

# Pengantar *Business intelligence System*

---

Business Intelligence Concept

Why BIS Needed?

Business Intelligence Tools

# *Intelligence?*

---

👉 *Intelligence* adalah kemampuan belajar, memahami, atau menyesuaikan situasi baru; kemampuan berargumentasi, mengaplikasikan pengetahuan untuk memanipulasi lingkungannya atau berpikir abstrak. [Michael H. Brackett](#)

# *Business intelligence?*

---

- ❖ *Business intelligence* adalah seperangkat *konsep*, metode dan proses untuk meningkatkan keputusan bisnis dengan menggunakan berbagai sumber informasi dan mengaplikasikan pengalaman, asumsi untuk mengembangkan akurasi pemahaman bisnis secara dinamik.
- ❖ Mencakup mendapatkan, mengelola dan menganalisis data untuk menghasilkan informasi dan mendistribusikan keseluruhan organisasi untuk meningkatkan keputusan taktis dan strategis.

[Michael H. Brackett](#)

# Business Intelligence? (Lanj)

---

Pengetahuan tentang pelanggan Anda, pesaing Anda, mitra bisnis Anda, lingkungan kompetitif Anda, dan operasi internal Anda sendiri - yang memberikan Anda kemampuan untuk membuat keputusan bisnis yang efektif, penting, dan sering strategis.

# Manfaat Business Intelligence?

---

Membantu kita mendapatkan informasi yang akurat sehingga dapat membuat keputusan bisnis yang berdasarkan data, bukan berdasarkan insting.

Mampu mengidentifikasi peluang baru dengan data informasi yang dihasilkan. Anda akan memperoleh informasi baru tentang peluang serta potensial yang selama ini dimiliki oleh bisnis Anda.

Business intelligence juga dapat membantu kita dalam mengelola manajemen persediaan bisnis dengan sangat baik.

Solusi yang ditawarkan dan diberikan oleh business intelligence dapat menghemat banyak waktu Anda karena dapat memberikan laporan bisnis yang Anda butuhkan hanya dalam hitungan detik.

Mengalokasikan sumber daya yang kita miliki dengan lebih akurat.

# Business Intelligence (BI)

---

BI merupakan payung dari kombinasi arsitektur, alat-alat, database, alat-alat analisis, aplikasi, dan metodologi

Seperti DSS, konten BI bebas ekspresi, jadi setiap orang bisa mengartikan berbeda

Tujuan utama BI adalah untuk memungkinkan akses data (dan model) yang mudah agar manajer bisnis mampu melakukan analisis

BI membantu mentransformasikan data, menjadi informasi (dan pengetahuan), untuk membuat keputusan dan tindakan

## *Business Intelligence Systems?*

---

BIS adalah sistem informasi yang menyediakan BI bagi pengambil keputusan pada setiap level organisasi (*operational, tactical, strategic levels*)

BIS adalah sistem informasi merubah data, informasi, and/or knowledge secara selektif menjadi *desired intelligence* untuk tujuan bisnis (Thierauf, 2001)

# Business Intelligence Systems(Lanj)

---

***Business intelligence (BI) systems*** – Aplikasi dan peralatan IT yang mendukung fungsi intelijen bisnis dalam organisasi.

***Competitive intelligence*** - business intelligence yang difokuskan pada lingkup kompetisi eksternal.



# Sejarah singkat BI

---

Istilah BI diciptakan oleh Gartner Group pada pertengahan 1990-an

Namun, konsep ini ada sejak lama

- 1970 – pelaporan MIS – laporan statis / periodik
- 1980 - Sistem Informasi Eksekutif (EIS)
- 1990 - OLAP, dinamis, multidimensi, pelaporan ad-hoc -> nilai dari "BI"
- 2005 + memasukkan AI dan Data / Teks mining, berbasis Web Portal / Panel Kontrol
- 2010-an - belum terlihat

# Evolusi kemampuan BI



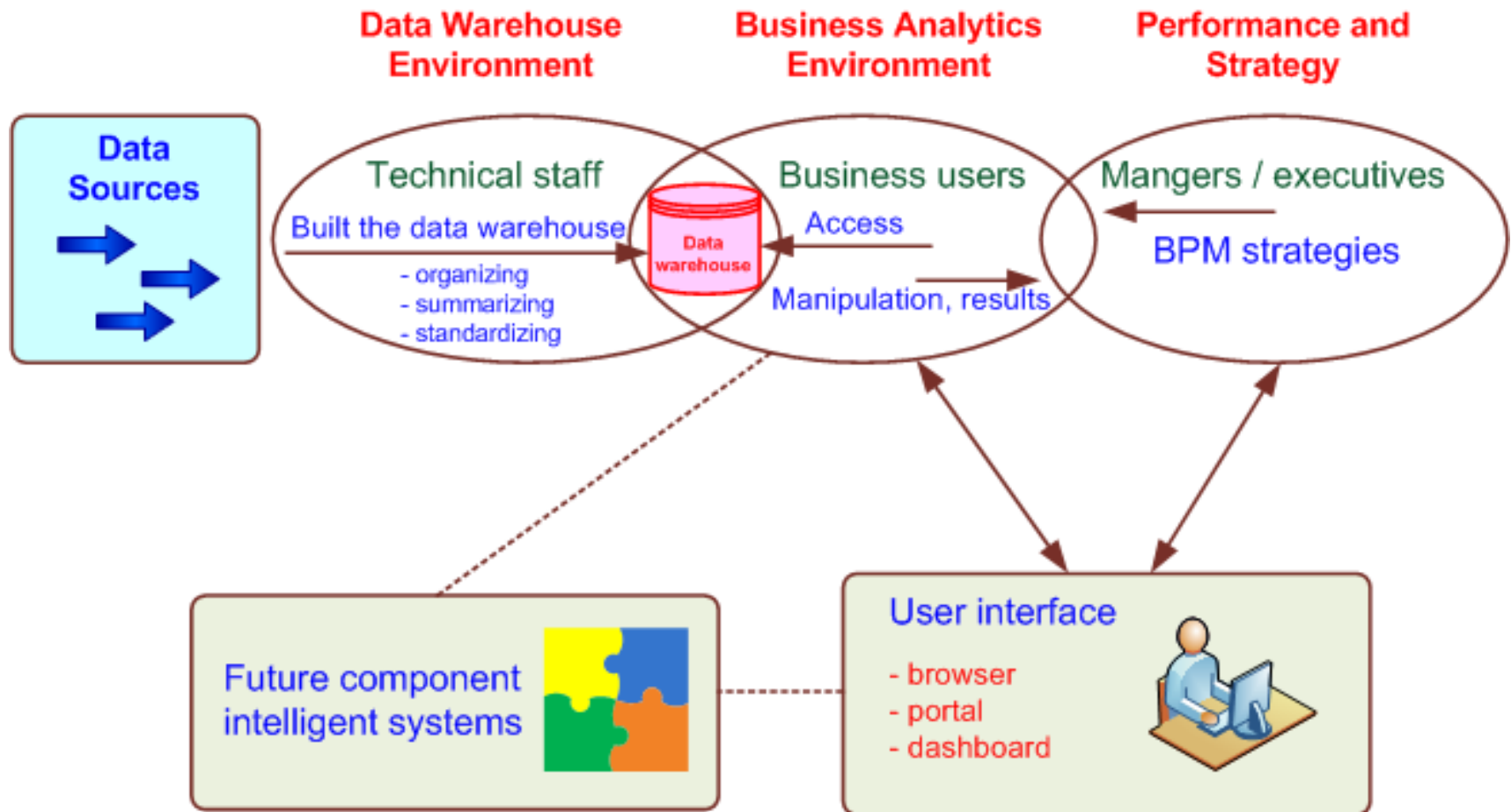
# Arsitektur BI

---

sistem BI memiliki empat komponen utama

- **data warehouse**, dengan sumber datanya
- **analisis bisnis**, kumpulan alat untuk memanipulasi, menggali, dan menganalisis data dalam data warehouse;
- **manajemen kinerja bisnis (BPM)** untuk memantau dan menganalisis kinerja
- **user interface** (misalnya, dashboard)

# Arsitektur tingkat tinggi BI



# Komponen dalam arsitektur BI

---

**data warehouse** adalah gudang data historis yang terorganisir dengan baik

**Analisis bisnis** adalah alat yang memungkinkan transformasi data menjadi informasi dan pengetahuan

**manajemen kinerja bisnis (BPM)** memungkinkan pemantauan, pengukuran, dan membandingkan indikator kinerja kunci

**User interface** (misalnya, dashboard) memungkinkan akses dan manipulasi komponen BI lainnya dengan mudah

# Ragam BI

---

MicroStrategy, Corp. membedakan lima gaya BI dan menawarkan alat untuk masing-masing

