

**IBM Cognos Performance**  
Better Intelligence. Better Business.

# Predictive Analytics

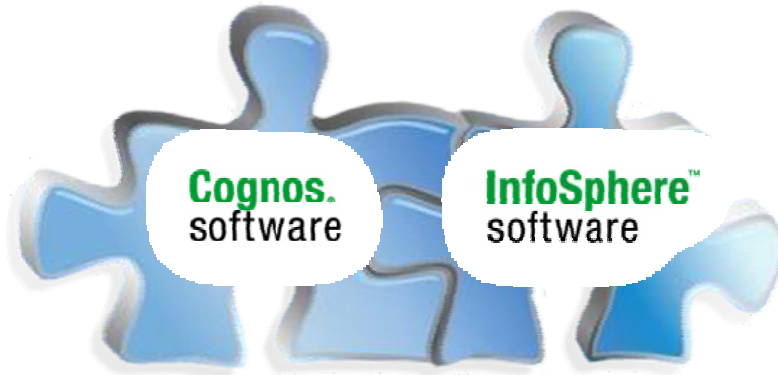


## *Agenda*

- Understand a bit about Data Mining/ Predictive Analytics
- Understand how IBM Cognos software and IBM InfoSphere Warehouse work together to answer your needs
- Use customer case studies to illustrate how
- Answer your questions



## Cognos Synergy with InfoSphere Warehouse



### Data Mining

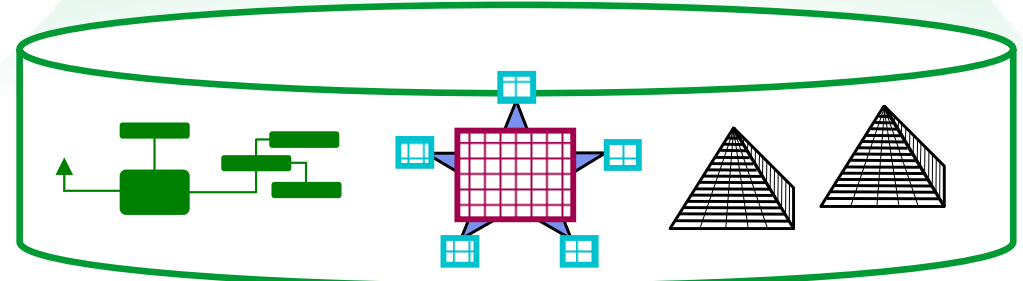
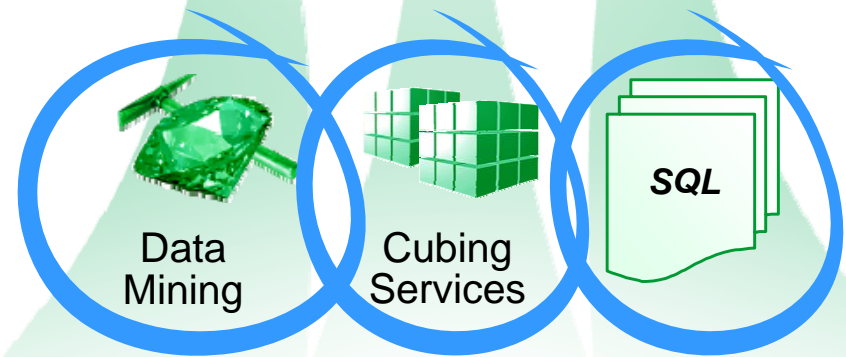
- Batch or Dynamic process integration and visualization
- Advanced DB2 features for data mining and unstructured (text) analytics

### Cubing Services

- Open cube access through XMLA

### Query Optimization

- Cognos Optimization
- Query optimization
- Performance tuning for AIX



InfoSphere Warehouse



## IBM Cognos 8 Integration with IBM InfoSphere Warehouse

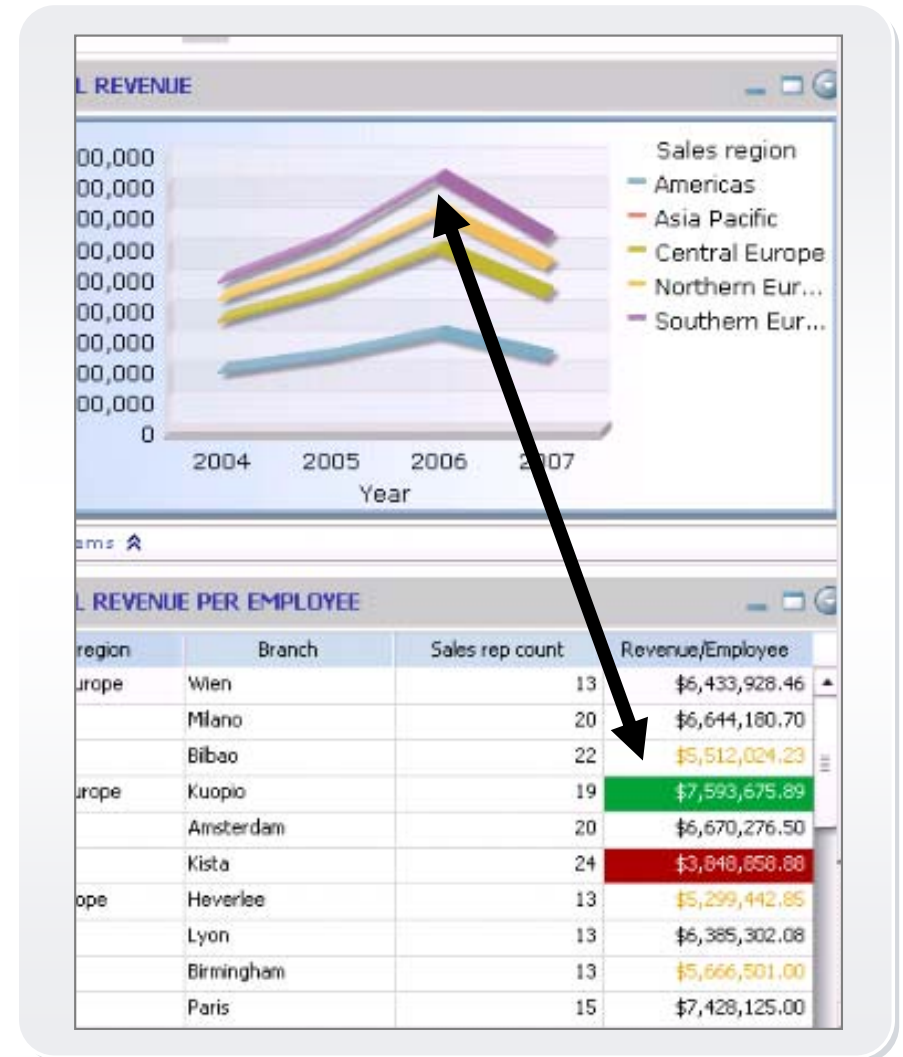
**Challenge:** Enable more sophisticated and predictive BI applications

### Solution

- Deliver data mining results to a broader audience of users with IBM Cognos 8 BI

### Benefits

- Integration and leverage of predictive models within IBM Cognos software provides necessary answers



## Breadth of Analysis styles

Broad Usage  
(Consumers)

Focused Usage  
(Specialist)

### Analytical Reporting *Drill*

- Top down view
- Drillable reports
- Sort top & bottom
- Review then query
- Market shifts
- Product ranking

### Trending Slice and Dice

- Personal exploration
- Compare & contrast
- Rotate and nest
- Work disconnected
- Sales trend analysis
- Market analysis

### Scenario Modeling What-if

- Model scenarios
- Reorganize, reshape
- Compare scenarios
- Save versions
- Financial analysis
- Profitability analysis

### Advanced Analytics What might be

- Uncover patterns
- Statistical algorithms
- Mine data and text
- Predict outcomes
- Fraud prevention
- Churn analysis



Casual  
Users



Executive



Business  
Manager



Financial &  
Business  
Analyst



Statistical  
Analyst



*The most important ingredient for meaningful predictive analytics:*

~~University Degree in Statistics~~

~~Computer Science Education~~

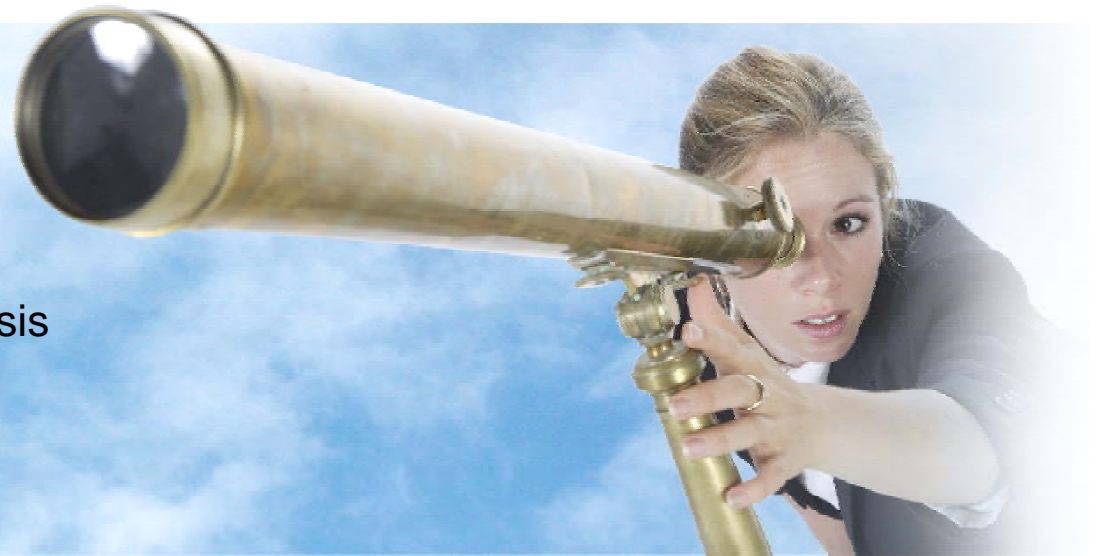
**Business Understanding**



## *Two Types of Data Mining – Discovery & Predictive*

### **Discovery**

- Automatically find trends and patterns
- Answer unasked questions
- Relatively undirected analysis
- Tool reports on findings
- In a word – “Easier”
- Useful for non-statisticians

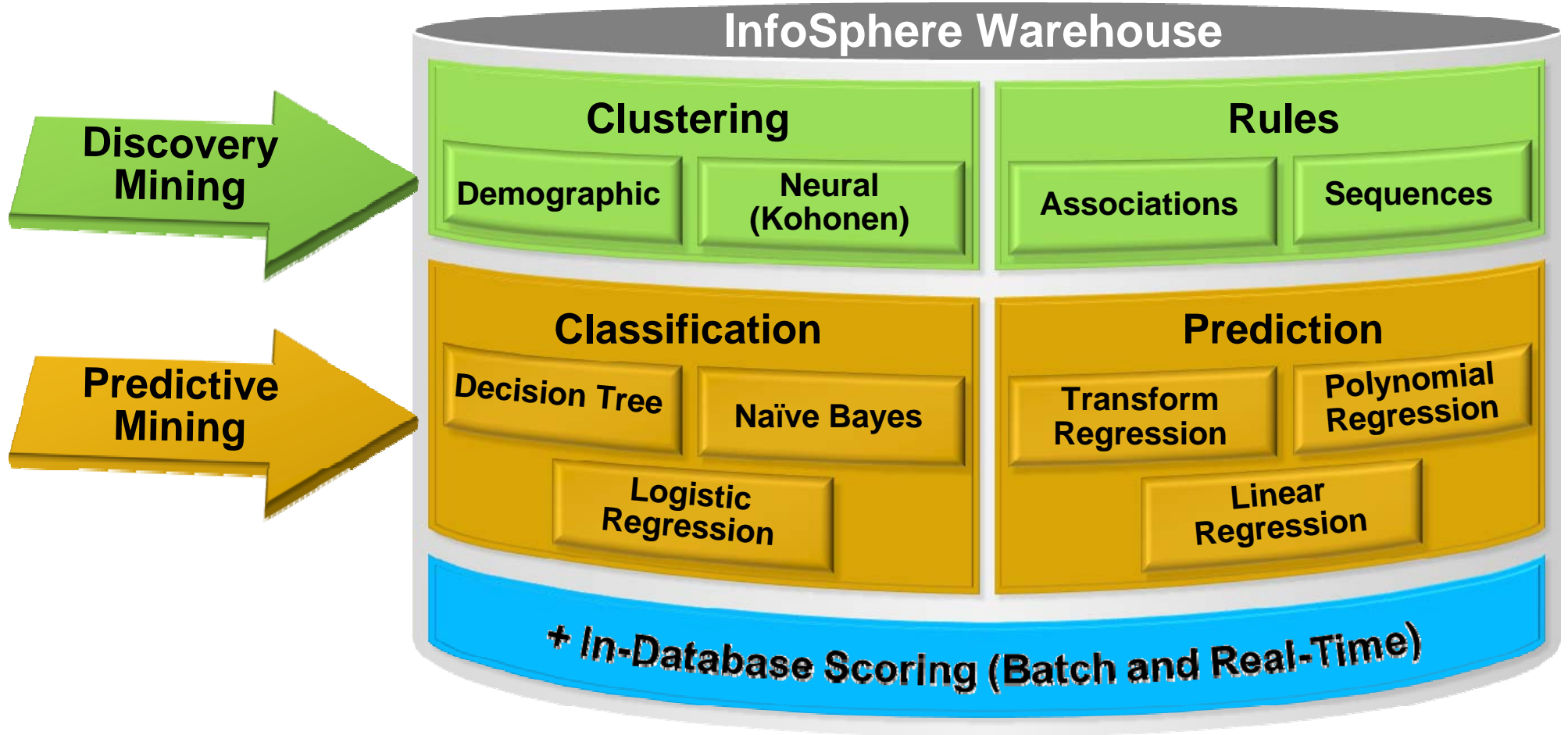


### **Predictive**

- Specific question
- Probability associated with outcomes
- Directed analysis
- Iterative process
  - Train, Test, Apply
- Apply model in database at customer touch points

