

# Learn Predictive Analytics in 2 Hours! Oracle Advanced Analytics Hands on Lab

Making Big Data + Analytics Simple—*On-Premise & Cloud*

Charlie Berger, MS Eng, MBA

Sr. Director Product Management, Data Mining and Advanced Analytics

[charlie.berger@oracle.com](mailto:charlie.berger@oracle.com) [www.twitter.com/CharlieDataMine](http://www.twitter.com/CharlieDataMine)

Brendan Tierney, President, Oralytics, Oracle ACE Director, Author

Tim Vlamis, VP & Analytics Strategist, Vlamis Software

Karl Rexer, Ph.D, President, Rexer Analytics



## BIWA SUMMIT 2016

The Oracle Big Data + Analytics User Conference

January 26–28, 2016

Oracle Spatial Summit + YesSQL Summit

ORACLE®

# Learn Predictive Analytics in 2 Hours!

## Oracle Advanced Analytics Hands on Lab



- **Lessons—Novices**

1. OAA Quick Overview
2. Data Mining Concepts Briefly
3. Quick Oracle Data Miner GUI Demo
4. Take off! Do as many Tutorials as you can in the 2 hours HOL
5. Ask questions! We're all here to help!



- **Take off!—Intermediate/Experts**

1. Environment

- Oracle 12c on the **Oracle Database Cloud**
- Need SQL Developer 4.1.3
  - [See roaming instructors for memory stick]
  - Set up/configure Oracle Data Miner extension (already done, but read instructions)

2. Do 3-5 Tutorials

- Instructors will walk around helping

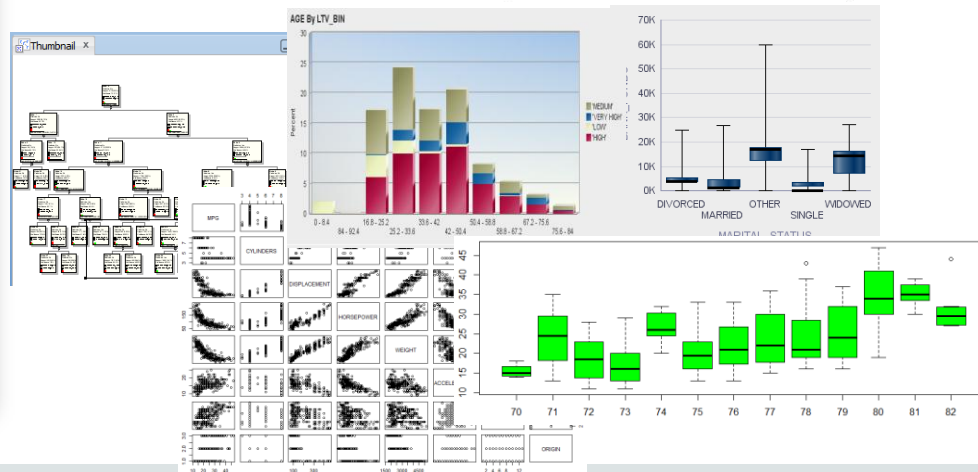
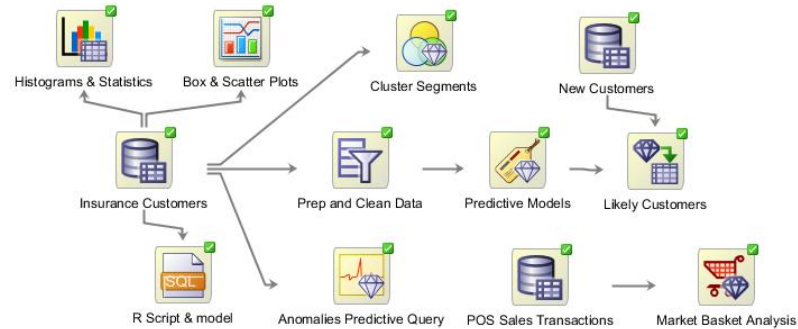
# Oracle's Advanced Analytics

## Fastest Way to Deliver Scalable Enterprise-wide Predictive Analytics



### Key Features

- Scalable in-Database + Hadoop data mining algorithms and R integration
- Powerful predictive analytics and deployment platform
- Drag and drop workflow, R and SQL APIs
- Data analysts, data scientists & developers
- Enables enterprise predictive analytics applications



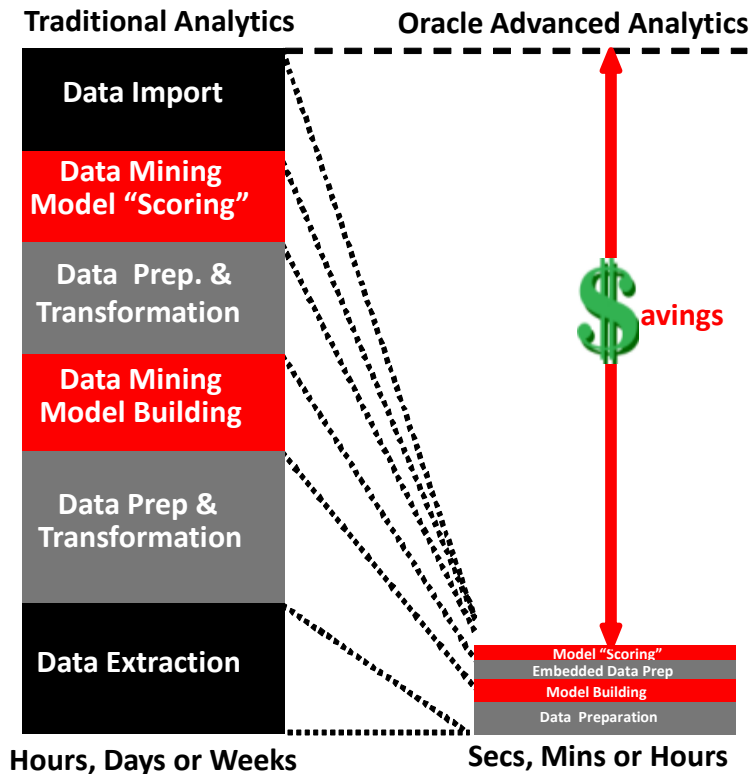
# Oracle's Advanced Analytics

Fastest Way to Deliver Scalable Enterprise-wide Predictive Analytics



## Major Benefits

- Data remains in Database & Hadoop
  - Model building and scoring occur in-database
  - Use R packages with data-parallel invocations
- Leverage investment in Oracle IT
  - Eliminate data duplication
  - Eliminate separate analytical servers
- Deliver enterprise-wide applications
  - GUI for predictive analytics & code generation
  - R interface leverages database as HPC engine



# Oracle's Advanced Analytics

Multiple interfaces across platforms — SQL, R, GUI, Dashboards, Apps

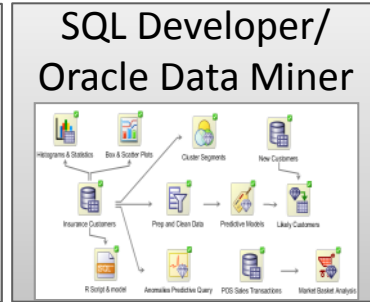
**Users**



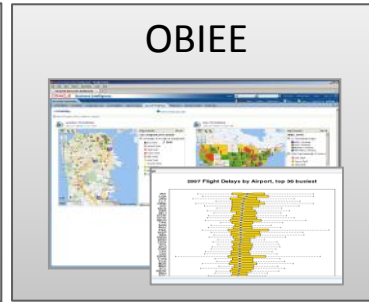
**R programmers**



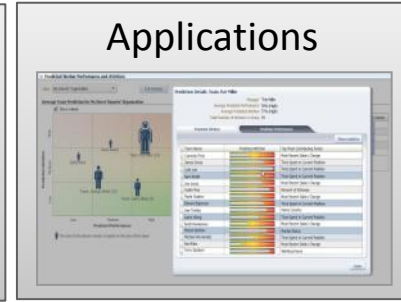
**Data & Business Analysts**



**Business Analysts/Mgrs**



**Domain End Users**



**Platform**



**Hadoop**

**ORAAH**  
Parallel,  
distributed  
algorithms

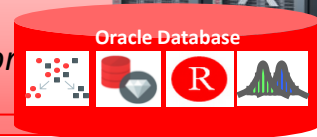
**Oracle Database Enterprise Edition**



**Oracle Advanced Analytics - Database Option**

*SQL Data Mining & Analytic Functions + R Integration  
for Scalable, Distributed, Parallel in-Database ML Execution*

**Oracle Cloud**



**ORACLE**





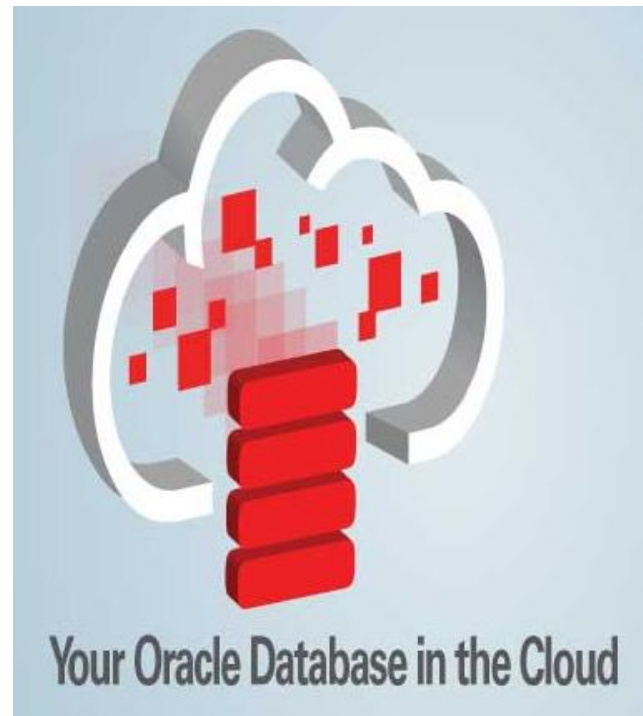
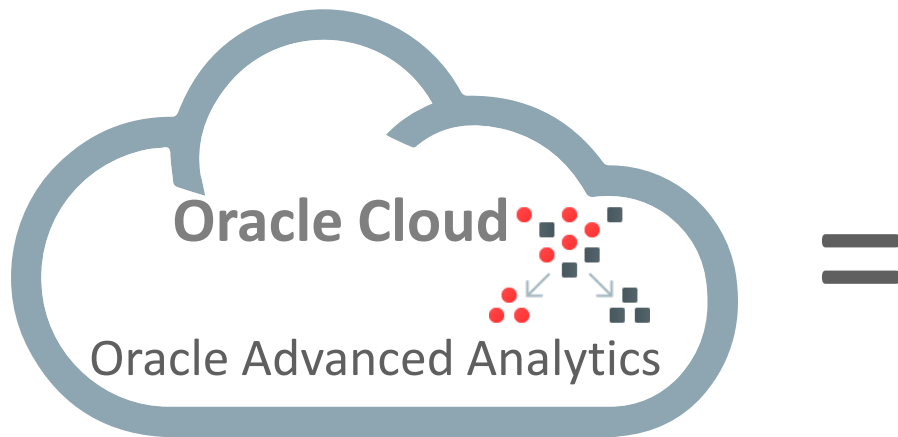
# Take off!—Intermediate/Experts

Quick Set up Overview

# Learn Predictive Analytics in 2 Hours!

## Oracle Advanced Analytics Hands on Lab

- We're using the Oracle Database Cloud!