1. laporan pengiriman dan mengingatkan
2. Laporan perusahaan (menggunakan dashboard dan scorecard)
3. kubus analisis (juga dikenal sebagai *slice-and-dice* analisis)
4. ad-hoc query
5. statistik dan data mining

# Keuntungan BI

---

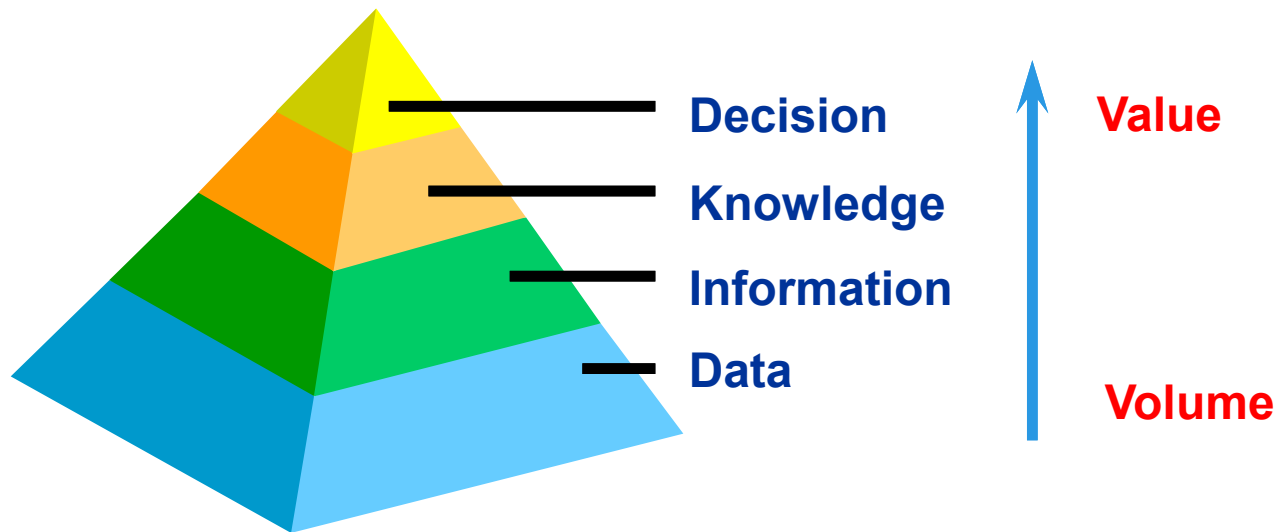
Kemampuan memberikan informasi yang akurat bila diperlukan, termasuk pandangan real-time dari kinerja perusahaan dan bagian-bagiannya

Sebuah survei yang dilakukan oleh Thompson (2004)

- Pelaporan lebih cepat, lebih akurat (81%)
- Peningkatan pembuatan keputusan (78%)
- Meningkatkan layanan pelanggan (56%)
- Meningkatkan pendapatan (49%)

# Tujuan Business Intelligence

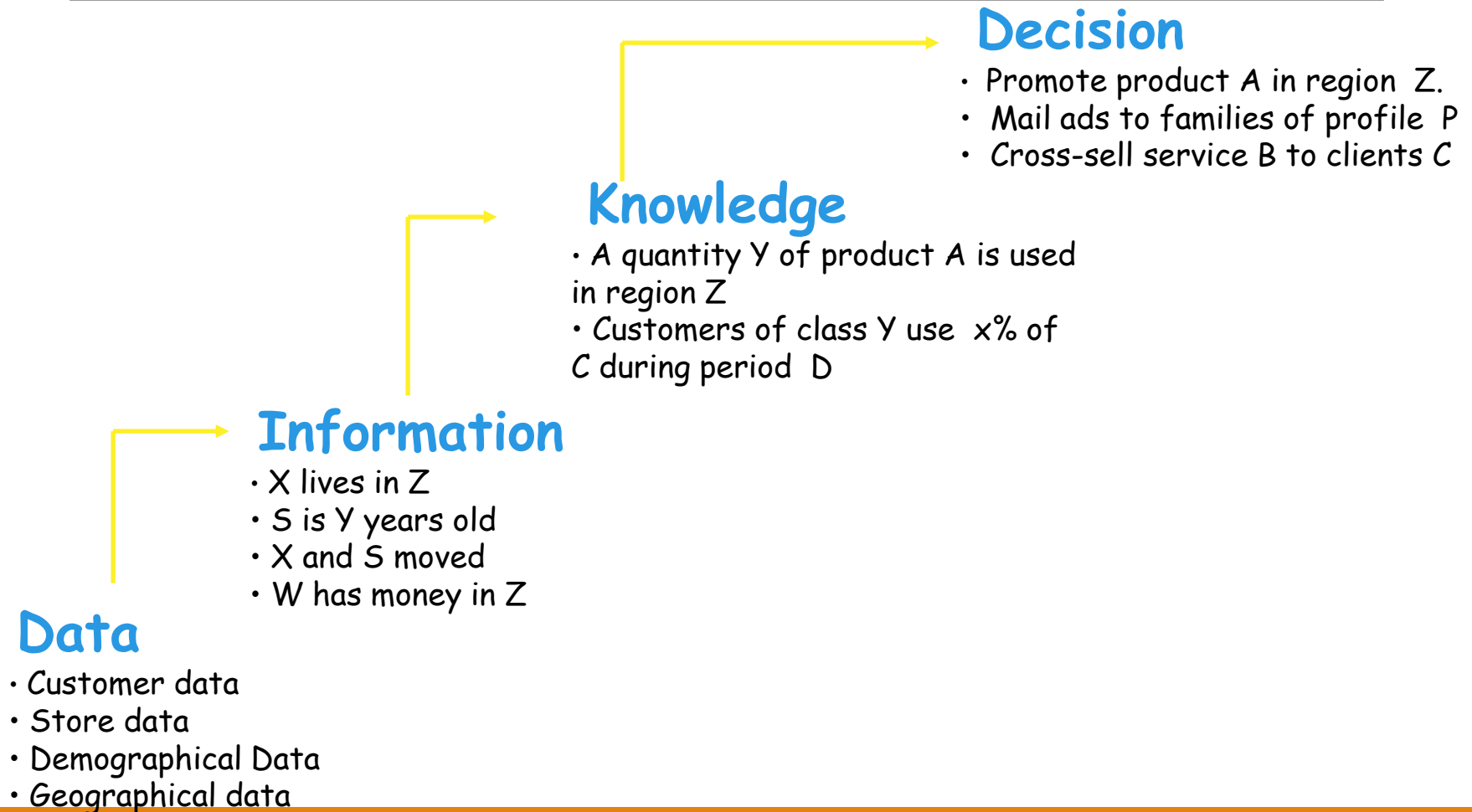
Tujuan *business intelligence* adalah mengkonversi sejumlah data menjadi nilai bagi pengguna.





