktu yang bisa ditangani oleh database
- Sumber data dan frekuensi update data

# Contoh BI: perusahaan grosir

- 300 toko cabang dengan POS (point of sales) system
- 30.000 produk per toko cabang
  - Rata-rata menjual 3.000 perhari
- Promosi untuk produk tertentu

# Desain Database



# Tabel fakta penjualan (*sales fact*)

- **Sales Fact:**

- Time key
- Store key
- Product key
- Promo key
- Quantity sold
- Dollar revenue
- Dollar cost
- Customer count



- Time dimension:

- Time key
- Date
- Day of week, Day of month
- Quarter, Fiscal period,
- Holiday flag, Weekday flag, Last in month flag
- Season
- Event



# Dimensi toko

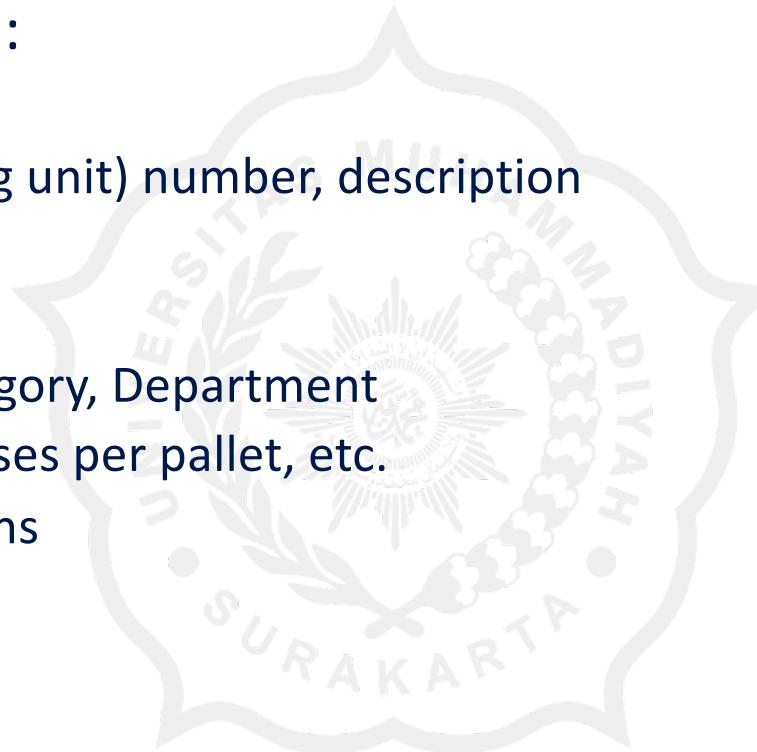
- **Store dimension:**

- Store key
- Store id
- Store city, Store county, Store state, Store zip
- Sales region
- Floor plan type, square footage, etc.
- Store open date



- **Product dimension:**

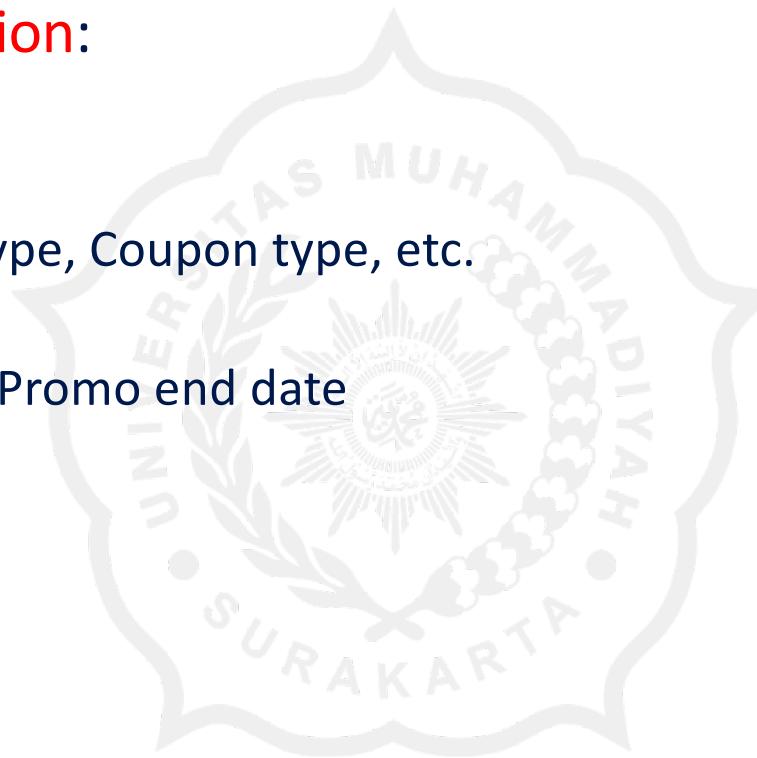
- Product key
- SKU (stock keeping unit) number, description
- Package size
- Brand
- Category, Subcategory, Department
- Units per case, Cases per pallet, etc.
- Weight, Dimensions



# Dimensi promosi

- Promotion dimension:

- Promo key
- Promotion name
- Ad type, Display type, Coupon type, etc.
- Promo cost
- Promo start date, Promo end date



# Perhitungan *storage* database yang diperlukan

- 2 tahun x 365 = 730 hari
- 300 toko dengan POS (Point of sale)
- 30,000 produk per toko
  - menjual 3,000 item per hari
- Row/baris data pada tabel fakta penjualan:  $730 \times 300 \times 3000 = 657.000.000$
- Each fact table row:
  - 8 kolom x 4 bytes=32 bytes
- Fact table size:  $657.000.000 \times 32$  sekitar 21 GB

Teşekkürler Salamat Gamsahabnida Gratias Ago Vos Dankie dat U Sas efcharistó **Thank You** Merci Asante Ašante Teşekkürler Grazie Tack Obrigado Hvala Köszönöm Kōsanzōm Arigato Ngiyabonga Sağ Olun Gràcies Xièxiè Nín Danke Dankon Dank U Khokhobchai Than Dziekujemy Spasibo Va Multumesc Go raibh maith agat **Gracias** Díolch yn fawr Terima Kasih Dank U