IBM Cognos Forum

Ignite knowledge, ideas, connections

Predictive Analytics

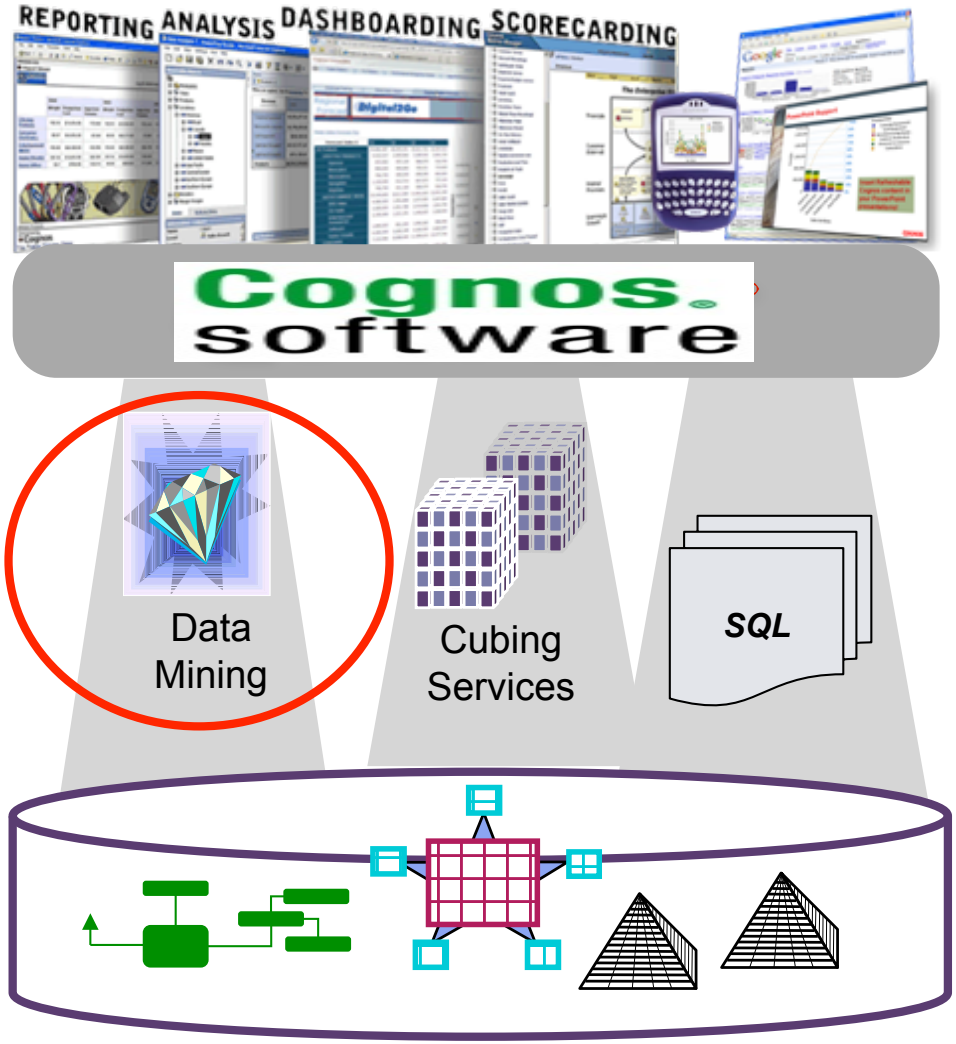
Information Management

Cognos. software

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

IBM Cognos 8 Integration with IBM InfoSphere Warehouse

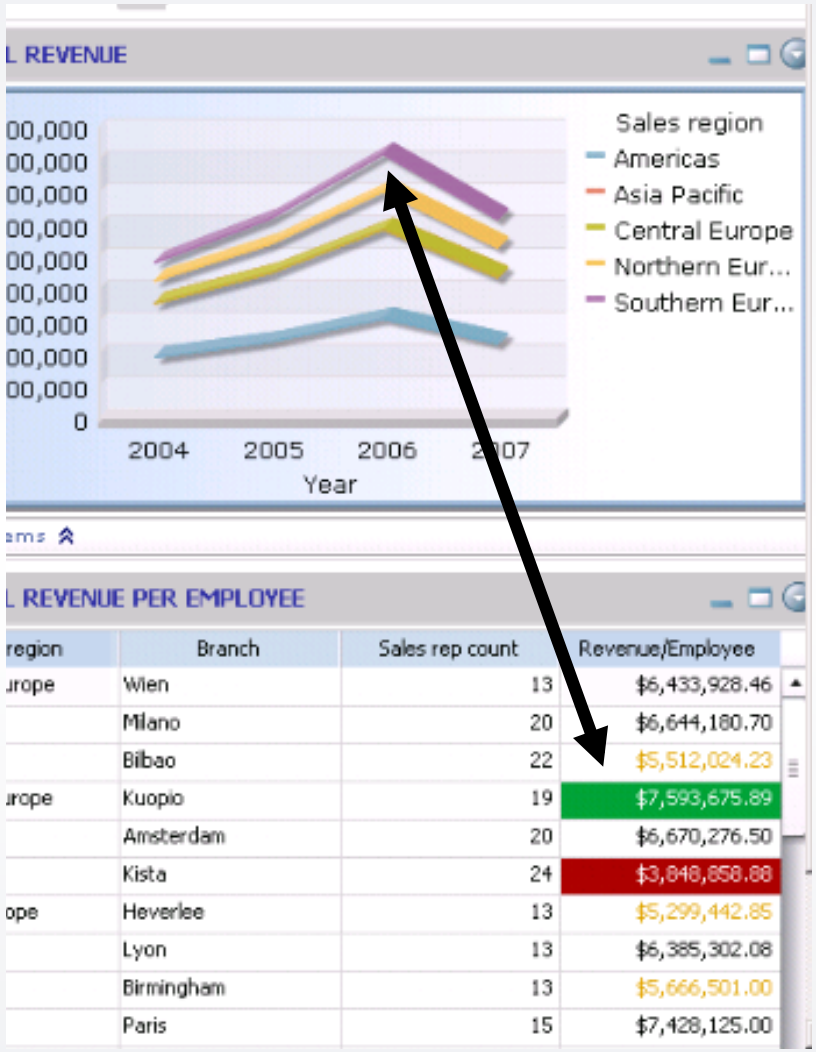
Challenge: Enable more sophisticated and predictive BI applications

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



The Four Styles of Analysis with IBM Cognos 8

Broad Usage (Consumers) ⟷ **Focused Usage** (Specialist)

Analytical Reporting *Drill*	Trending Slice and Dice	Scenario Modeling What-if	Predictive Analytics What might be
• Top down view • Drillable reports • Sort top & bottom • Review then query	• Personal exploration • Compare & contrast • Rotate and nest • Work disconnected	• Model scenarios • Reorganize, reshape • Compare scenarios • Save versions	• Uncover patterns • Statistical algorithms • Mine data and text • Predict outcomes
• Market shifts • Product ranking	• Sales trend analysis • Market analysis	• Financial analysis • Profitability analysis	• Fraud prevention • Churn analysis

Information Management

Cognos. software

The most important ingredient for meaningful predictive analytics:

~~University Degree in Statistics~~

~~Computer Science~~ Education

Business Understanding

Two Types of Data Mining – Discovery & Predictive

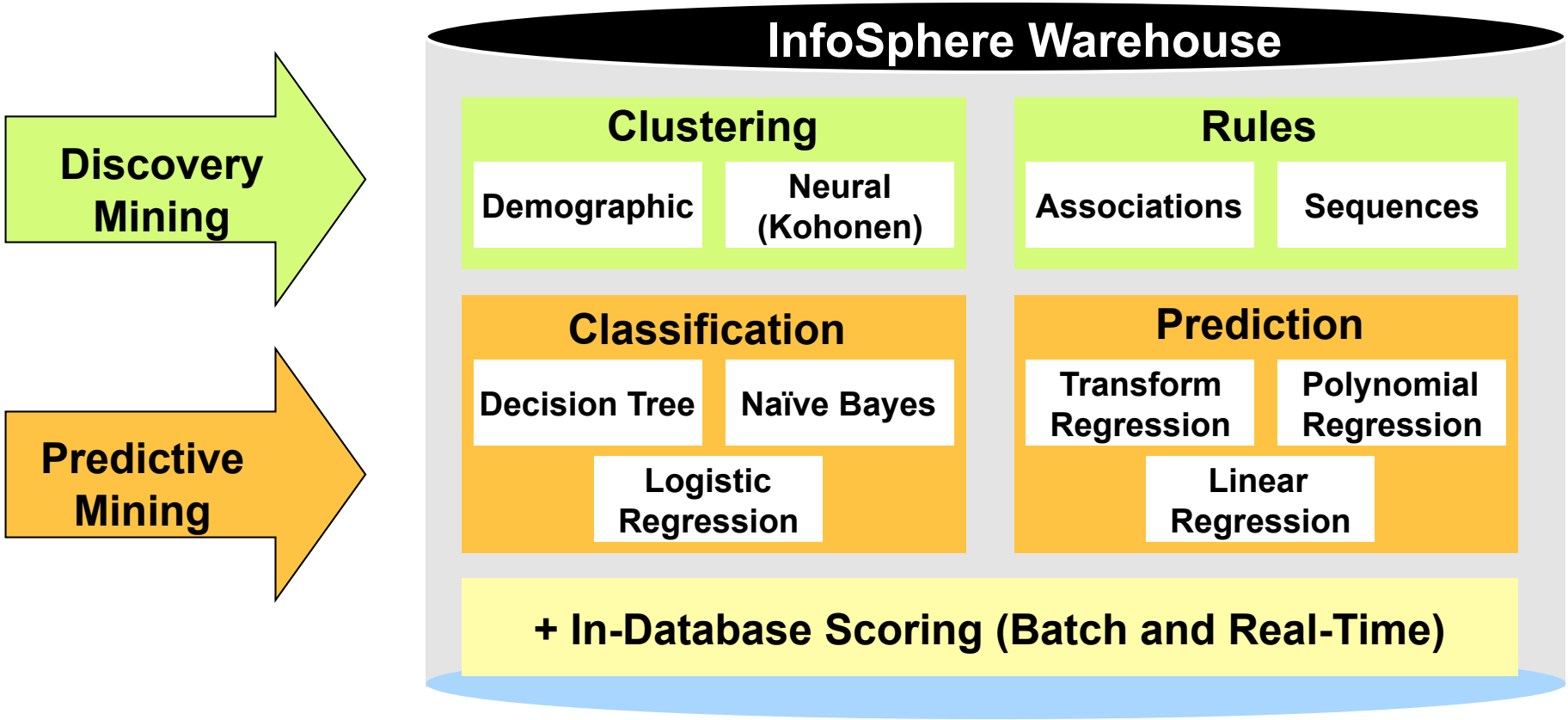
Discovery

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

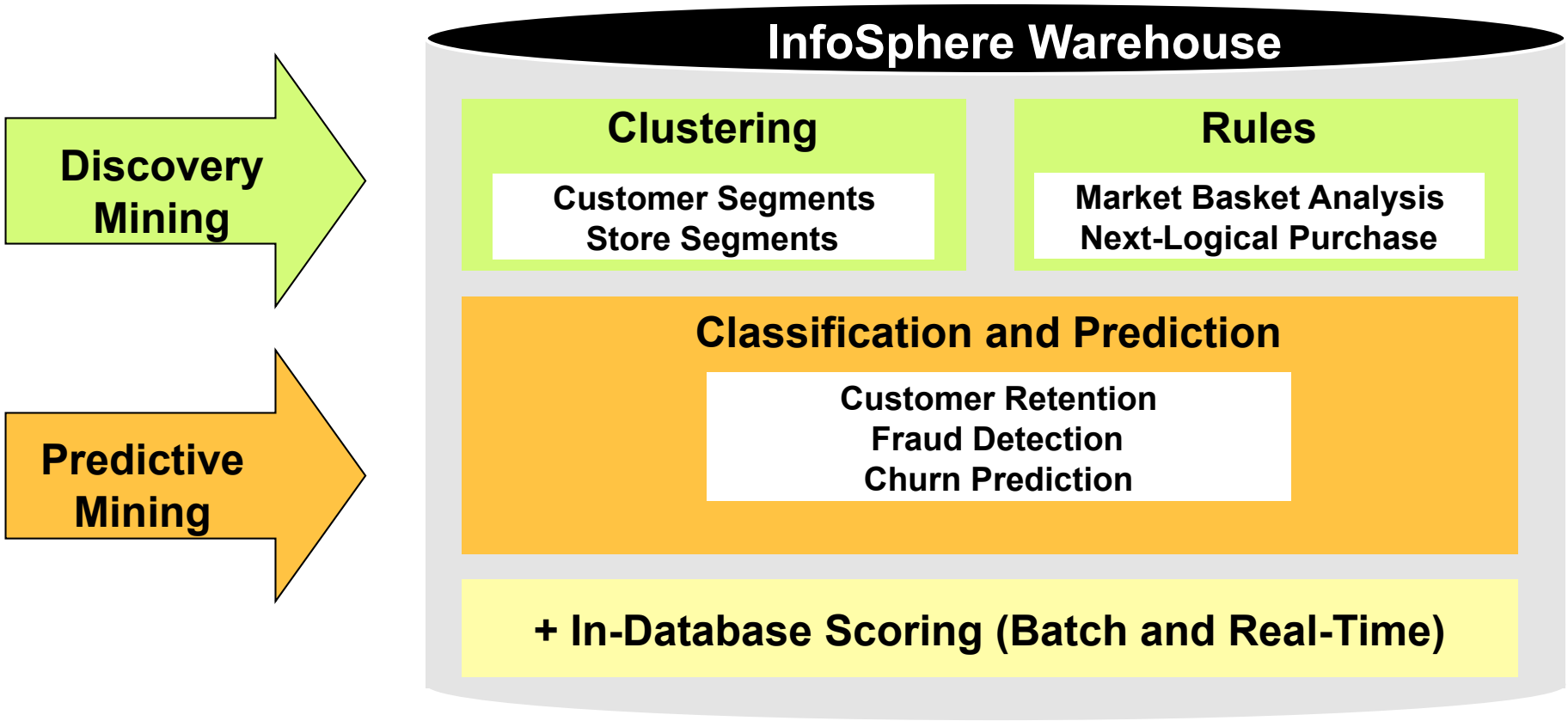
Predictive

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

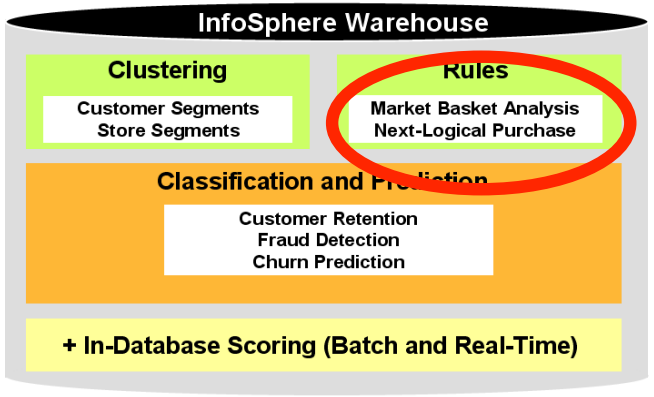
InfoSphere Data Mining Methods and Algorithms