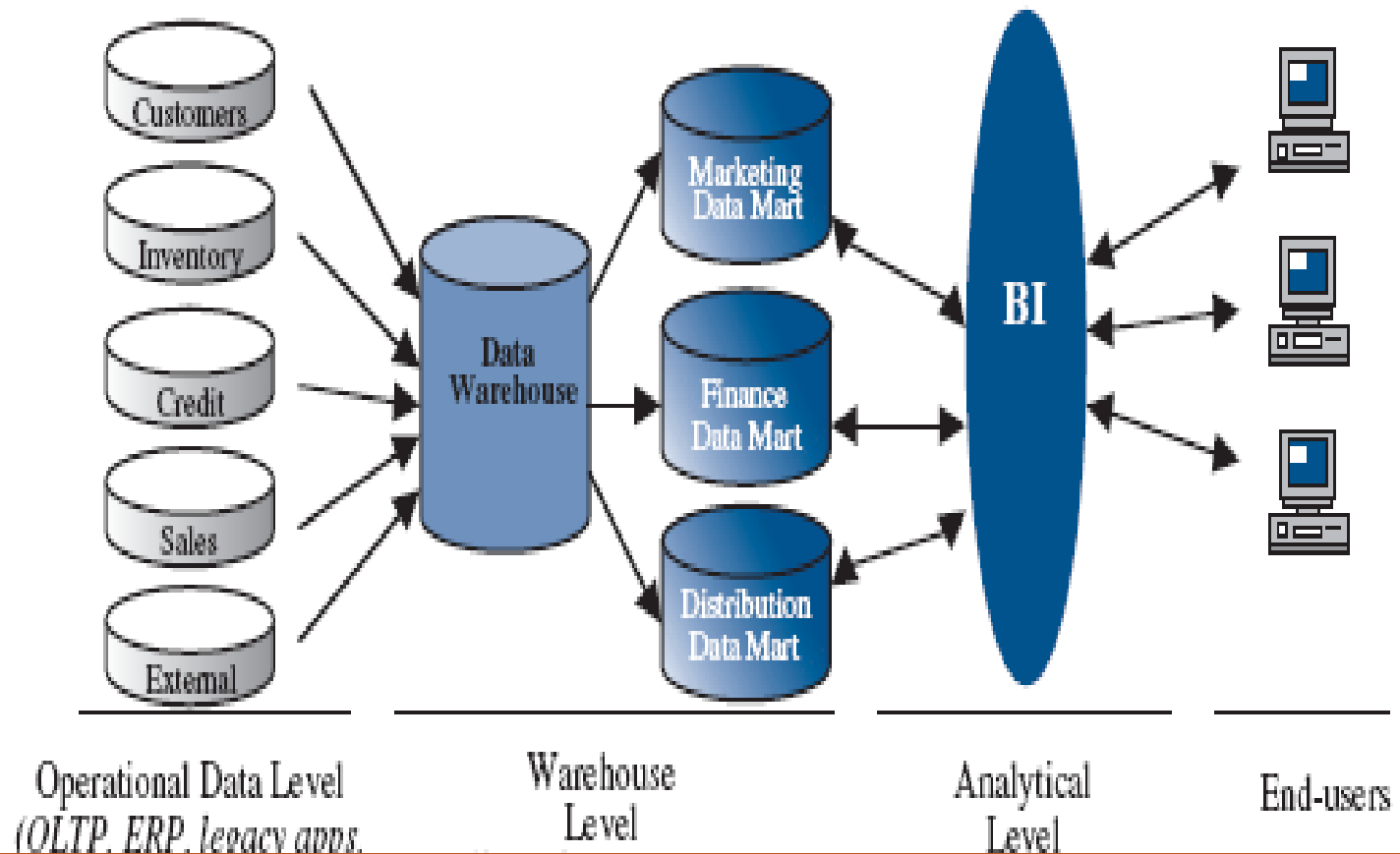
# The Data-decision *Value Chain*

---



# Business Intelligence

Business Intelligence Flow Chart



# MENGAPA BIS DIPERLUKAN ?

1. PARA EXECUTIVE MEMBUTUHKAN INFORMASI YANG BERSIFAT AGREGAT
  2. KEBUTUHAN PROSES KNOWLEDGE DAN ALAT ANALISIS DALAM MENGHASILKAN INFORMASI
  3. KETERBATASAN SISTEM SEBELUMNYA
  4. ORGANISASI INGIN LEBIH KOMPETITIF
-

# Kebutuhan Informasi Agregat (1)

SEORANG SALES EXECUTIVE :

1. INGIN MELIHAT SELURUH PENJUALAN SELAMA 3 TAHUN LALU, BULAN-BULAN YANG MEMILIKI PROFIT LEBIH BESAR XX PERSEN.
2. INGIN MENGETAHUI DAERAH MANA SAJA YANG PENJUALANNYA MELEBIHI YY PERSEN.
3. INGIN MELIHAT BAGAIMANA KEBERHASILAN TIM PENJUALAN PADA SUATU DAERAH TERTENTU DIBANDINGKAN DENGAN TIM PENJUALAN SEBELUMNYA. IA AKAN MEMASANGKAN ANTARA TERITORIAL DENGAN TIM YANG COCOK YANG MENGAKIBATKAN SUKSESNYA PENJUALAN.

# Kebutuhan Informasi Agregat (2)

SEORANG SALES EXECUTIVE :

4. INGIN MELIHAT KECENDERUNGAN PROFIT DARI PRODUK-PRODUK YANG DIJUAL UNTUK BEBERAPA TAHUN TERAKHIR.
  5. INGIN MENGETAHUI LIMA PRODUK YANG MEMILIKI RANGKING PENINGKATAN PROFIT TERTINGGI.
-

# Kebutuhan Proses dan Alat Analisis Bisnis

---

Perlunya alat yang menyediakan data agregat seperti *sum, min, max, and count*.

Perlunya alat yang memiliki kemampuan menghitung prosentase.

Perlunya alat yang menyediakan *trend analysis*.

Perlunya alat yang mampu *time-based analysis of data* (e.g., *year, month, week, day, special groupings*).

# Kebutuhan Sistem Baru

---

Membutuhkan tools, database, proses, dan fungsi yang mampu mentransformasi data dari satu bentuk ke bentuk lain yang sesuai dengan **kebutuhan** analisis bisnis.

Memerlukan Informasi pesaing:

- ❖ Produk/jasa pesaing, saat ini dan yang akan datang, khususnya barang yang sama atau serupa dengan produk/jasa kita.
- ❖ Kebijakan harga pesaing.
- ❖ Hubungan pesaing dengan pelanggan (*customers - their own, your shared*)
- ❖ *Market share* pesaing
- ❖ Pengeluaran biaya oleh pesaing (*manufacturing costs, advertising*)
- ❖ Hubungan pesaing dengan perusahaan kita.



# Organisasi ingin lebih kompetitif (2)

---

Memerlukan Informasi pelanggan:

- Consumer taste
- Kebutuhan konsumen
- Keinginan konsumen
- Kesukaan konsumen
- Sikap/Perilaku konsumen
- Kepuasan konsumen
- Tingkat konsumsi konsumen

Informasi Lingkungan :

Teknologi : teknik produksi terbaru.

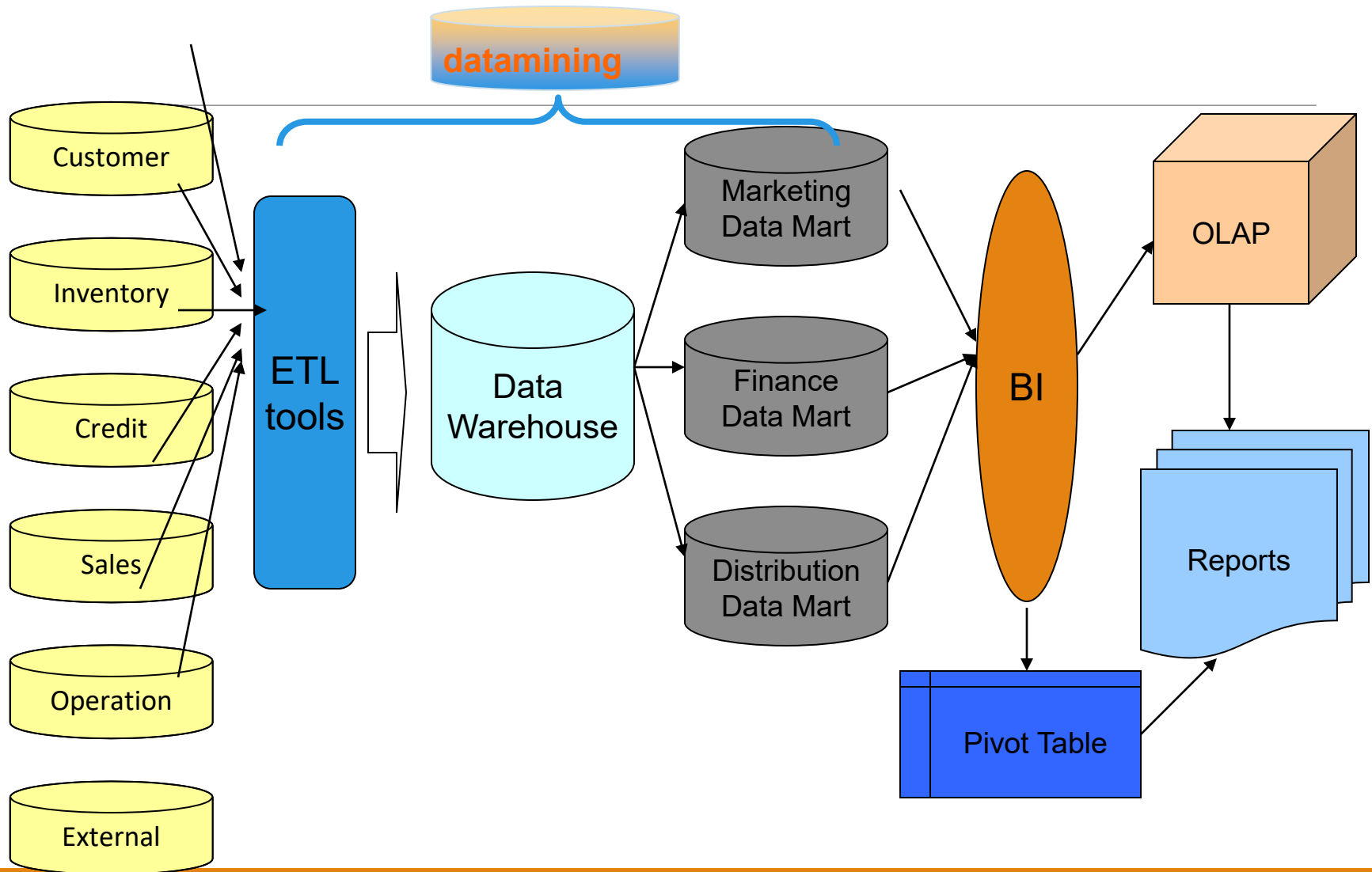
Peraturan : aturan cukai, kadar nikotin rokok, kadar alkohol minuman.

