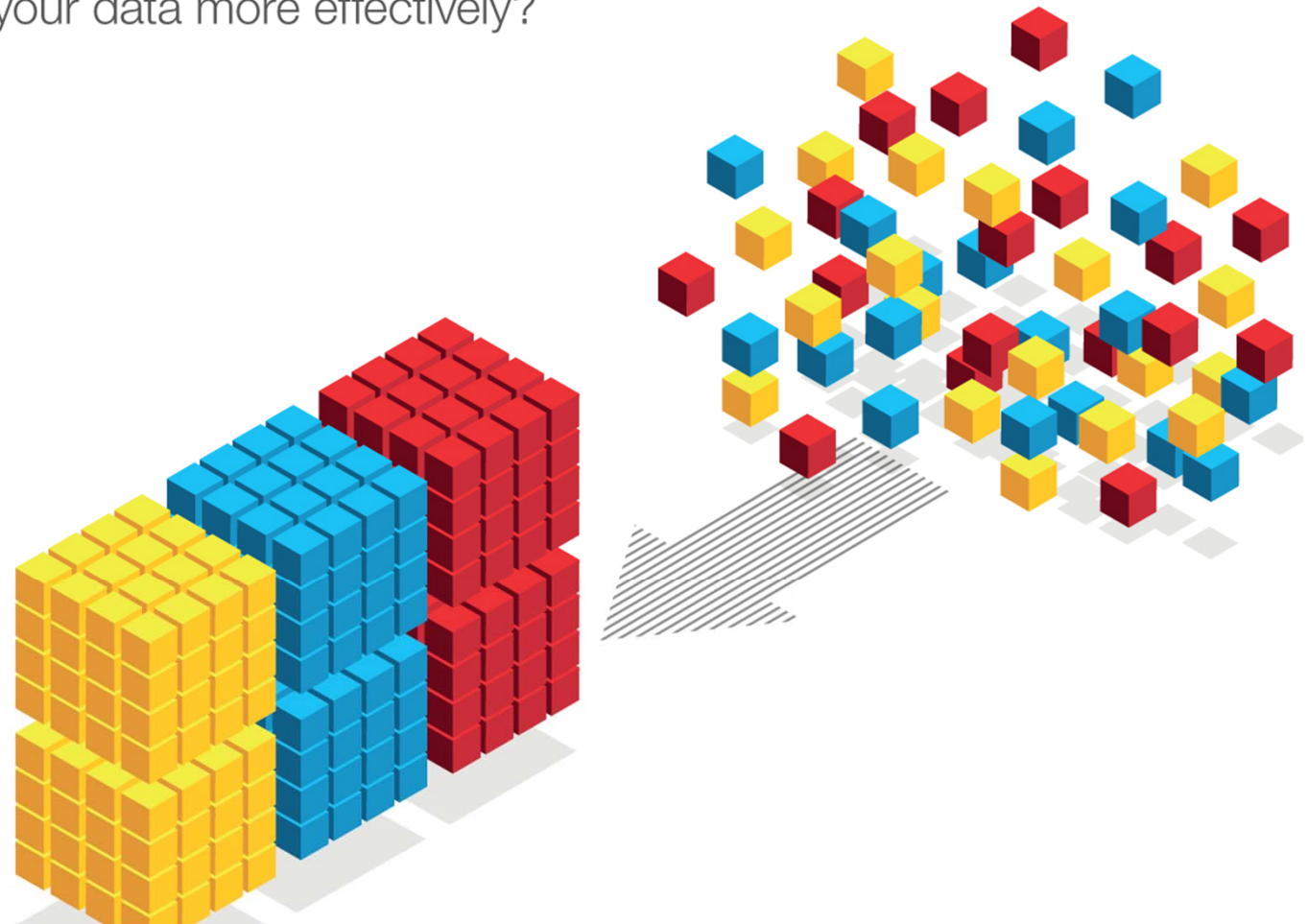


# Turn the data you have into the information you need

How will you manage your data more effectively?