InfoSphere Data Mining Application Examples

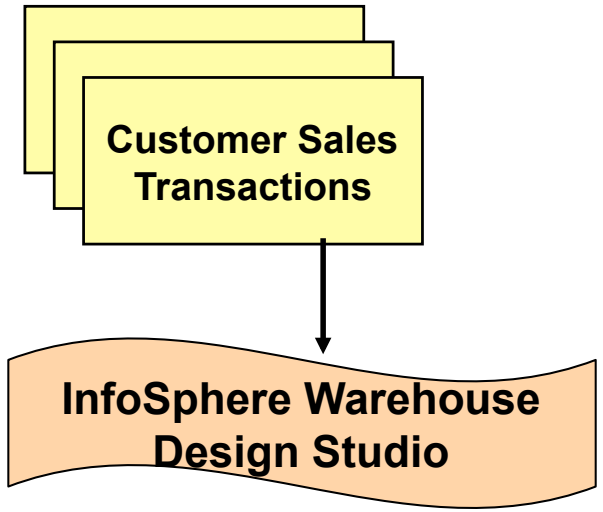


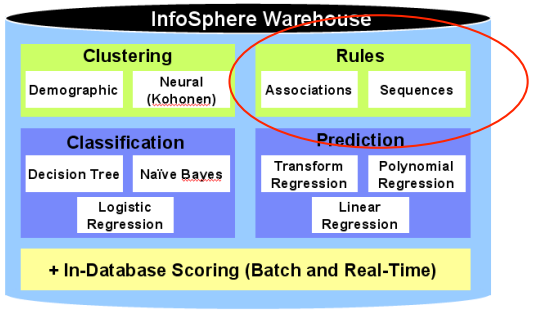
Market Basket Analysis and Next Logical Purchase



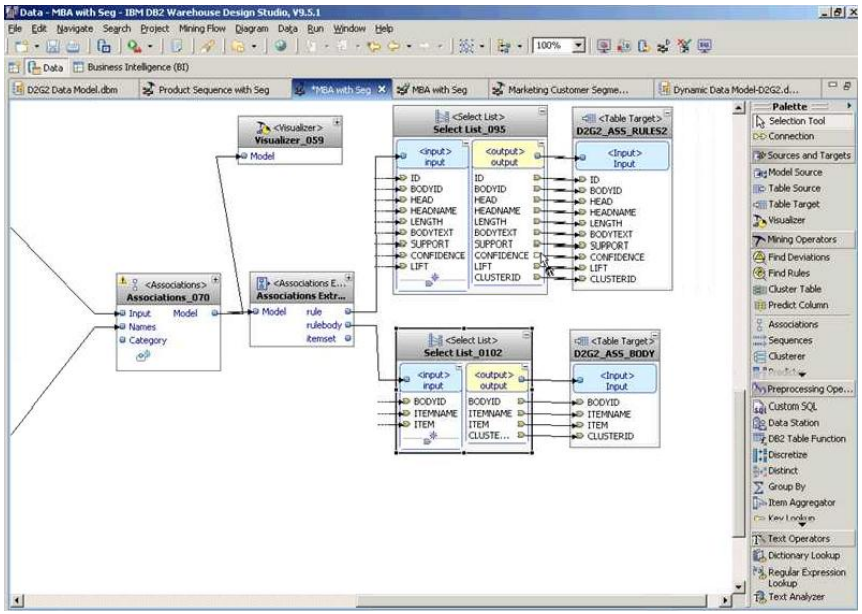
- Customer Background
  - Electronics Retailer
- Business Problem
  - Need to better understand what products are sold together and
  - predict the next, most-logical product a customer will purchase
- Business Goals
  - Develop a data mining flow to use advanced analytics to show driver and implied 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

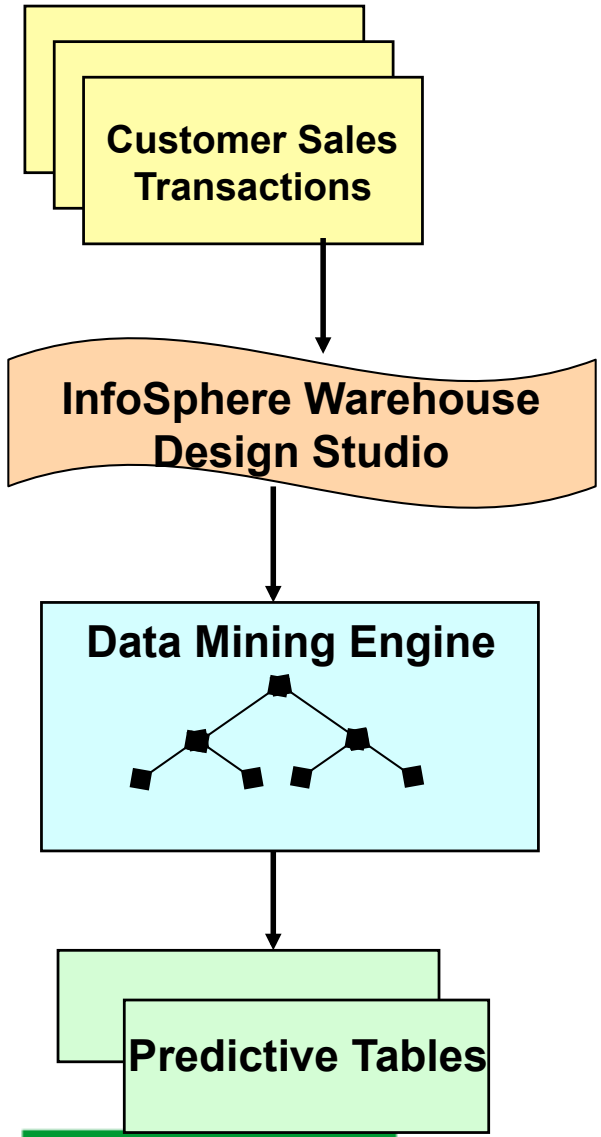




- **Step 1 – Create Mining Flow in InfoSphere Warehouse using Design Studio operators via a drag-and-drop palette approach. In this case we use the associations and sequences algorithms.**

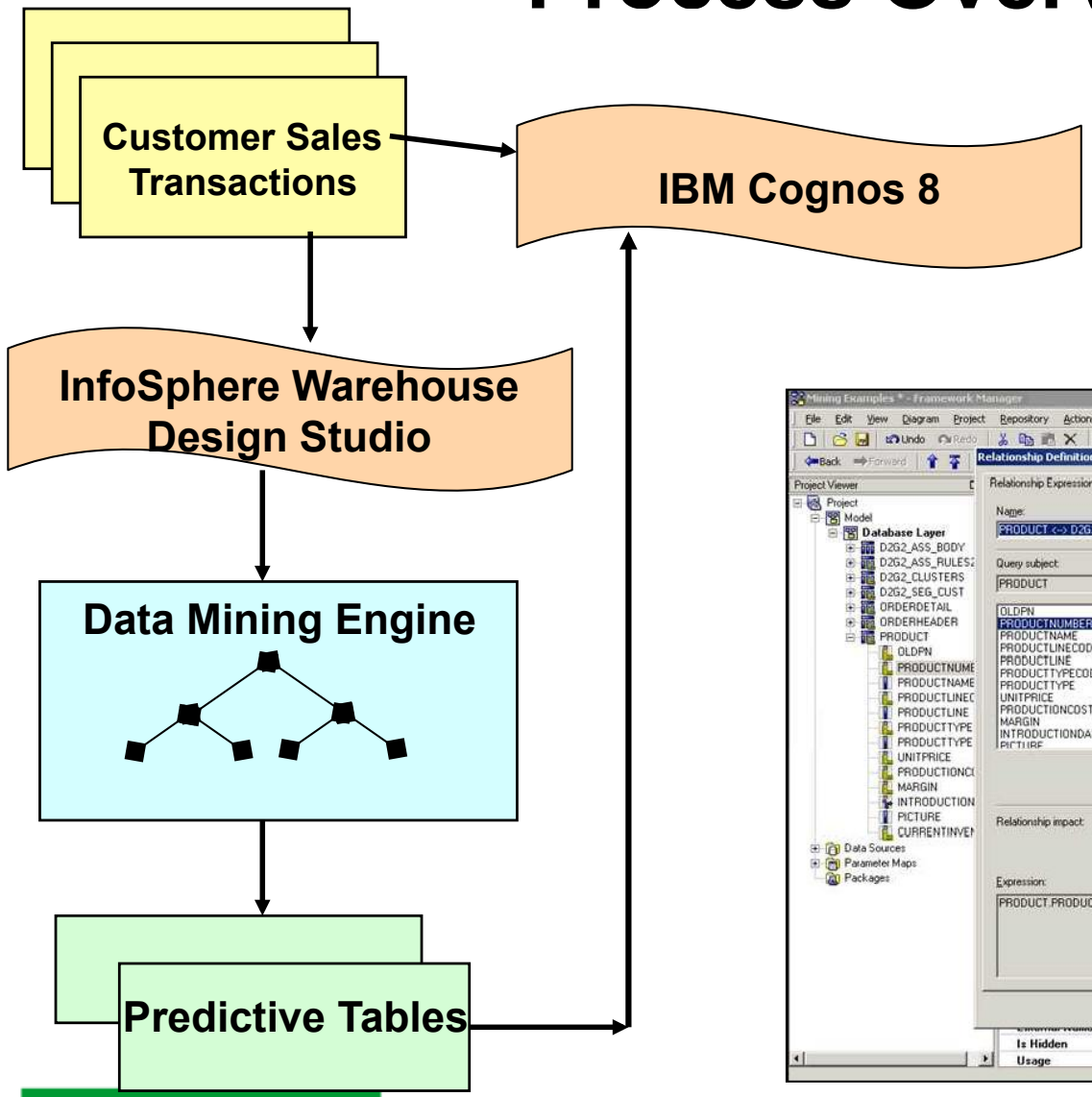


Process Overview

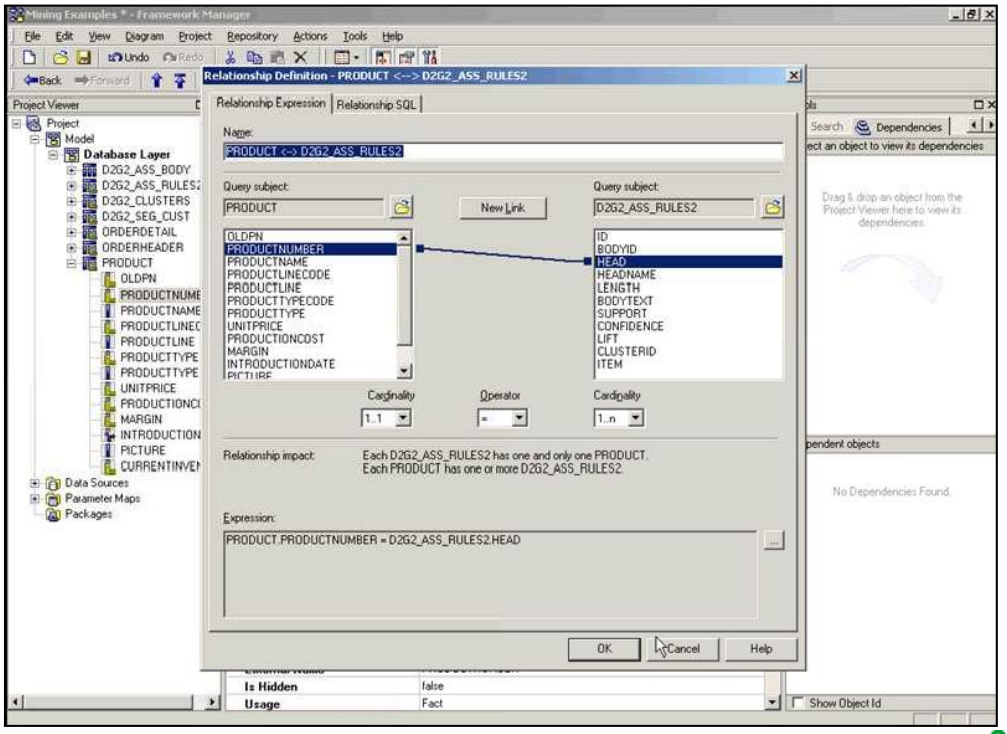


- Step 1 – Create Mining Flow in InfoSphere Warehouse using Design Studio operators via a drag-and-drop palette approach
- **Step 2 – Run the data mining engine to create the predictive tables containing information about which products are purchased together (market basket analysis) or over time (sequencing)**