Dinamika pasar: dimana dan seberapa besar pasar baru yang ada?

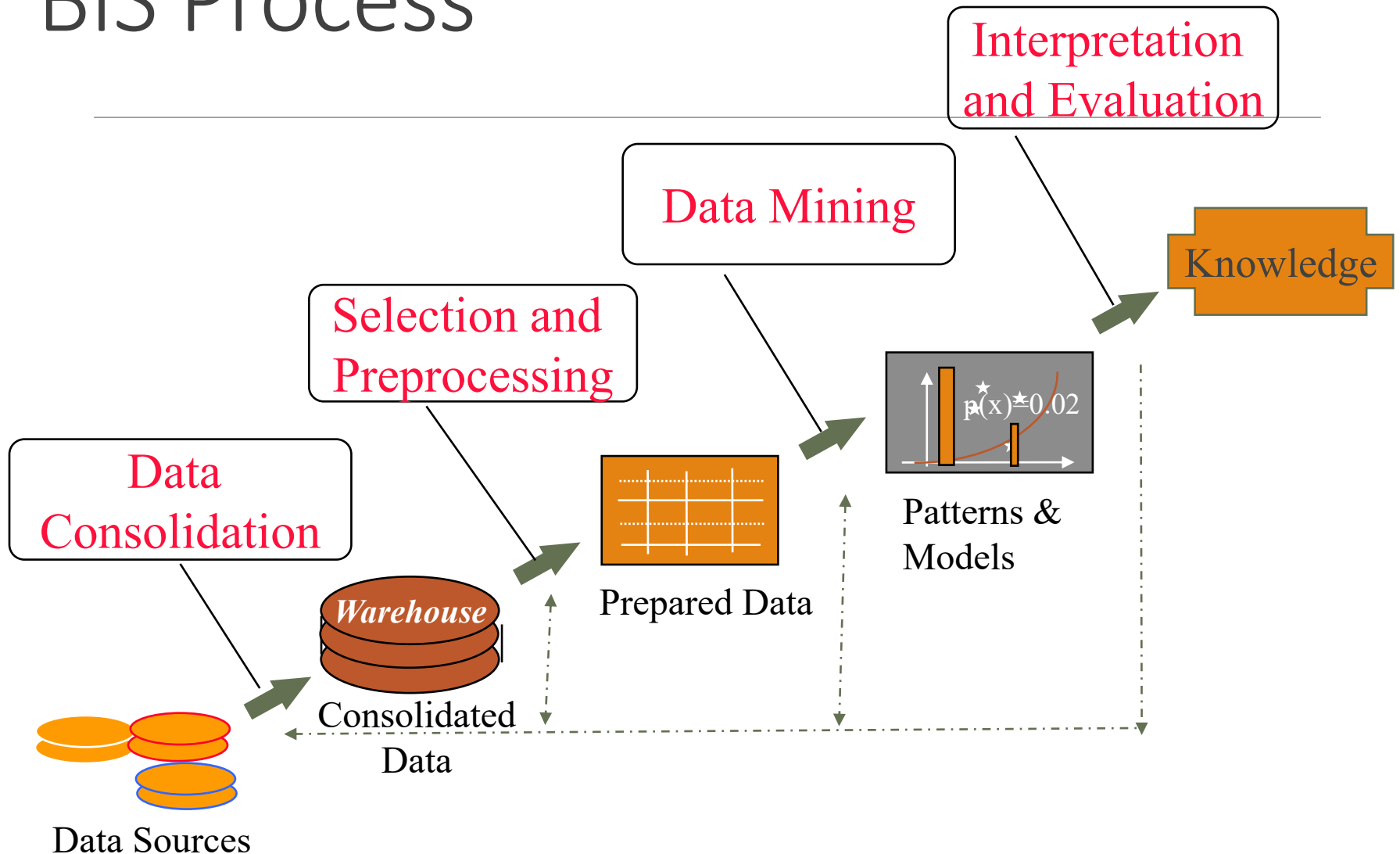
Perekonomian: tingkat bunga, tingkat pengangguran.

Lingkungan social: kondisi demografi.

# Business Intelligence Tools



# BIS Process



# Business Intelligence Tools

---

Operational Data Source

ETL tools (Extract, Transform, Load)

Data Warehouse

Data mart

Datamining

OLAP

# Operational Data Source

---

*Business Intelligence system* mengkoleksi data dari berbagai sumber termasuk operation database, OLTP, ERP, external database and etc.

# ETL tools (Extract, Transform, Load)

---

Digunakan untuk menarik data dari *source database*, mentransformasi data sehingga kompatibel dengan *data warehouse* kemudian menempatkan ke *data warehouse*.

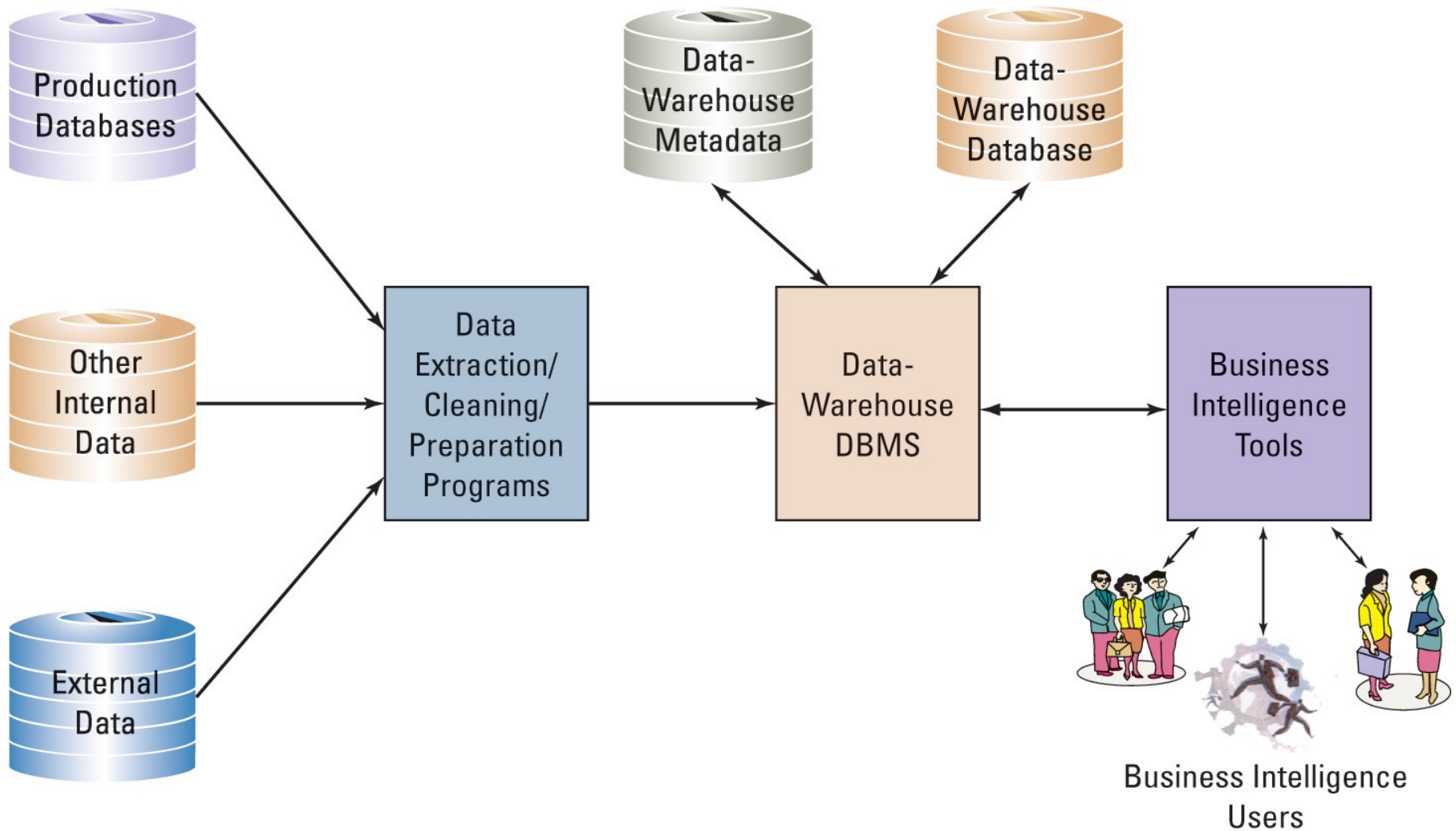
# What is Data Warehouse

---

**Data Warehouse** adalah koleksi data yang memiliki sifat *subject-oriented, integrated, time-variant, non-volatile* dalam mendukung proses pengambilan keputusan manajemen.