# Learn Predictive Analytics in 2 Hours!

## Oracle Advanced Analytics Hands on Lab

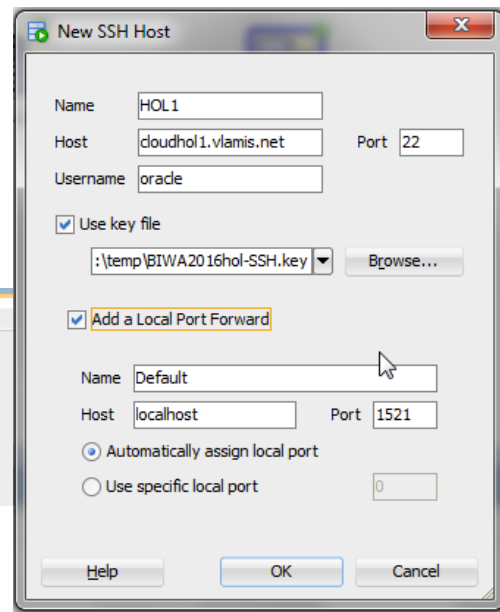
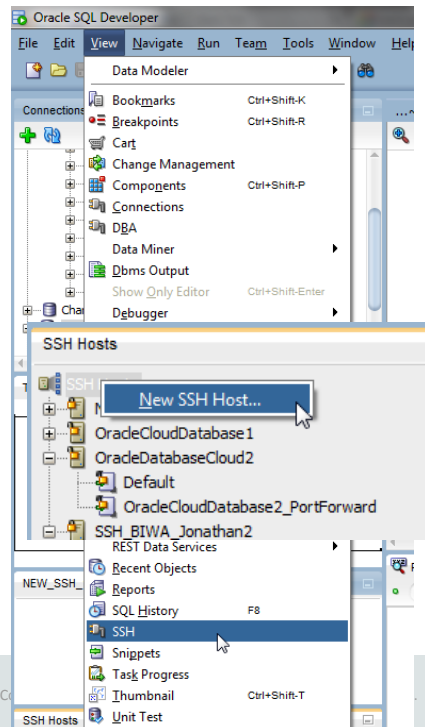
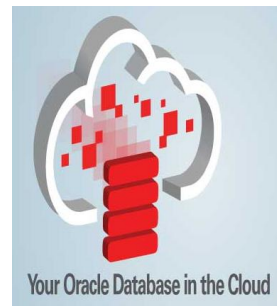
- Step 1—Install SQLDEV 4.1.3
- Step 2—Connect to Oracle Database Cloud

- 1. Create SSH using “key”

- [www.vlamis.com/storage/biwa2016/BIWA2016hol-SSH.key](http://www.vlamis.com/storage/biwa2016/BIWA2016hol-SSH.key)
    - [www.vlamis.com/storage/biwa.key](http://www.vlamis.com/storage/biwa.key)
    - Cloudhol”N”.vlamis.net
    - Username is “oracle”
    - Add Port Forward and accept defaults

Cloud Computing  
**ORACLE**  
SQL Developer  
4.1.3

Exchange Forums Download





# Learn Predictive Analytics in 2 Hours!

## Oracle Advanced Analytics Hands on Lab

- Step 1—Install SQLDEV  
4.1.3

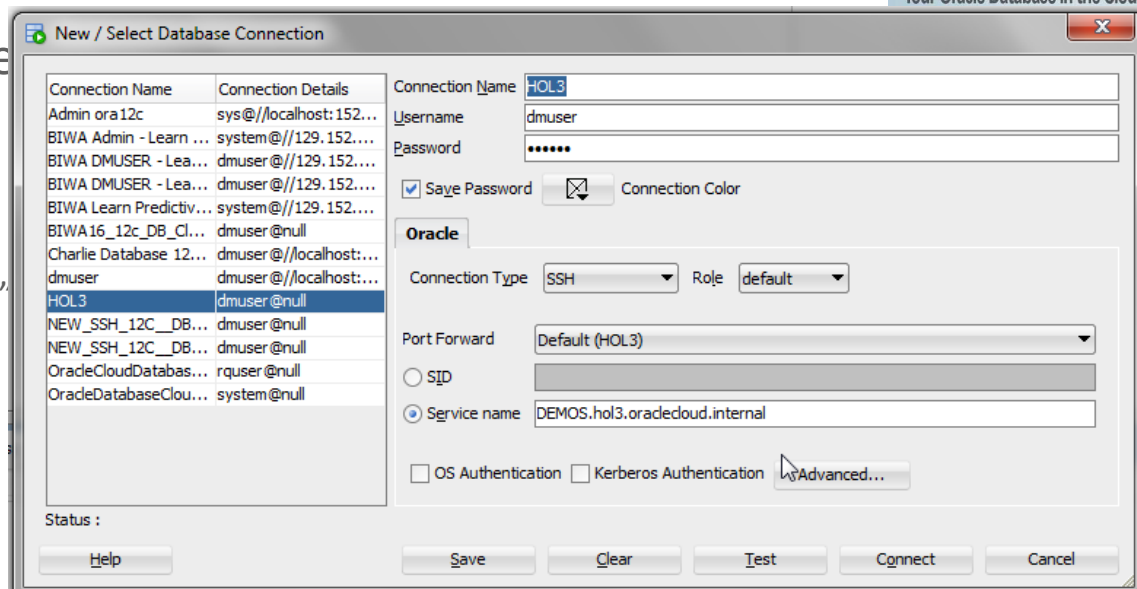
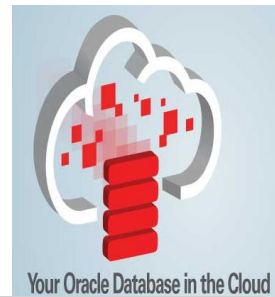
- Step 2—Connect to Oracle Database Cloud

- 1. Create SQL Database Connection

- (Each student take unique HOL"N" assigned)
    - dmuser/dmuser (save password)
    - SSH
    - Port Forward (Default HOL"N")
    - Service ID (e.g. DEMOS.hol"N".oraclecloud.internal)

Cloud Computing  
**ORACLE**  
SQL Developer  
4.1.3

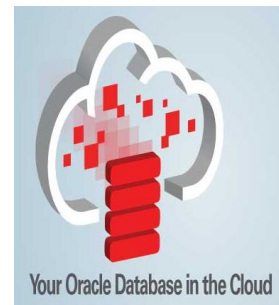
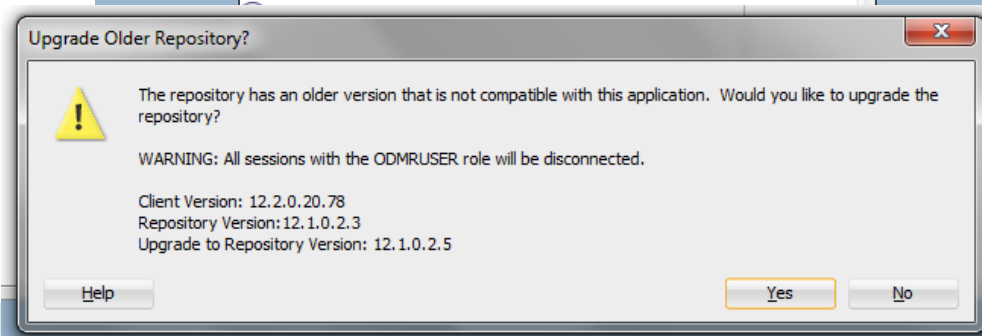
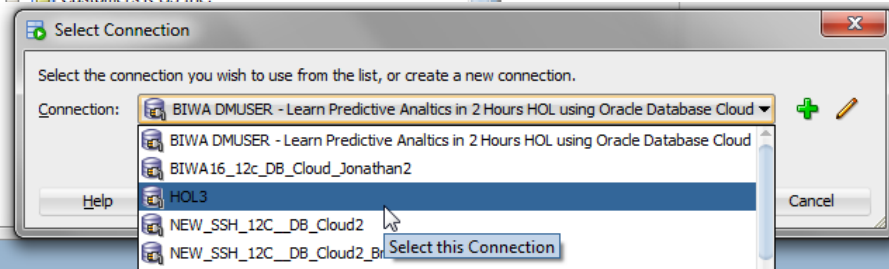
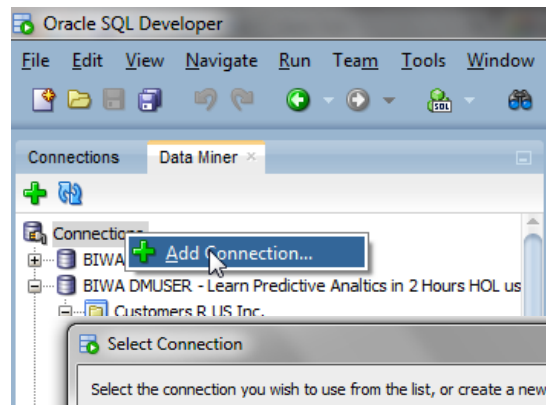
Exchange Forums Download



# Learn Predictive Analytics in 2 Hours!

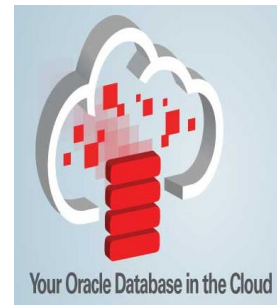
## Oracle Advanced Analytics Hands on Lab

- Step 1—Install SQLDEV 4.1.3
- Step 2—Connect to Oracle Database Cloud
  - 1. Go to Oracle Data Miner & create a NEW Connection e.g. HOL"N"
  - Select HOL"N" from drop down menu
  - Optionally may need to upgrade older Data Mining repository (may take 3 mins)
  - You are done!