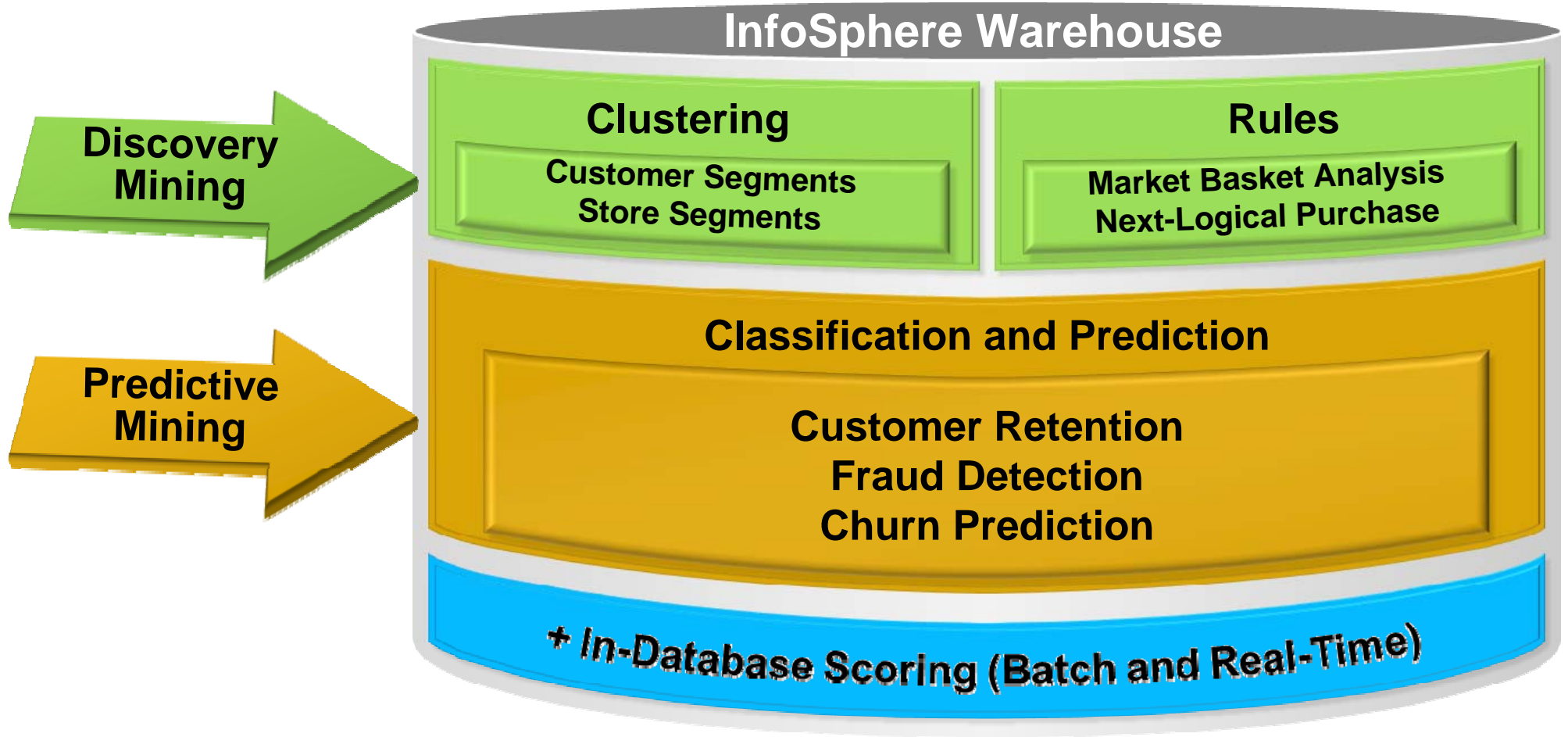


# InfoSphere Data Mining Methods and Algorithms





# InfoSphere Data Mining Application Examples



## Next Logical Purchase and Customer Segmentation

### ■ Customer Background

- Electronics Retailer

### ■ Business Problem

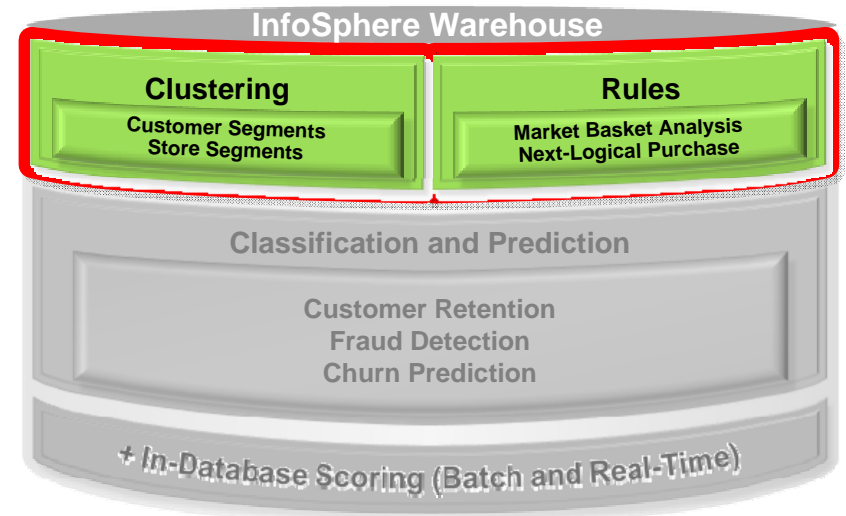
- Need to predict the next, most-logical product a customer will purchase by type of customer

### ■ Business Goals

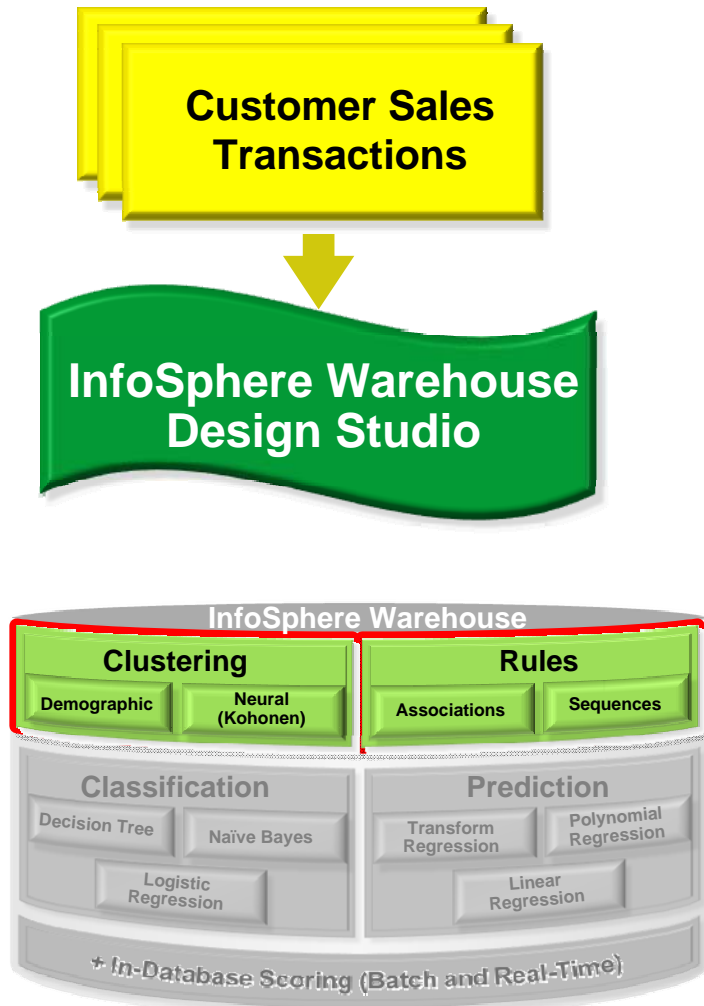
- Develop a data mining flow to use advanced analytics to show driver and implied items within a Market Basket analysis
- Develop a platform for delivering the “next purchase” prediction results in support of marketing

### ■ Results

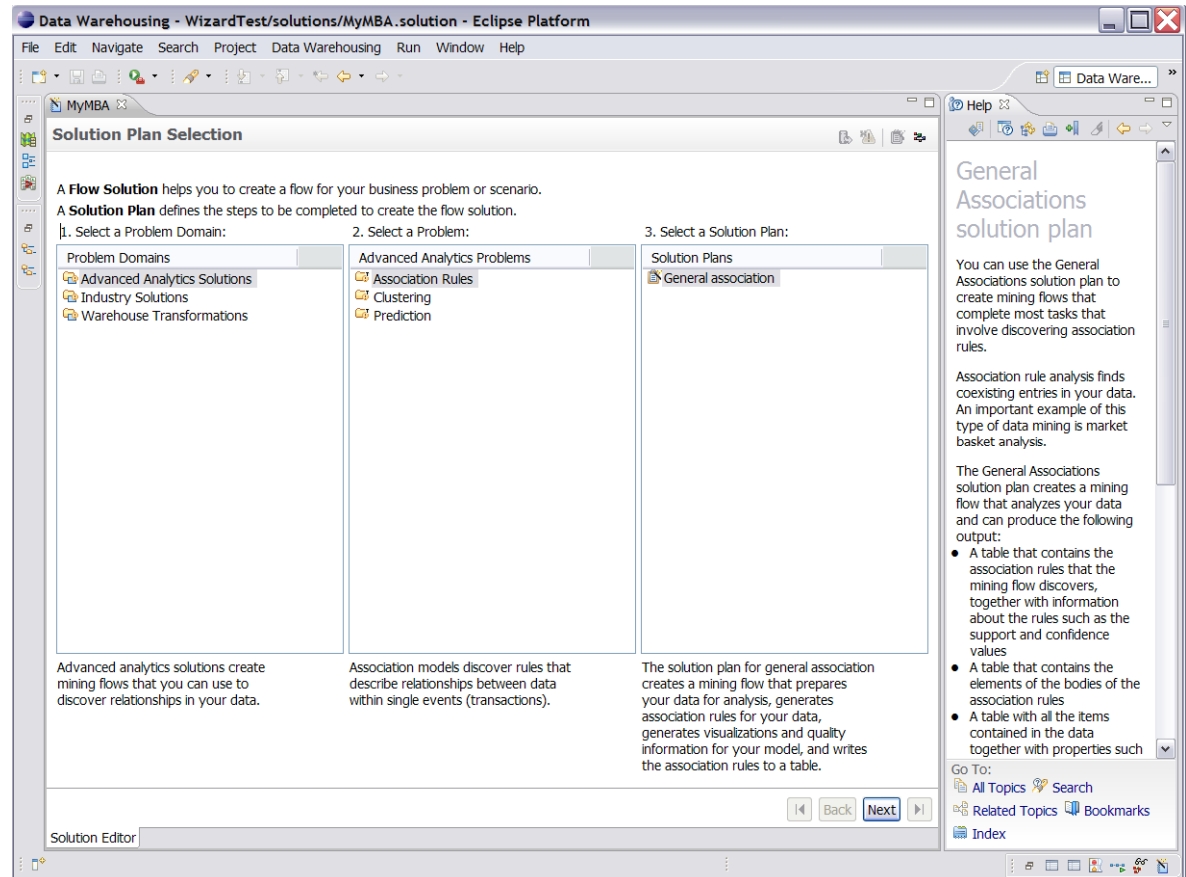
- Created a set of tables with the results of data mining to allow more team members within the retailer to take proper action
- Report shows the average amount of time from one purchase to the next of a given set of items and driver items