# Components of a Data Warehouse



# Data Mining

---

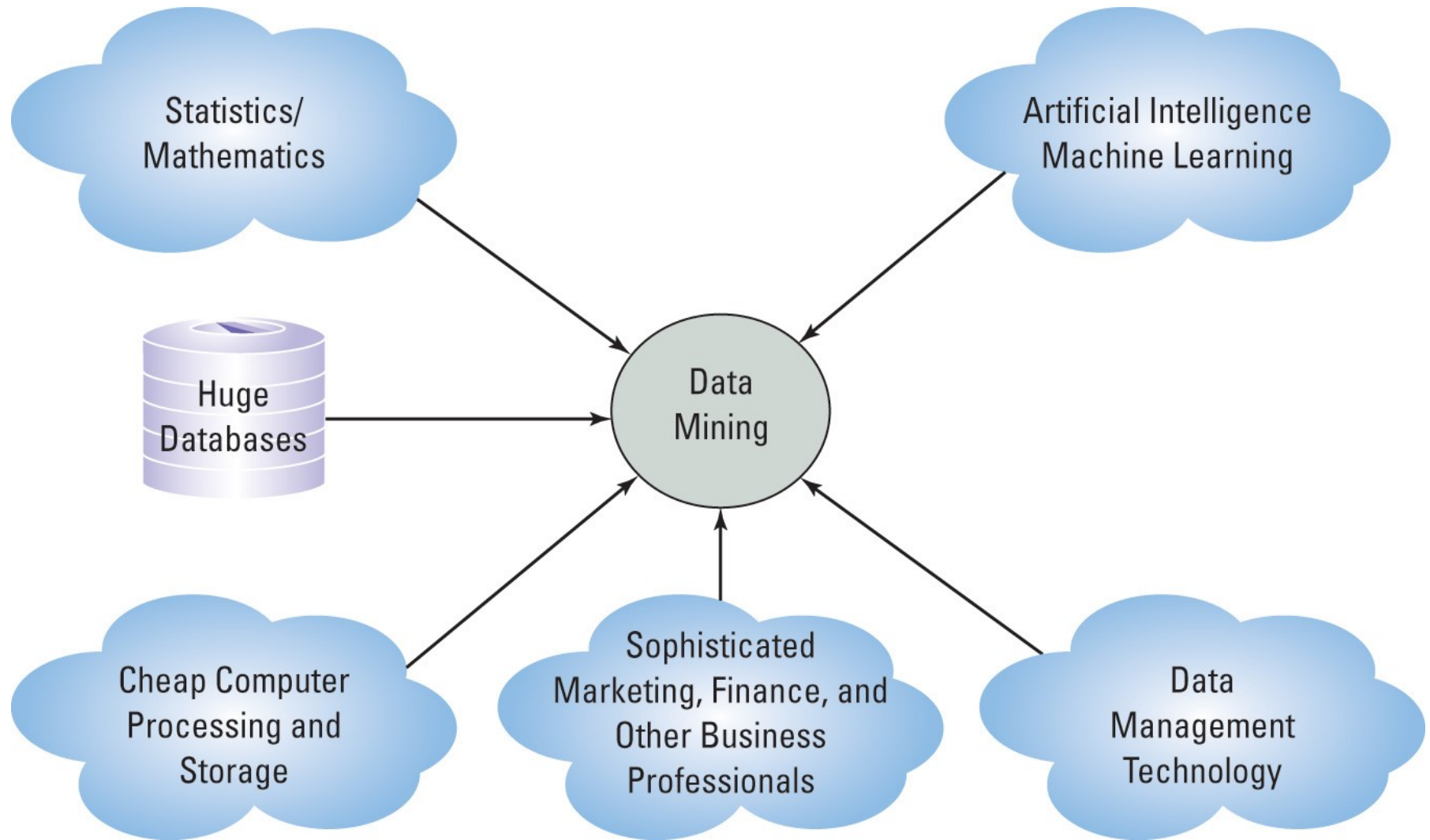
The application of statistical and mathematic techniques to find patterns and relationships among data for classifying and predicting

From artificial intelligence and machine-learning

Type

- Unsupervised data mining
- Supervised data mining

# Convergence Disciplines for Data Mining





## ... subject-oriented ...

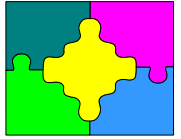
---

Data dalam warehouse didefinisikan dan diorganisasikan dalam terminologi bisnis, dan lebih dikelompokkan berdasar subyek bisnis, seperti:

- pelanggan
- produk
- sales
- daerah

dari pada transaksi individu.

*Normalization* is not relevant.



# integrated

---

Data warehouse bersisi data silang berbagai operasional fungsi/bagian perusahaan dan data eksternal

- Data dalam warehouse harus
  - Clean
  - validated
  - properly integrated

# An Example of Data Integration

## Checking Account System

Jane Doe (name)

Female (gender)

Bounced check #145 on 1/5/95

Opened account 1994

## Savings Account System

Jane Doe

F (gender)

Opened account 1992

## Investment Account System

Jane Doe

Owns 25 Shares Exxon

Opened account 1995

← *Operational data*

## Customer

Jane Doe

Female

Bounced check #145

Married

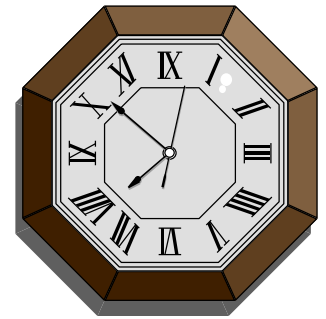
Owns 25 Shares Exxon

Customer since 1992

↑ *data warehouse*

## ... time-variant ...

---



Semua data dalam data warehouse ditandai waktu kapan dientry atau kapan diringkas dalam *warehouse*.

Kronologi perekaman data menunjukkan historis dan memungkinkan dilakukan *trend analysis*.

Memungkinkan untuk mengklasifikan berdasar waktu (bulan, tahun fiskal, periode, dsb.)

# ... nonvolatile ...



---

Sekali direkam kedalam data warehouse, data tidak di update.

Data menunjukkan sebagai sumber yang stabil, untuk pelaporan dan analisis perbandingan yang konsisten.

Kontradiksi dengan data operational yang bisa diupdate (*inserted, deleted, modified*).