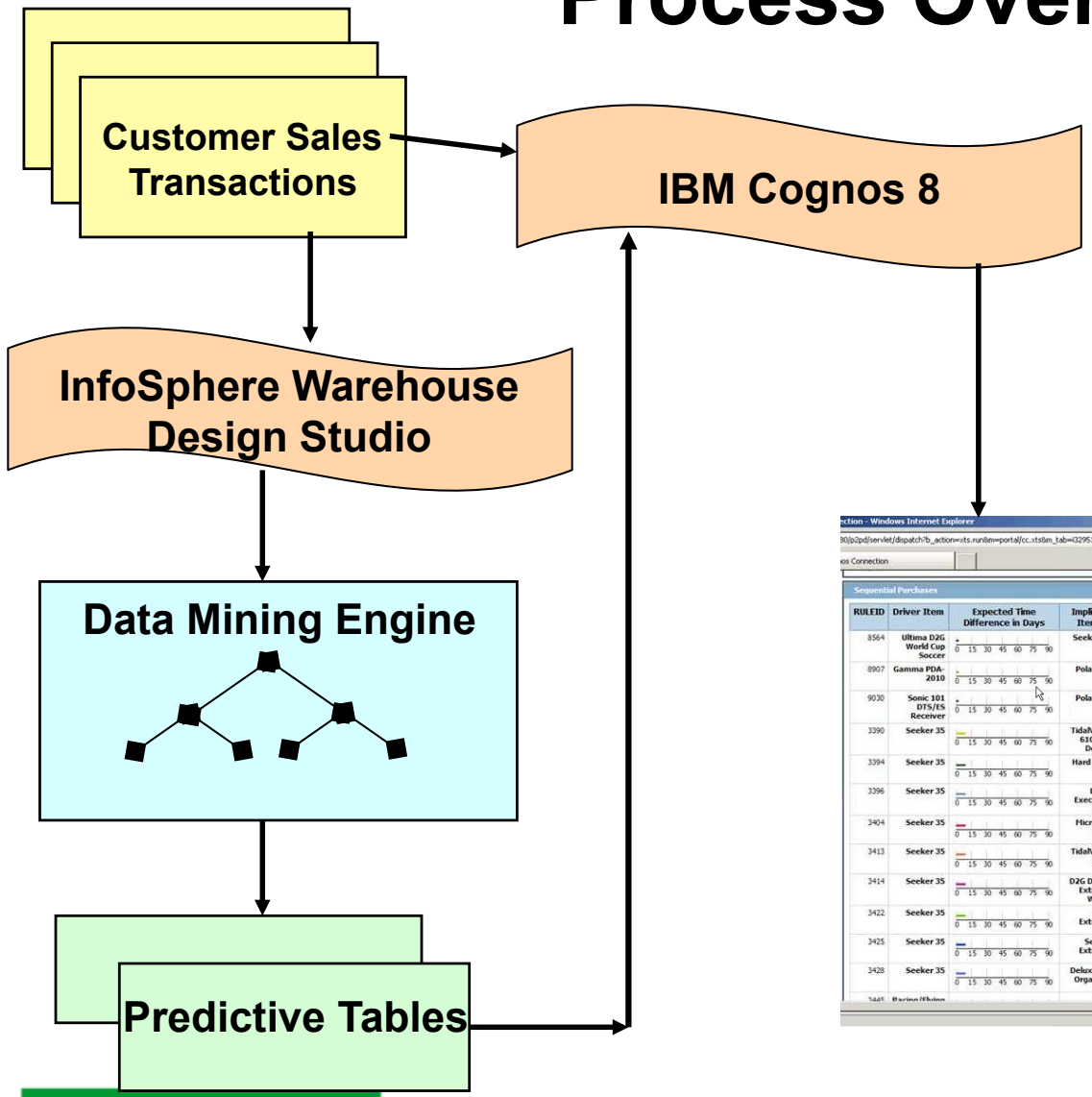
Process Overview



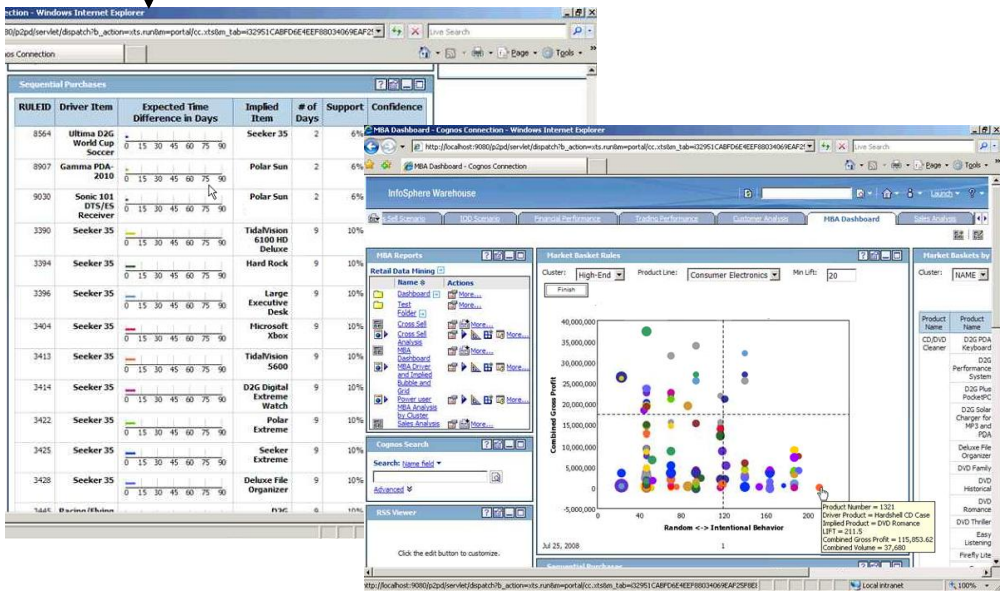
- **Step 3 - Use IBM Cognos 8 Framework Manager to setup table relationships with existing tables and the predictive results tables. Can also define functions and stored procedures as req'd.**