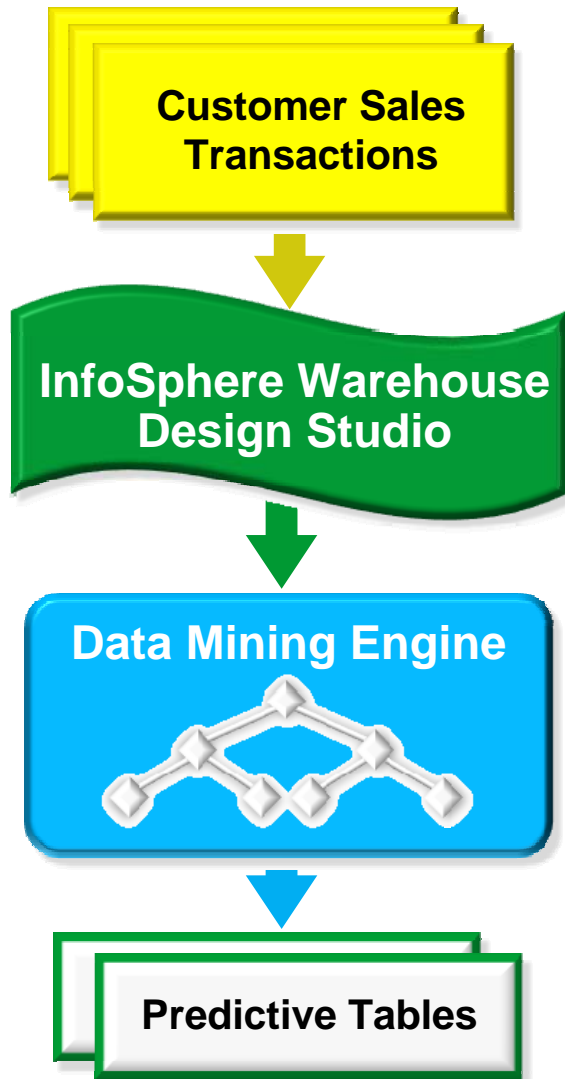
## Process Overview



- Step 1 – Create Mining Flow in InfoSphere Warehouse using Design Studio Wizards.



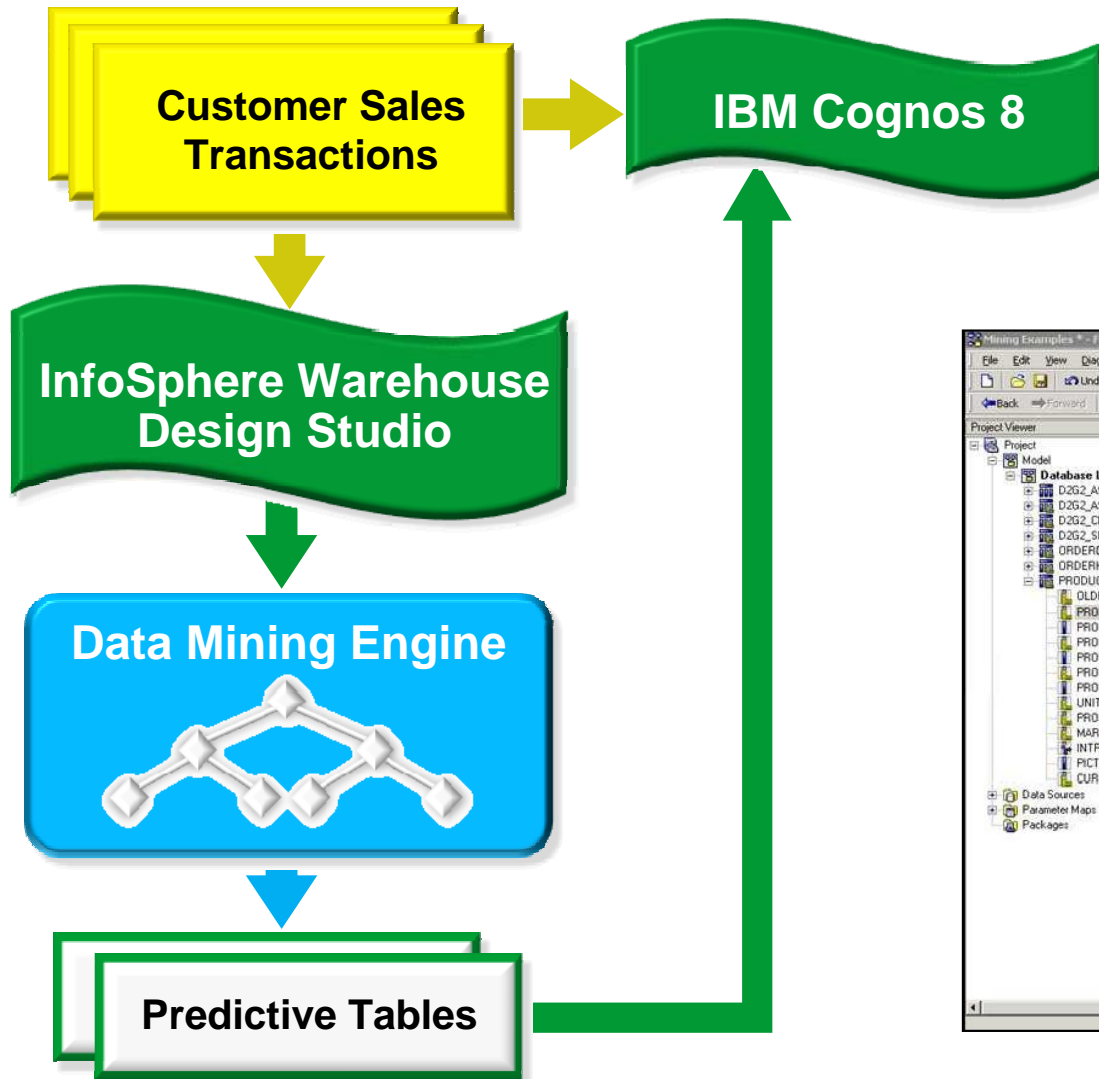
## Process Overview



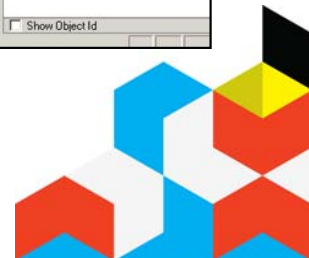
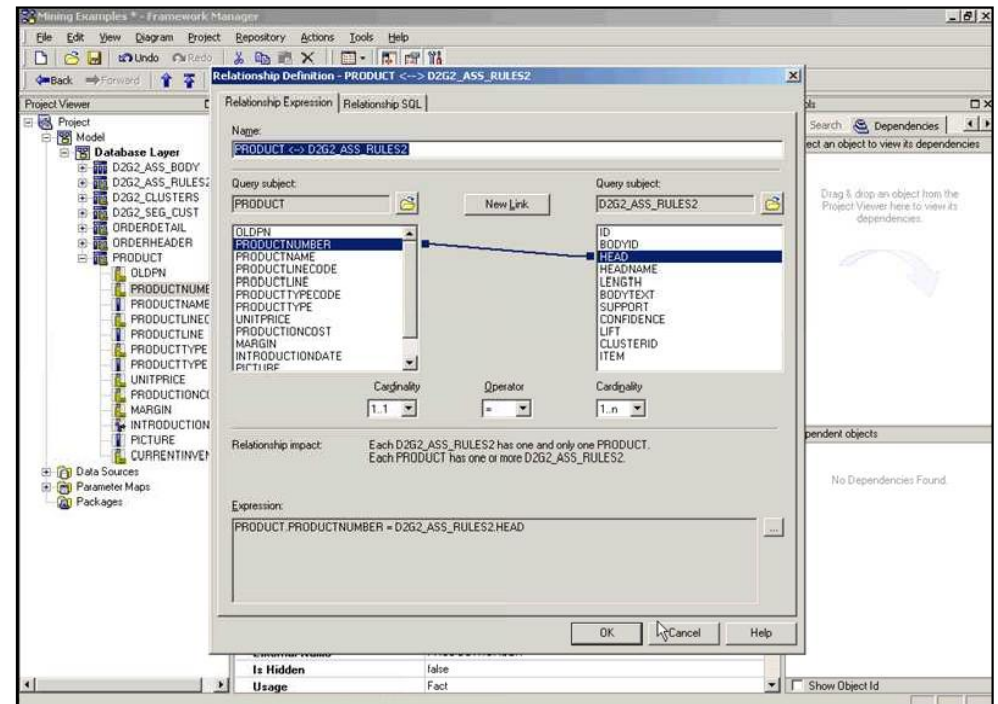
- Step 1 – Create Mining Flow in InfoSphere Warehouse using Design Studio Wizards.
- Step 2 – Run the data mining engine to create the predictive tables containing information about which products are over time (sequencing) and by what type of customer (clustering)