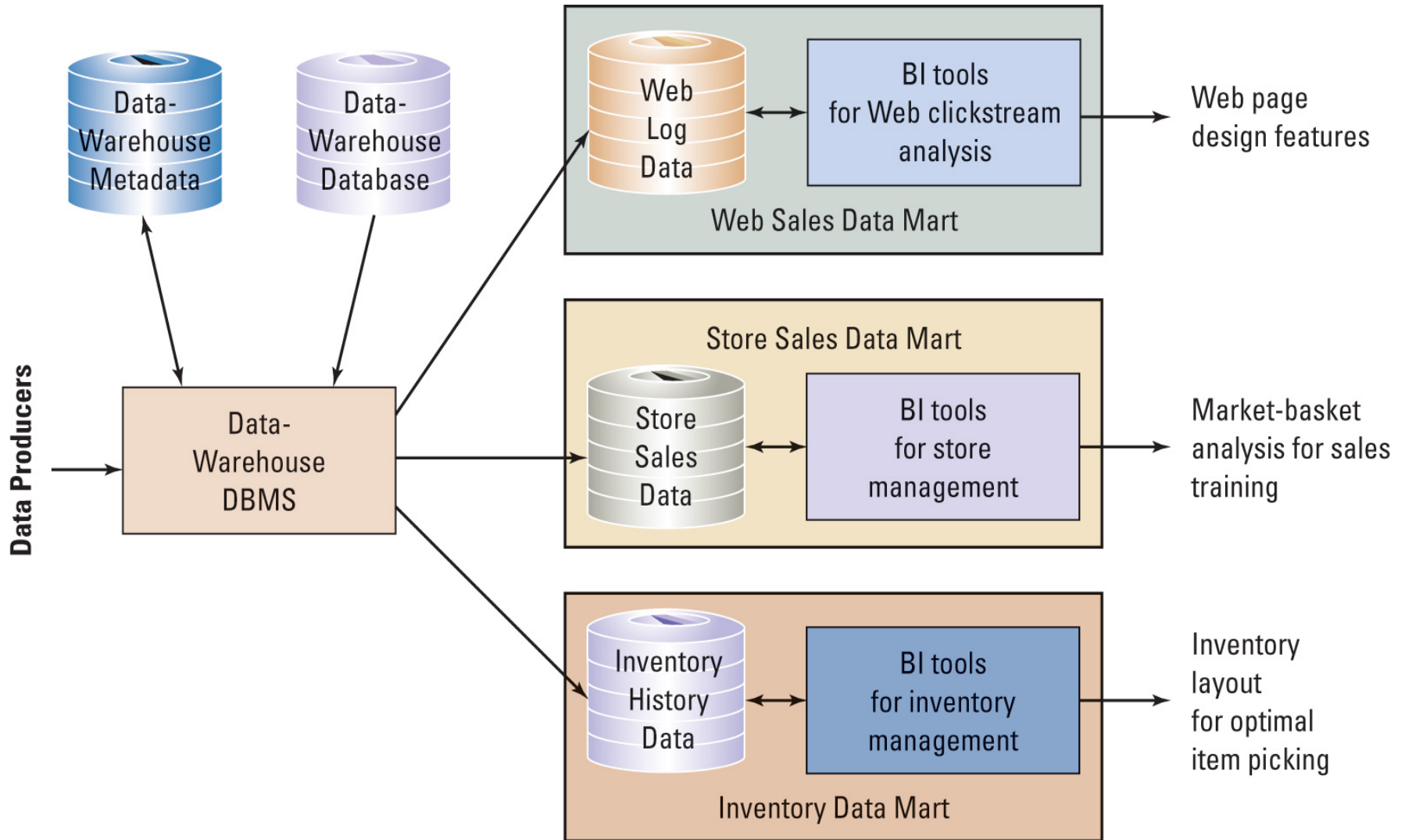


# Data Mart

---

Data mart adalah tempat penyimpanan data hasil dari penurunan data warehouse dan sumber lainnya yang didesain untuk melayani secara khusus komunitas atau fungsi tertentu.

# Data Mart Examples



# On-Line Analytical Processing (OLAP)

---

OLAP dikenalkan E.F. Codd (1993) untuk membedakan dengan *On-Line Transaction Processing* (OLTP)

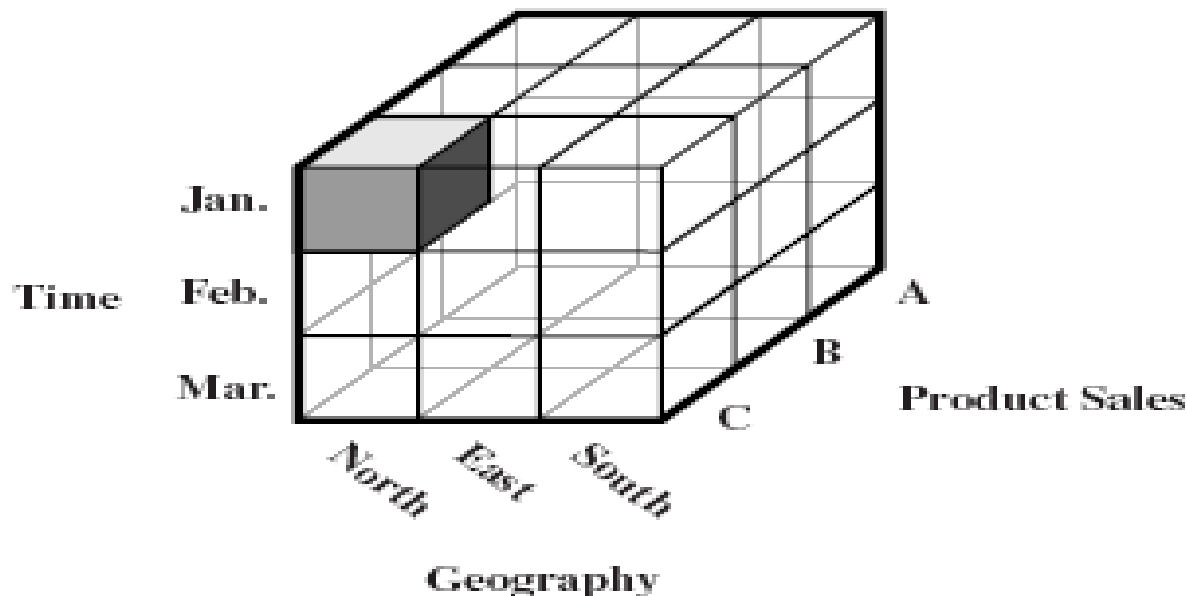
Definisi OLAP :

***“Kategori dari aplikasi dan teknologi yang melayani para analis, manajer dan eksekutif untuk mendapatkan informasi secara cepat, konsisten, akses interaktif, berbagai sudut pandang yang luas hasil transformasi raw data yang merefleksikan multidimensi riil perusahaan”***

# On-Line Analytical Processing (OLAP)

Ide Dasar: user ingin dapat memanipulasi model data perusahaan ke berbagai dimensi(multidimensi) untuk memahami perubahan-perubahan yang sedang terjadi.

## OLAP or multidimensional analysis



# Data-Driven BIS

---

Model aliran akses sistem informasi BI :

- “drill down” for more detailed information
- “drill up” for broader, more summarised view
- “slice and dice” for a change in data dimensions

# Data-Driven BIS

[illegible]

# Data-Driven BIS

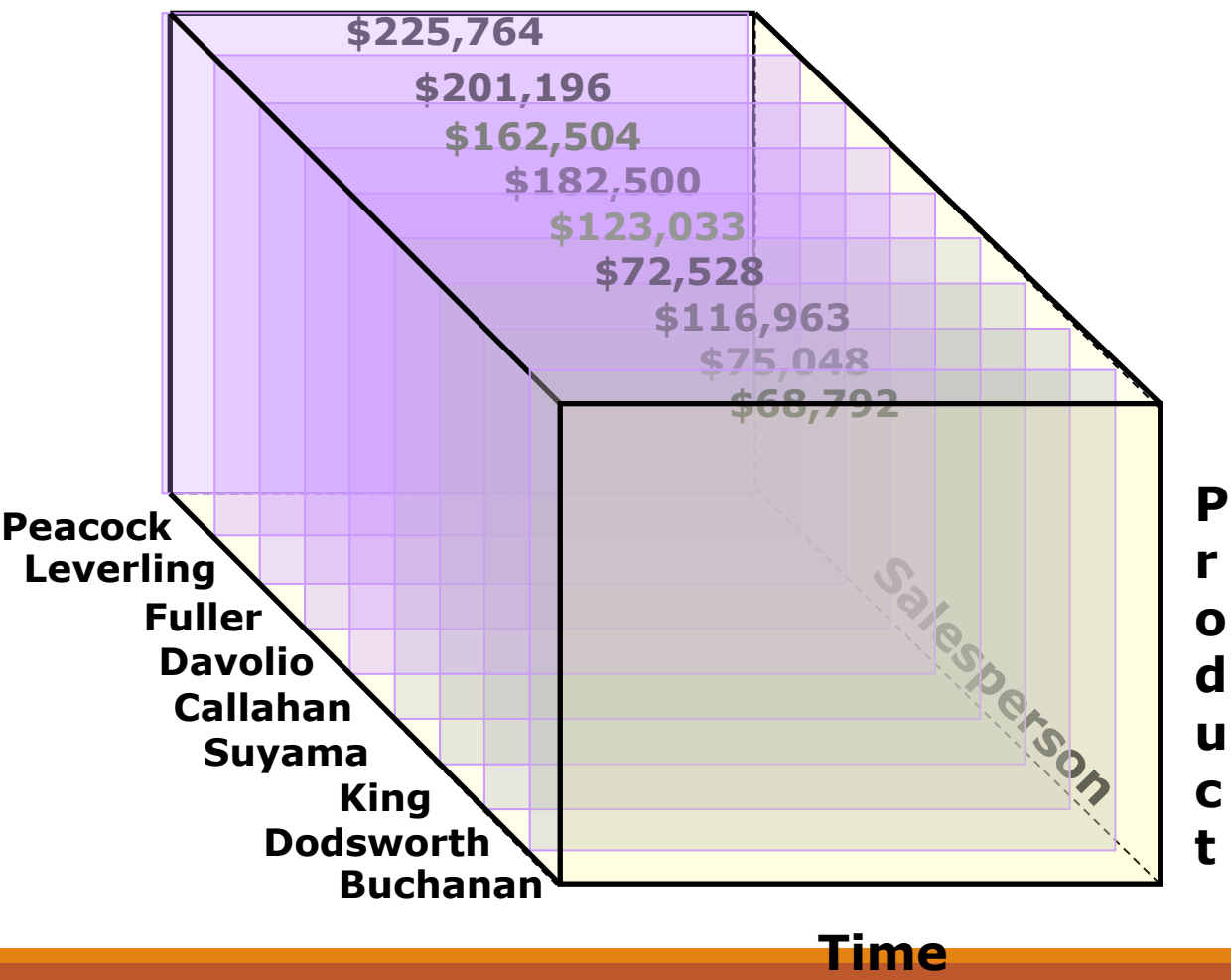
A	B	C	D	E	F	G	H	I	Formula B
Country	Salesperson	Order Date	OrderID	Order Amount					
UK	Buchanan	16/07/2003	10248	\$440.00					
UK	Suyama	10/07/2003	10249	\$1,863.40					
USA	Peacock	12/07/2003	10250	\$1,552.60					
USA	Leverling	15/07/2003	10251	\$654.06					
USA	Peacock	11/07/2003	10252	\$3,597.90					
USA	Leverling	16/07/2003	10253	\$1,444.80					
UK	Buchanan	23/07/2003	10254	\$556.62					
UK	Dodsworth	15/07/2003	10255	\$2,490.50					
USA	Leverling	17/07/2003	10256	\$517.80					
USA	Peacock	22/07/2003	10257	\$1,119.90					
USA	Davolio	23/07/2003	10258	\$1,614.88					
USA	Peacock	25/07/2003	10259	\$100.80					
USA	Peacock	29/07/2003	10260	\$1,504.65					
USA	Peacock	30/07/2003	10261	\$448.00					
USA	Callahan	25/07/2003	10262	\$584.00					
UK	Dodsworth	31/07/2003	10263	\$1,873.80					
UK	Suyama	23/08/2003	10264	\$695.62					
USA	Fuller	12/08/2003	10265	\$1,176.00					
USA	Leverling	31/07/2003	10266	\$346.56					
USA	Peacock	6/08/2003	10267	\$3,536.60					
USA	Callahan	2/08/2003	10268	\$1,101.20					
UK	Buchanan	9/08/2003	10269	\$642.20					
USA	Davolio	2/08/2003	10270	\$1,376.00					
UK	Suyama	30/08/2003	10271	\$48.00					
UK	Suyama	6/08/2003	10272	\$1,456.00					
USA	Leverling	12/08/2003	10273	\$2,037.28					