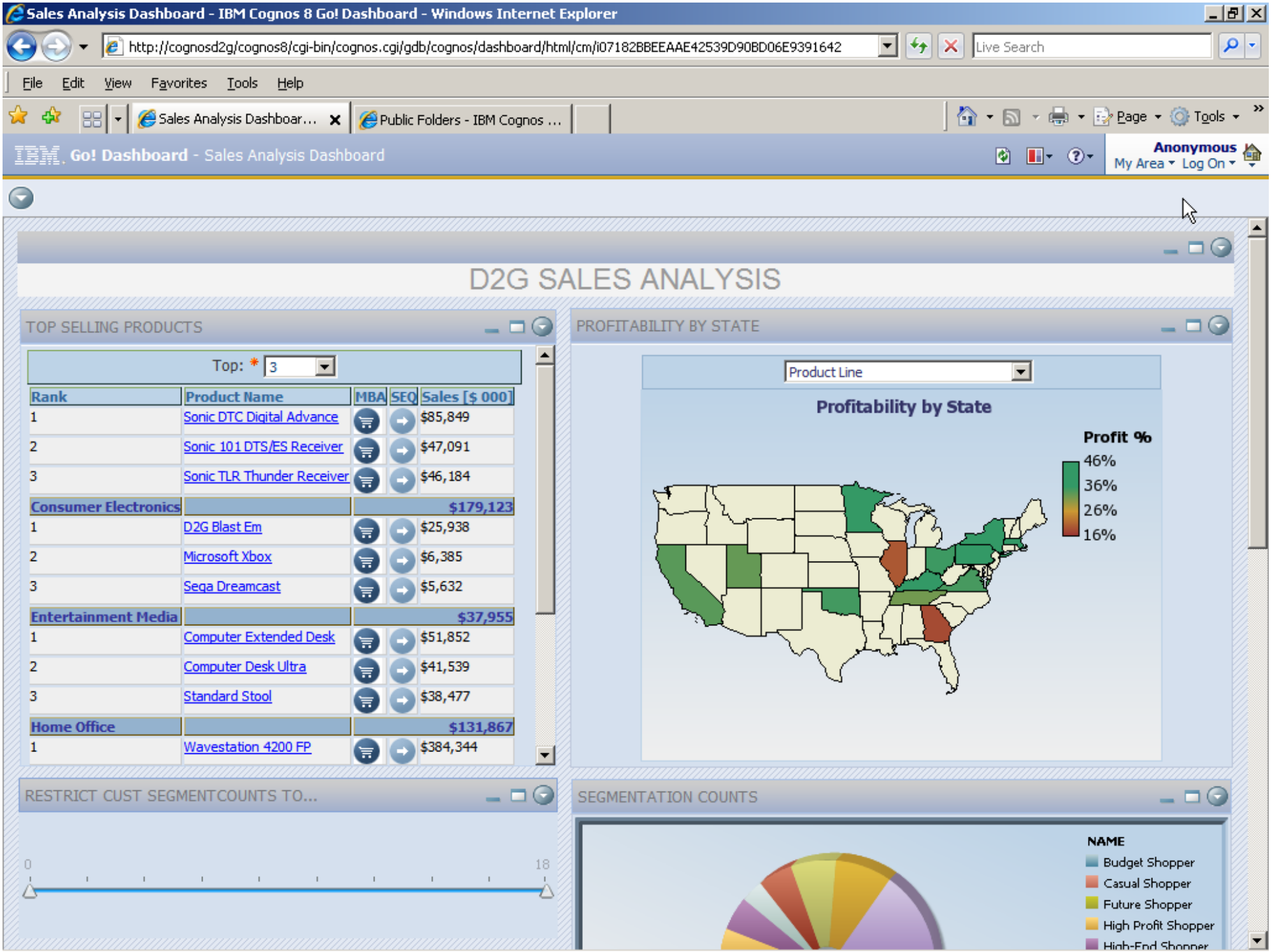


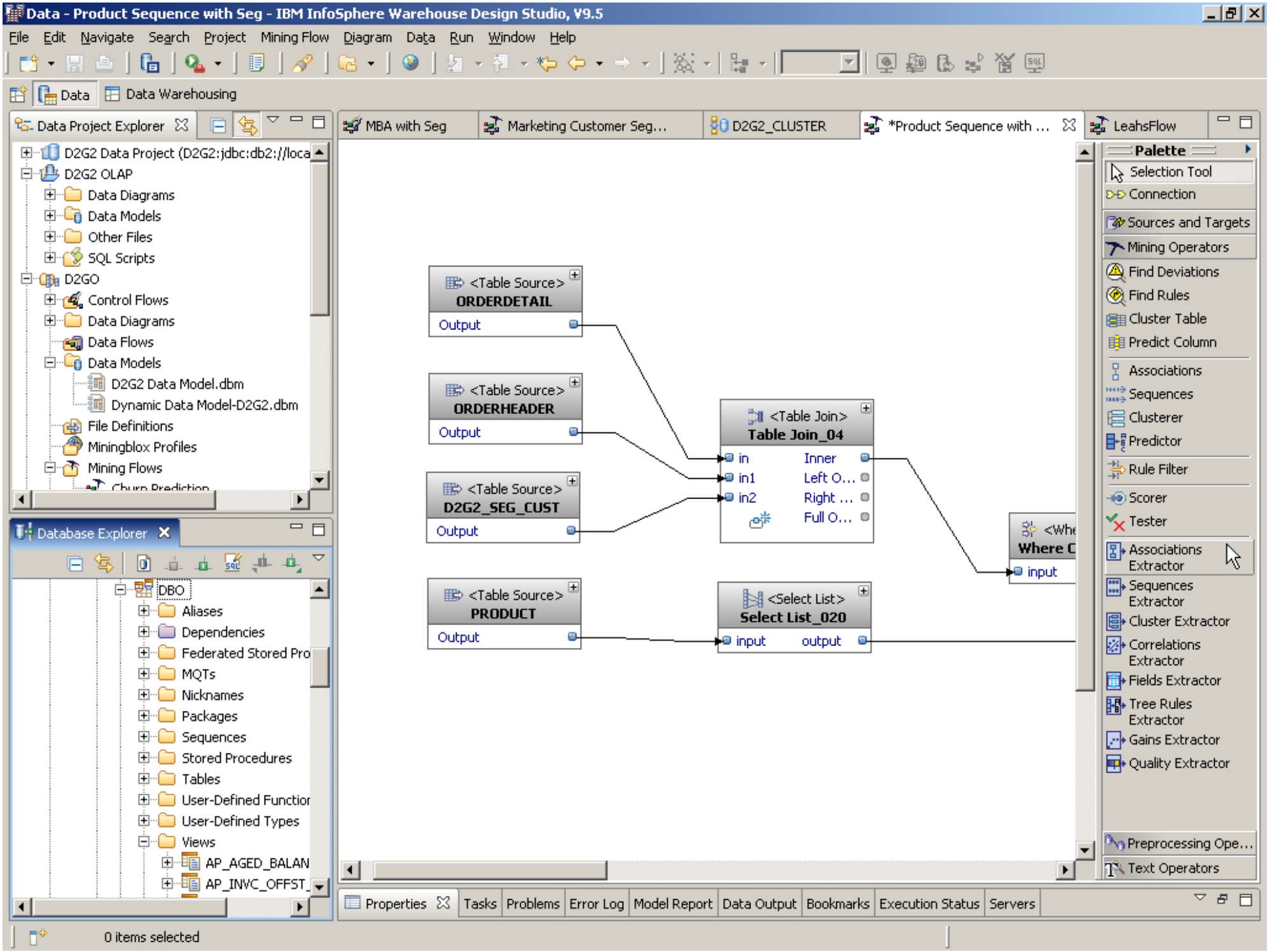
Process Overview



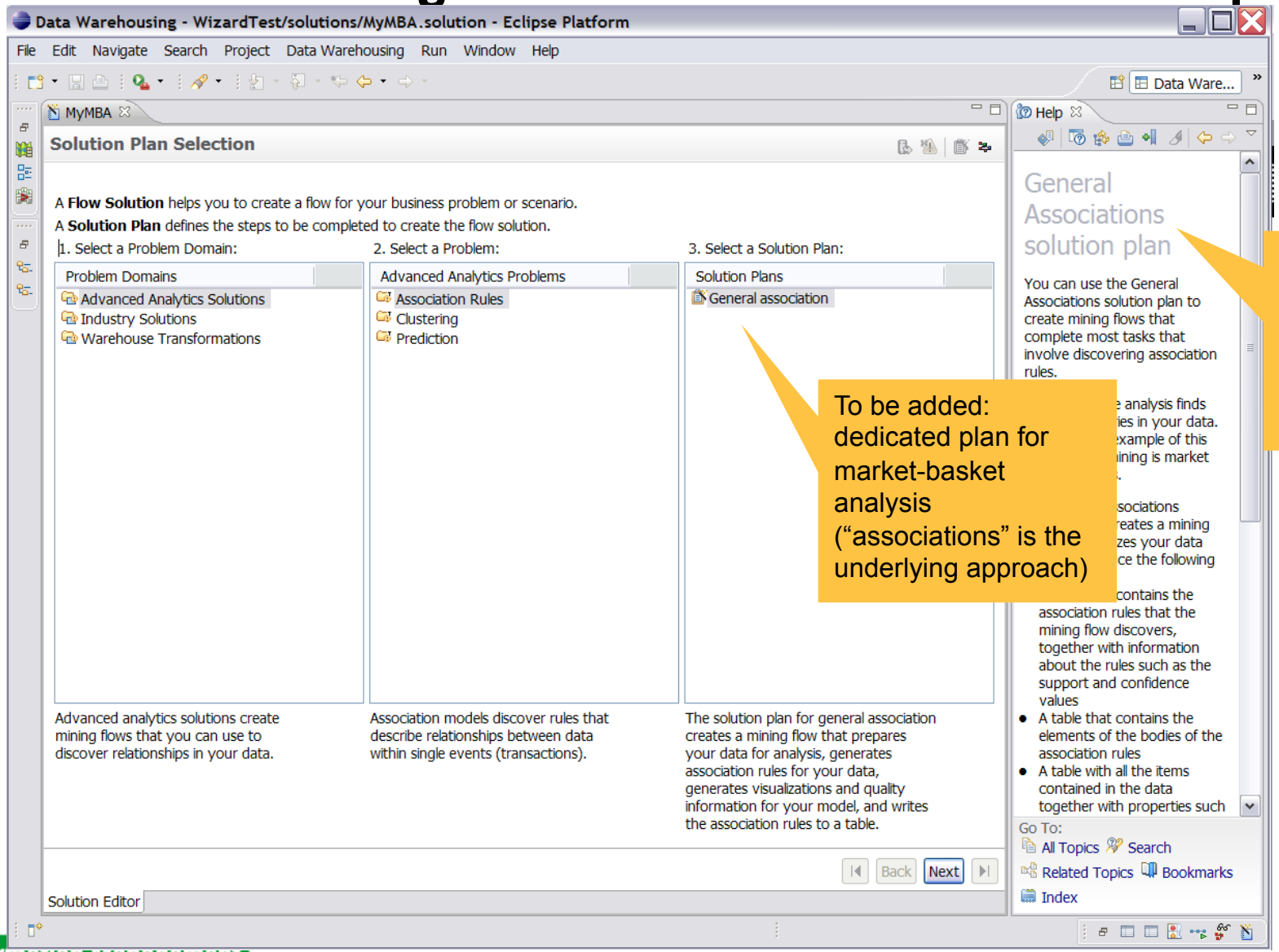
- **Step 4 - Use IBM Cognos 8 to create Dashboards, Reports etc to communicate the information broadly**




Sales Analysis Dashboard - IBM Cognos 8 Go! Dashboard - Windows Internet Explorer
http://cognosd2g/cognos8/cgi-bin/cognos.cgi/gdb/cognos/dashboard/html/cm/i07182BBEEAAE42539D90BD06E9391642
Live Search
File Edit View Favorites Tools Help
Sales Analysis Dashboar...
Public Folders - IBM Cognos ...
Page Tools
Go! Dashboard - Sales Analysis Dashboard
Anonymous
My Area Log On
D2G SALES ANALYSIS
TOP SELLING PRODUCTS
Top: * 3
Rank Product Name MBA SEQ Sales [$ 000]
1 Sonic DTC Digital Advance $85,849
2 Sonic 101 DTS/ES Receiver $47,091
3 Sonic TLR Thunder Receiver $46,184
Consumer Electronics $179,123
1 D2G Blast Em $25,938
2 Microsoft Xbox $6,385
3 Sega Dreamcast $5,632
Entertainment Media $37,955
1 Computer Extended Desk $51,852
2 Computer Desk Ultra $41,539
3 Standard Stool $38,477
Home Office $131,867
1 Wavestation 4200 FP $384,344
PROFITABILITY BY STATE
Product Line
Profitability by State
Profit %
46%
36%
26%
16%
RESTRICT CUST SEGMENTCOUNTS TO...
0
18
SEGMENTATION COUNTS
NAME
Budget Shopper
Casual Shopper
Future Shopper
High Profit Shopper
High-End Shopper


Data - Product Sequence with Seg - IBM InfoSphere Warehouse Design Studio, V9.5
File Edit Navigate Search Project Mining Flow Diagram Data Run Window Help
Data
Data Warehousing
Data Project Explorer
D2G2 Data Project (D2G2:jdbc:db2://loca
D2G2 OLAP
Data Diagrams
Data Models
Other Files
SQL Scripts
D2GO
Control Flows
Data Diagrams
Data Flows
Data Models
D2G2 Data Model.dbm
Dynamic Data Model-D2G2.dbm
File Definitions
Miningblox Profiles
Mining Flows
Database Explorer
DBO
Aliases
Dependencies
Federated Stored Pro
MQTs
Nicknames
Packages
Sequences
Stored Procedures
Tables
User-Defined Functior
User-Defined Types
Views
AP_AGED_BALAN
AP_INVC_OFFST_
MBA with Seg
Marketing Customer Seg...
D2G2_CLUSTER
*Product Sequence with ...
LeahsFlow
<Table Source>
ORDERDETAIL
Output
<Table Source>
ORDERHEADER
Output
<Table Source>
D2G2_SEG_CUST
Output
<Table Source>
PRODUCT
Output
<Table Join>
Table Join_04
in
in1
in2
Inner
Left O...
Right ...
Full O...
<Select List>
Select List_020
input
output
input
Palette
Selection Tool
Connection
Sources and Targets
Mining Operators
Find Deviations
Find Rules
Cluster Table
Predict Column
Associations
Sequences
Clusterer
Predictor
Rule Filter
Scorer
Tester
Associations Extractor
Sequences Extractor
Cluster Extractor
Correlations Extractor
Fields Extractor
Tree Rules Extractor
Gains Extractor
Quality Extractor
Preprocessing Ope...
Text Operators
Properties
Tasks
Problems
Error Log
Model Report
Data Output
Bookmarks
Execution Status
Servers
0 items selected

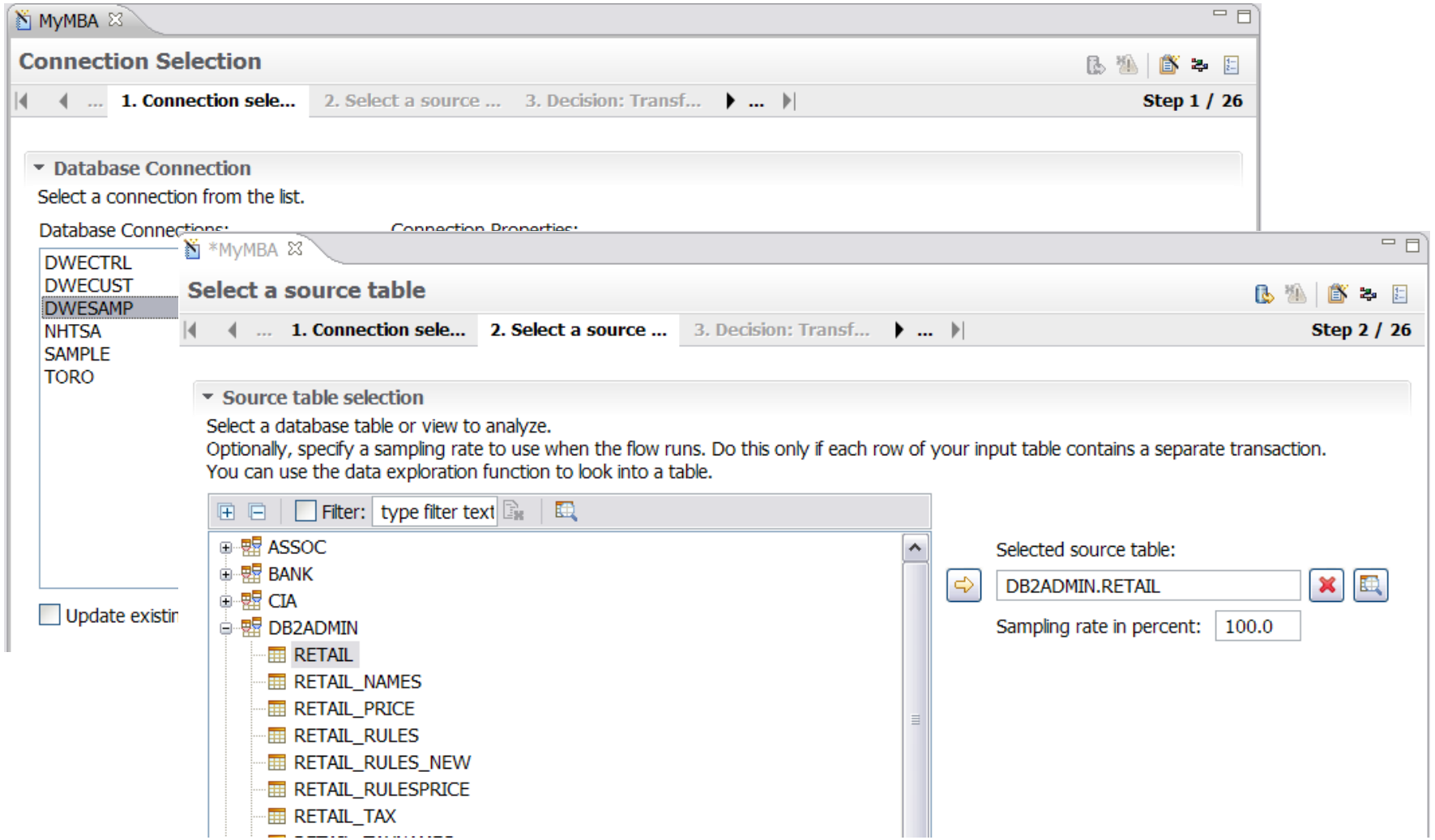
Launch Design Studio and select a solution plan



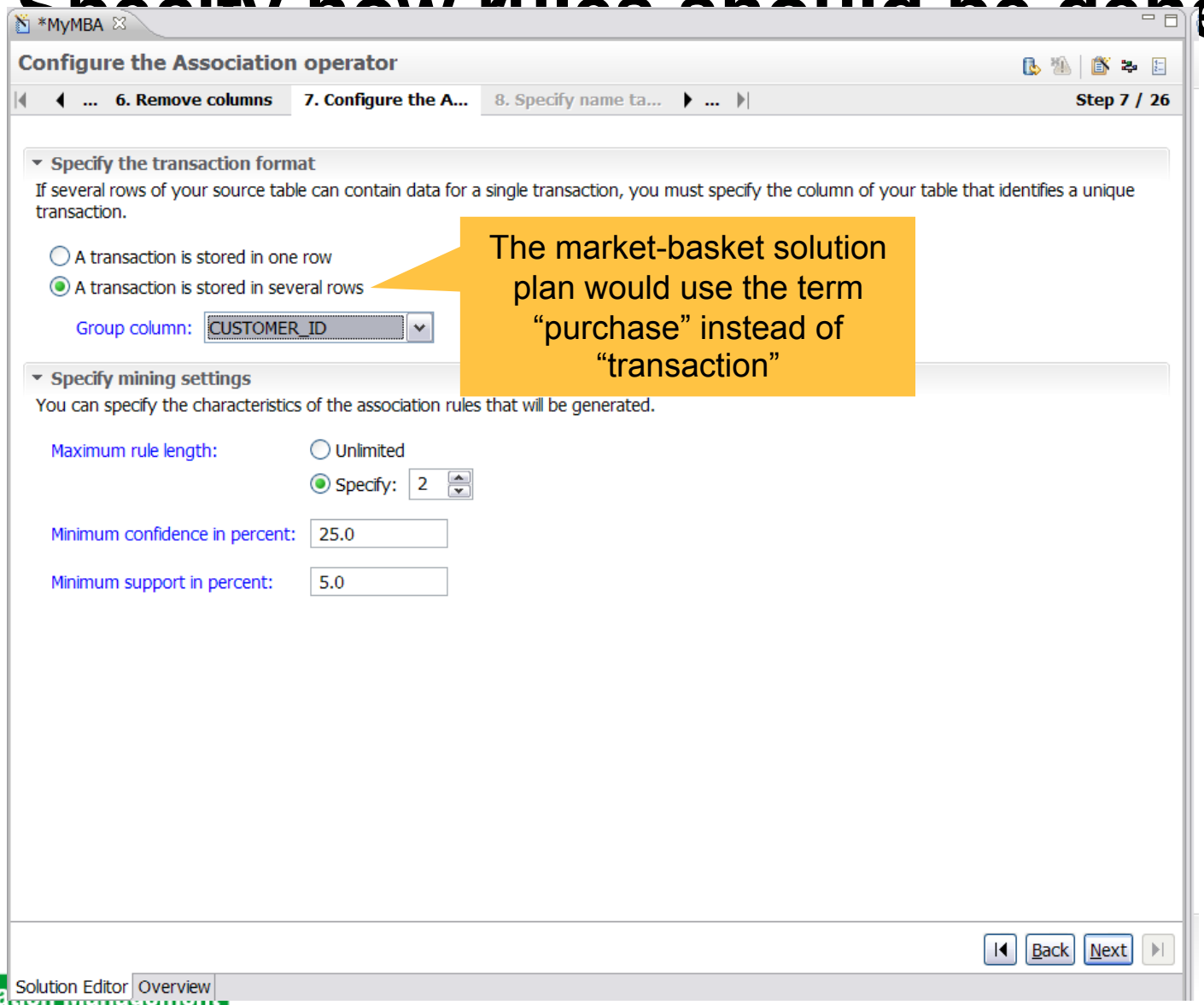
To be added: dedicated plan for market-basket analysis ("associations" is the underlying approach)