**DEMO**

**Analytics Solution and Cubing Services  
with IBM InfoSphere Warehouse**

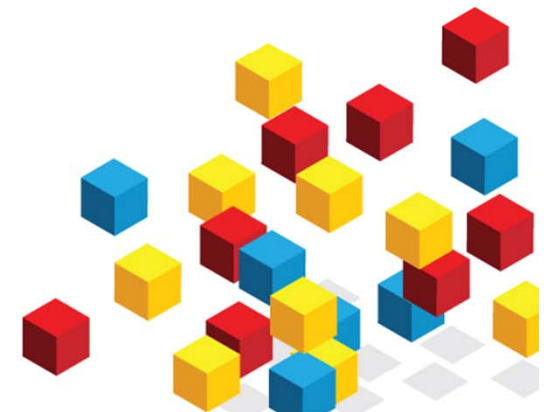


Vinnie Cardoso

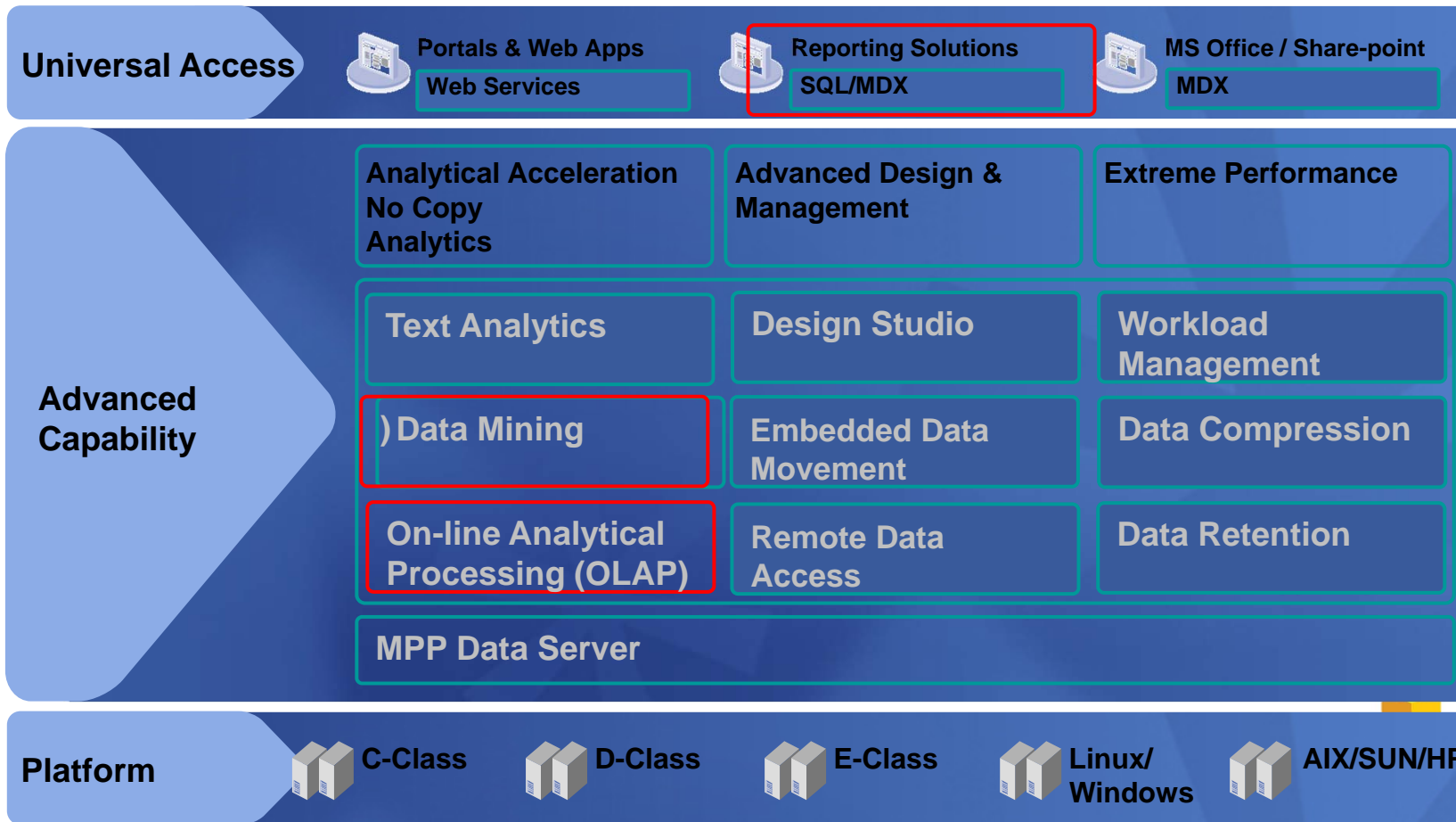
Data Management Technical Pre-sales

IBM SWG Group – A/NZ

13/07/2011

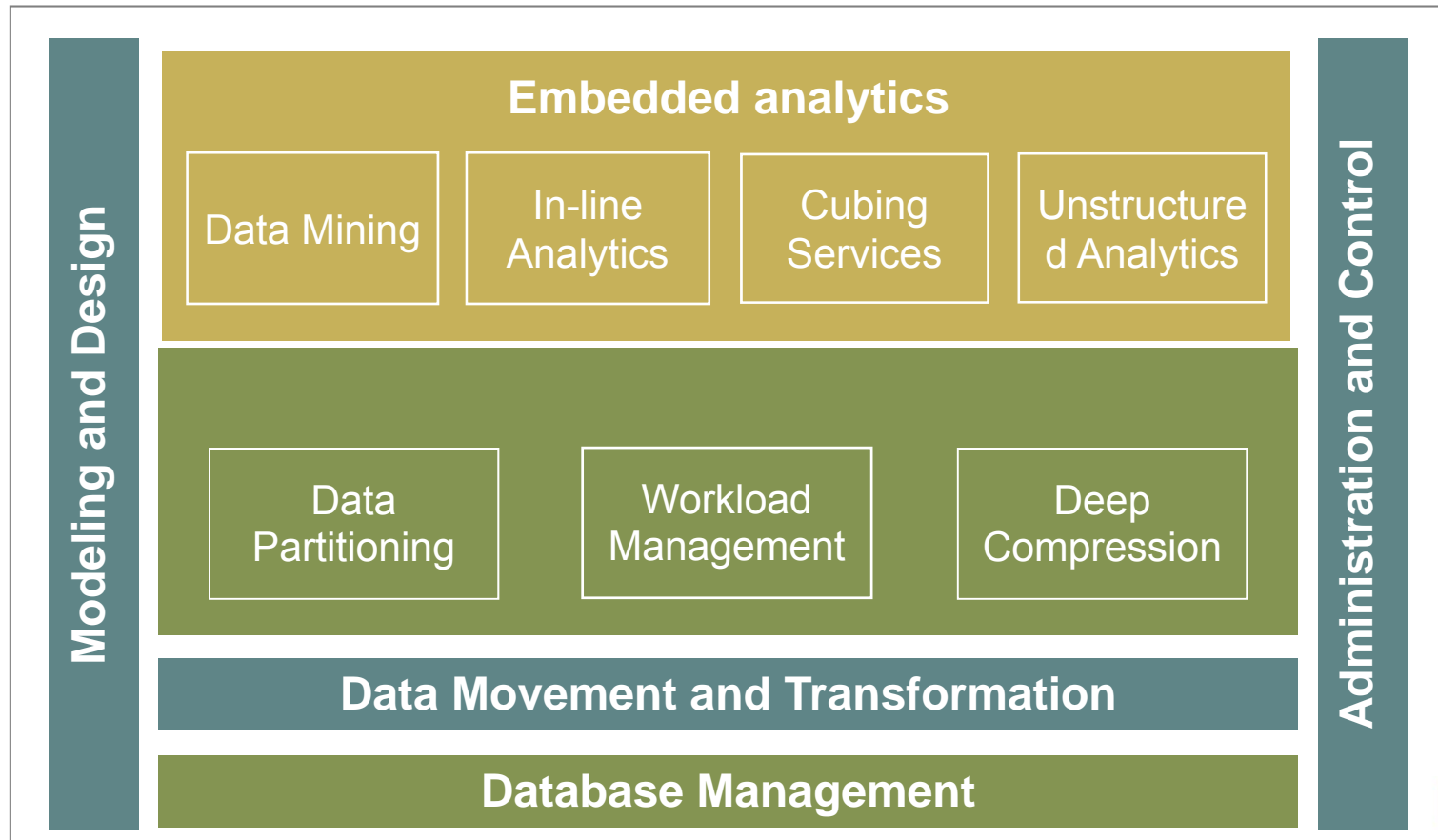


# InfoSphere Warehouse : Supporting Advanced Analytics

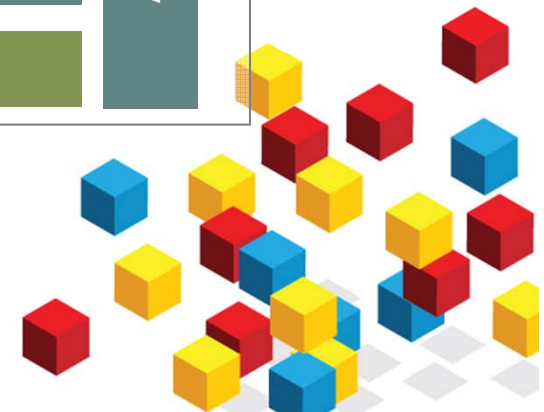


# IBM InfoSphere Warehouse

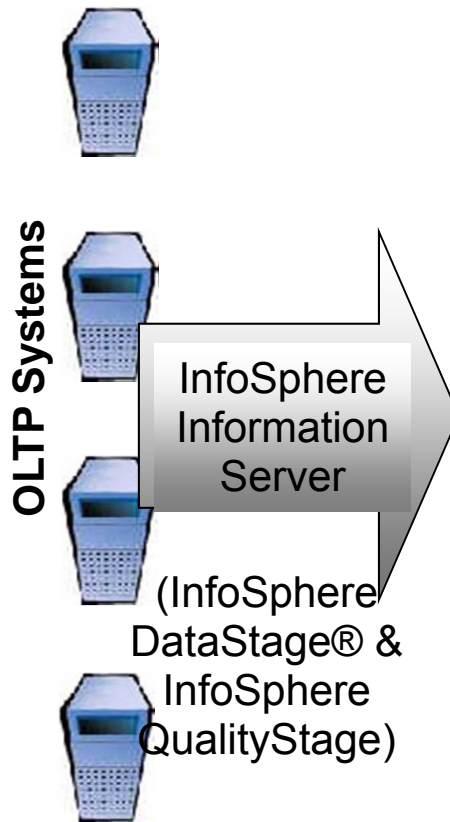
A comprehensive database and analytics solution



**IBM InfoSphere Warehouse v9.7**



# InfoSphere Warehouse Architecture



**IBM InfoSphere  
Warehouse**



# What is Data Mining?

**Data Mining is the art and science of discovering significant information and patterns in large databases.**



## Discovery mining

- **Automatically find trends and patterns**
  - ▶ Verify what is known now
  - ▶ Observe and interpret new insights
- **Combine mining methods for powerful insights**

## Methods

- **Clustering**
  - ▶ Find groups
- **Associations**
  - ▶ What entities are associated?
- **Sequences**
  - ▶ What entities follow others?

## Predictive mining

- **Predict an outcome based on historical behavior**
- **Directed analysis answering a specific question**
- **Readily interpretable prediction ingredients**

## Methods

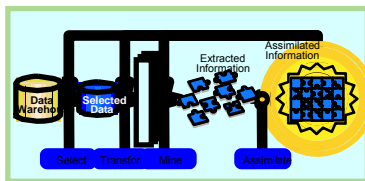
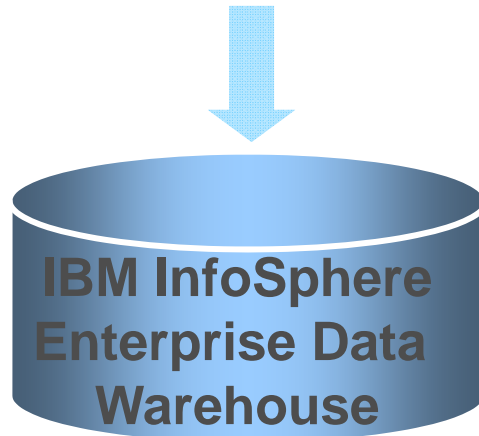
- **Classification**
  - ▶ Categorical (will event occur?)
- **Regression**
  - ▶ Numeric (predict cost?)