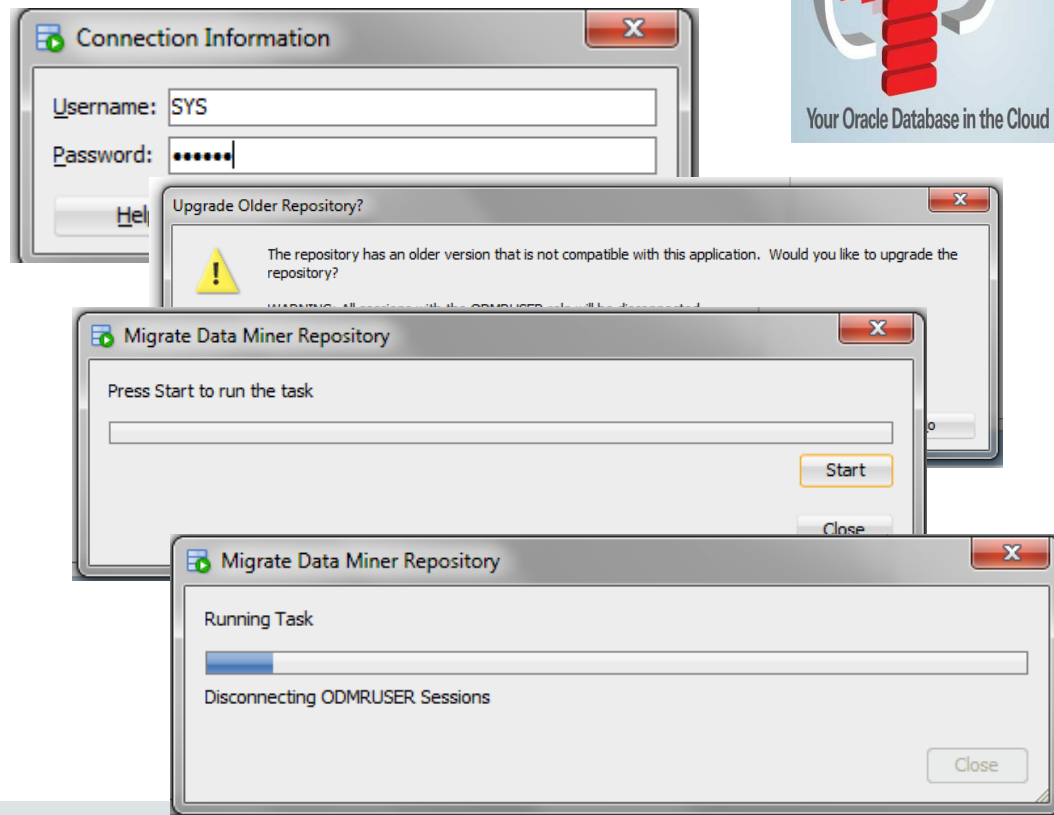


# Learn Predictive Analytics in 2 Hours!

## Oracle Advanced Analytics Hands on Lab



- Step 1—Install SQLDEV 4.1.3
- Step 2—Connect to Oracle Database Cloud
  - Connect as SYS/Welcome#1
  - Start to run task
  - Running task (may take 3 mins)



# Learn Predictive Analytics in 2 Hours!

## Oracle Advanced Analytics Hands on Lab

- Step 1—Install SQLDEV 4.1.3
- Step 2—Connect to Oracle Database Cloud
- Step 3—Start HOL!

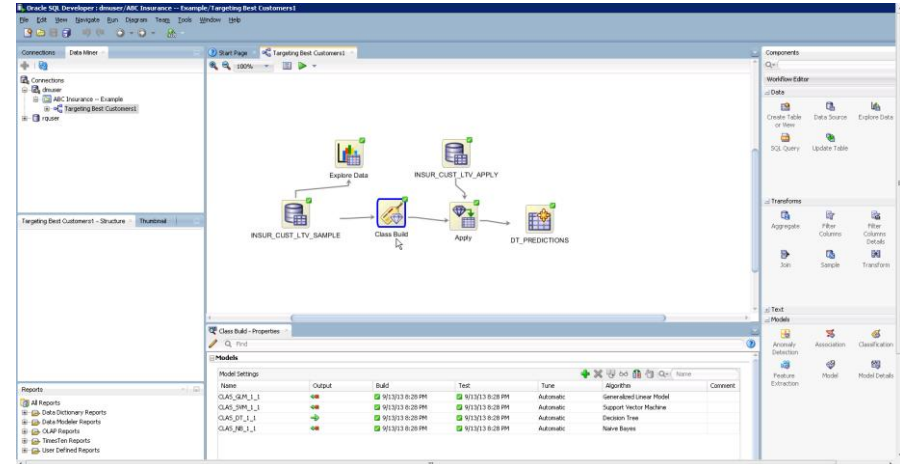
- dmuser/dmuser
- Demo data for learning
- Follow 3-5 OBE Online Tutorials

1. [Using Oracle Data Miner 4.1](#)
2. [Star Schema Mining Using Oracle Data Miner 4.1](#)
3. [Text Mining with an EM Clustering Model Using Data Miner 4.1](#)
4. **Anomaly Detection (CLAIMS)** *See Instructor for assistance*
5. **Market Basket Analysis (SH.SALES)** *See Instructor for assistance*

Cloud Computing  
**ORACLE**  
SQL Developer  
4.1.3



Exchange Forums **Download** f y d



# OAA/Oracle Data Miner 4.1 HOL

We're Using the Oracle by Example Free Online Tutorials

- Google “Oracle Data Miner”
- Scroll down to bottom of page & launch tutorials

– [https://apexapps.oracle.com/pls/apex/f?p=44785:24:::NO::P24\\_CONTENT\\_ID,P24\\_PREV\\_PAGE:11925,2](https://apexapps.oracle.com/pls/apex/f?p=44785:24:::NO::P24_CONTENT_ID,P24_PREV_PAGE:11925,2)

## Demos and Training



2-Day Oracle University Course



Oracle Data Miner Workflow GUI Tutorials

Oracle Data Miner on Big Data Lite VM Tutorial



OAA YouTube demos



Sample Programs

## Oracle Data Miner/ SQL Developer Extension



Drag & Drop Workflow for Creating and Deploying in-  
Database Analytical Methodologies

## Oracle Data Miner GUI

The Oracle Data Miner is an Oracle SQL Developer extension that enables data analysts, business analysts and data scientists to work directly with data inside the database using in-database data mining algorithms of **Oracle Advanced Analytics** and it's **Oracle Data Mining** component SQL functions. Data analyst can use Oracle Data Miner's "drag and drop" workflow and component pallet to explore the data in graphs, build and evaluate data mining models, apply the predictive models to new data and then deploy the OAA/Oracle Data Mining's predictions and insights



# OAA/Oracle Data Miner 4.1 HOL

## Uses Oracle by Example Free Online Tutorials

- There are 6 Tutorials

- The first tutorial is already done for you

- Recommend doing 3-5 Tutorials

1. [Using Oracle Data Miner 4.1](#)

2. [Star Schema Mining Using Oracle Data Miner 4.1](#)

3. [Text Mining with an EM Clustering Model Using Data Miner 4.1](#)

4. **Anomaly Detection (CLAIMS)** *See Instructor for assistance*

5. **Market Basket Analysis (SH.SALES)** *See Instructor for assistance*

### Content List (click links below to view content)

08-SEP-2015  
15 mins  
★★★★★

#### [Setting Up Oracle Data Miner 4.1](#)

This tutorial covers the process of setting up Oracle Data Miner for use within Oracle SQL Developer 4.1.

08-SEP-2015  
45 mins  
★★★★★

#### [Using Oracle Data Miner 4.1](#)

This tutorial covers the use of Oracle Data Miner 4.1 to perform data mining against Oracle Database 12c. In this lesson, you examine the Oracle Data Miner GUI. The Oracle Data Miner GUI is included as an extension of Oracle SQL Developer, version 4.1.

08-SEP-2015  
30 mins  
★★★★★

#### [Star Schema Mining Using Oracle Data Miner 4.1](#)

This tutorial covers the use of Oracle Data Miner 4.1 to perform star schema mining activities against Oracle Database 12c Release 2.

08-SEP-2015  
30 mins  
★★★★★

#### [Text Mining with an EM Clustering Model Using Data Miner 4.1](#)

In this lesson, you learn how to use the EM algorithm in a clustering model while leveraging text mining enhancements that are included in Oracle Database 12c Release 2.

08-SEP-2015  
30 mins  
★★★★★

#### [Using Logistic Regression Models \(GLM\) to Predict Customer Affinity](#)

This tutorial covers the use of Oracle Data Miner 4.1 to leverage enhancements to the Oracle implementation of Generalized Linear Models (GLM) for Predictive Analytics.

08-SEP-2015  
30 mins  
★★★★★

#### [Using the SQL Query Node With Oracle Data Miner 4.1](#)

This tutorial covers the use of the new SQL Query Node in an Oracle Data Miner 4.1 workflow.

08-SEP-2015  
30 mins  
★★★★★

#### [Using Predictive Queries With Oracle Data Miner 4.1](#)

This tutorial covers the use of Predictive Queries against mining data by Oracle Data Miner 4.1.