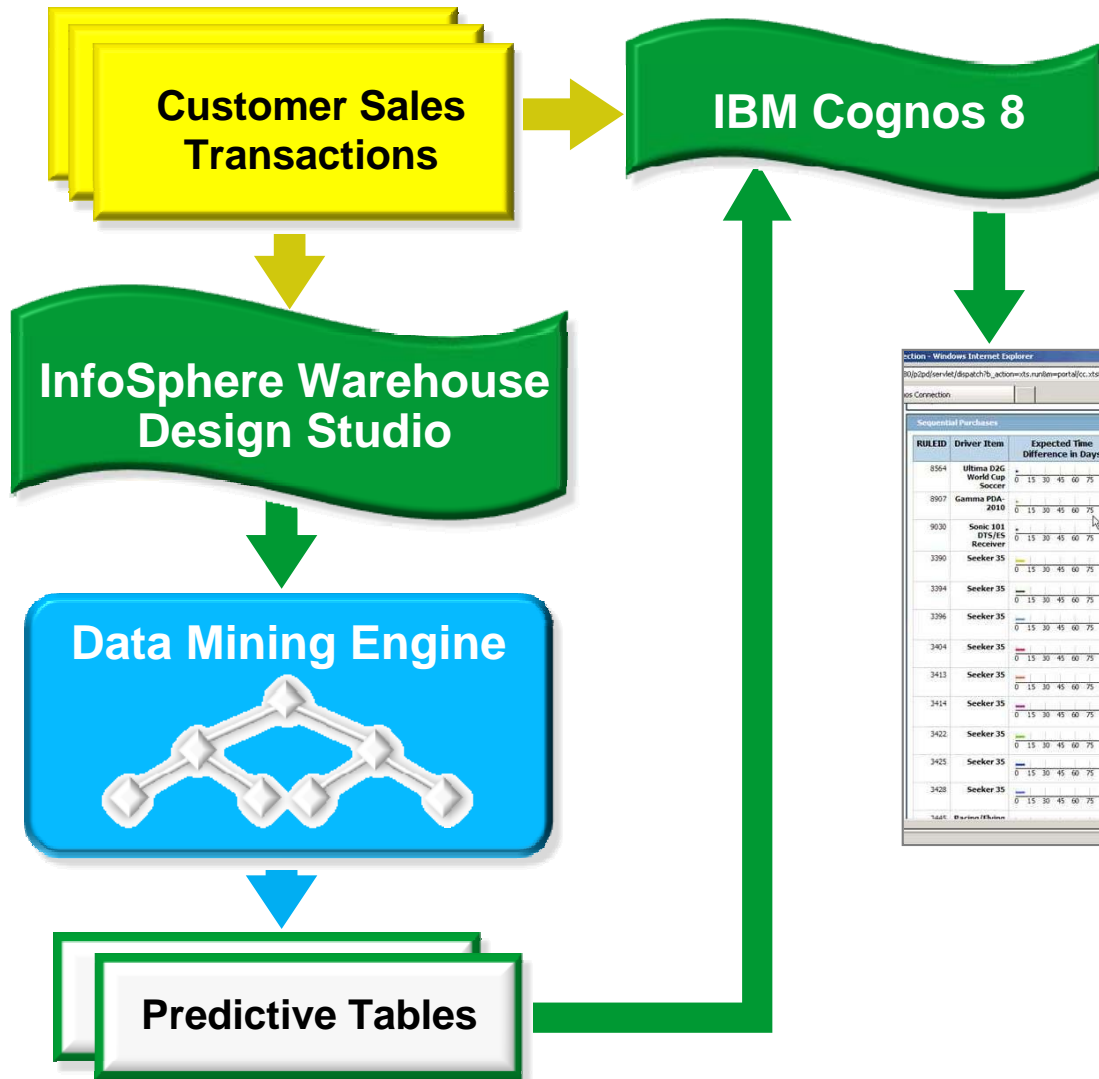
## Process Overview



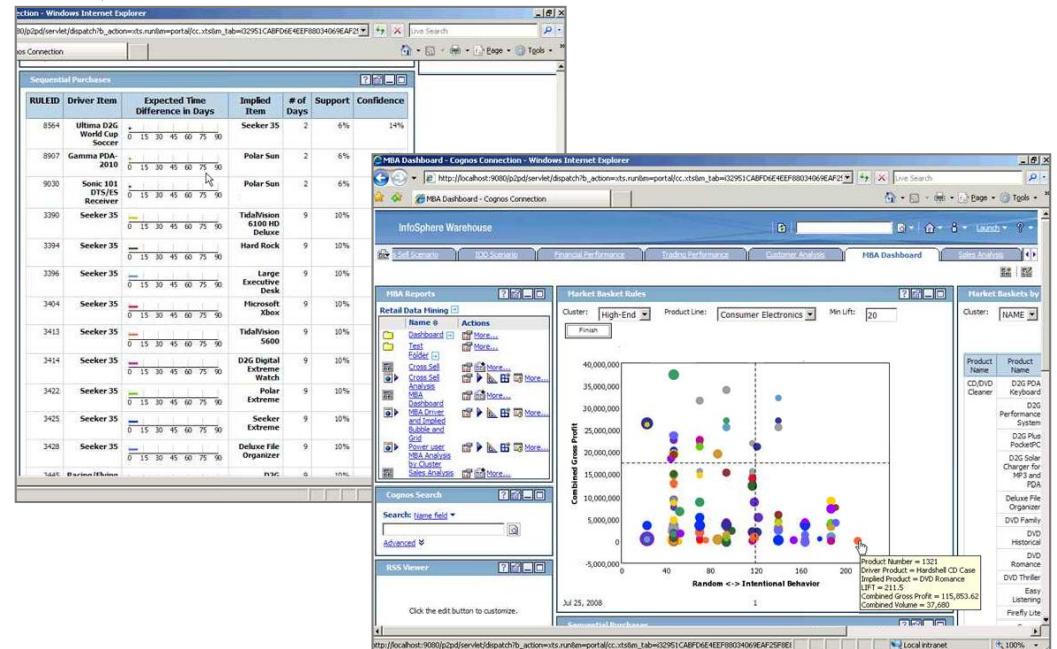
- Step 3 - Use IBM Cognos 8 Framework Manager to setup table relationships with existing tables and the predictive results tables. Can also define functions and stored procedures as req'd.



## Process Overview



- Step 4 - Use IBM Cognos 8 to create Dashboards, Reports etc to communicate the information broadly



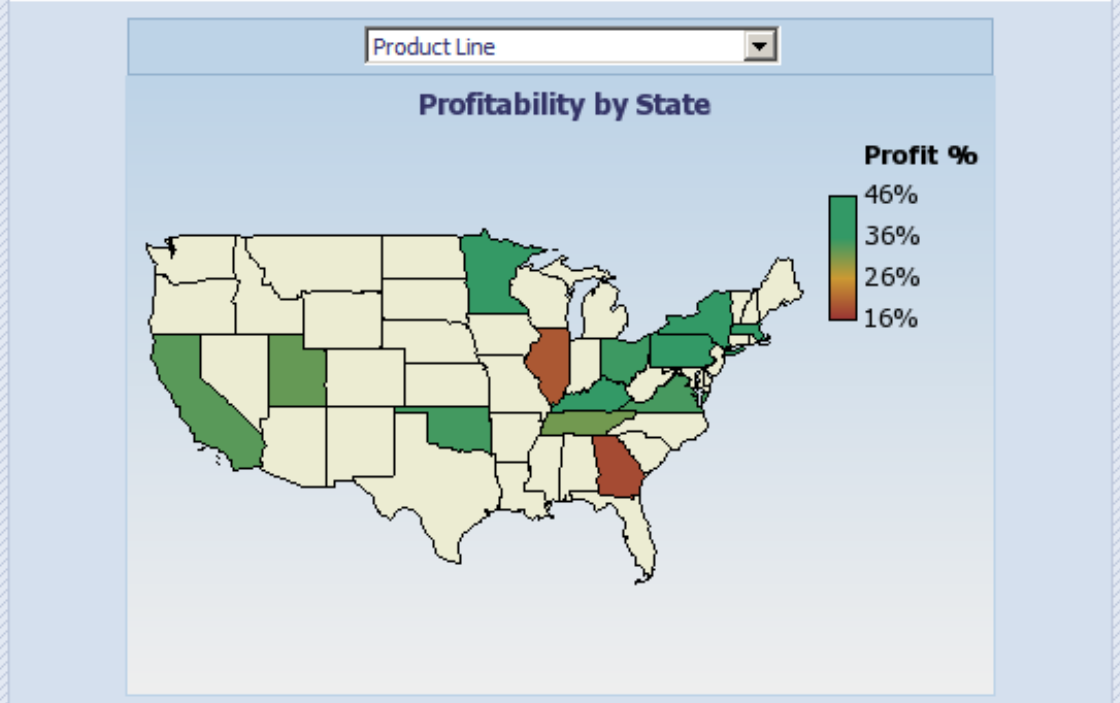
# D2G SALES ANALYSIS

## TOP SELLING PRODUCTS

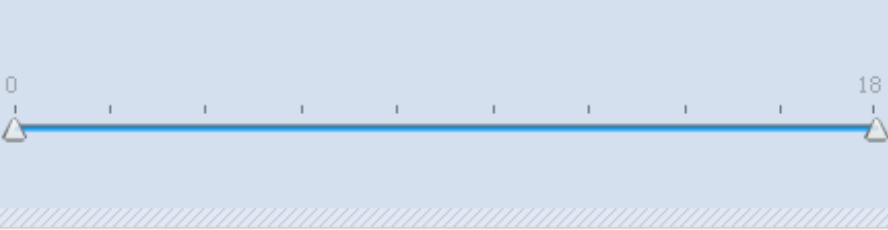
Top: \* 3

Rank	Product Name	MBA	SEQ	Sales [\$ 000]
1	<a href="#">Sonic DTC Digital Advance</a>	☰	➔	\$85,849
2	<a href="#">Sonic 101 DTS/ES Receiver</a>	☰	➔	\$47,091
3	<a href="#">Sonic TLR Thunder Receiver</a>	☰	➔	\$46,184
<b>Consumer Electronics</b>				<b>\$179,123</b>
1	<a href="#">D2G Blast Em</a>	☰	➔	\$25,938
2	<a href="#">Microsoft Xbox</a>	☰	➔	\$6,385
3	<a href="#">Sega Dreamcast</a>	☰	➔	\$5,632
<b>Entertainment Media</b>				<b>\$37,955</b>
1	<a href="#">Computer Extended Desk</a>	☰	➔	\$51,852
2	<a href="#">Computer Desk Ultra</a>	☰	➔	\$41,539
3	<a href="#">Standard Stool</a>	☰	➔	\$38,477
<b>Home Office</b>				<b>\$131,867</b>
1	<a href="#">Wavestation 4200 FP</a>	☰	➔	\$384,344

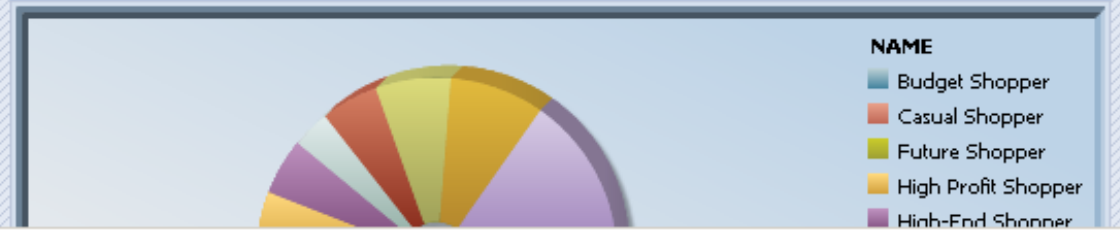
## PROFITABILITY BY STATE



## RESTRICT CUST SEGMENTCOUNTS TO...



## SEGMENTATION COUNTS

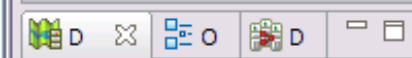




Data Warehousing

Data... x Navi...

- [-] DemoProject
  - Application Misc Files
  - Control Flows
  - [+] Data Diagrams
  - Data Flows
  - [+] Data Models
  - [+] Data Preparation
  - File Definitions
  - Flow Solutions
  - Miningbox Profiles
  - Mining Flows
  - [+] Other Files
  - Resources
  - Run Profiles
  - [+] SQL Scripts
  - [+] Subflows
  - Subprocesses
  - [+] Text Analysis
  - [+] Variables
  - Warehouse Application Pro



Database Connections

- [-] Database Connections
  - BIRT Classic Models Sa
  - COGSAM83 [DB2 Alias]
  - Derby Sample Connect
  - DWESAMP [DB2 Alias]
  - HOL97 (DB2 for Linux,
  - [+] HOL97