# Data mining methods in IBM InfoSphere Warehouse



Business Analyst



Statistician & Data Mining Workbench

- “Easy Mining” algorithms

- ▶ **Associations**

- Which item affinities (“rules”) are in my data?
- *[Beer => Diapers] ...single transaction*

- ▶ **Sequences**

- Which sequential patterns are in my data?
- *[Love] => [Marriage] => [Baby Products] ...sequential transactions*

- ▶ **Clustering**

- Which interesting groups are in my data?
- *...customer profiles, store profiles*

- ▶ **Classification**

- How to predict categorical values in my data?
- *...will the patient be cured, harmed, or unaffected by this treatment?*

- ▶ **Regression**

- How to predict numerical values in my data?
- *...how likely a customer will respond to the promotion*
- *...how much will each customer spend this year?*

- **Score data directly in IBM InfoSphere, scalable and real time**

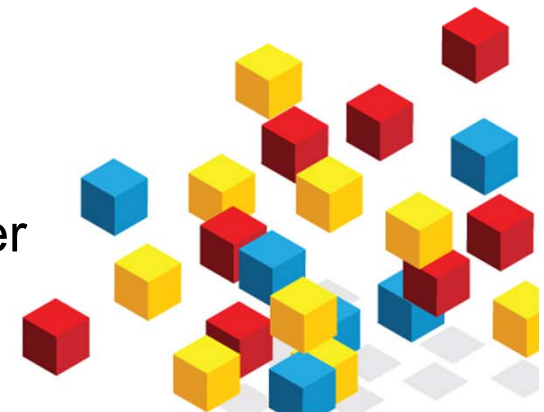


# Data Mining Benefits

**Data mining identifies valuable information quickly and efficiently to support business operations and decisions.**

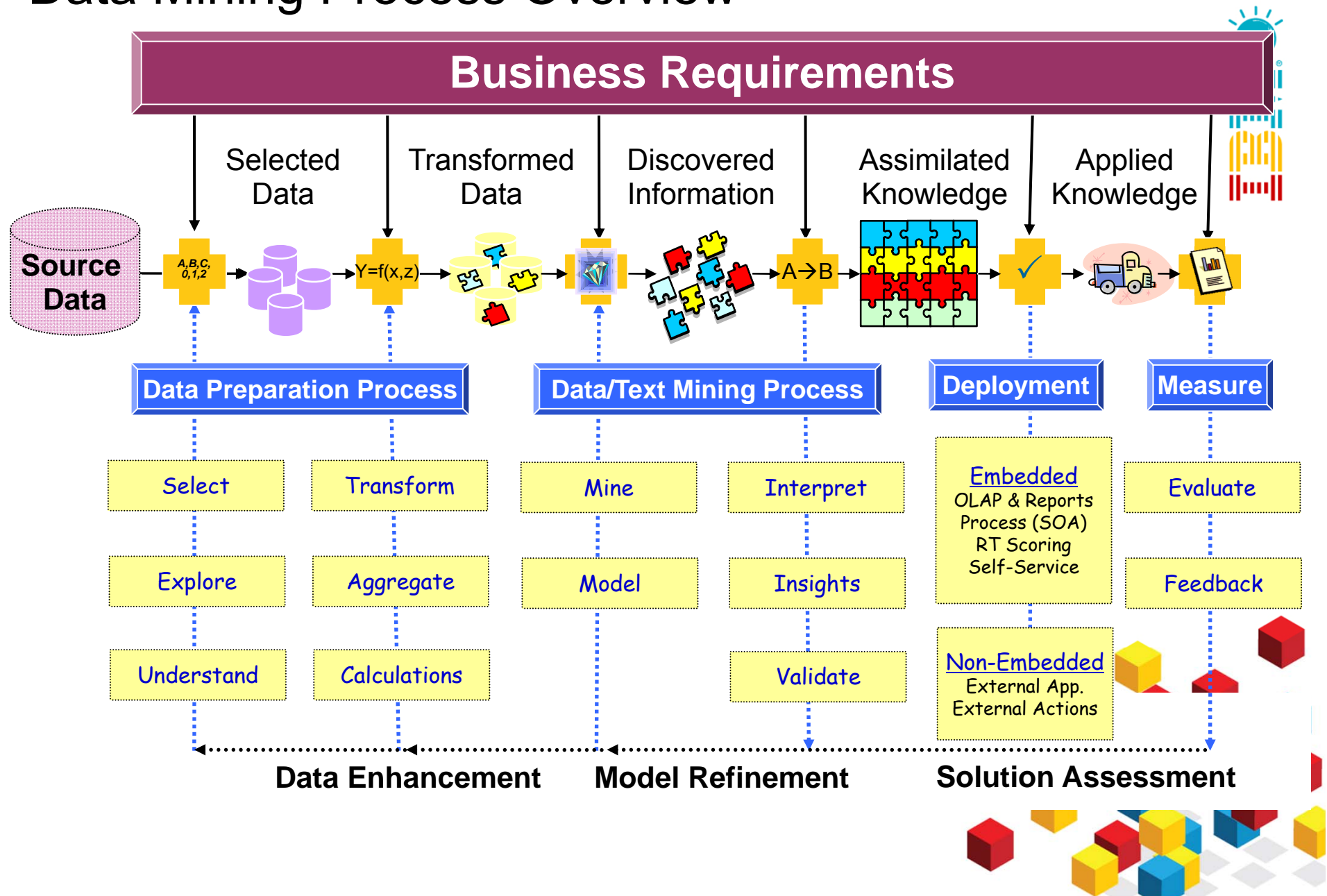


- **Analyze vast amounts of data**
  - Real-time, accurate, efficient
  - Automate the expert knowledge of the organization
  - Quickly identify rich set of data for further analysis
- **Create business value:**
  - Accelerate time-to-value
  - Expand your organization's analytics capabilities
  - Quickly discover new business opportunities
- **Democratize business information**
  - Enable more employees to make more and better decisions

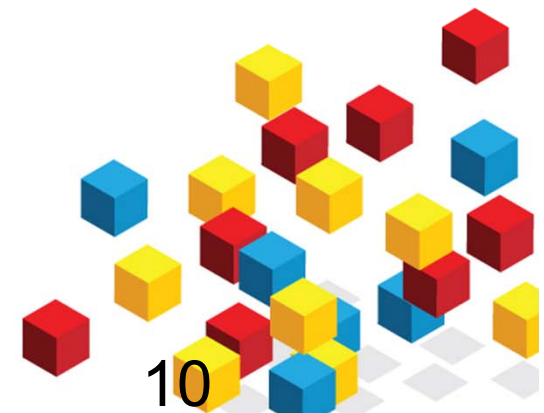
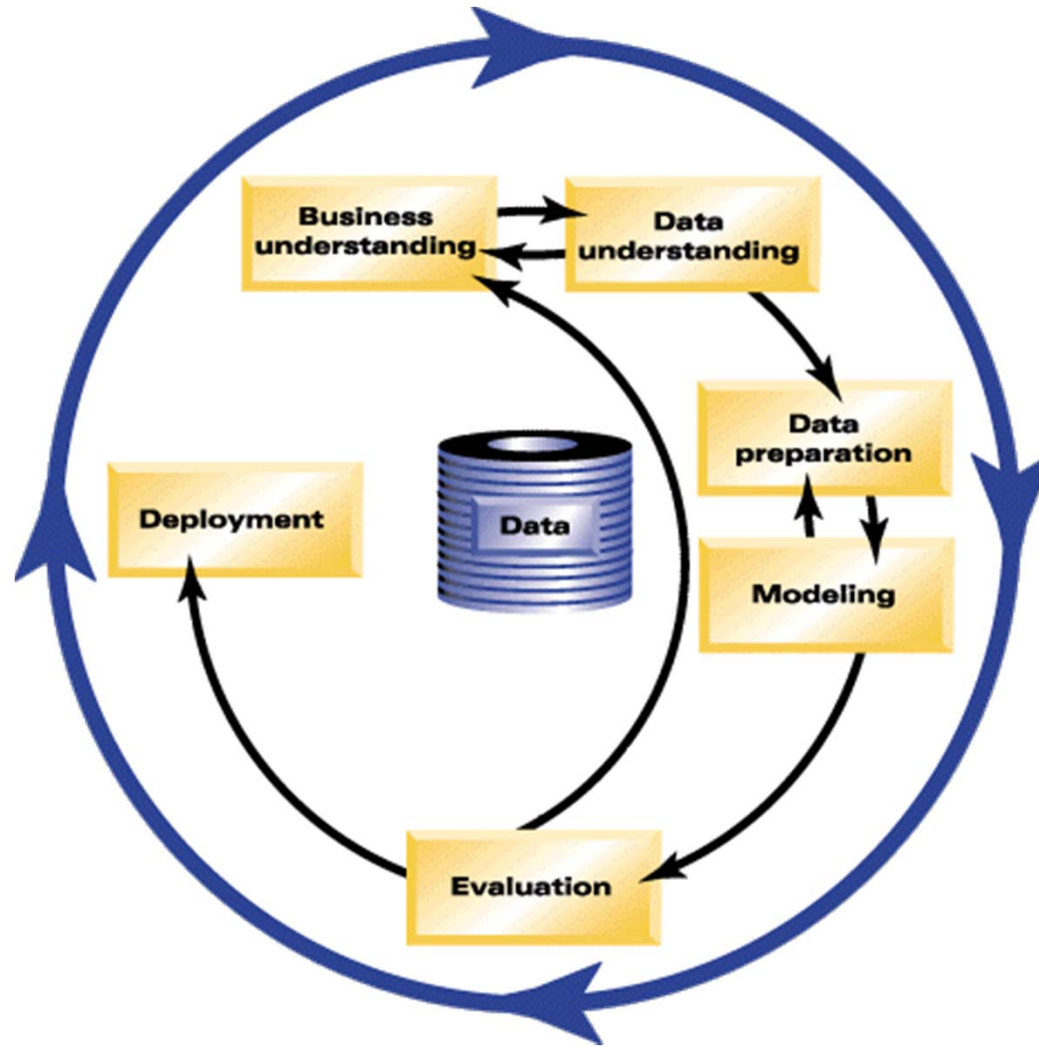