08-SEP-2015  
30 mins  
★★★★★

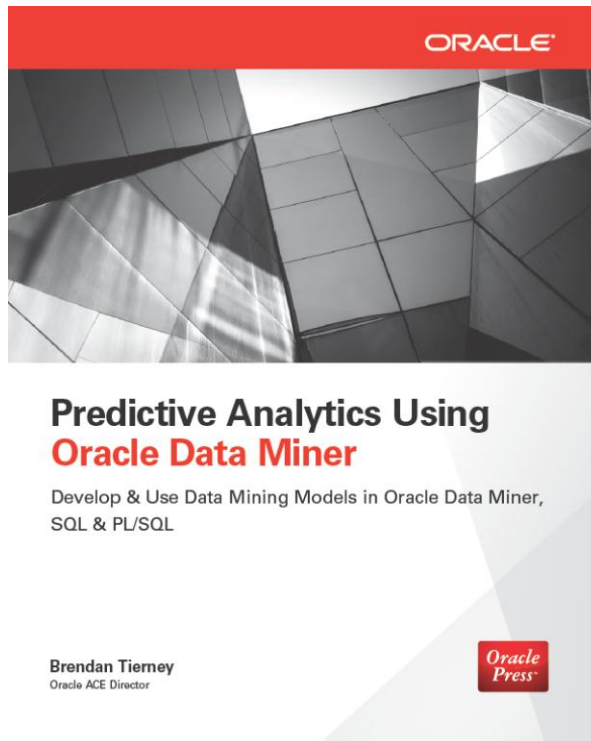
#### [Mining JSON Data Using Oracle Data Miner 4.1](#)

This tutorial covers the use of the JSON Query Node in an Oracle Data Miner 4.1 workflow in order to mine this Big Data format.

# Great book on Oracle Advanced Analytics

Available on Amazon or from Author

- Predictive Analytics Using Oracle Data Miner: Develop for ODM in SQL & PL/SQL



# OAA/Oracle Data Miner 4.1 HOL

## Setting Up Oracle Data Miner

Setting Up Oracle Data Miner 4.1

**Done ✓**

The screenshot shows the Oracle Data Miner 4.1 setup wizard interface. On the left, there is a sidebar with 'Options' (Expand All Topics, Hide All Media) and a list of steps: 'Before You Begin', 'Create a Data Miner User Account', 'Create a SQL Developer Connection for the Data Miner User', 'Install the Data Miner Repository', and 'Want to Learn More?'. A red arrow points from the 'Before You Begin' step to the 'New / Select Database Connection' dialog box. The dialog box is titled 'New / Select Database Connection' and contains fields for 'Connection Name' (admin), 'Username' (SYS), and 'Password' (masked). It also has checkboxes for 'Save Password' and 'Connection Color'. Below these, the 'Oracle' tab is selected, showing 'Connection Type' (Basic) and 'Role' (SYSDBA). The 'Hostname' is localhost, 'Port' is 1521, and 'SID' is orcl. There are also checkboxes for 'OS Authentication' and 'Kerberos Authentication', and an 'Advanced...' button. At the bottom, there are buttons for 'Save', 'Clear', 'Test' (highlighted with a yellow border), 'Connect', and 'Cancel'. In the top right corner, there is a 'Connections' panel with a '+ New Connection...' button, which is also highlighted with a red arrow.

# OAA/Oracle Data Miner 4.1 HOL

## Setting Up Oracle Data Miner

Setting Up Oracle Data Miner 4.1

*Done ✓*

**Options**

[Expand All Topics](#)

[Hide All Media](#)

- Before You Begin
- Create a Data Miner User Account
- Create a SQL Developer Connection for the Data Miner User
- Install the Data Miner Repository
- Want to Learn More?

### Create User

User | Granted Roles | System Privileges | Quotas | SQL

User Name:

New Password:

Confirm Password:

☐ Password Expired (user must change next login)

☐ Operating System User

☐ Account is Locked

☐ Edition Enabled

Default Tablespace:

Temporary Tablespace:

### Connections

[New Connection...](#)

[Import Connections...](#)

[Export Connections...](#)

[Create Local Connections](#)

[Disconnect Connections](#)

### Install Data Miner Repository

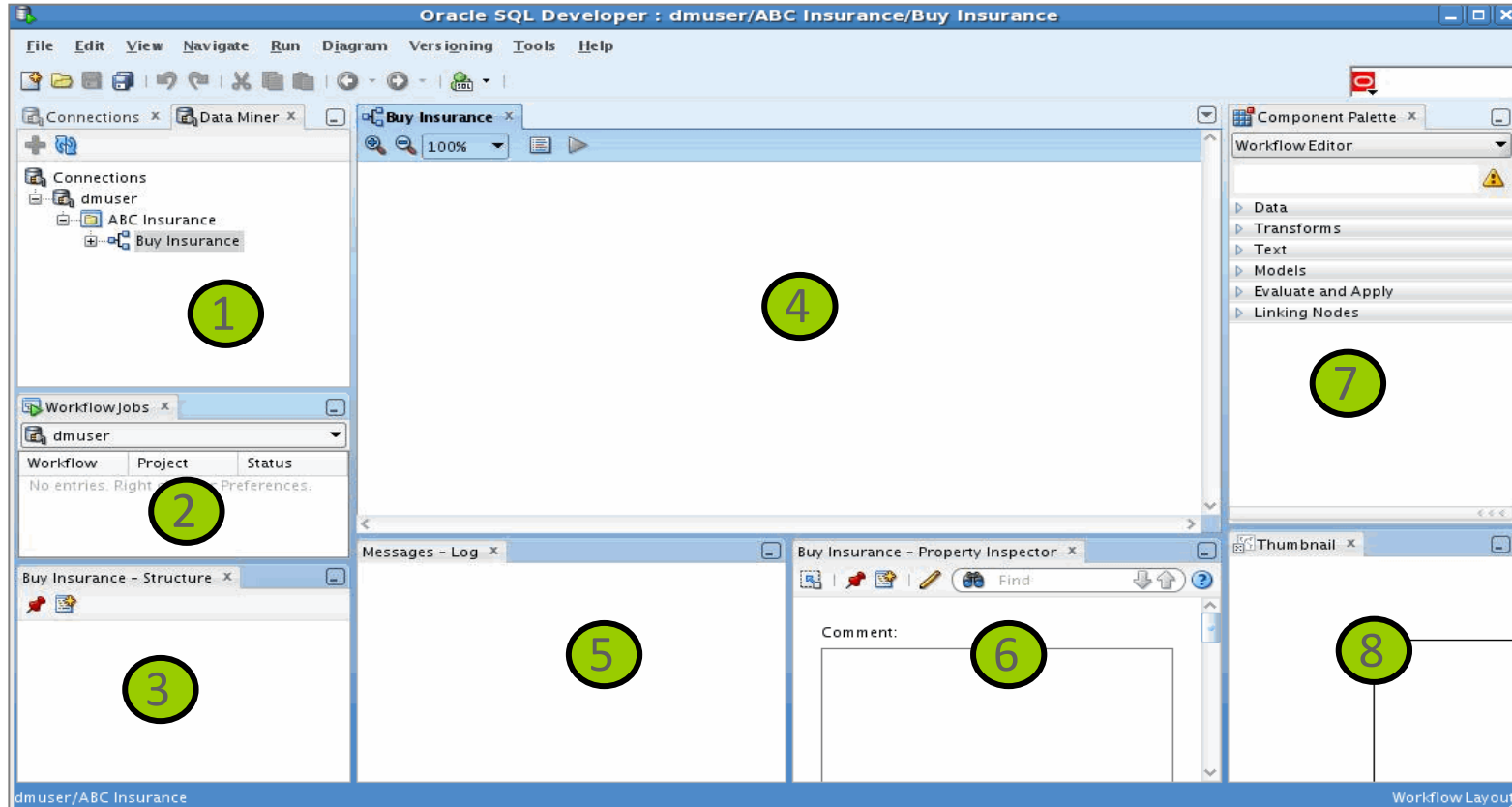
Running Task

Installing Schema

☒ Install Demo Data

[Close](#)