1 items selected

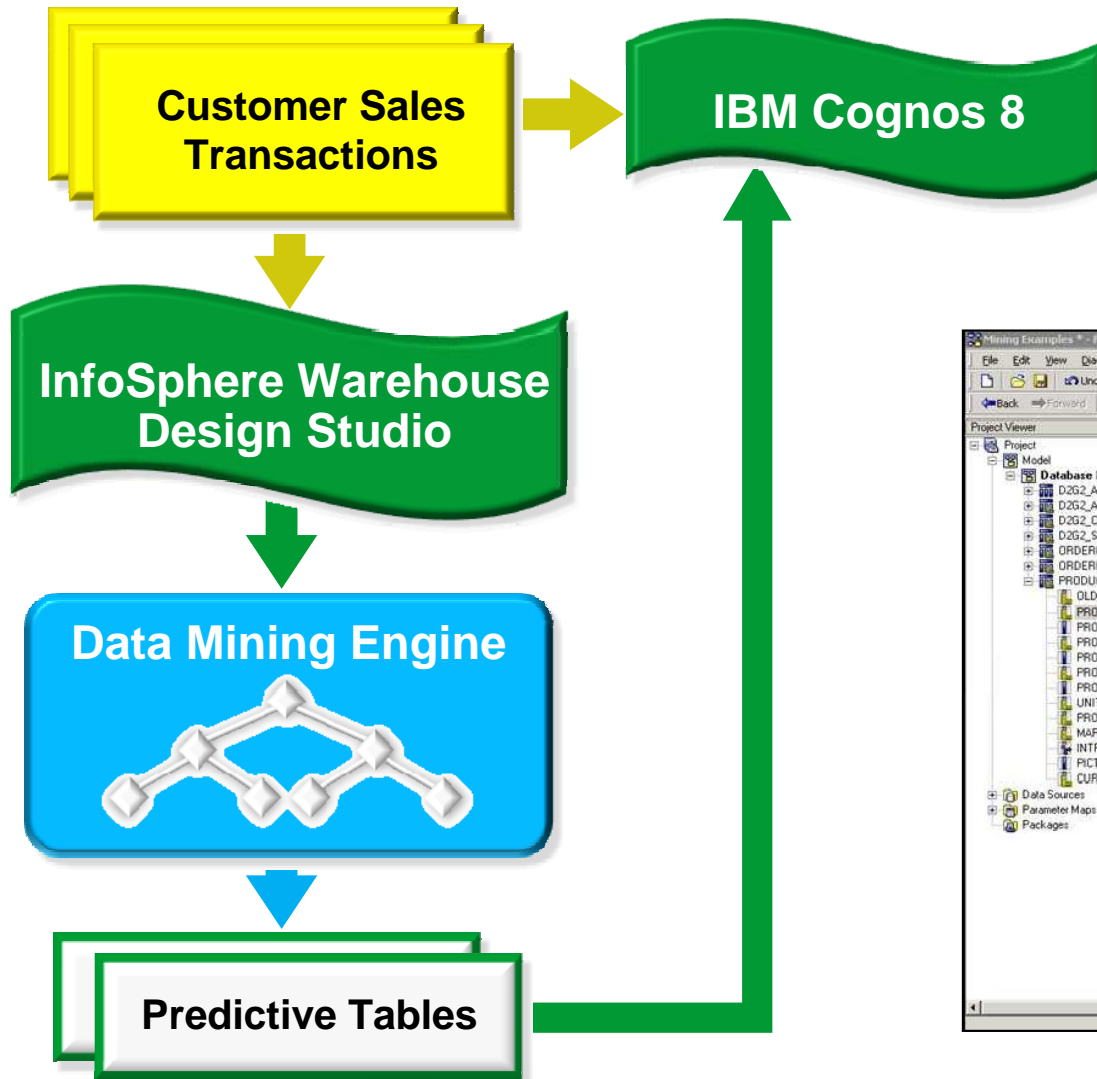


Properties SQL Results Problems Execution Status Job Status Data Exploration

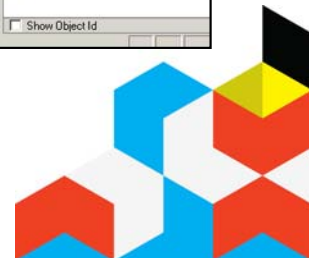
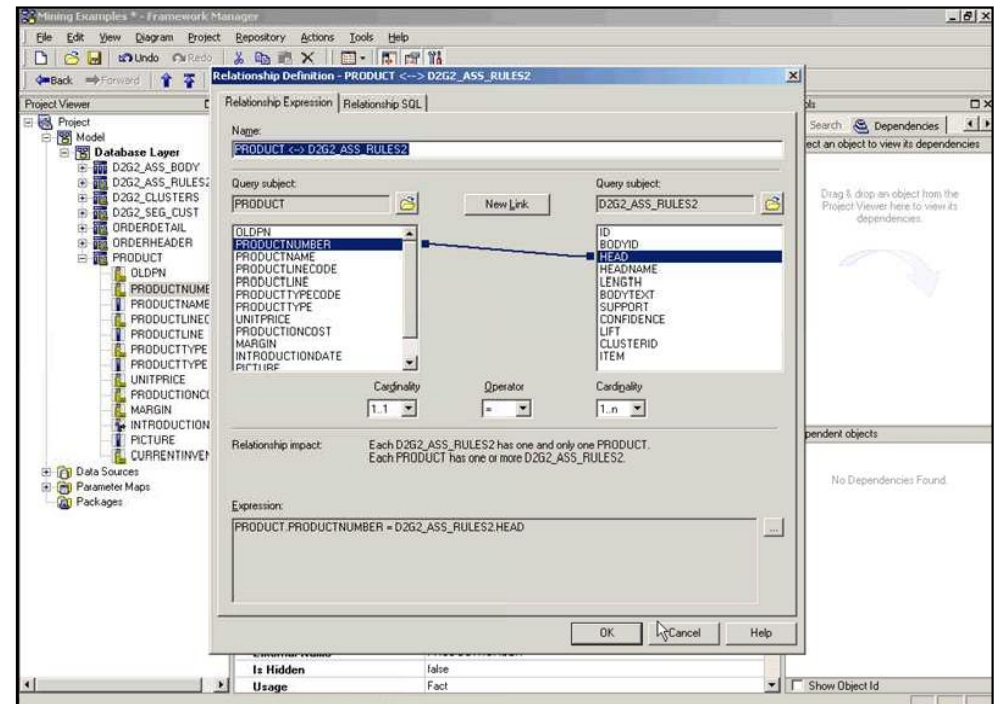
Property	Value
General	



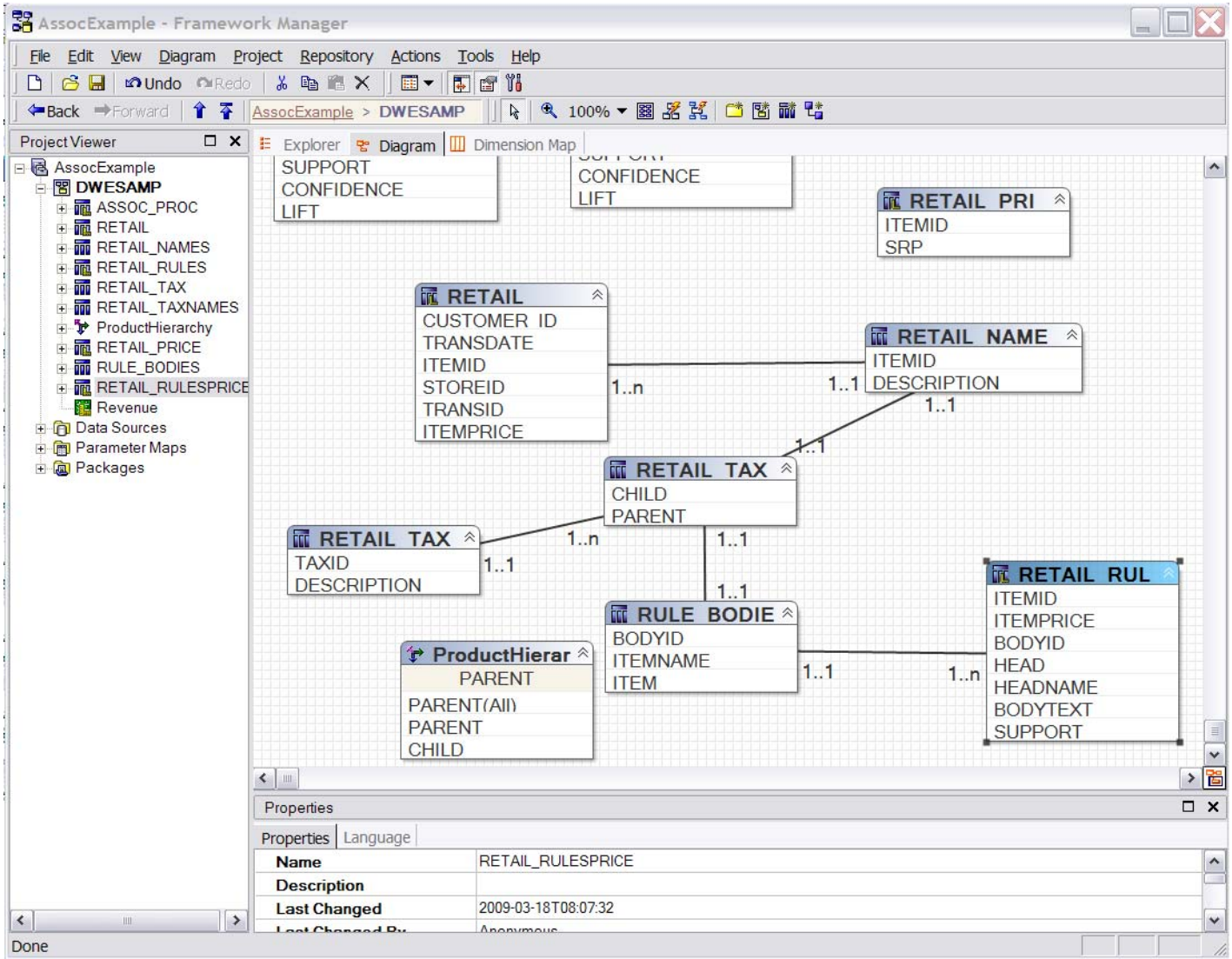
## Process Overview



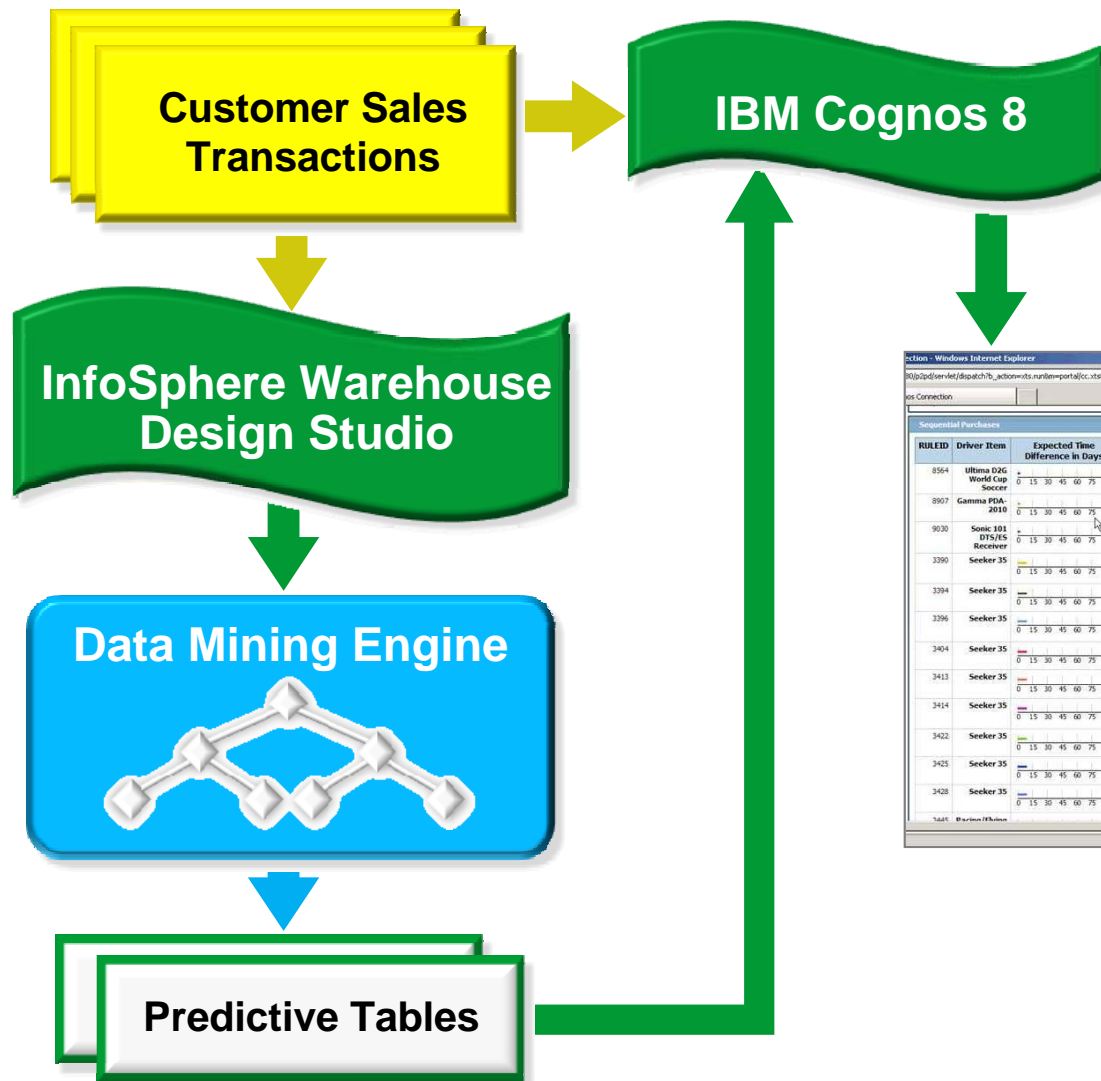
- Step 3 - Use IBM Cognos 8 Framework Manager to setup table relationships with existing tables and the predictive results tables. Can also define functions and stored procedures as req'd.