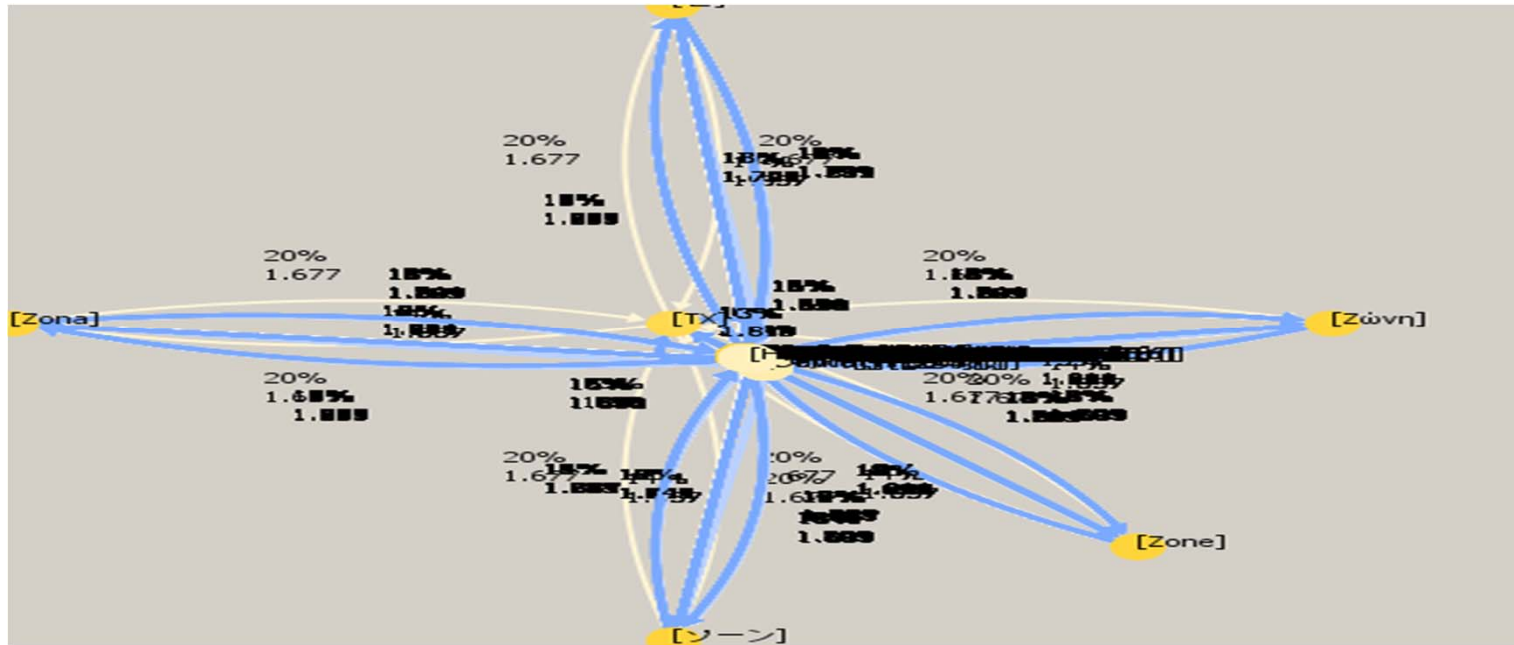
# Data Mining Process Overview



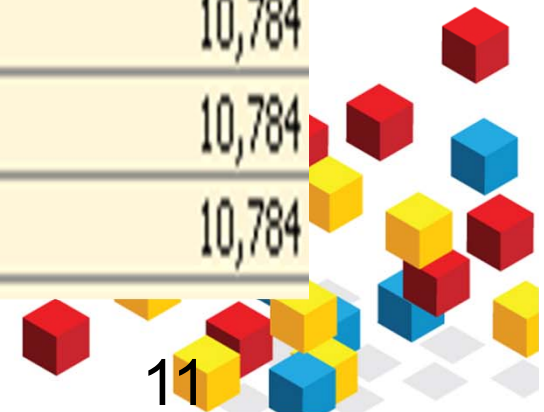
# Demo : CRISP-DM



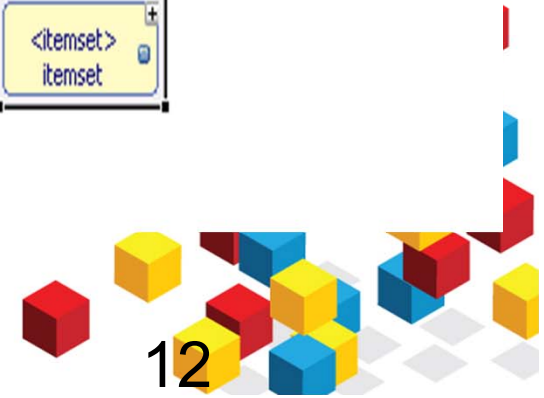
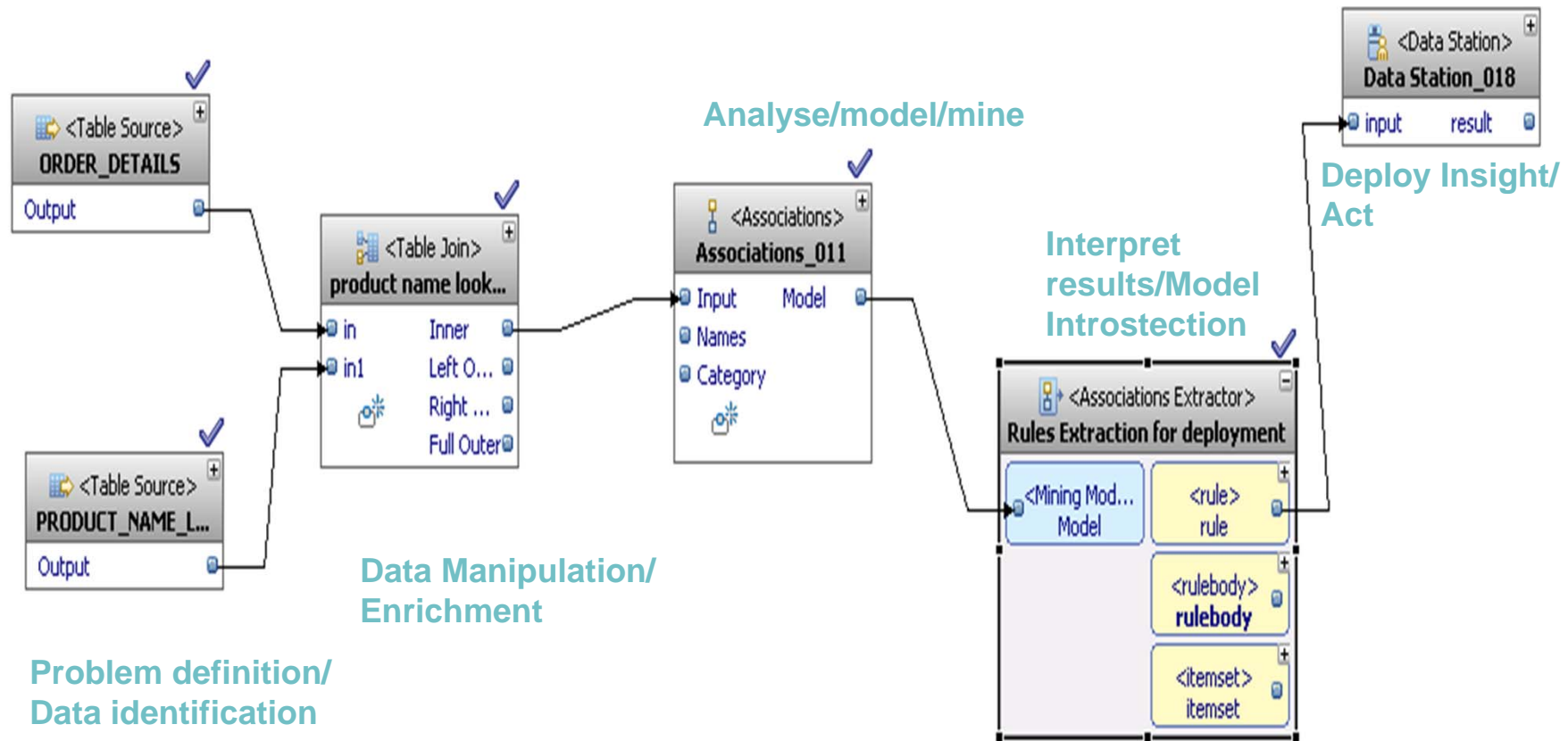
# Demo : Data Mining :Insight ->Action