Context-sensitive help throughout the solution

Select the data source with customer purchases

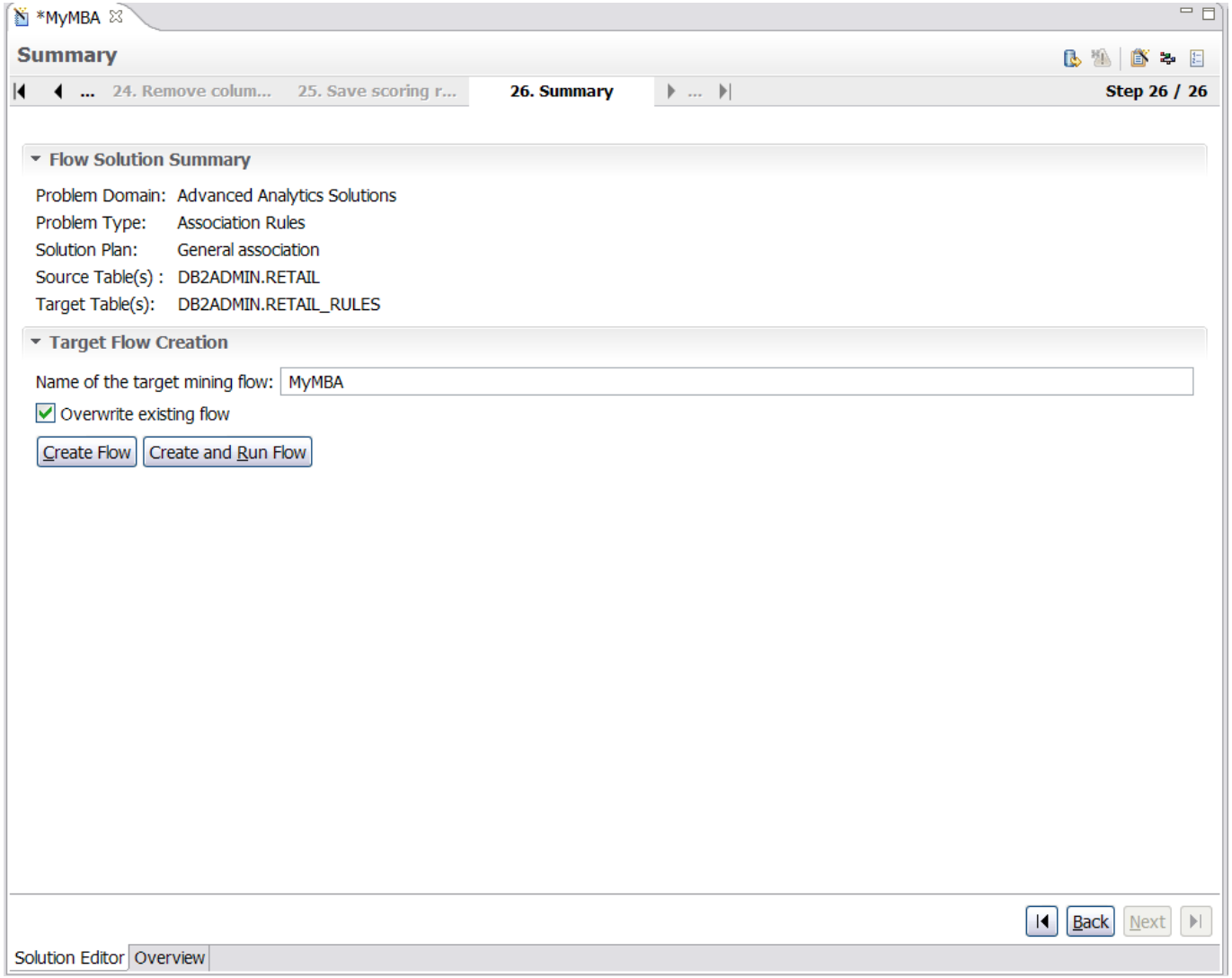


Specify how rules should be generated

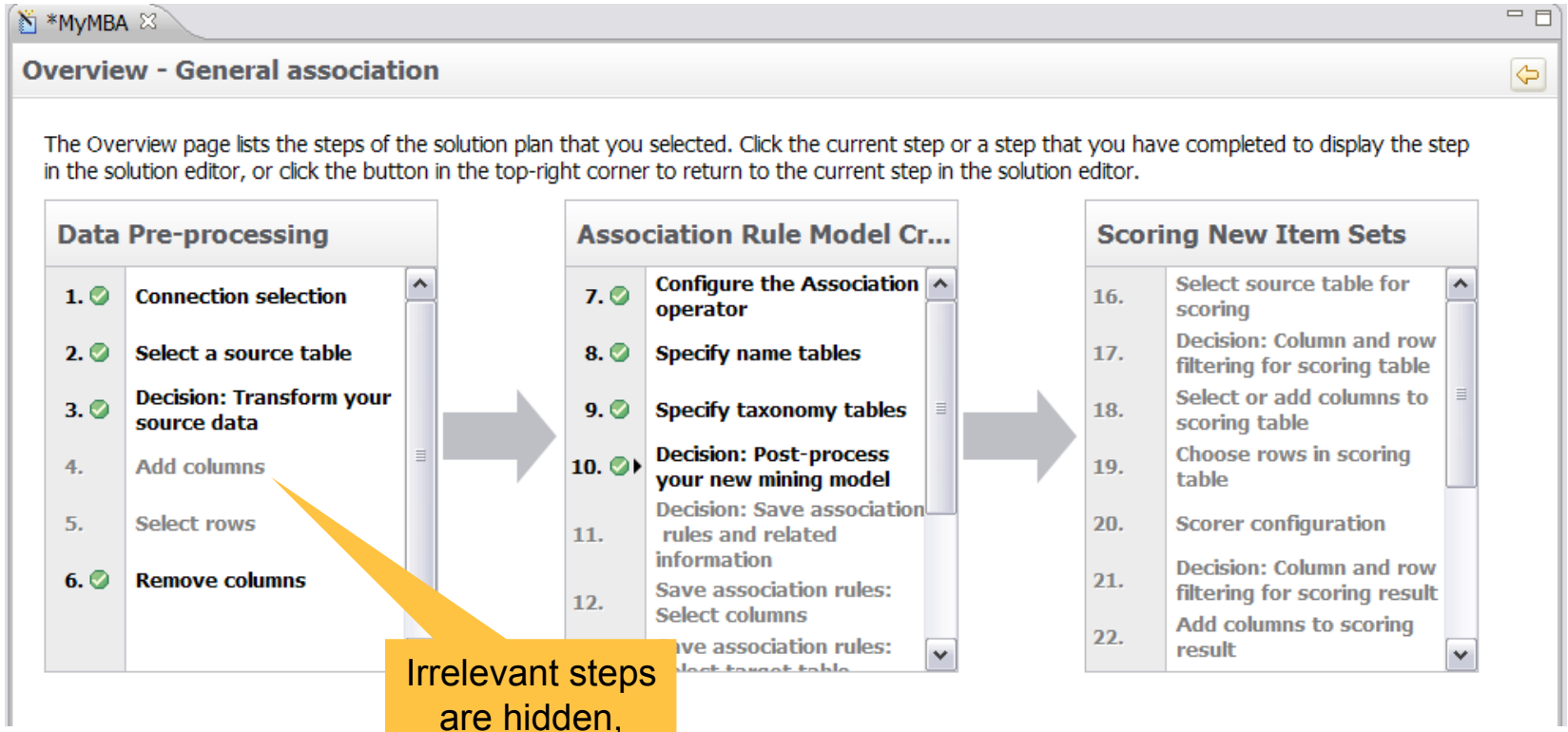


The market-basket solution plan would use the term "purchase" instead of "transaction"

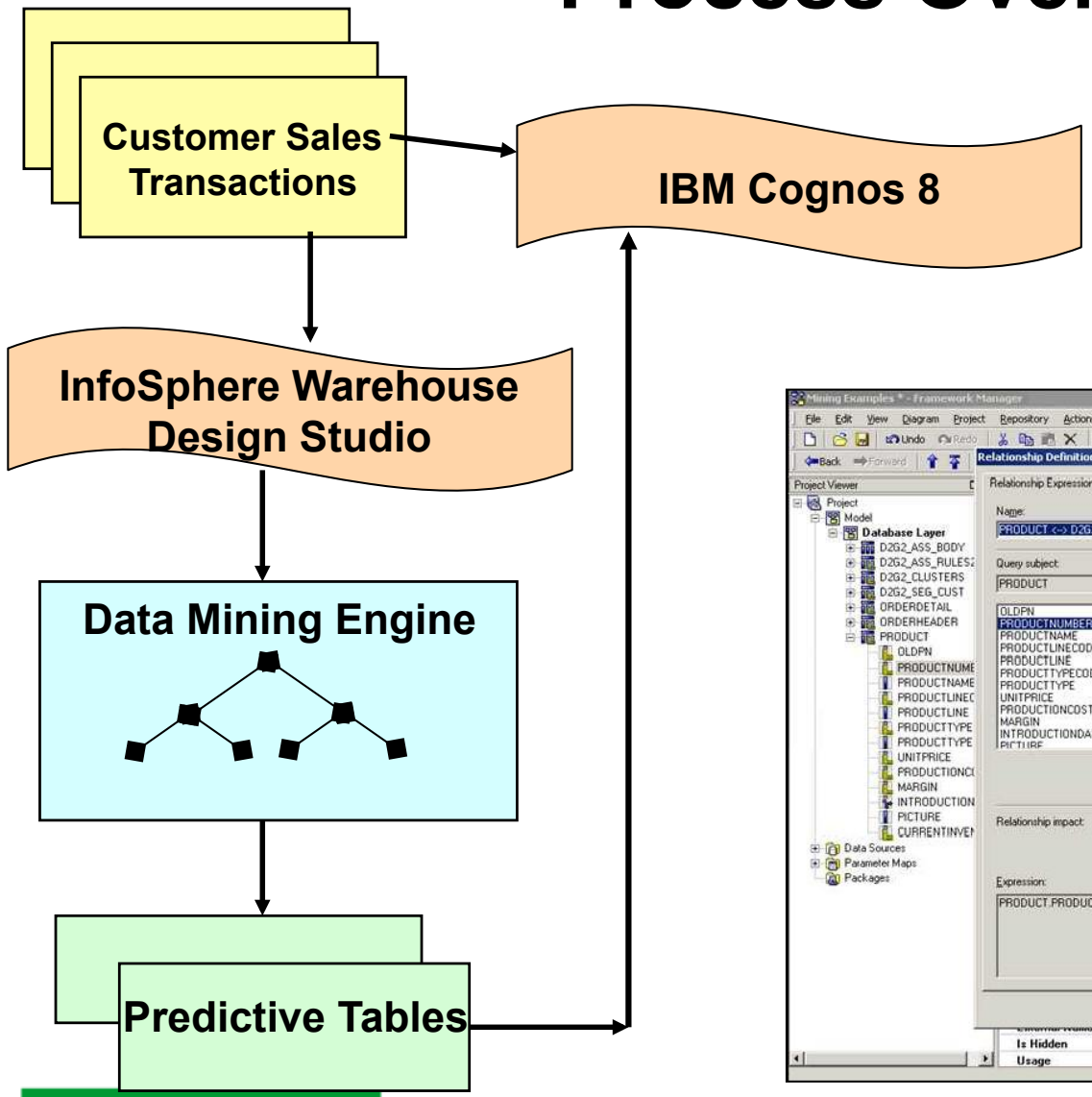
Create (and run) “analytic flow” that generates rules