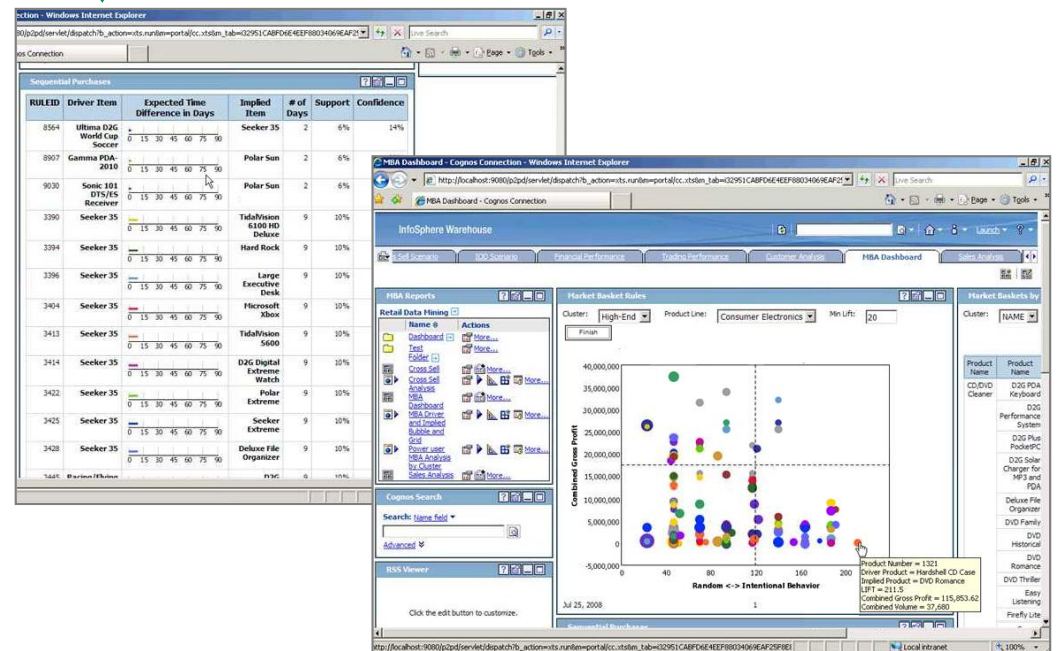
Framework Manager: Model the data so users can create reports easily



## Process Overview



- Step 4 - Use IBM Cognos 8 to create Dashboards, Reports etc to communicate the information broadly



## Build Reports, Dashboards, etc

DraggedRevenueForProduct - Report Studio - Microsoft Internet Explorer

File Edit View Structure Table Data Run Tools Help

Font Size 1 pt

Insertable Objects

- AssociationRulePackage
  - ASSOC\_PROC
  - RETAIL
  - RETAIL\_NAMES
  - RETAIL\_RULES
  - RETAIL\_TAX
  - RETAIL\_TAXNAMES
  - ProductHierarchy
  - RETAIL\_PRICE
  - RULE\_BODIES
  - RETAIL\_RULESPRICE
    - ITEMID
    - ITEMPRICE
    - BODYID
    - HEAD
    - HEADNAME
    - BODYTEXT
    - SUPPORT
  - Revenue

Page Explorer

Page layers: Drop members here to create page layers

Context filter: Drop members here to create a context filter

**Drag-on revenue for <%ParamDisplay...%>**

Default measure (y-axis): <ITEMPRICE>

Series: Drop item here

(Default Legend Title)

100  
75  
50  
25

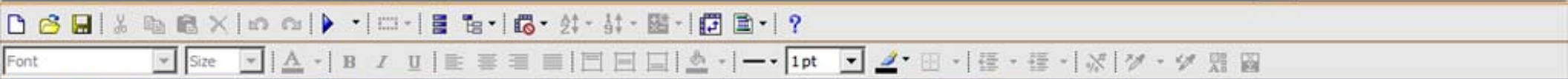
abc abc abc abc

Axis titles:

Category (x-axis): <#HEADNAME#>

Date 1 Time





Insertable Objects

- Great Outdoors Company
  - Years
  - Products
  - Sales region
  - Retailers**
  - Margin range
  - Measures
  - Currency

Page layers: Drop members here to create page layers

Context filter: Drop members here to create a context filter (slicer)

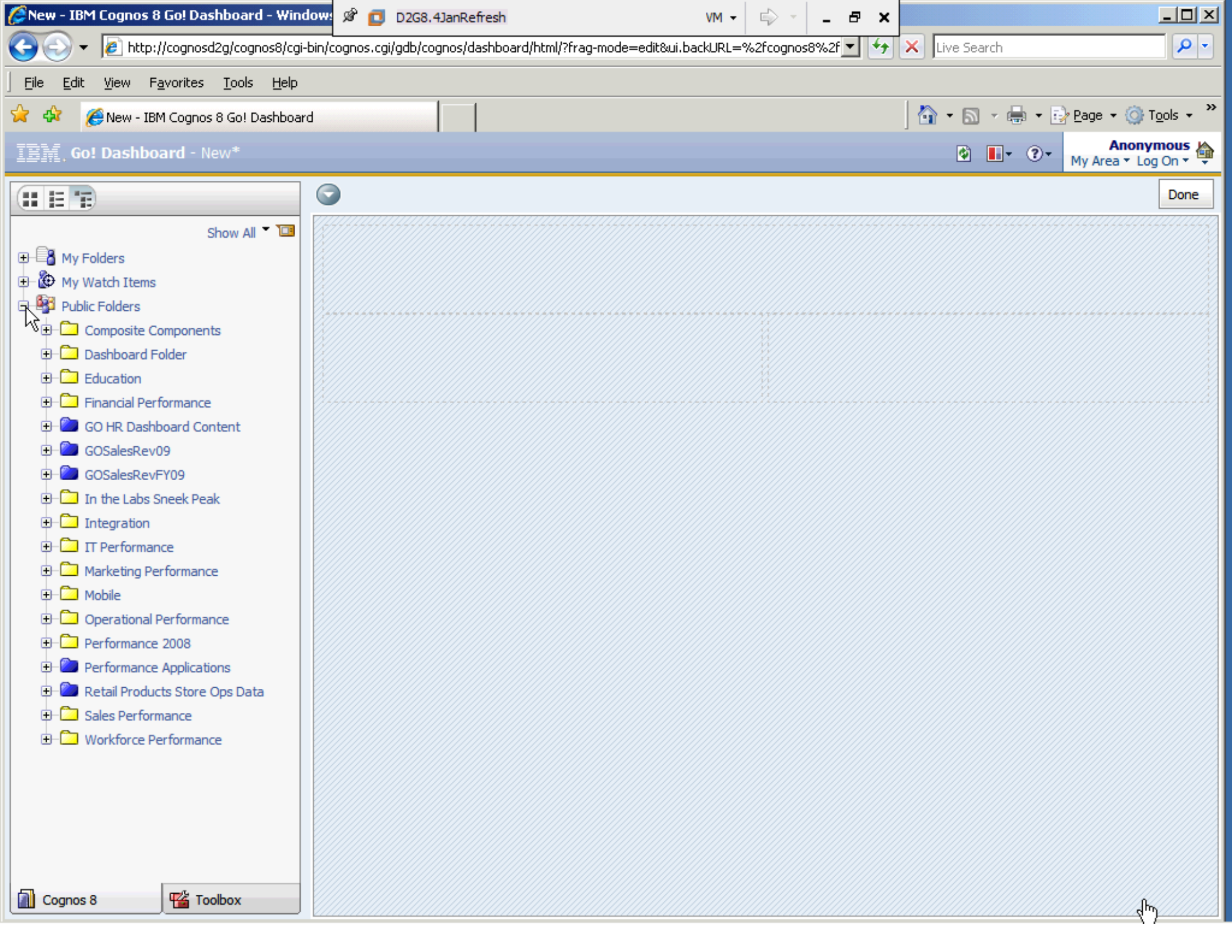
**Double click to edit text**

---

Columns

---

Rows Measures



Layout icons: Grid, List, Split

Show All

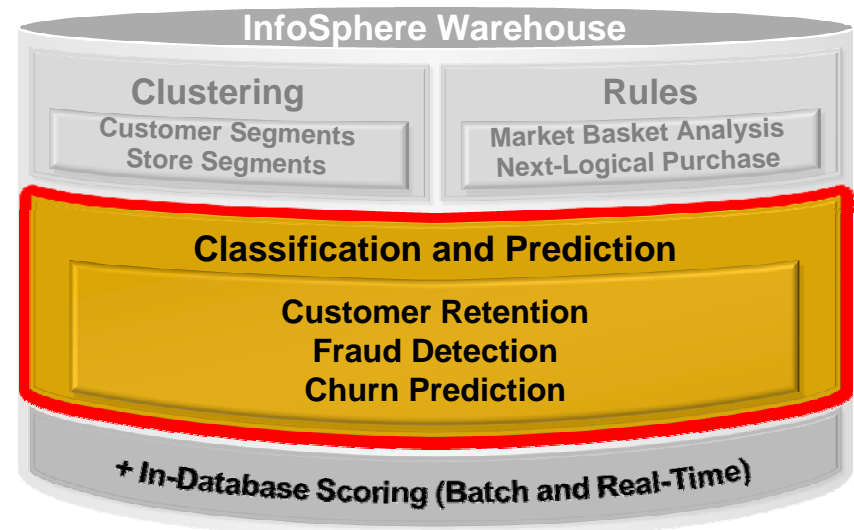
- + My Folders
- + My Watch Items
- + Public Folders
  - + Composite Components
  - + Dashboard Folder
  - + Education
  - + Financial Performance
  - + GO HR Dashboard Content
  - + GOSalesRev09
  - + GOSalesRevFY09
  - + In the Labs Sneek Peak
  - + Integration
  - + IT Performance
  - + Marketing Performance
  - + Mobile
  - + Operational Performance
  - + Performance 2008
  - + Performance Applications
  - + Retail Products Store Ops Data
  - + Sales Performance
  - + Workforce Performance

Done

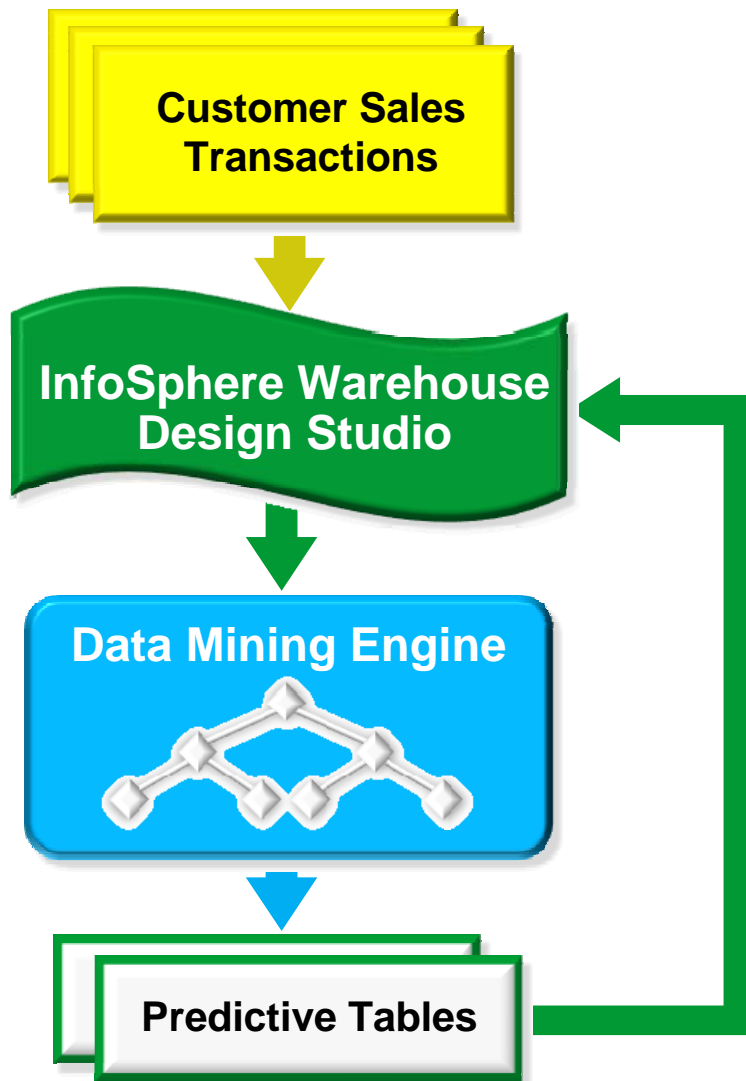
Main workspace area with a light blue background and a dashed border, indicating a drag-and-drop area.