7,975	[TX] ==> [Zone]	20.2494%	60.5843%	1.68	10,784
7,982	[존] ==> [TX]	20.2494%	56.0412%	1.68	10,784
7,978	[Zώνη] ==> [TX]	20.2494%	56.0412%	1.68	10,784
7,977	[TX] ==> [존]	20.2494%	60.5843%	1.68	10,784

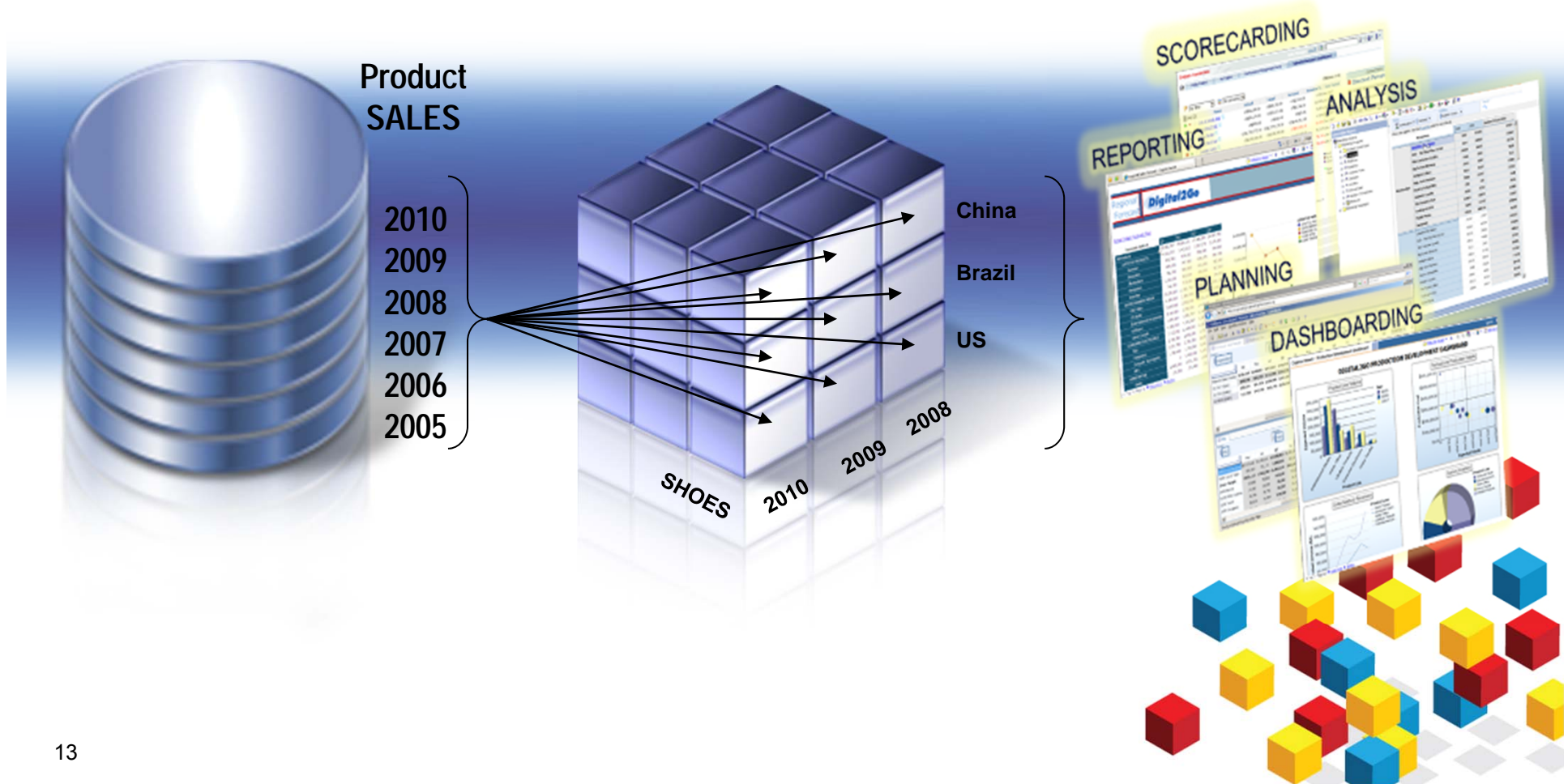


# Demo : Advanced Analytics



# Demo: Cubing Services

*Enables OLAP applications to access large data volumes stored inside the warehouse*



# What is OLAP? – Online Analytical Processing



## Gives structure to the schema data through metadata

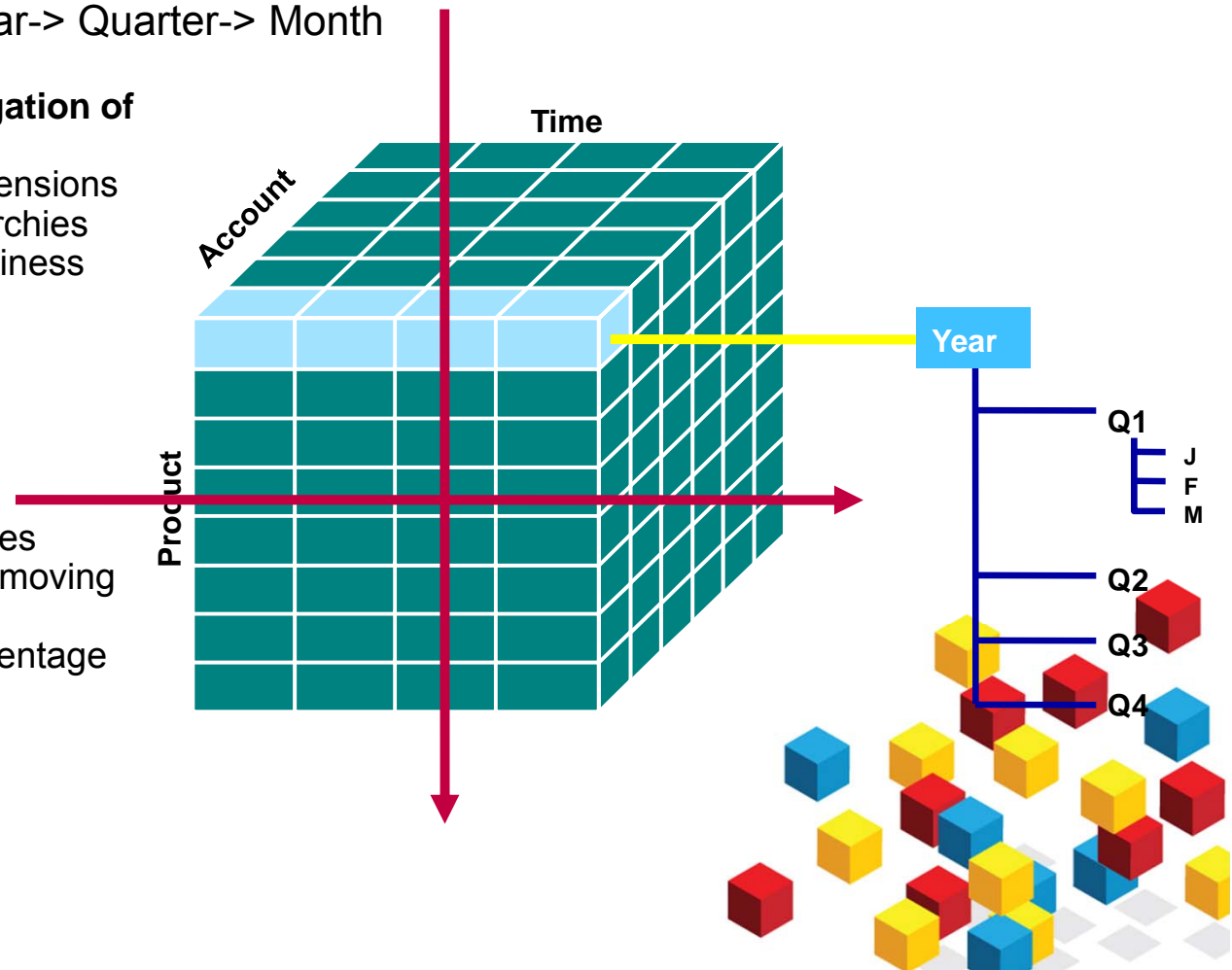
- Dimensions: Product, Geography, Time
- Dimensions have attributes: Products have colors, sizes, price ranges
- Dimensions have levels: Year-> Quarter-> Month