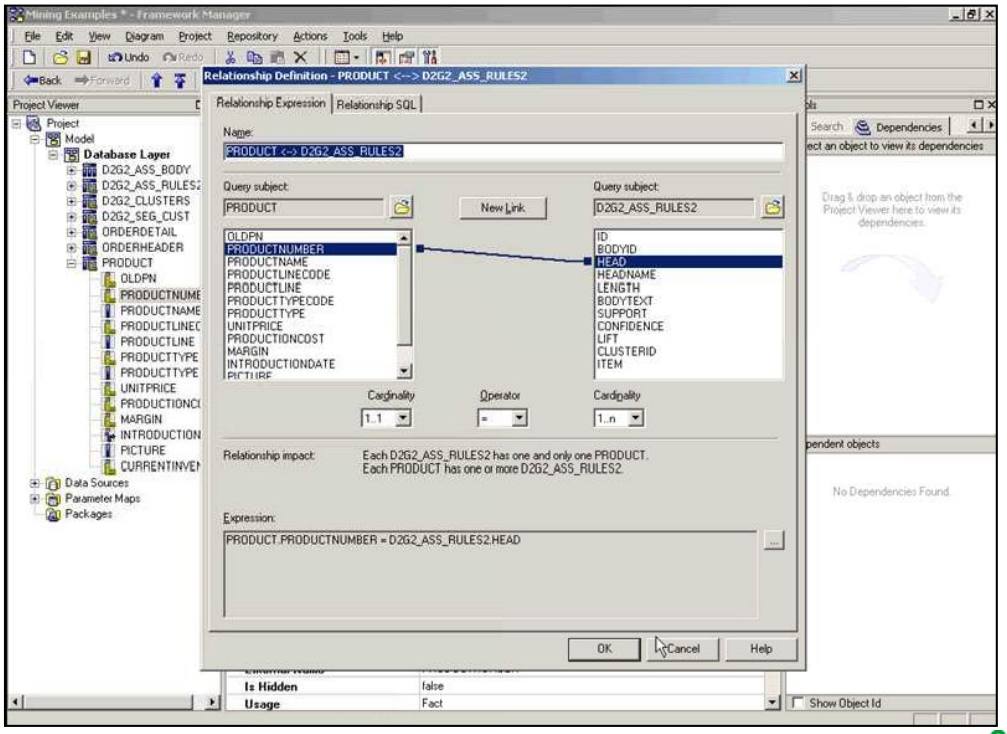
Keep track of where you are



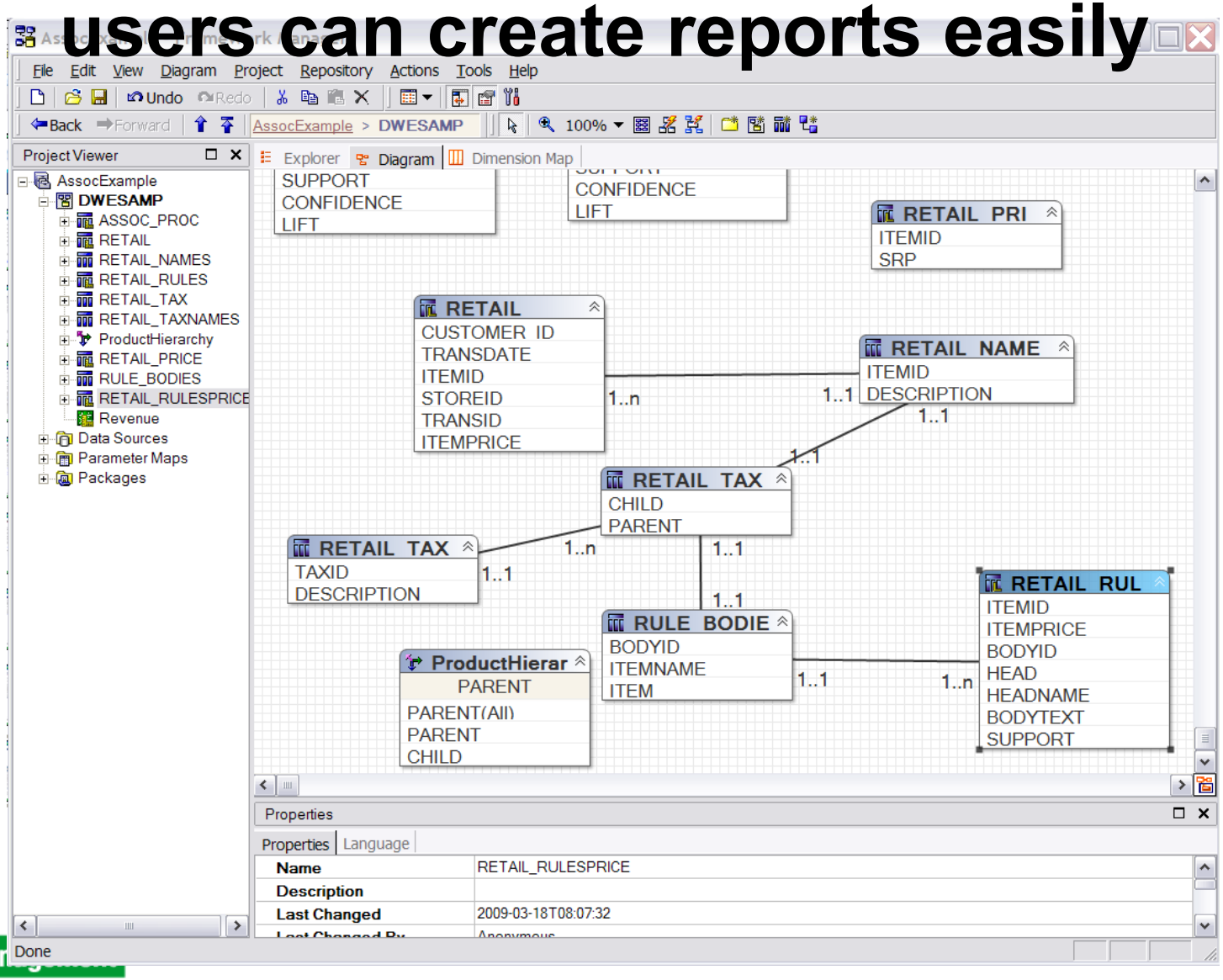
Process Overview



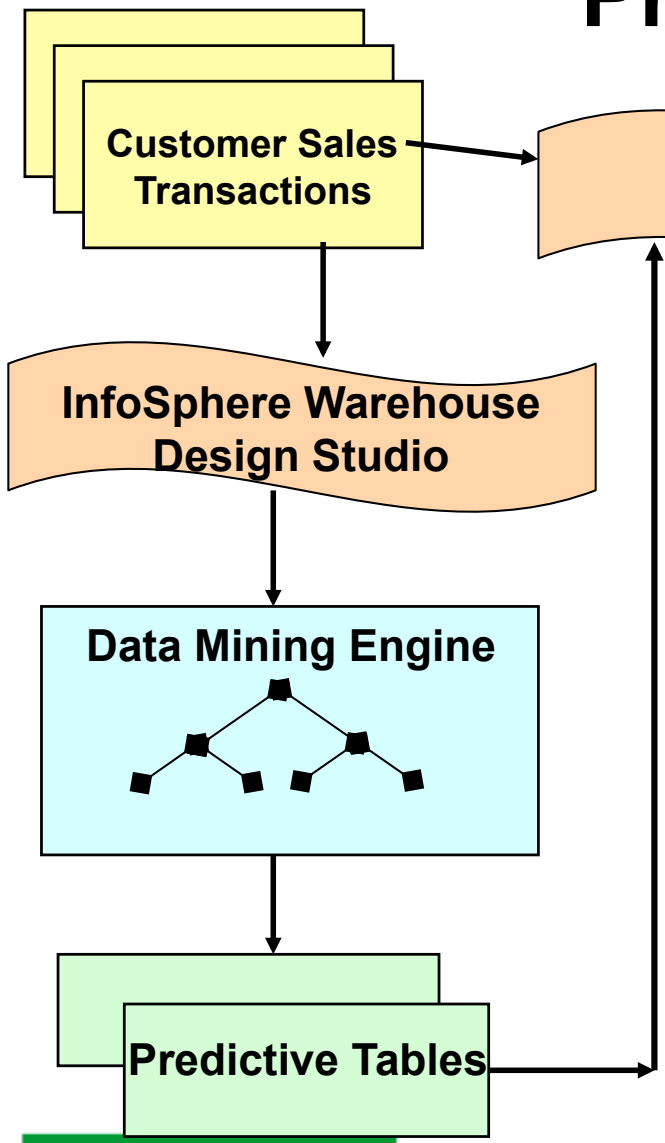
> Step 3 - Use IBM Cognos 8 Framework Manager to setup table relationships with existing tables and the predictive results tables. Can also define functions and stored procedures as req'd.



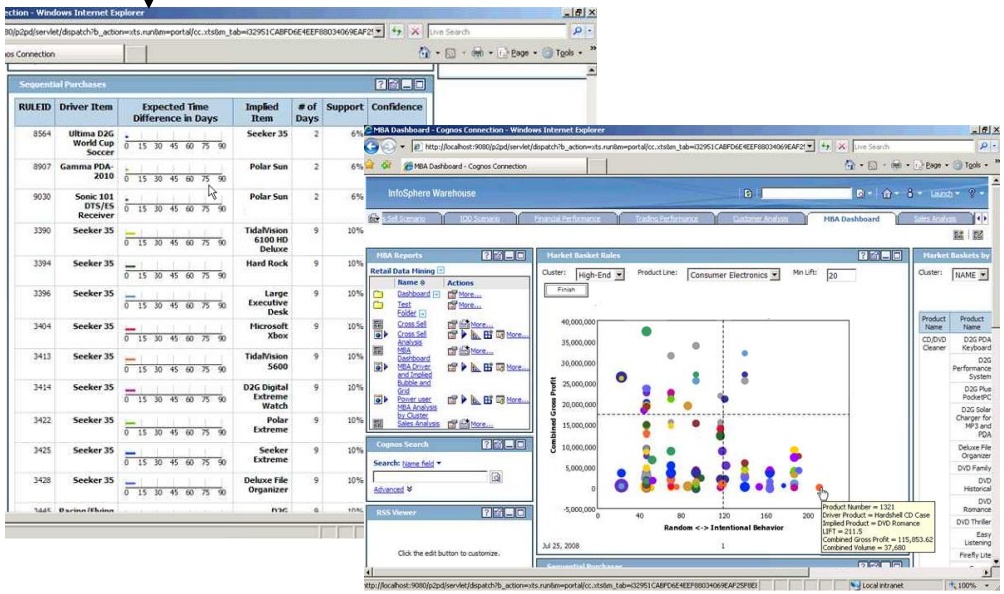
Framework Manager: Model the data so users can create reports easily