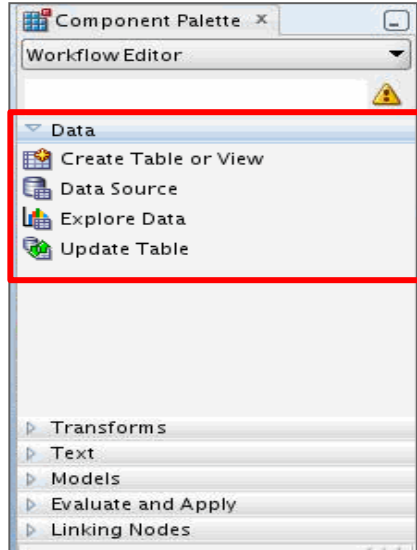
# Introducing the Data Miner Interface



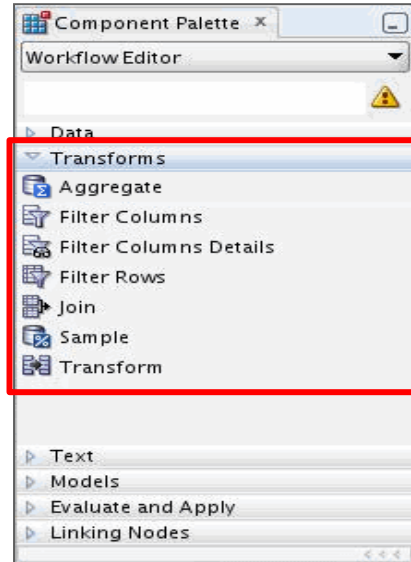


# Examining Oracle Data Miner Nodes

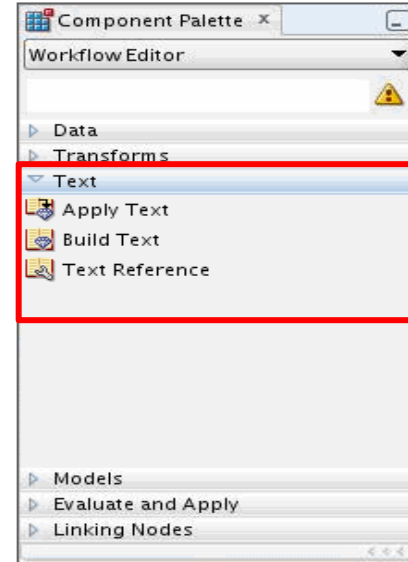
## Data



## Transforms

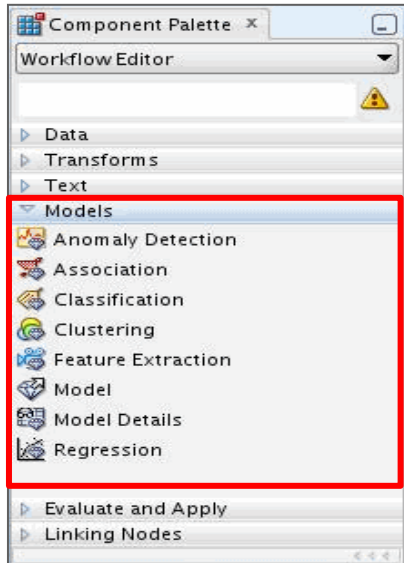


## Text

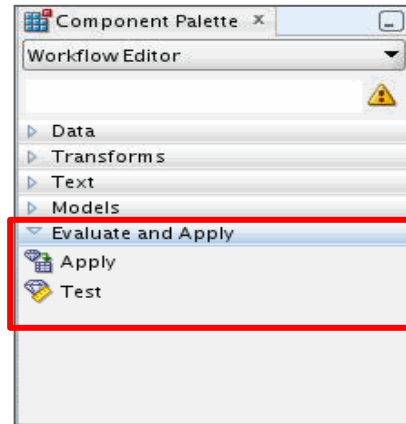


# Examining Oracle Data Miner Nodes

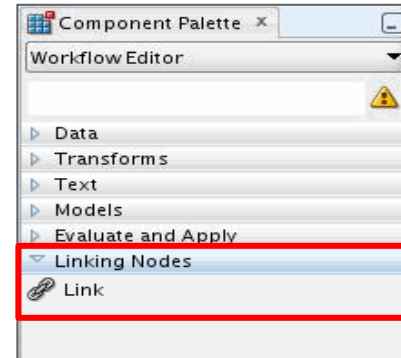
## Models



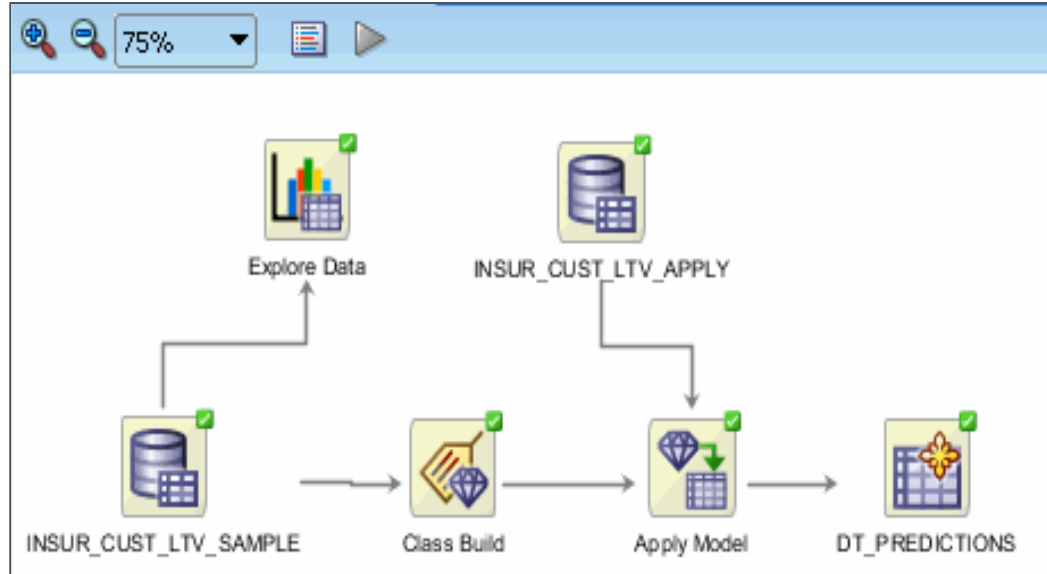
## Evaluate and Apply



## Linking



# Previewing a Data Miner Workflow



# Multiple Data Sources/Types with Predictive Modeling

## Ease of Deployment through SQL Script Generation

*SQL Joins and arbitrary SQL transforms & queries – power of SQL*

*Transactional POS data*



*Generates SQL scripts for deployment*

*Inline predictive model to augment input data*

*Unstructured data also mined by algorithms*

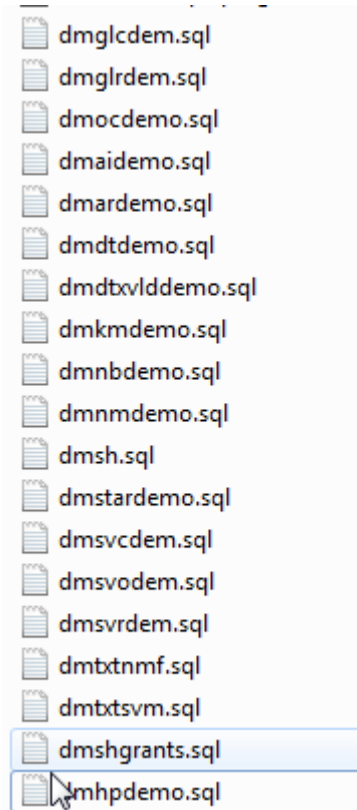
Consider:

- Demographics
- Past purchases
- Recent purchases
- Customer comments & tweets

# The Data Mining Sample Programs

- The [Data Mining Sample Programs 12c Documentation](#)

- You can learn a great deal about the Oracle Data Mining API from the data mining sample programs. The programs illustrate typical approaches to data preparation, algorithm selection, algorithm tuning, testing, and scoring.
- The programs are easy to use. They include extensive inline comments to help you understand the code. They delete all temporary objects on exit; you can run the programs repeatedly without setup or cleanup.
- The data mining sample programs are installed with Oracle Database Examples in the demo directory under Oracle Home. The demo directory contains sample programs that illustrate many features of Oracle Database. You can locate the data mining files by doing a directory listing of dm\*.sql.





# The Data Mining Sample Programs

## Attribute Importance Sample Code

```
5.09600019 seconds

Worksheet Query Builder

-----
BUILD THE MODEL
-----

-- Cleanup old output table for repeat runs
BEGIN EXECUTE IMMEDIATE 'DROP TABLE ai_explain_output';
EXCEPTION WHEN OTHERS THEN NULL; END;
/

-----

-- Run the EXPLAIN routine to get attribute importance results
BEGIN
  DBMS_PREDICTIVE_ANALYTICS.EXPLAIN(
    data_table_name      => 'mining_data_build_v',
    explain_column_name => 'affinity_card',
    result_table_name    => 'ai_explain_output');
END;
/

-----

-- DISPLAY RESULTS
```

Script Output x

Task completed in 5.096 seconds

ATTRIBUTE_NAME	EXPLANATORY_VALUE	RANK
HOUSEHOLD_SIZE	.195	1
CUST_MARITAL_STATUS	.194	2
YRS_RESIDENCE	.115	3
EDUCATION	.106	4
AGE	.104	5
OCCUPATION	.092	6
Y_BOX_GAMES	.077	7
HOME_THEATER_PACKAGE	.069	8
CUST_GENDER	.043	9
BOOKKEEPING_APPLICATION	.024	10
BULK_PACK_DISKETTES	.000	11
COUNTRY_NAME	.000	11
CUST_ID	.000	11
CUST_INCOME_LEVEL	.000	11
FLAT_PANEL_MONITOR	.000	11
OS_DOC_SET_KANJI	.000	11
PRINTER_SUPPLIES	.000	11

17 rows selected

# Learn Predictive Analytics in 2 Hours!

## Oracle Advanced Analytics Hands on Lab

- Recommend doing 3-5 Tutorials

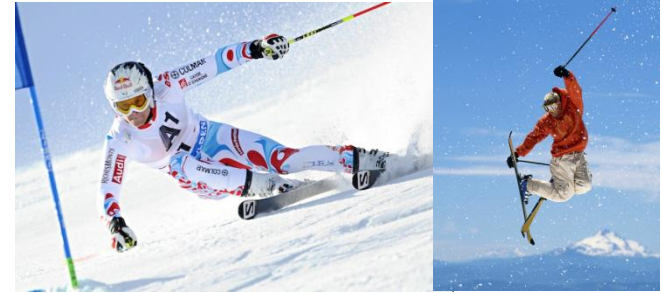
1.Using Oracle Data Miner 4.1

2.Star Schema Mining Using Oracle Data Miner 4.1

3.Text Mining with an EM Clustering Model Using Data Miner 4.1

4.**Anomaly Detection (CLAIMS)** *See Instructor for assistance*

5.**Market Basket Analysis (SH.SALES)** *See Instructor for assistance*



# Novice/Introductory/Overviews

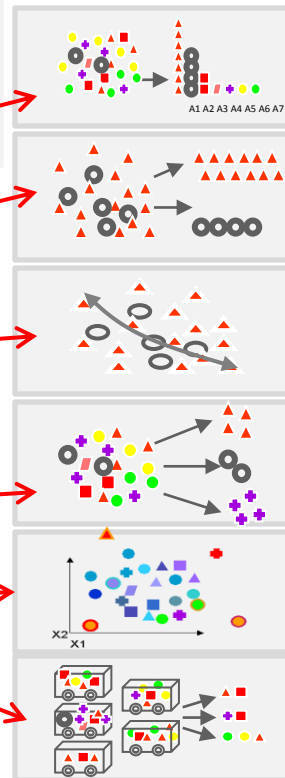
Quick Overview of Concepts, Process and Use Cases



# What is Data Mining & Predictive Analytics?

**Automatically** sifting through **large amounts** of data to create models that **find previously hidden patterns**, **discover valuable new insights** and **make predictions**