## ***Sales Example – Account Abandonment Prediction***

- Customer Background
  - Multi-billion \$ bank
- Business Problem
  - Some customers who develop negative account balances subsequently abandon their accounts, resulting in a loss (charge-off) to the bank.
- Business Goals
  - Develop a data mining model to predict account abandonment
  - Develop a platform for delivering the prediction results in support of bank operations
- Results
  - Created a classification model to predict account abandonment
  - Model shows that account balances and customer interactions with the bank are good predictors of account abandonment



## Process Overview

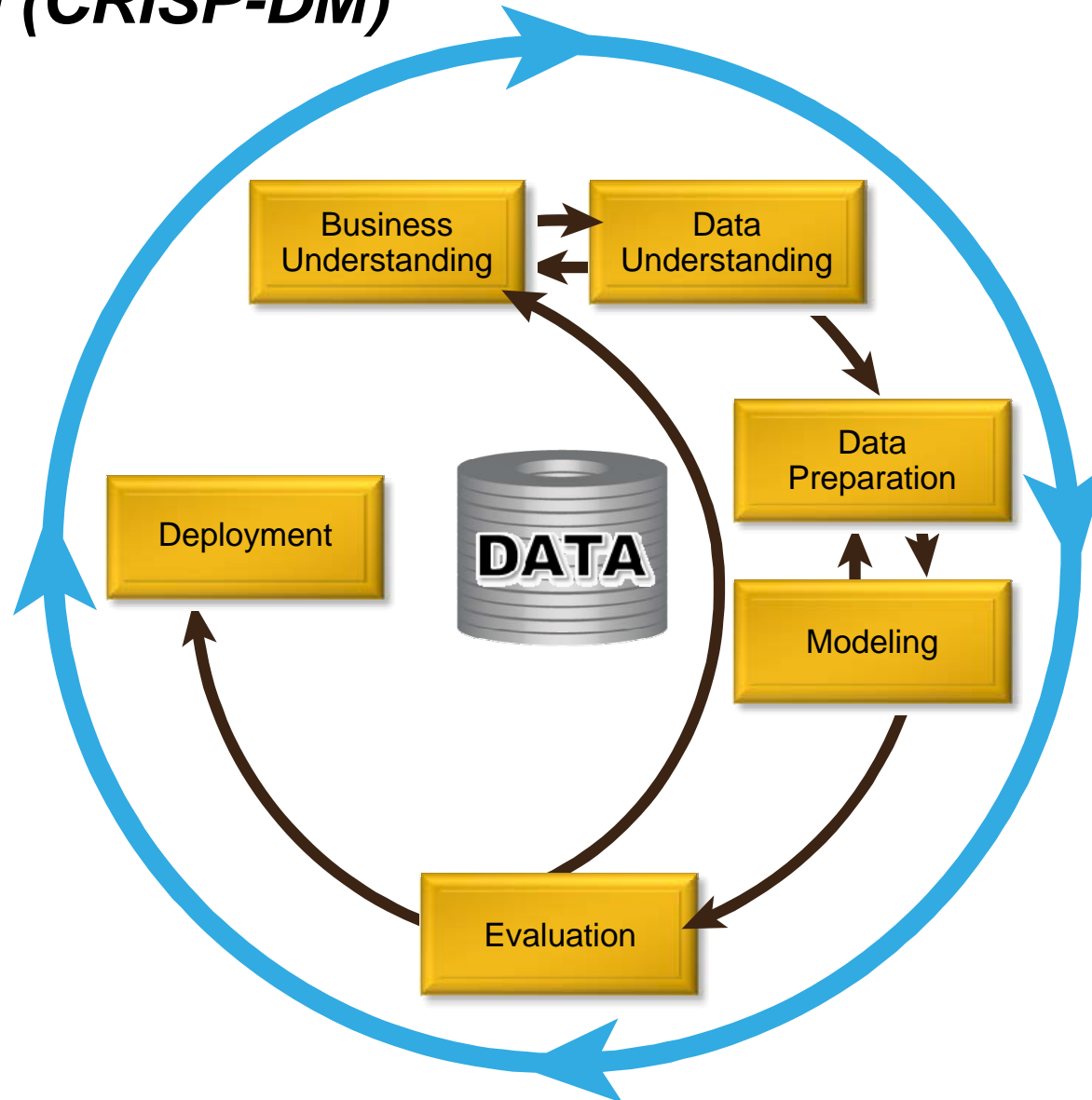


- Step 1 – Create Mining Flow in InfoSphere Warehouse using Design Studio Wizards.
- Step 2 – Run the data mining engine to create the predictive tables containing information about which products are over time (sequencing) and by what type of customer (clustering)
- Step N+1 – Run the data mining engine to create the predictive tables containing information about which customers are likely to abandon their accounts (scoring)

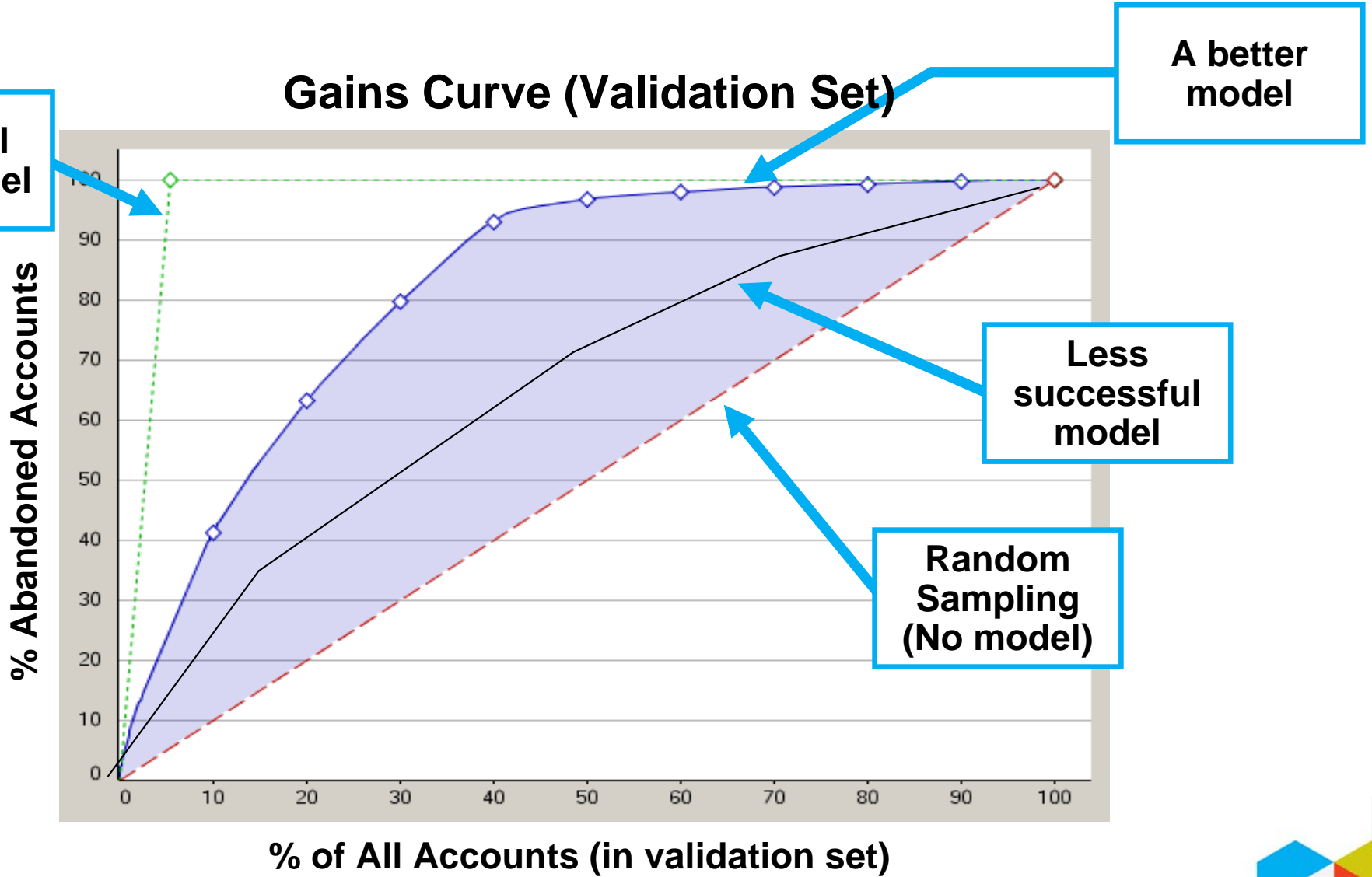




# CRoss Industry Standard Process for Data Mining (CRISP-DM)



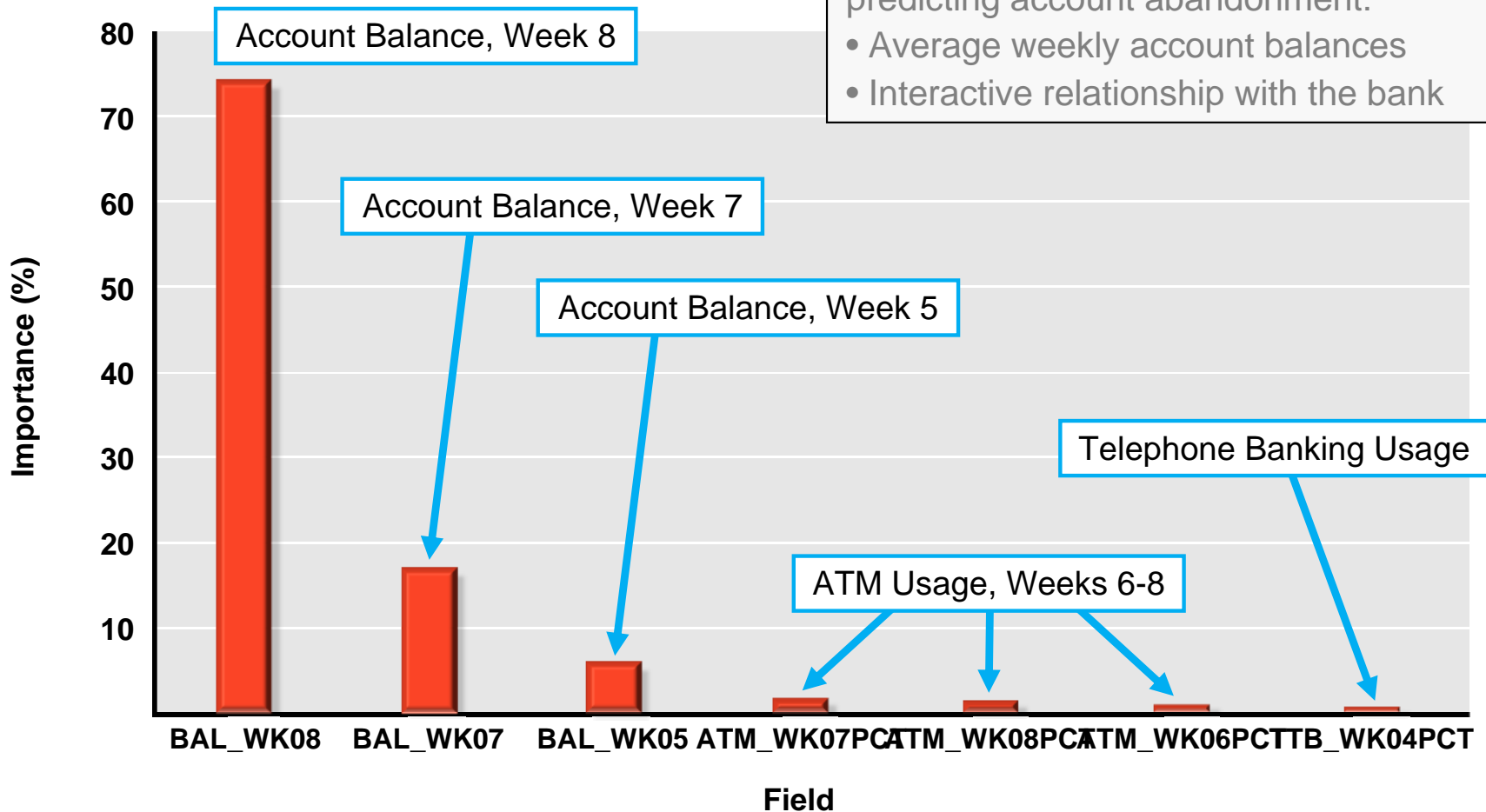
# Account Abandonment Prediction – Gains curve