## Enables dynamic and easy navigation of the data

- Express queries in terms of dimensions
- Navigate using dimension hierarchies
- Identify key indicators using business terminology

## Enables definition and efficient calculations over the data

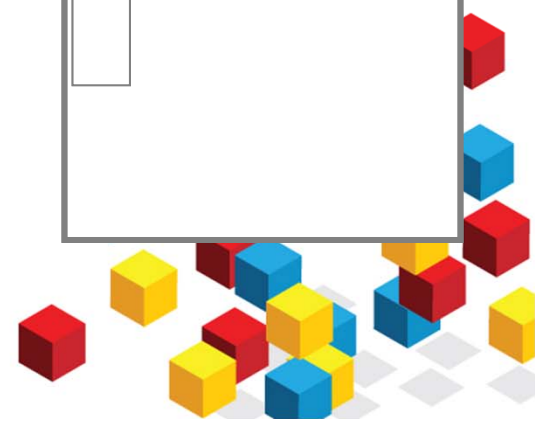
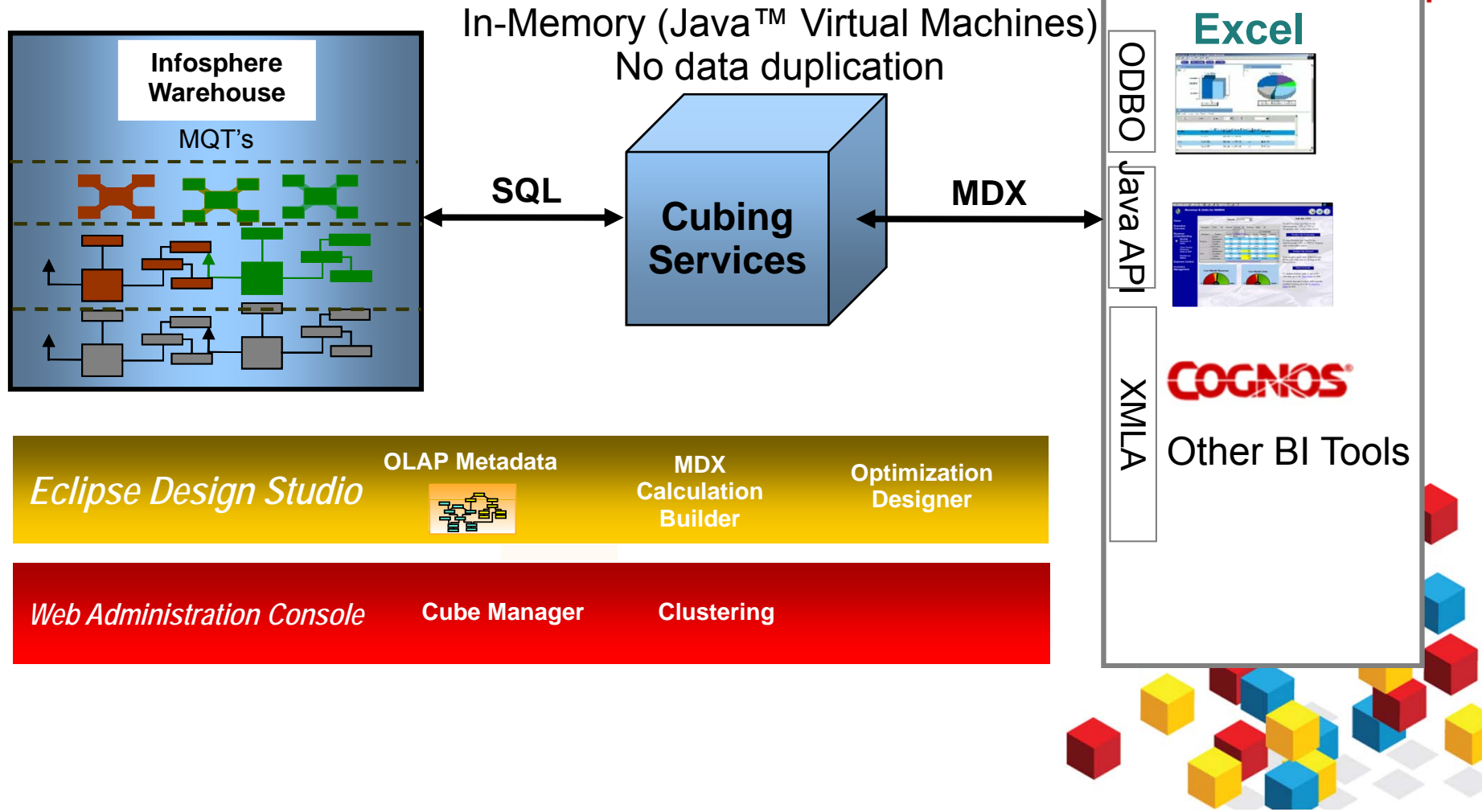
- Simple aggregations: sums, averages
- Time based calculations – 3 month moving averages
- Multi-pass calculations – rank, percentage of total



# What is InfoSphere Warehouse Cubing Services?



Multidimensional Analysis Server that enables Analytical applications access to Terabyte data volumes via industry standard OLAP connectivity



# Benefits Of Using Cubing Services

High performance OLAP layer for easy navigation through aggregates and ease of implementing complex calculations.

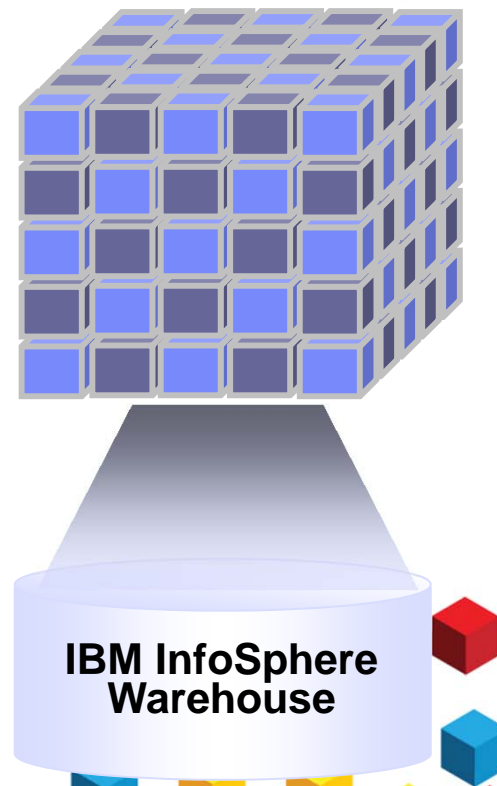
Uses industry standard MDX query language for access to a wide variety of tools for presentation and reporting

Empowers users with ad hoc access to business information.

What is the profitability for Product A across the Branches X,Y,Z?

Need-of-thought access to OLAP data:

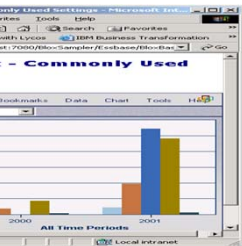
- OLAP and SQL shared access to the same information
- Single point of Management
- Single point of Maintenance
- Single point of Performance Tuning





# Supporting Services within BI Architecture.

Reporting Tool



Alphablox



Excel

A screenshot of an Excel spreadsheet. The table has columns for months (Jun, Jul, Aug, Sep, Oct, Nov, Dec) and a 'Totals' column. The rows represent different categories: 'Work Breakdown', 'SQL Warehouse', 'SQL Warehouse', 'Alphablox', 'Warehouse Manager', and 'Project Management'. Each cell contains numerical data.