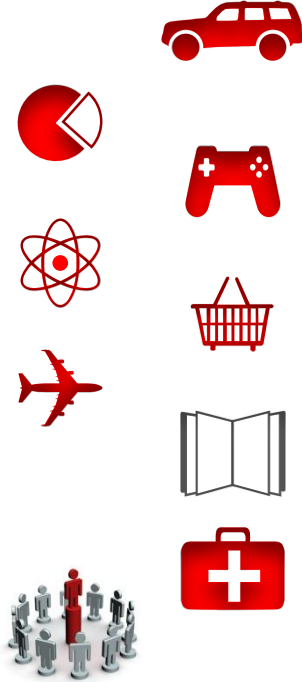
- Identify most important factor (*Attribute Importance*)
- Predict customer behavior (*Classification*)
- Predict or estimate a value (*Regression*)
- Find profiles of targeted people or items (*Decision Trees*)
- Segment a population (*Clustering*)
- Find fraudulent or “rare events” (*Anomaly Detection*)
- Determine co-occurring items in a “baskets” (*Associations*)



# Predictive Analytics & Data Mining

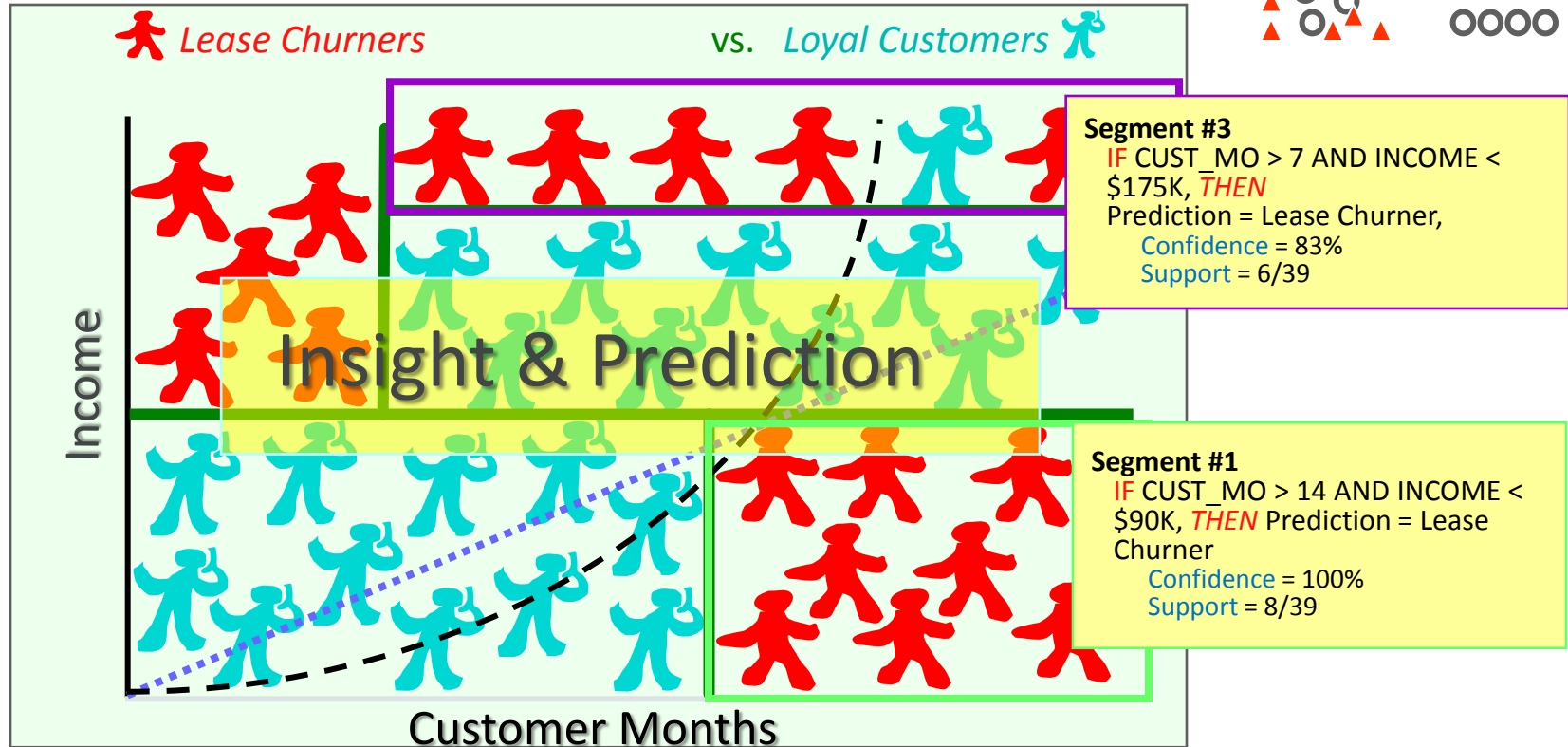
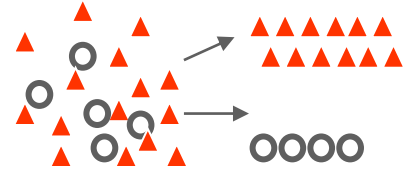
## Typical Use Cases

- Targeting the right customer with the right offer
- How is a customer likely to respond to an offer?
- Finding the most profitable growth opportunities
- Finding and preventing customer churn
- Maximizing cross-business impact
- Security and suspicious activity detection
- Understanding sentiments in customer conversations
- Reducing medical errors & improving quality of health
- Understanding influencers in social networks





# Data Mining Provides Better Information, Valuable Insights and Predictions



# Oracle's Advanced Analytics

In-Database Data Mining Algorithms\*—SQL &  & GUI Access



## Classification



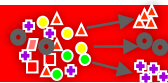
- Decision Tree
- Logistic Regression (GLM)
- Naïve Bayes
- Support Vector Machine (SVM)
- Random Forest

## Regression



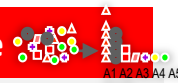
- Multiple Regression (GLM)
- Support Vector Machine (SVM)
- Linear Model
- Generalized Linear Model
- Multi-Layer Neural Networks
- Stepwise Linear Regression

## Clustering



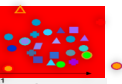
- Hierarchical k-Means
- Orthogonal Partitioning Clustering
- Expectation-Maximization

## Attribute Importance



- Minimum Description Length
- Unsupervised pair-wise KL div.

## Anomaly Detection



- 1 Class Support Vector Machine

## Time Series

- Single & Double Exp. Smoothing

## Predictive Queries

- Clustering
- Regression
- Anomaly Detection
- Feature Extraction

## Feature Extraction & Creation

- Nonnegative Matrix Factorization
- Principal Component Analysis
- Singular Value Decomposition

## Market Basket Analysis



- Apriori – Association Rules

## Open Source R Algorithms

- Ability to run any R package via Embedded R mode



\* supports partitioned models, text mining

# Oracle University's Learn Predictive Analytics Using Oracle Data Mining Course Agenda

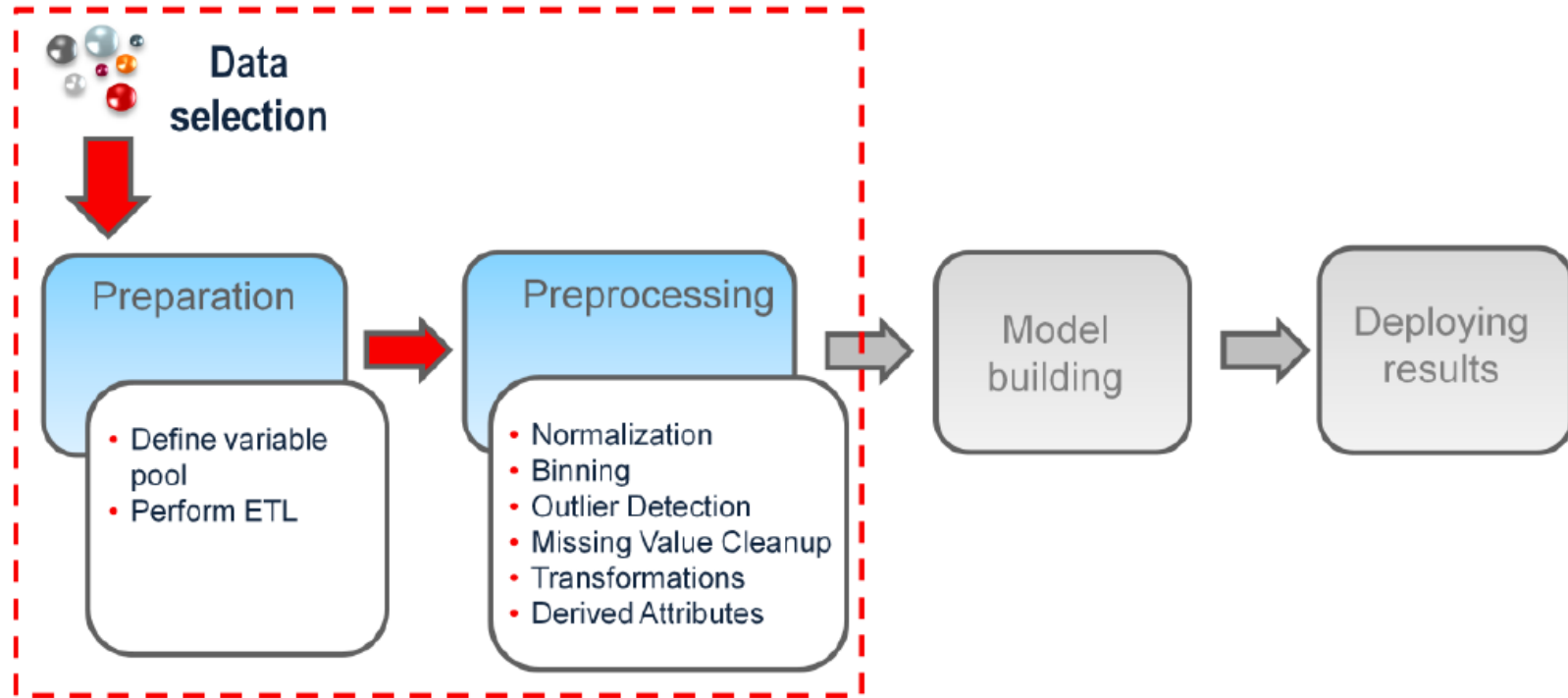
## Day 1

1. Introduction
2. Predictive Analytics and Data Mining Concepts
3. The Data Mining Process
4. Introducing Oracle Data Miner 4.1
5. Using Classification Models
6. Using Regression Models