Salesperson	Order Amount
Buchanan	\$68,792
Callahan	\$123,033
Davolio	\$182,500
Dodsworth	\$75,048
Fuller	\$162,504
King	\$116,963
Leverling	\$201,196
Peacock	\$225,764
Suyama	\$72,528
<b>Grand Total</b>	<b>\$1,228,327</b>

# Data-Driven BIS

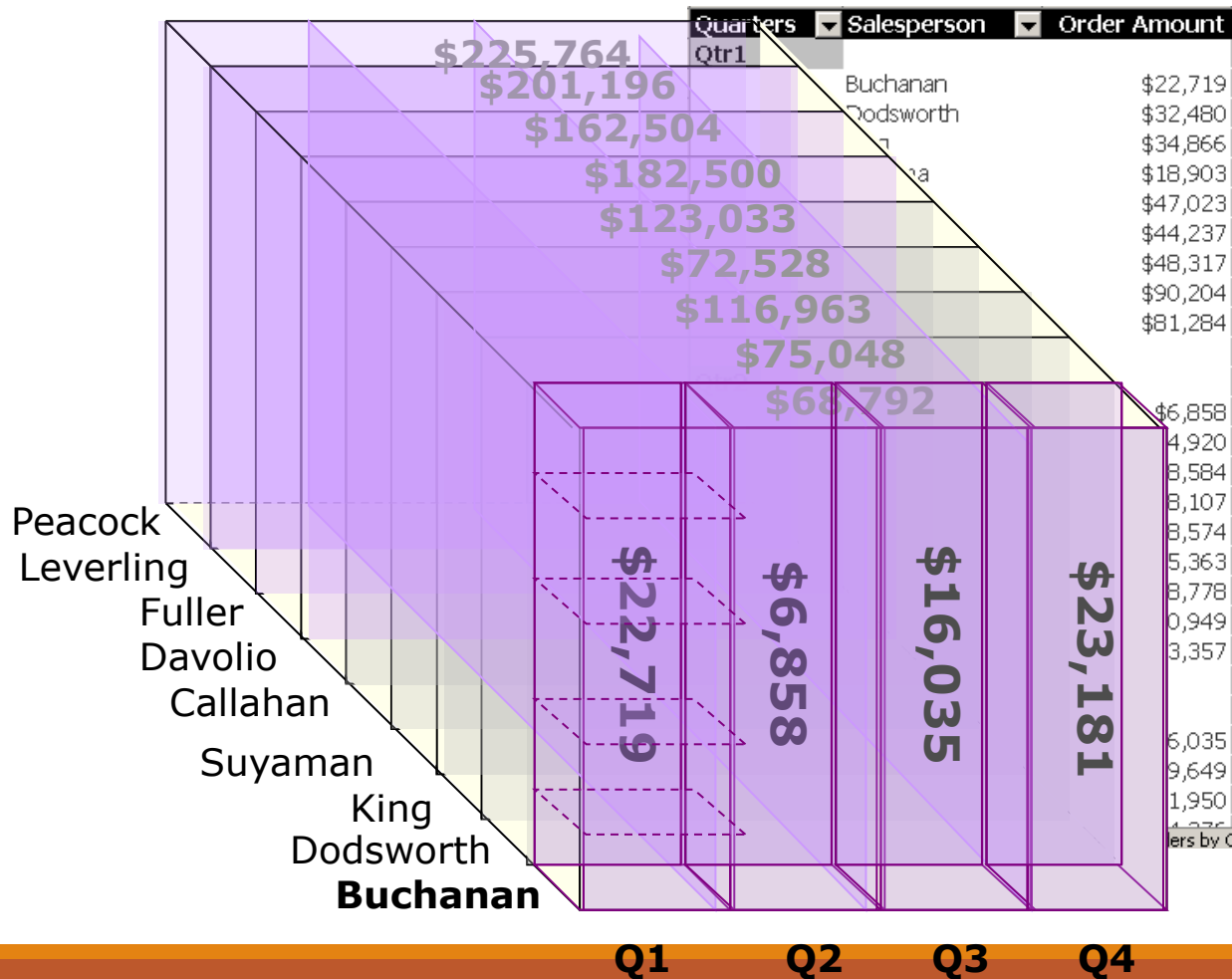
*"Slicing" the cube*



Salesperson ▼	Order Amount
Peacock	\$225,764
Leverling	\$201,196
Fuller	\$162,504
Davolio	\$182,500
Callahan	\$123,033
Suyama	\$72,528
King	\$116,963
Dodsworth	\$75,048
Buchanan	\$68,792
Grand Total	\$1,228,327



# Data-Driven BIS



# Dimensional Hierarchies

---

Setiap dimensi memiliki struktur hirarki



# What is Data Mining?

---

- Serangkaian aktivitas yang digunakan untuk menemukan sesuatu yang baru atau tersembunyi, atau sesuatu yang tidak terduga dari suatu data.
- Knowledge Discovery : Menggali data untuk mendapatkan temuan-temuan baru.

# Data Mining Tasks...

---

Classification [Predictive]

Clustering [Descriptive]

Association Rule Discovery [Descriptive]

Sequential Pattern Discovery [Descriptive]

Regression [Predictive]

Deviation Detection [Predictive]

# Classification: Definition

---

Diberikan koleksi *records* (*training set* )

- Masing-masing record berisi sejumlah atribut, salah satu atribut sebagai *class (dependen)*.

Temukan model untuk *class attribute* sebagai fungsi dari nilai-nilai atribut lain.

Tujuan: record yang sebelumnya tidak nampak dapat dimunculkan seakurat mungkin.