Work Breakdown	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Totals
SQL Warehouse	8	8	8	8	8	8	8	64
SQL Warehouse	2	2	2	2	2	2	2	16
SQL Warehouse	6	6	6	6	6	6	6	48
Alphablox	2	2	2	2	2	2	2	16
Warehouse Manager	0	0	0	0	0	0	0	0
Project Management	2	2	2	2	2	2	2	16
Totals	18	18	18	18	18	18	18	144



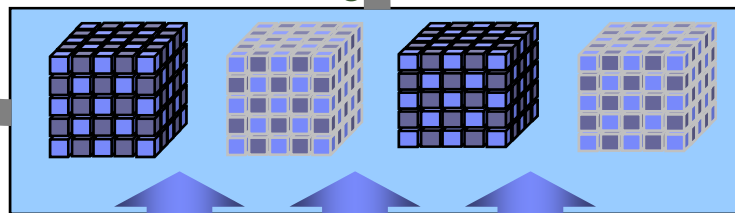
MDX

MDX

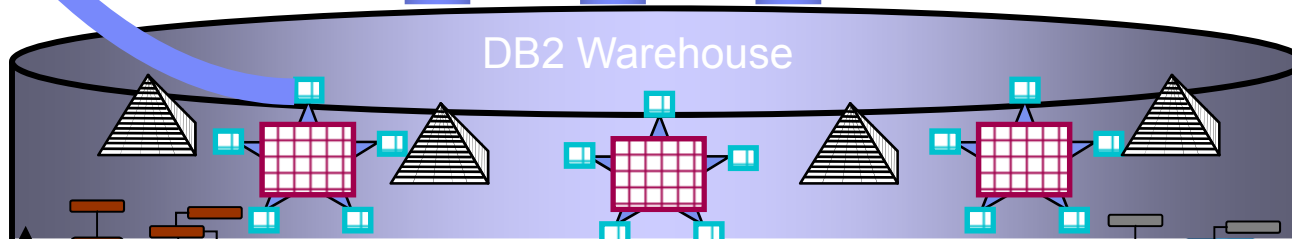
MDX

SQL

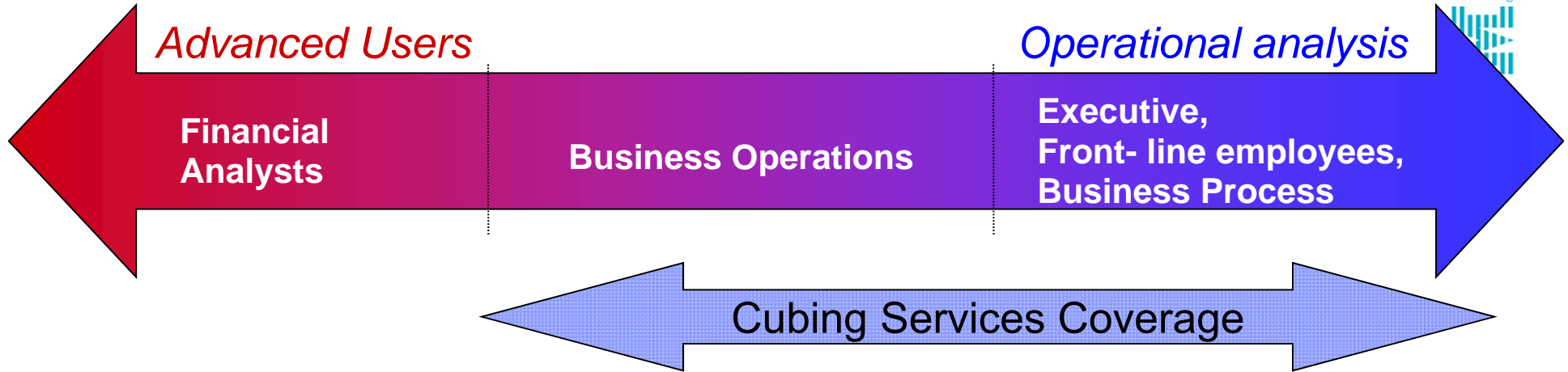
Cubing Services



DB2 Warehouse



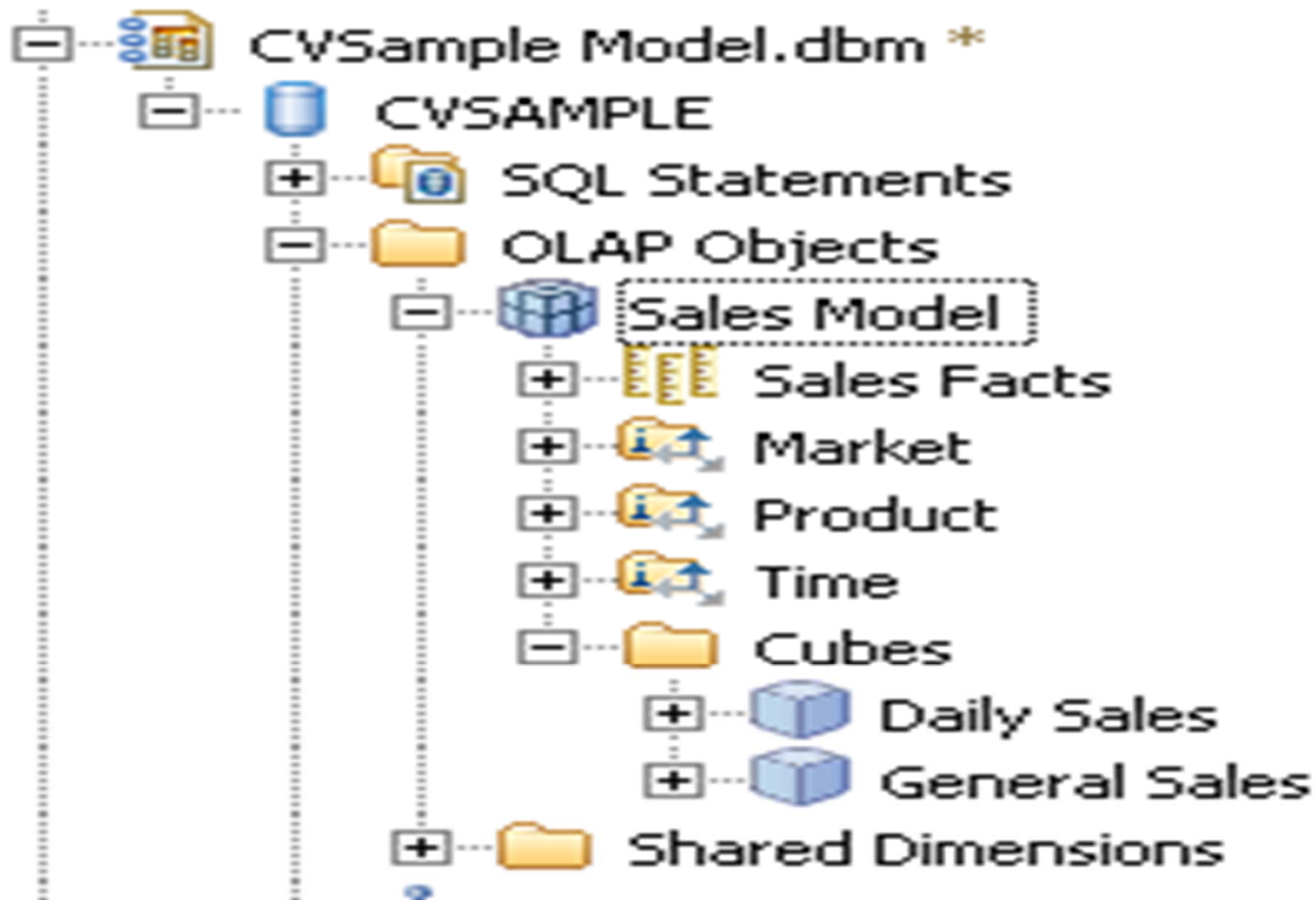
# Who Uses Cubing Services?