## Day 2

7. Using Clustering Models
8. Performing Market Basket Analysis
9. Performing Anomaly Detection
10. Performing Star Schema Mining
11. Using Predictive Queries
12. Deploying Data Mining Results

# The Data Mining Process



# Data Mining Attributes

Data mining attributes can be:

- Categorical
- Numerical
- Text
- Target
- Active or Inactive

OCCUPATION
Prof.
Sales
Machine
Crafts

Categorical

AGE
40
47
32
63
40

Numerical

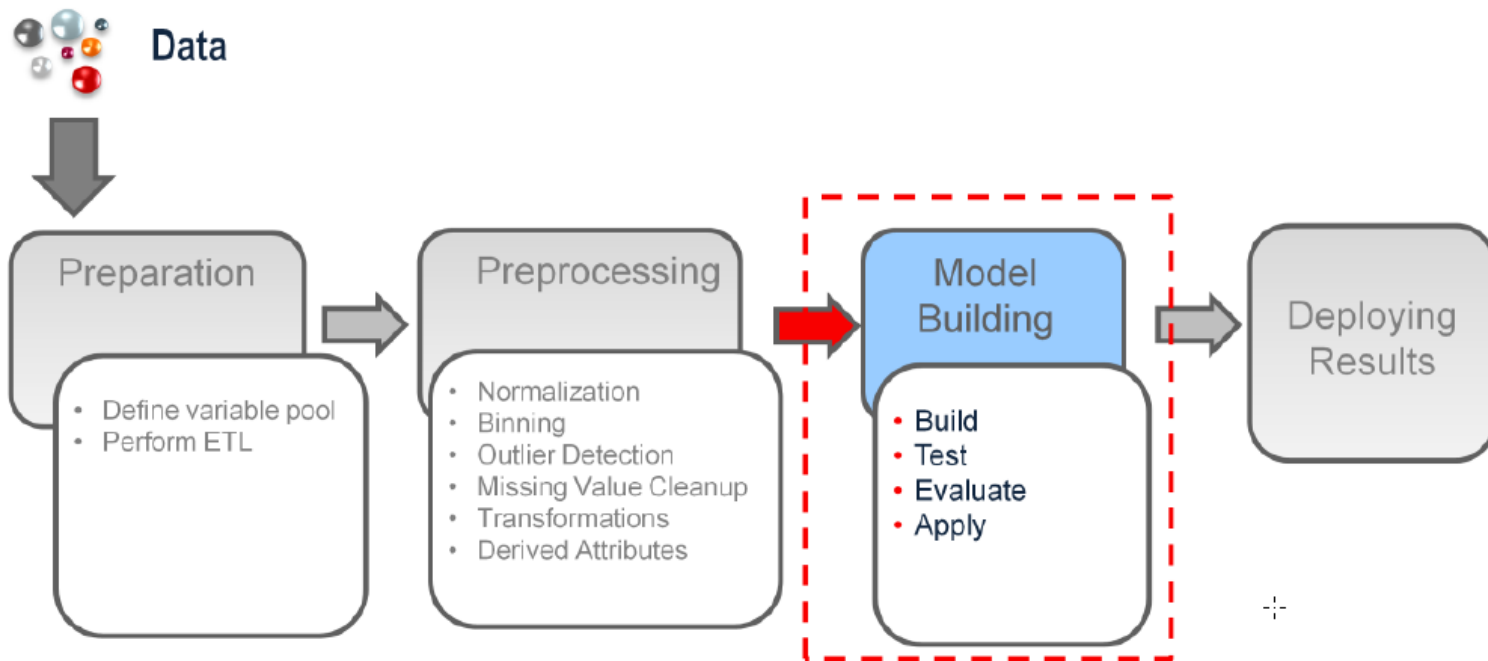
CLASS
1
0
0
0
0

Target

EDUCATION
Bach.
5th-6th
HS-grad
Bach.
HS-grad

Active

# Building and Evaluating Models

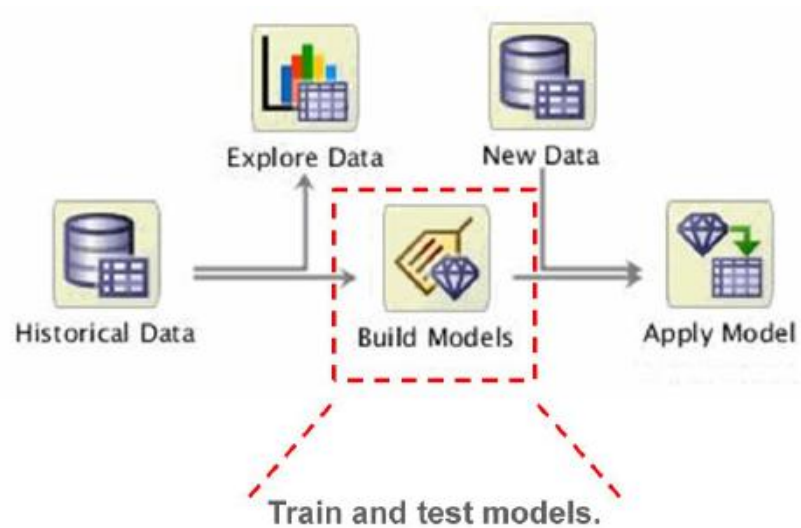


# Model Building Tasks

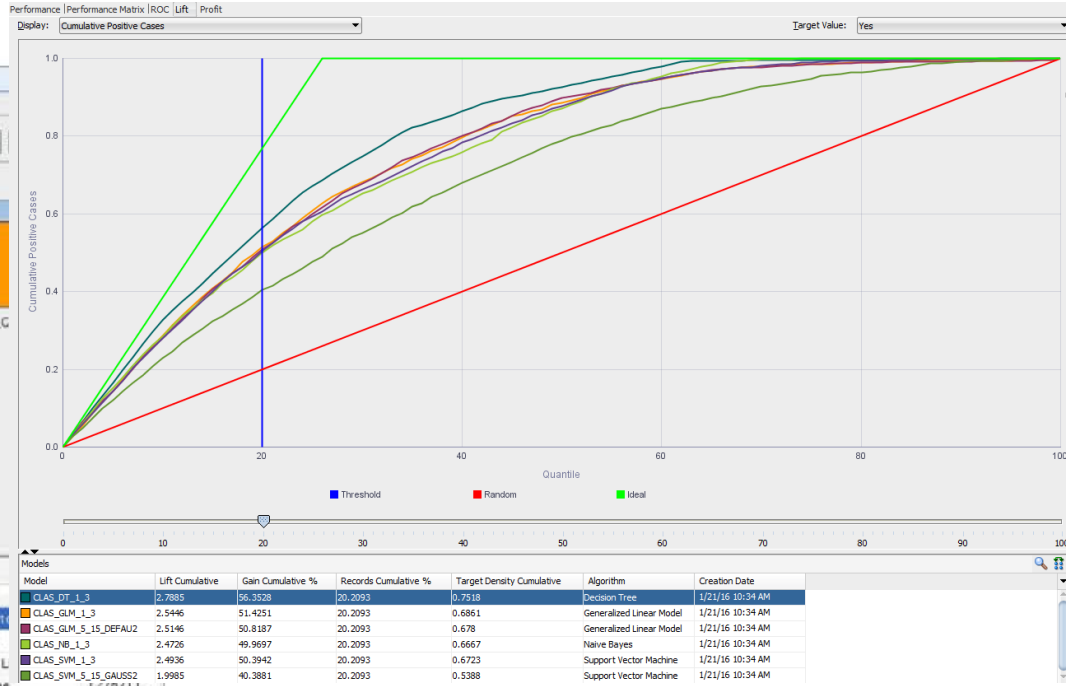
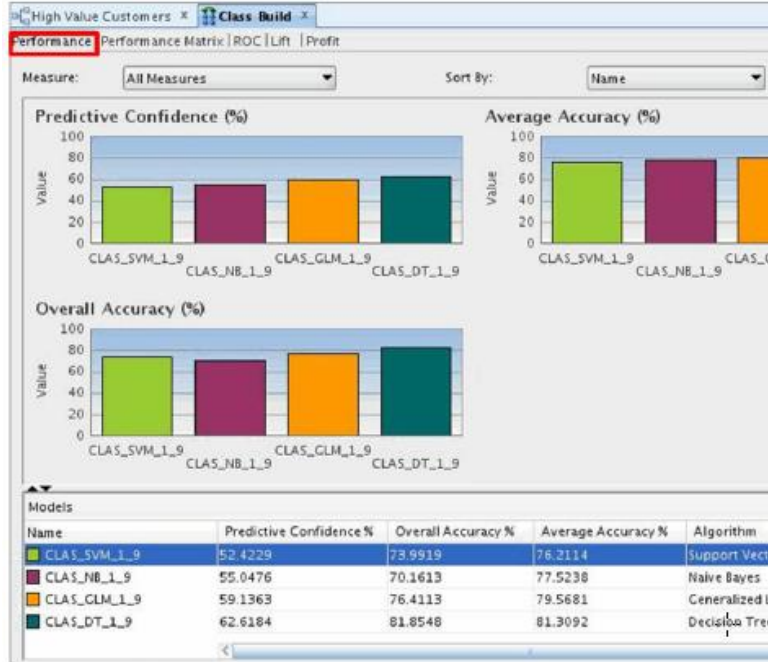




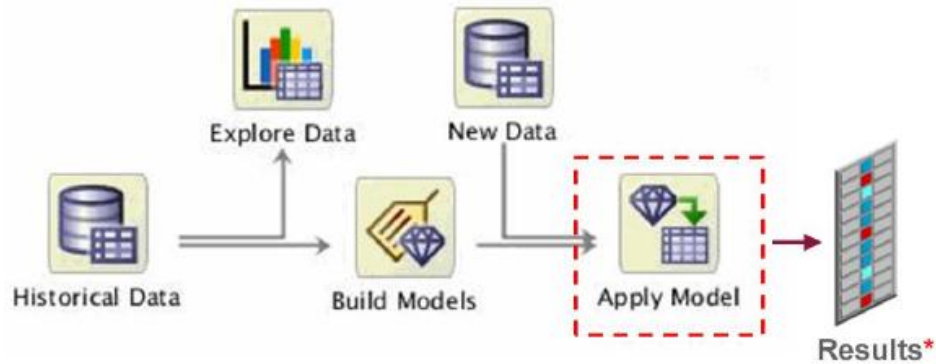
# Model Train & Test: Supervised Learning



# Model Evaluation: Supervised Learning



# Applying the Selected Model(s)



(\* Output settings may be specified by the user.)