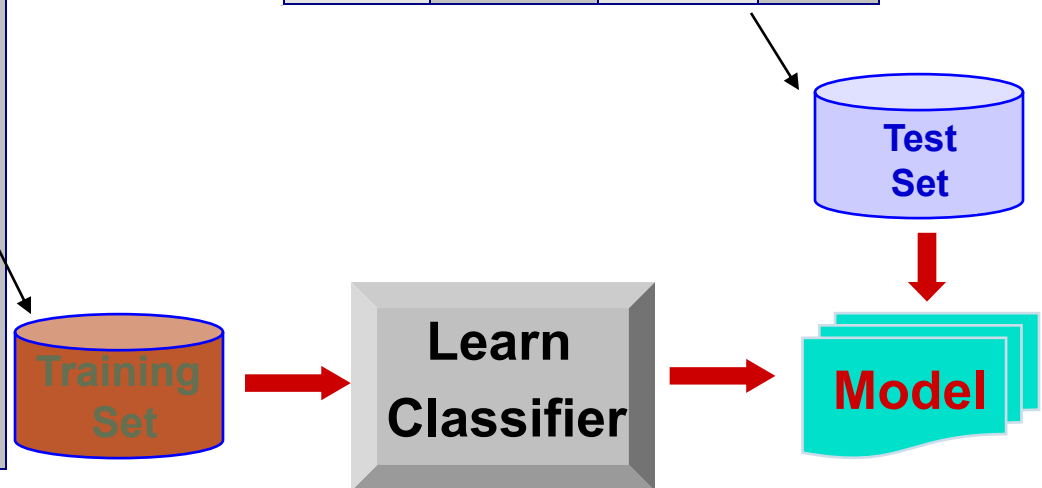
- sebuah *test set* digunakan untuk menguji akurasi model. Yaitu dengan menggunakan simulasi dengan memasukkan data pada *training set*,

# Classification Exam

categorical  
categorical  
continuous  
class

Tid	Refund	Marital Status	Taxable Income	Cheat
1	Yes	Single	125K	No
2	No	Married	100K	No
3	No	Single	70K	No
4	Yes	Married	120K	No
5	No	Divorced	95K	Yes
6	No	Married	60K	No
7	Yes	Divorced	220K	No
8	No	Single	85K	Yes
9	No	Married	75K	No
10	No	Single	90K	Yes

Refund	Marital Status	Taxable Income	Cheat
No	Single	75K	?
Yes	Married	50K	?
No	Married	150K	?
Yes	Divorced	90K	?
No	Single	40K	?
No	Married	80K	?



# Classification: Application 1

---

## Direct Marketing

- Tujuan: Mengurangi biaya pengenalan produk dengan menetapkan target konsumen yang memiliki kemungkinan besar membeli produk
- Approach:
  - Gunakan data pengalaman pengenalan produk yang serupa sebelumnya
  - Tetapkan keputusan konsumen *{buy, don't buy}* sebagai *class attribute*.
  - Kumpulkan data seperti *demographic, lifestyle*, tipe bisnis dan informasi lain yang berhubungan dengan konsumen.
    - Tipe bisnis, daerah tinggal, pendapatan, waktu transaksi dsb.
  - Gunakan informasi tersebut sebagai atribut input untuk *classifier model*.

# Classification: Application 2

---

## Fraud Detection

- Tujuan : Memprediksi pemalsuan transaksi kartu kredit.
- Approach:
  - Gunakan transaksi credit card dan information account-holder sebagai atribut.
  - Kapan nasabah biasa melakukan pembelian, apa saja yang dibeli, berapa sering membayar tepat waktu, dsb
  - Beri label historis transaksi yang palsu dan yang benar sebagai *class attribute*.
  - Buat model class dari transaksi.
  - Gunakan model untuk mendeteksi pemalsuan dengan observasi transaksi-transaksi credit card pada sebuah account.



# Clustering Definition

---

Diberikan sekelompok titik data yang masing-masing memiliki atribut, dan sebuah ukuran yang serupa, dapat ditemukan *clusters* sebagai berikut:

- Kelompokkan point data yang mirip menjadi satu cluster.
- Dan point data terpisah menjadi beberapa cluster lain.

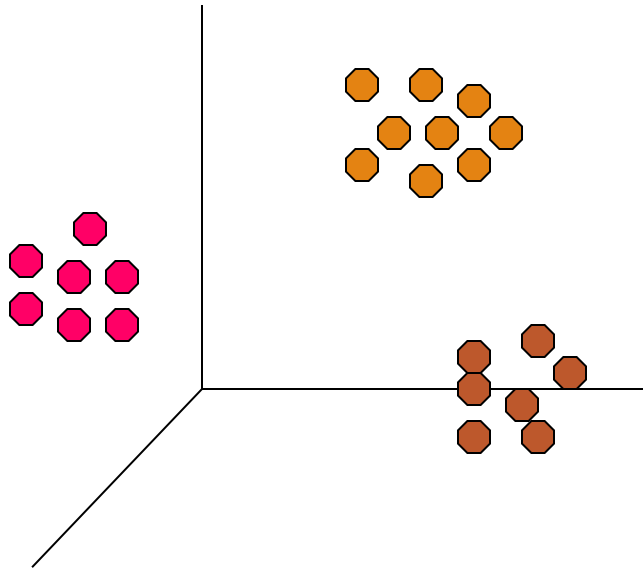
.

# Illustrating Clustering

---

Intracuster distances  
are minimized

Intercluster distances  
are maximized



# Clustering: Application

---

## Market Segmentation:

- Tujuan : memisah-misah sebuah pasar menjadi kelompok-kelompok konsumen, dimana beberapa kelompok terpilih sebagai target pemasaran.
- Approach:
  - Kelompokkan atribut yang berbeda dari konsumen berdasarkan geografi, gaya hidup dan informasi yang berkaitan.
  - Temukan cluster dari konsumen yang serupa.
  - Mengukur kualitas pemodelan dengan mengobservasi pola pembelian konsumen dalam cluster yang sama dan dibandingkan dengan cluster yang berbeda.

# Association Rule Discovery: Definition

---

Jika diberikan sekelompok records yang masing-masing berisi sejumlah item, maka;

- Diperoleh ketergantungan kejadian item yang berdasarkan pada kejadian item lainnya.

<i>TID</i>	<i>Items</i>
1	Bread, Coke, Milk
2	Beer, Bread
3	Beer, Coke, Diaper, Milk
4	Beer, Bread, Diaper, Milk
5	Coke, Diaper, Milk

Rules Discovered:

**{Milk} --> {Coke}**

**{Diaper, Milk} --> {Beer}**

# Association Rule Discovery: Application

---

## Manajemen Rak Supermarket

- Tujuan: Mengidentifikasi item yang yang dibeli secara bersama-sama oleh konsumen.
- Pendekatan: Memproses koleksi data penjualan untuk menemukan antar item yang saling ketergantungan.
- A classic rule --
  - Jika konsumen membeli *diaper dan milk*, maka ia juga membeli beer.
  - Sehingga perlu menempatkan rak beer secukupnya dekat Rak diaper maupun rak milk.

# Regression

---

Memprediksi suatu nilai dari variabel bernilai kontinu berdasarkan nilai variabel-variabel lainnya.

Examples:

- Memprediksi jumlah penjualan produk baru berdasarkan pengeluaran iklan.
- Memprediksi secara *time series* tentang persediaan

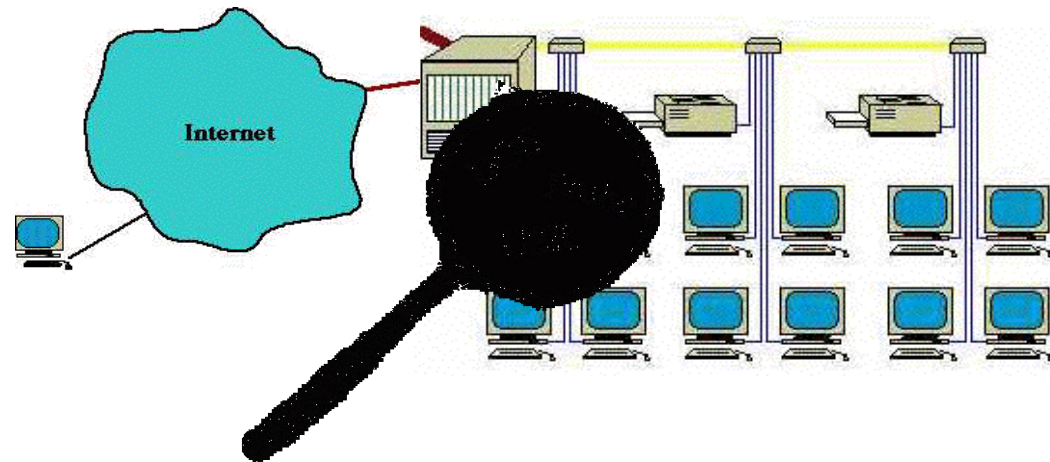
# Deviation/Anomaly Detection

---

Mendeteksi penyimpangan yang signifikan dari perilaku normal

Aplikasi :

- Deteksi penipuan (*Fraud*) Credit Card
- Deteksi penyusup jaringan



*Typical network traffic at University level may reach over 100 million connections per day*