Financial Analysis	Sales and Marketing Analysis	Key Performance Indicators (KPIs)
Profitability Analysis	Click-stream Analysis	Enterprise Information Systems
Forecasting with Write-back	Quality Analysis	Dashboards, Embedded & Process driven analytics
Budgeting with Write-back	Database Marketing	Business Process Management
<b>Special uses ~ 25%</b>	<b>Primary uses ~ 75%</b>	

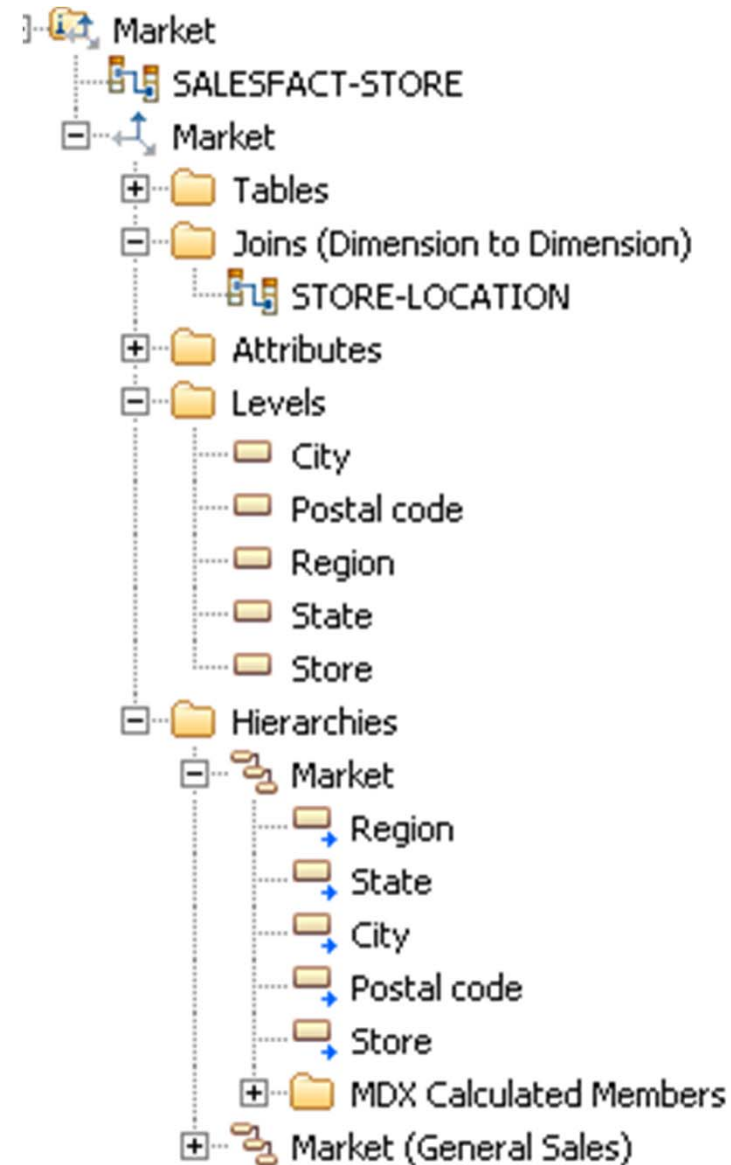
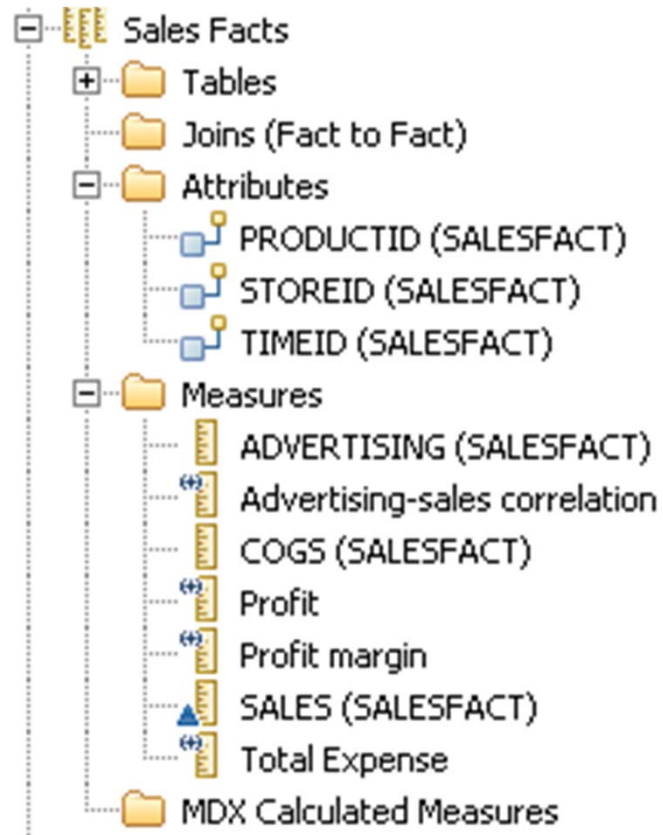


## Demo : OLAP Models





## Demo : OLAP Models



# Demo : Cube Deployment



Warehouse Administration Console - Windows Internet Explorer

http://localhost:9090/bm/warehouse/PAW/LUN/adminconsole/welcome.jsp

InfoSphere Warehouse: Administration Console

Welcome Configuration Manage Logs Manage Connections Manage System Resources SQL Warehousing Cubing Services Mining

Manage Cube Servers Manage Cube Models Manage Roles

The list shows the cube servers that are defined.

Cube Server Name	Host Name	Port Number
CVSample	Windows-XP-Prof	9610
FW97Lab	Windows-XP-Prof	9500

### Edit ConversionCube Cube

Edit the ConversionCube cube configuration on the cube server FW97Lab.

Enable this cube

---

Refresh settings

Automatically refresh the cube caches

Refresh cube caches after the cube runs for (hours, minutes) 12 0

Refresh cube caches every 1 days at 12 00 PM

---

Member cache settings

Enable dynamic caching

Do not limit the number of cached members

Limit the number of cached members to 100000

---

Data cache settings

Cache all cells

Limit the number of cached cells to 10000

Populate the cache using the following MDX query

# Demo : Cognos Integration



### Metadata Wizard - Select Cube

Select the cube you want to import.

- CubingServices97
  - Cubing Services
    - CSSalesMarketing
    - ConversionCube
    - LocalCurrencySales
    - RegionalSalesAnalysis

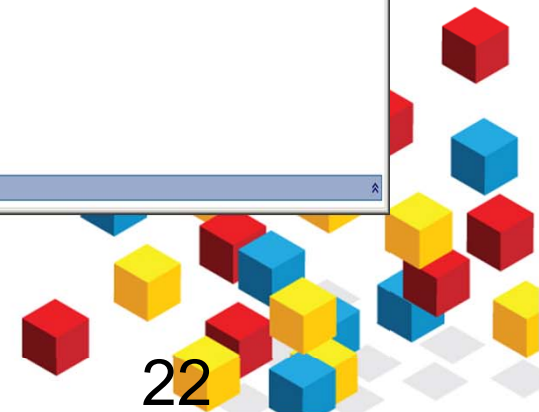
Insertable Objects

- RegionalSalesAnalysis
  - Retailer (20 of 21)
    - Australia
    - Austria
    - Belgium
    - Brazil
    - Canada
    - China
    - Denmark
    - Finland
    - France
    - Germany
    - Italy
    - Japan
    - Korea
    - Mexico
    - Netherlands
    - Singapore
    - Spain
    - Sweden
    - Switzerland
    - United Kingdom
    - More
  - Time
    - 2004
    - 2005
    - 2006
    - 2007
  - Measures
    - Unit sale price

Unit sale price

Unit sale price	2004	2005	2006	2007	All
Australia		386,707.17	482,006.67	320,863.36	1,189,577.2
Austria	473,504.29	525,190.79	657,369.84	399,529.23	2,055,594.15
Belgium	328,062.33	420,071.22	538,503.57	326,580.27	1,613,217.39
Brazil	376,630.45	431,576.76	496,968.37	304,834.15	1,610,009.73
Canada	738,614.56	1,003,116.99	1,208,935.48	702,059.45	3,652,726.48
China	485,010.07	662,982.19	820,942.9	524,766	2,493,701.16
Denmark	173,841.69	251,912.56	331,061.89	165,299.22	922,115.36
Finland	326,195.33	450,814.89	528,509.8	314,059.6	1,619,579.62
France	917,320.39	1,122,518.67	1,232,683.21	743,331.82	4,015,854.09
Germany	750,299.37	922,651.65	1,075,008.11	703,090.36	3,451,049.49
Italy	333,742.11	522,457.73	612,724.44	425,214.4	1,894,138.68

Help    Cancel    < Back    Next >    Finish



# InfoSphere Warehouse : Supporting Advanced Analytics