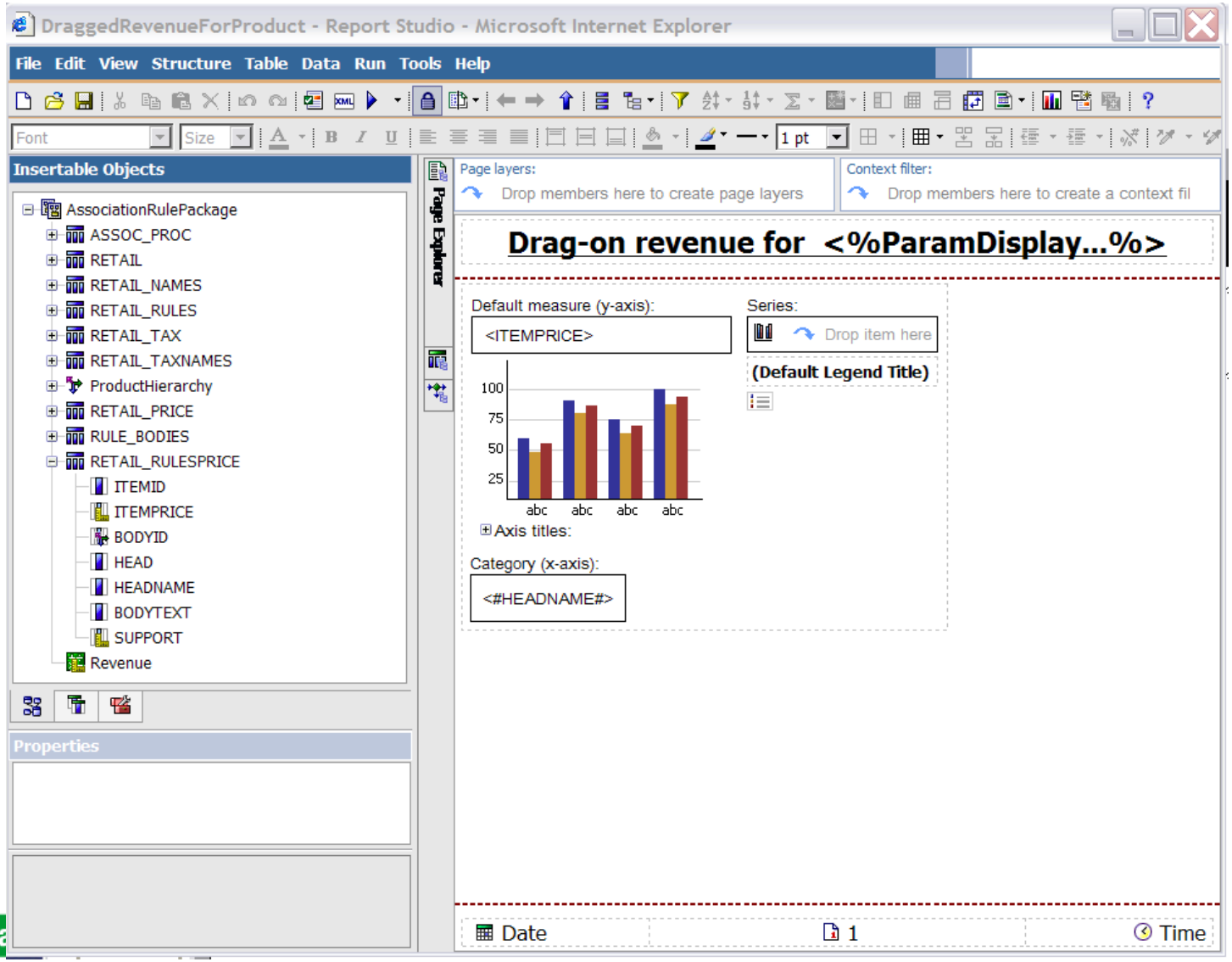
Process Overview

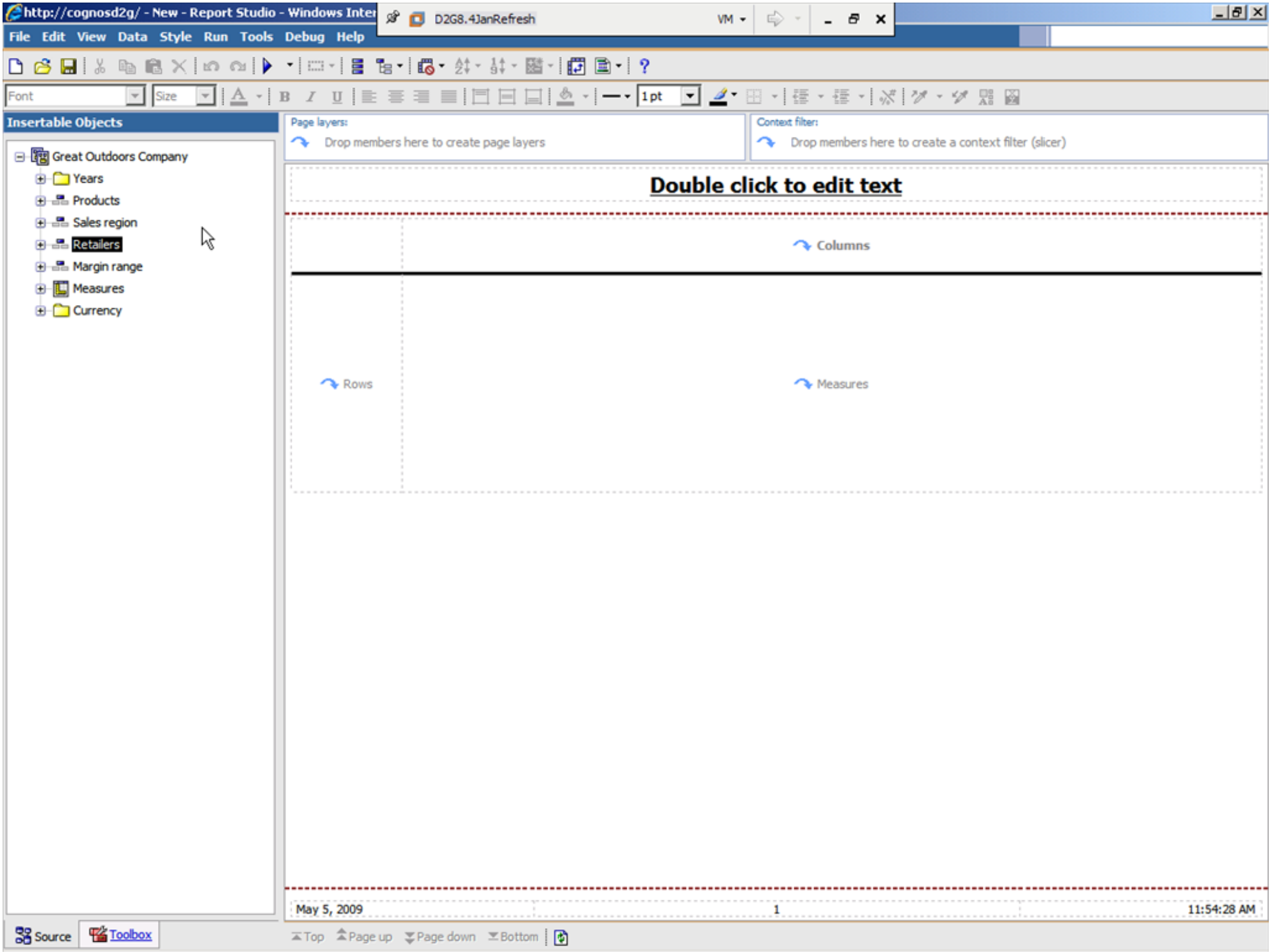


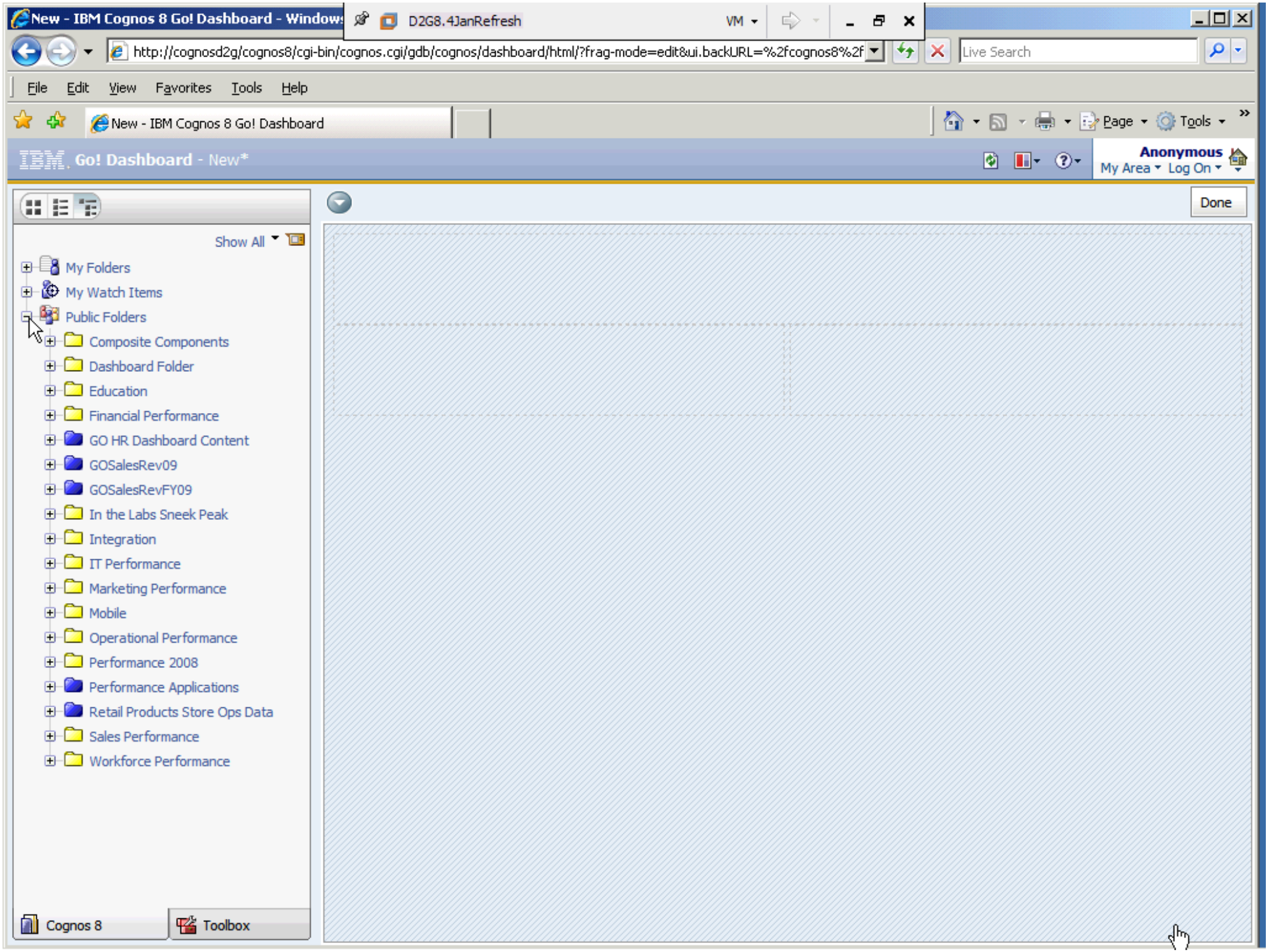
- Step 4 - Use IBM Cognos 8 to create Dashboards, Reports etc to communicate the information broadly



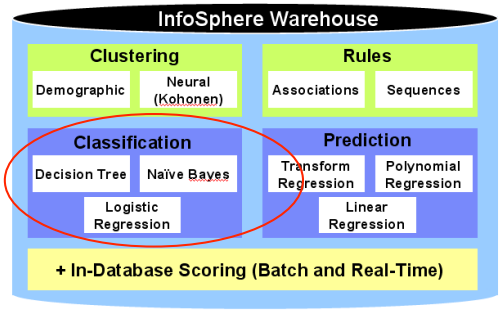
Build Reports, Dashboards, etc




http://cognosd2g/ - New - Report Studio - Windows Inter
D2G8.4JanRefresh
VM
File Edit View Data Style Run Tools Debug Help
Font
Size
1 pt
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
May 5, 2009
1
11:54:28 AM
Source
Toolbox
Top
Page up
Page down
Bottom

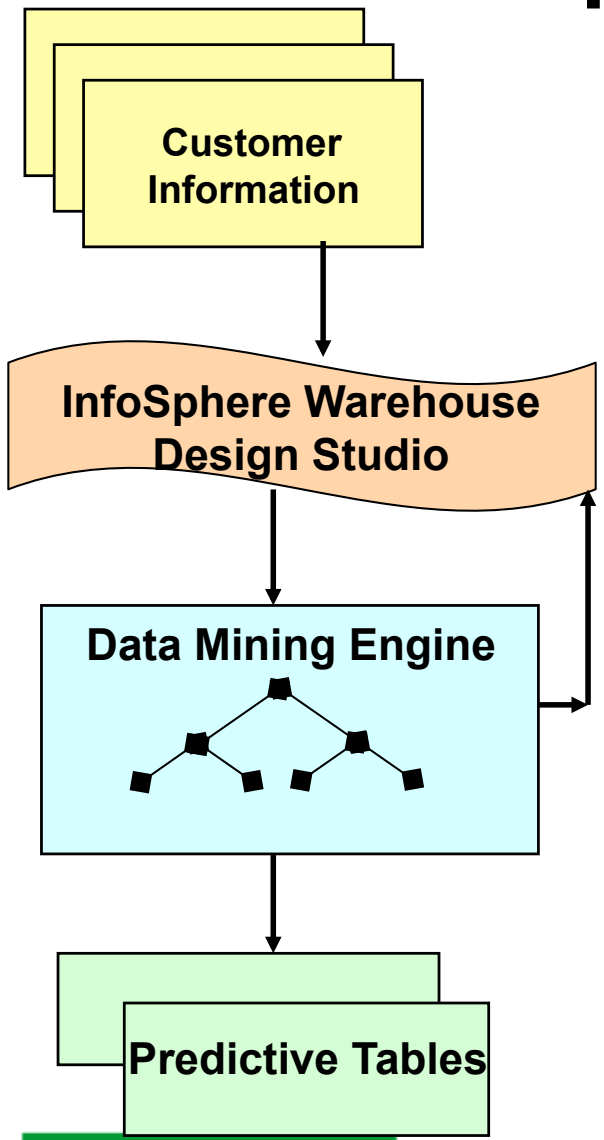

New - IBM Cognos 8 Go! Dashboard - Windows
D2G8.4JanRefresh
VM
http://cognosd2g/cognos8/cgi-bin/cognos.cgi/gdb/cognos/dashboard/html/?frag-mode=edit&ui.backURL=%2fcognos8%2f
Live Search
File Edit View Favorites Tools Help
New - IBM Cognos 8 Go! Dashboard
Page Tools
IBM Go! Dashboard - New*
Anonymous
My Area Log On
Done
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
Cognos 8
Toolbox

Sales Example – Account Abandonment Prediction



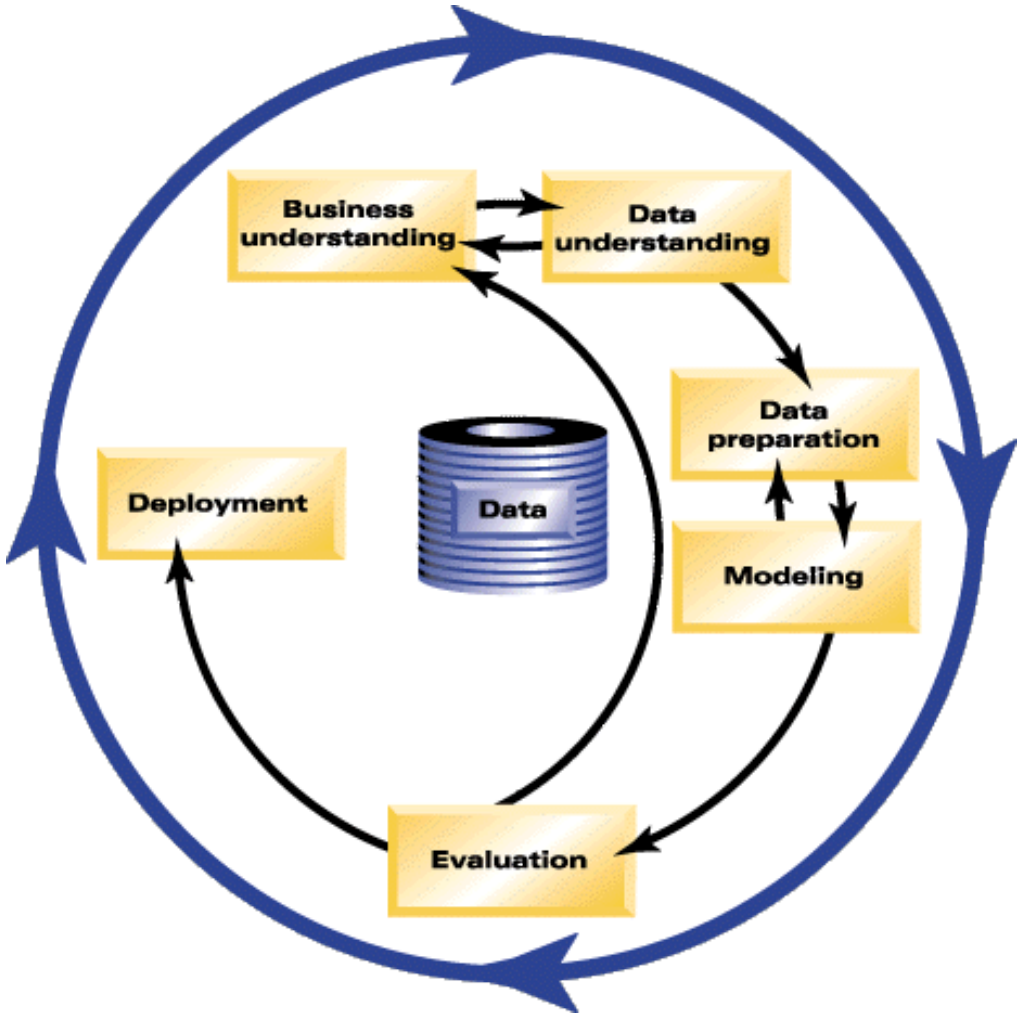
- Customer Background
  - **Multi-billion $ bank**
- Business Problem
  - **Some customers who develop negative account balances subsequently abandon their accounts, resulting in a loss (charge-off) to the bank.**
- Business Goals
  - **Develop a data mining model to predict account abandonment**
  - **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