## Example: Account Abandonment Prediction

### ▼ Field Importance Chart

Target Field CO\_WK15

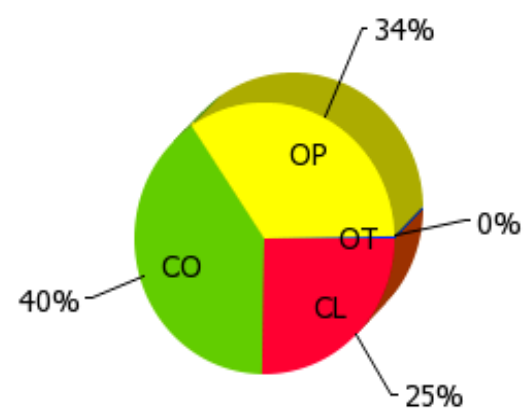


B5\_Dashboard1

# Account Summary for Region: **Dallas** Region

\* Dallas Region

**Account Status**  
 CL  
 CO  
 OP  
 OT



Total No. Accounts	3,673
# Accounts Open	1,254
# Accounts Charged Off	1,471
# Accounts Closed	933
# <b>Accounts At Risk</b>	37

[Account Summary](#)

% of Accounts

OP-Open, CL-Closed, CO-Charged Off, OT-Other

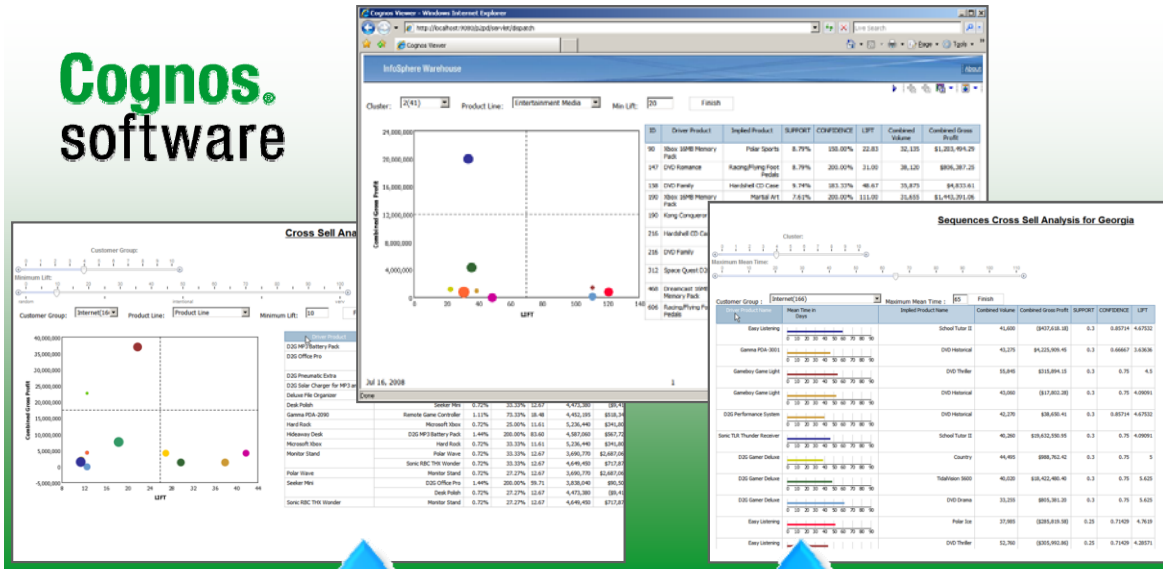
Branch Region	Account Status	# Accounts	% Accounts	Fees Generated
All Regions	CL	7,485	22%	\$118,755.30
	CO	14,072	40%	\$828,523.90

\* Alabama \* All Regions

**REGION**  
 All Regions  
 Alabama

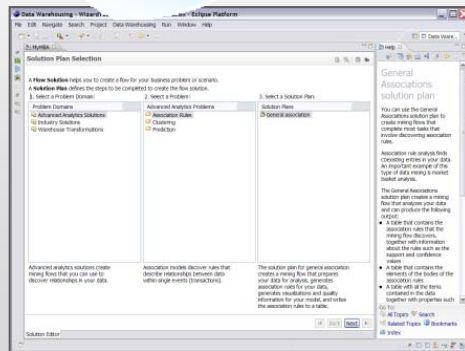
## Partnering with SPSS

Cognos software

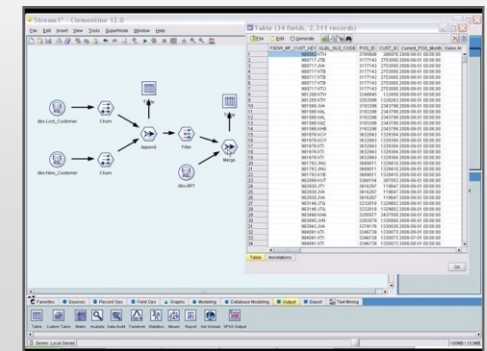


InfoSphere™ warehouse

- In-Database Data Mining
- Enterprise Scalability
- Direct data source for Cognos



- Predictive Data Mining
- Powerful Workbench
- Leverage in-database capabilities



Cognos Viewer - Customer Retention Dashboard

Cognos Navigator

**Alerts**  
 Clementine Customer Churn Process is Complete!  
[Read](#)  
 The monthly SPSS Clementine Customer Churn process has completed. 26 customers have had an increase in their churn score from July to August.

Cognos Navigator

**Relational Reports**

Name	Actions
Drill Through Target Reports	More...
Churn Score by Customer with Opportunity Amount	More...
Churn Score by Customer with Opportunity Amount - TEST	More...
Churn Score with Opportunity Amount	More...
Lost Business with Opportunity Amount and Churn Score	More...
New Activity Report	More...
New Customers with Churn Scores	More...
New Customers with Opportunity Amounts	More...

Cognos Search

Search:

Advanced

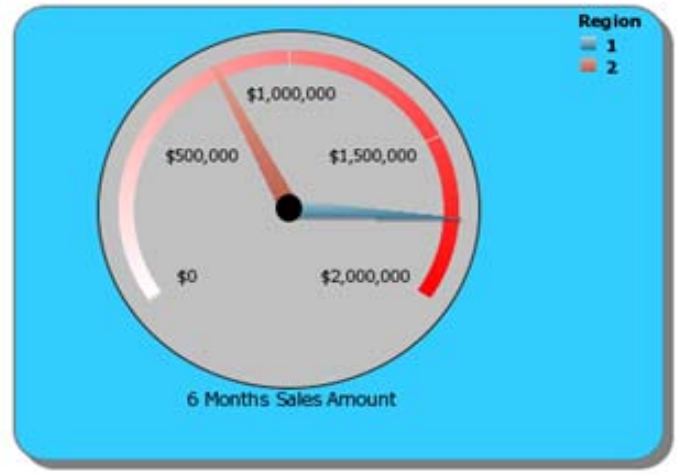
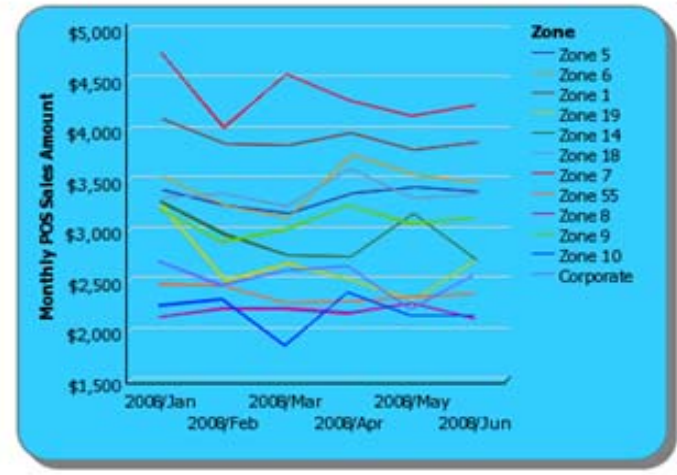
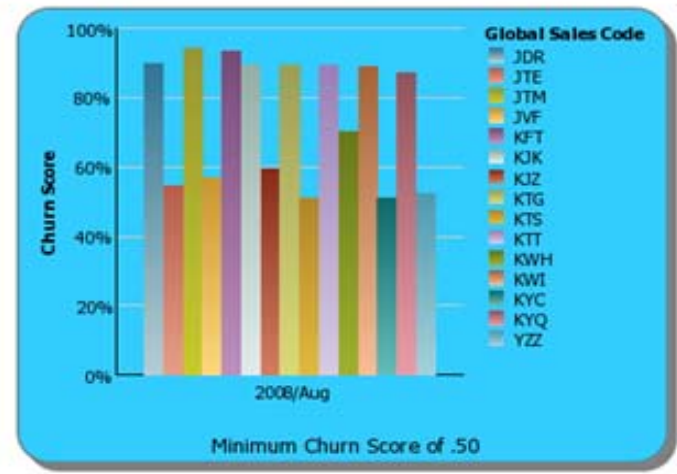
Method: Contains the exact string

Type: Any

Modified: Any

Scope: Public Folders

Cognos Viewer



## ***Customer references***

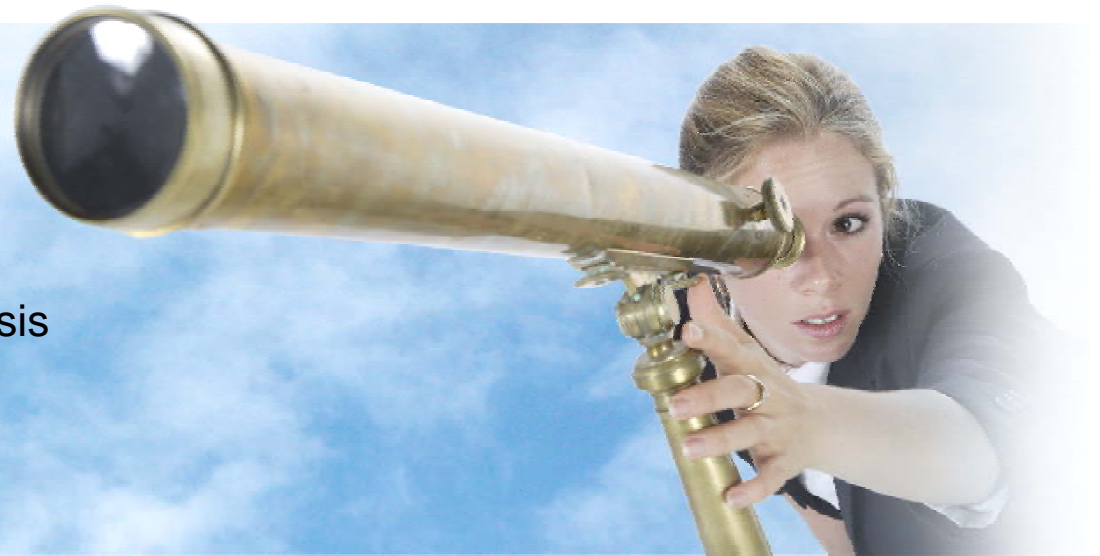
- German Car Manufacturer (Automotive)
  - Identify problematic combination of parts/defects early
  - ⇒ Reduce warranty expenses
  - ⇒ Improve customer satisfaction
  - Association Rule Mining (Sequential Pattern Mining)
- United Health Group (Healthcare)
  - Analysis of End Stage Renal Disease (ESRD)
  - Sequential Pattern Mining, Clustering
- Dillards (Retail)
  - Increase revenue by identifying distinct segments of high-profit customers (and others who behave like them) and targeting them with more personalized offers
  - Clustering, Association Rule Mining
- University Hospital Freiburg (Clinical Research)
  - Analysis of relations between weather and mental symptoms
  - Clustering (Classification)
- Woodforest Bank (Finance)
  - Better understand the characteristics of charge-off customers as a first step to developing predictive models for use in the approval process
  - Clustering, (Classification)



## *Two Types of Data Mining – Discovery & Predictive*

### **Discovery**

- Automatically find trends and patterns
- Answer unasked questions
- Relatively undirected analysis
- Tool reports on findings
- In a word – “Easier”
- Useful for non-statisticians



### **Predictive**

- Specific question
- Probability associated with outcomes
- Directed analysis
- Iterative process
  - Train, Test, Apply
- Apply model in database at customer touch points