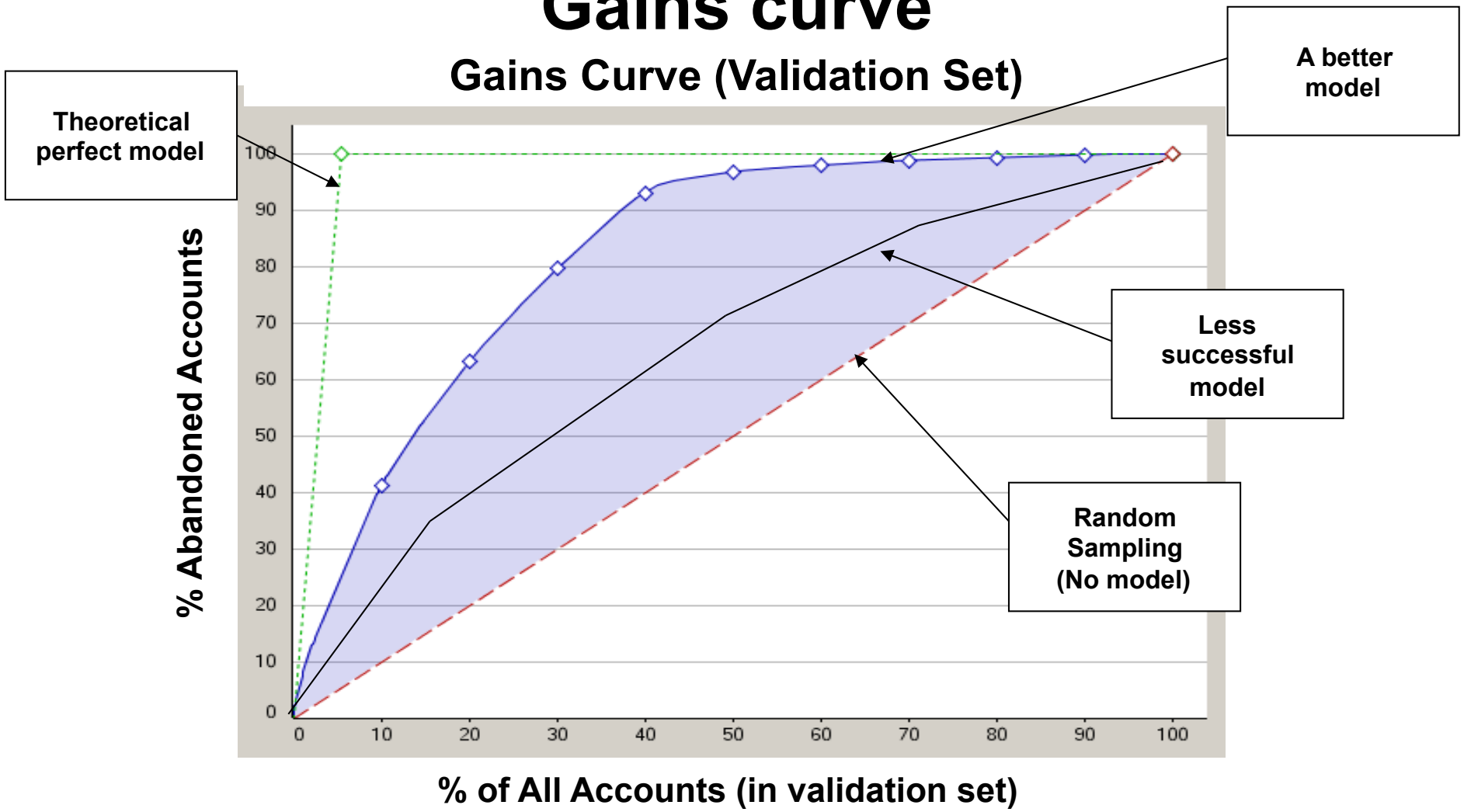


- **Step 1 – Create Mining Flow in InfoSphere Warehouse using Design Studio operators via a drag-and-drop palette approach**
- **Step 2-N Iterate to develop a useful model**
- **Step N+1 – Run the data mining engine to create the predictive tables containing information about which customers are likely to abandon their accounts (scoring)**

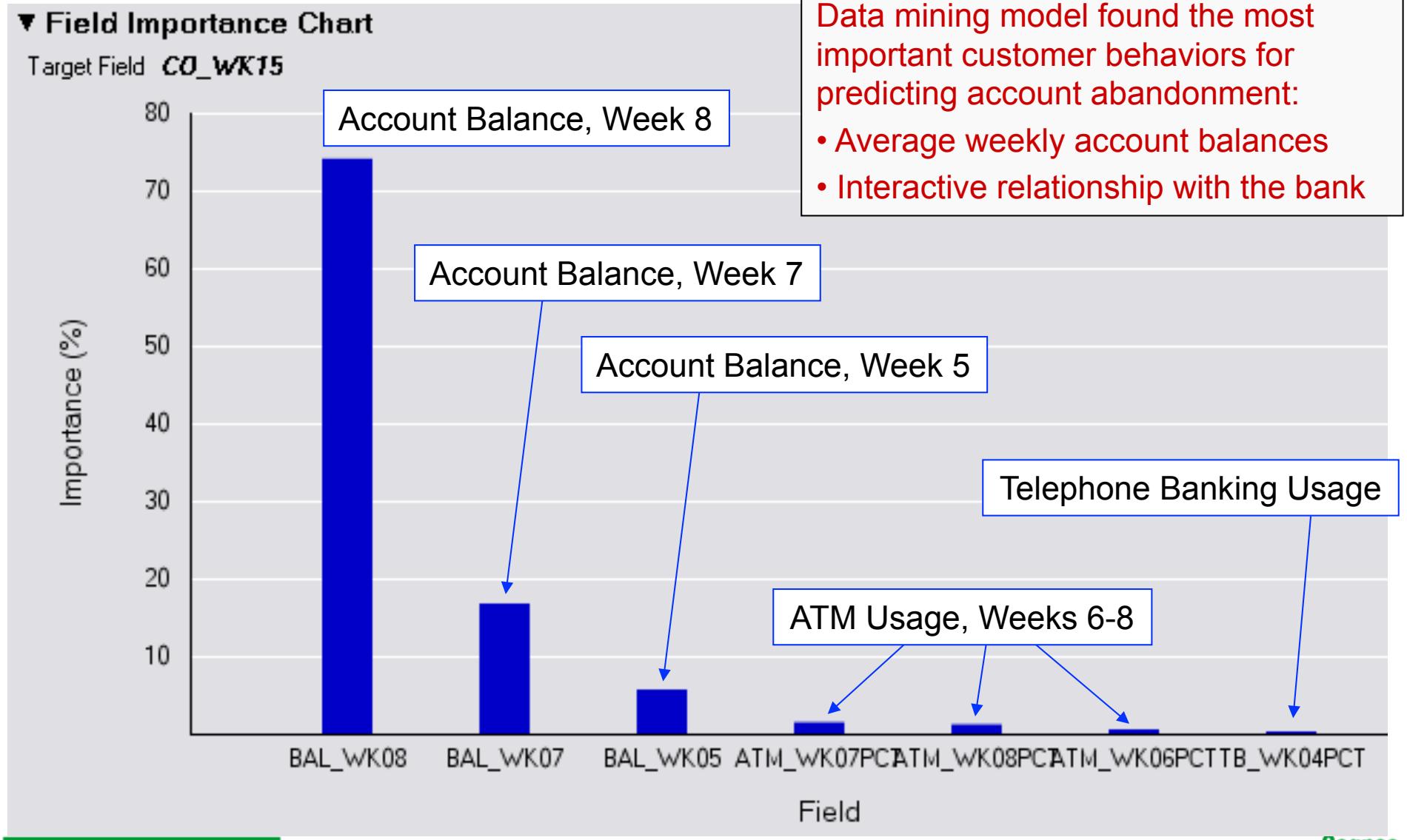
CRoss Industry Standard Process for Data Mining (CRISP-DM)



Account Abandonment Prediction – Gains curve

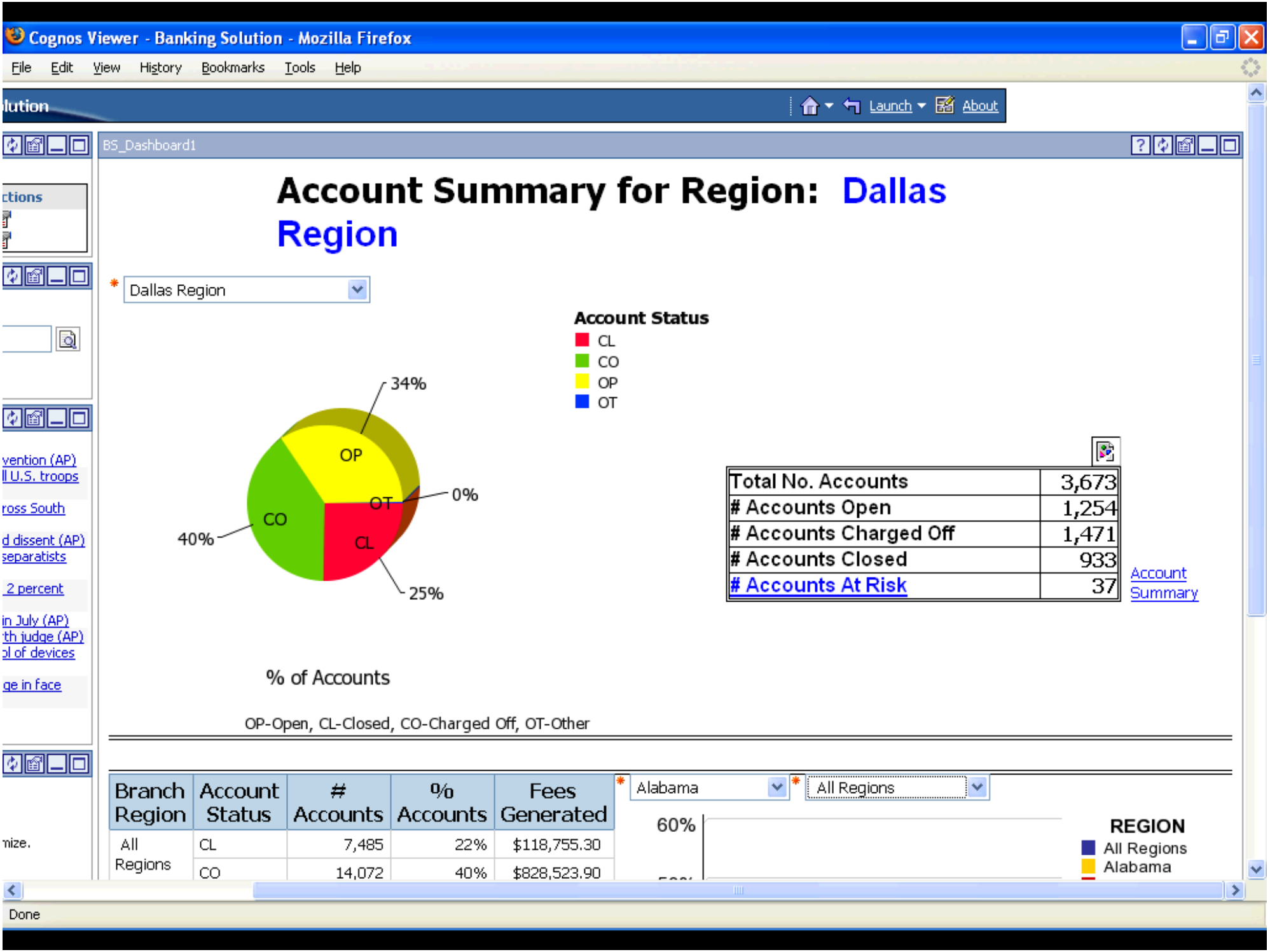


Example: Account Abandonment Prediction



Data mining model found the most important customer behaviors for predicting account abandonment:

- Average weekly account balances
- Interactive relationship with the bank


Cognos Viewer - Banking Solution - Mozilla Firefox
File Edit View History Bookmarks Tools Help
Launch About
BS_Dashboard1
Account Summary for Region: Dallas Region
Dallas Region
Account Status
CL
CO
OP
OT
34%
OP
0%
OT
40%
CO
CL
25%
% of Accounts
OP-Open, CL-Closed, CO-Charged Off, OT-Other
Total No. Accounts 3,673
Accounts Open 1,254
Accounts Charged Off 1,471
Accounts Closed 933
Accounts At Risk 37
Account Summary
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
60%
REGION
All Regions
Alabama
Done

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

Summary

- You can get more value from your existing information assets
- Start with discovery
- Work towards Predictive as you get more comfortable
- We can help

Qs and As