# Oracle Advanced Analytics

Brief Demos

ORACLE®

Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

FileEditViewNavigateRunDiagramTeamToolsWindowHelp

ConnectionsData MinerBUY\_INSURANCE\_WORKFLOW

Connections

dmuser

ACME Mfg Paint Project

BERGERS R US

Fun with Gov

Predictive Analytics WF

Chicago Crime

Customers R Us Project

A+ Students OAA analytics

Big Data Analytics w JSON

Big Data Analytics

BUY\_INSURANCE\_WORKFLOW

CARS\_DATA\_MINING

Churners01 work flow

Claims Fraud Clustering + SVM2

Customer Analytics ind RFM

Employees\_attrition

Insurance Customer Analytics

Manufacturing Painting Ops

Market Basket Analysis

OOV14\_CUSTOMER\_ANALYTICS\_360\_I

Predictive Oueries

Thumbnail

BUY\_INSURANCE\_WORKFLOW - ...

Reports

Clustering Segmentation

Explore/Profile Data

Prediction Queries by Region

Prediction Query

Filter Cols\_Attrib Importance

Predictive Models for Student

Class Build

Predictive models ind TEXT

Filter Columns Details

Scatter Box plots etc.

Graph

Anomaly Detection Query

CUST\_INSUR\_LTV

CUST\_INSUR\_LTV\_APPLY

MINING\_DATA\_TEXT\_BUILD\_V

Start Page

100%

Parallel Query Off

CUST\_INSUR\_LTV1

CUST\_INSUR\_LTV1 - Properties

Find

Data

Cache

Details

Source Table: DMUSER.CUST\_INSUR\_LTV

Name	Alias	Data Type
AGE		NUMBER
BANK_FUNDS		NUMBER
BUY_INSURANCE		VARCHAR2
CAR_OWNERSHIP		NUMBER
CHECKING_AMOUNT		NUMBER

Components

Workflow Editor

Data

Create Table or View

Data Source

Explore Data

Graph

SQL Query

Update Table

Transforms

Aggregate

Filter Columns

Filter Columns Details

Filter Rows

Join

JSON Query

Sample

Transform

Text

Models

Anomaly Detection

Association

Classification

Clustering

Feature Extraction

Model

Model Details

Regression

Predictive Queries

Evaluate and Apply

Linking Nodes

FileEditViewNavigateRunDiagramTeamToolsWindowHelp

ConnectionsData MinerBUY\_INSURANCE\_WORKFLOW

Connections

dmuser

ACME Mfg Paint Project

BERGERS R US

Fun with Gov

Predictive Analytics WF

Chicago Crime

Customers R Us Project

A+ Students OAA analytics

Big Data Analytics w JSON

Big Data Analytics

BUY\_INSURANCE\_WORKFLOW

CARS\_DATA\_MINING

Churners01 work flow

Claims Fraud Clustering + SVM2

Customer Analytics ind RFM

Employees\_attrition

Insurance Customer Analytics

Manufacturing Painting Ops

Market Basket Analysis

OOW14\_CUSTOMER\_ANALYTICS\_360\_I

Predictive Queries

Thumbnail

BUY\_INSURANCE\_WORKFLOW - Reports

Clustering Segmentation

Explore/Profile Data

Prediction Queries by Region

Prediction Query

Filter Cols\_Attrb Importance

Predictive Models for Student

Class Build

Predictive models ind TEXT

Filter Columns Details

Scatter Box plots etc.

Graph

Anomaly Detection Query

CUST\_INSUR\_LTV

CUST\_INSUR\_LTV\_APPLY

MINING\_DATA\_TEXT\_BUILD\_V

Start Page

100%

Parallel Query Off

CUST\_INSUR\_LTV1 - Properties

Find

Data

Cache

Details

Source Table: DMUSER.CUST\_INSUR\_LTV

Name	Alias	Data Type
AGE		NUMBER
BANK_FUNDS		NUMBER
BUY_INSURANCE		VARCHAR2
CAR_OWNERSHIP		NUMBER
CHECKING_AMOUNT		NUMBER

Connect

Run

Force Run

Edit...

View Data

Generate Apply Chain

Show Event Log

Validate Parents

Deploy

Save SQL

Cut

Copy

Paste

Extended Paste...

Select All

Parallel Query ...

Copy Image to Clipboard

Save Image As...

Go to Properties

Components

Workflow Editor

Data

Create Table or View

Data Source

Explore Data

Graph

SQL Query

Update Table

Transforms

Aggregate

Filter Columns

Filter Columns Details

Filter Rows

Join

JSON Query

Text

Models

Anomaly Detection

Association

Classification

Clustering

Feature Extraction

Model

Model Details

Regression

Predictive Queries

Evaluate and Apply

Linking Nodes

dmuser/Customers R Us Project/BUY\_INSURANCE\_WORKFLOW

FileEditViewNavigateRunTeamToolsWindowHelp

ConnectionsData Miner

Start PageBUY\_INSURANCE\_WORKFLOWCUST\_INSUR\_LTV1

Connections

dmuser

ACME Mfg Paint Project

BERGERS R US

Fun with Gov

Predictive Analytics WF

Chicago Crime

Customers R Us Project

A+ Students OAA analytics

Big Data Analytics

Big Data Analytics w JSON

BUY\_INSURANCE\_WORKFLOW

CARS\_DATA\_MINING

Churners01 work flow

Claims Fraud Clustering + SVM2

Customer Analytics ind RFM

Employees\_attrition

Insurance Customer Analytics

Manufacturing Painting Ops

Market Basket Analysis

OOO14\_CUSTOMER\_ANALYTICS\_360\_I

Predictive Queries

Public Sector Tax Audit

R\_regression\_models

Star Schema360 Degree Customer

Fun with Fraud

CUST\_INSUR\_LTV1 - Structure

Reports

No Structure

DataColumnsSQL

View: Actual DataSort...Parallel Query Off...Filter: Enter Where Clause

CUST_ID	N_TRANS_ATM	CHECKING_AMOUNT	MARITAL_STATUS	SEX	N_TRANS_KIOSK	STATE	HOUSE_OWNERSHIP	MONTHLY_CHECKS_WRITTEN	LTV	SALARY	BANK_FUNDS	BUY_INSURANCE	CREDIT_BALANCE	N_MORTGAI
1	CU2404	6	25 WIDOWED	F	9 DC	1		25,370...	63,883	25,000	No		549	
2	CU2405	0	25 MARRIED	M	2 MI	1		29,594.5	59,978	0	No		0	
3	CU2406	5	25 WIDOWED	M	2 CA	1		26,473.5	67,894	17,200	Yes		13,859	
4	CU2407	2	3,226 DIVORCED	F	2 MI	1		25,976...	67,107	7,600	No		0	
5	CU2408	4	130 DIVORCED	F	2 NY	1		28,446.5	70,986	1,500	No		23,525	
6	CU2409	4	6,008 DIVORCED	F	2 CA	0		13,921...	58,887	10,900	No		0	
7	CU2411	4	10,943 MARRIED	M	2 NY	1		23,833...	69,335	0	No		4,734	
8	CU2412	5	25 MARRIED	F	2 FL	1		22,485.5	58,342	2,600	Yes		0	
9	CU2413	4	293 SINGLE	M	1 NY	0		18,865...	67,061	1,614	Yes		2,249	
10	CU2414	2	25 DIVORCED	M	2 CA	1		29,672...	67,489	0	No		0	
11	CU2416	2	25 SINGLE	M	2 NY	0		20,190...	63,963	0	No		0	
12	CU2417	2	2,420 SINGLE	M	2 NY	0		19,941	64,964	1,156	No		963	
13	CU2418	2	1,089 DIVORCED	F	2 MI	1		27,892...	58,771	6,850	No		0	
14	CU2420	5	25 DIVORCED	M	2 NY	1		19,536.5	60,946	5,400	No		44,020	
15	CU2421	5	882 DIVORCED	F	2 CA	1		29,123.5	67,694	3,950	No		0	
16	CU2422	3	157 MARRIED	F	2 MI	1		29,145...	61,781	1,850	Yes		0	
17	CU2423	4	9,894 WIDOWED	F	8 NY	1		25,868...	63,875	3,300	No		6,707	
18	CU3100	6	227 WIDOWED	F	2 MI	1		17,539...	65,359	4,280	No		0	
19	CU3101	4	40 MARRIED	F	2 MI	1		23,178...	61,513	800	Yes		0	
20	CU3102	4	21,094 SINGLE	M	1 NV	0		21,644...	65,379	8,129	No		0	
21	CU3104	5	25 MARRIED	F	2 NY	1		16,525.5	66,902	17,001	No		92,408	
22	CU3105	7	638 OTHER	F	2 NY	1		24,184.5	63,538	20,200	Yes		0	
23	CU3106	4	372 DIVORCED	M	2 MI	1		26,580.5	69,522	0	No		0	
24	CU3107	0	25 SINGLE	M	2 CA	0		19,374	64,296	0	No		0	
25	CU3108	3	25 MARRIED	F	2 NY	1		14,843...	60,575	700	Yes		0	
26	CU3109	4	25 MARRIED	M	2 MI	1		30,839.5	74,158	0	No		0	
27	CU3110	4	25 MARRIED	F	2 CA	1		21,882...	59,929	2,300	No		36,102	
28	CU3111	4	25 DIVORCED	M	2 MN	1		27,128...	55,713	3,650	No		0	
29	CU3113	1	25 SINGLE	M	2 MS	0		15,834.5	64,138	0	No		0	
30	CU3115	2	25 MARRIED	M	3 CA	1		27,130	66,120	1,600	No		0	
31	CU3116	4	17,173 DIVORCED	F	2 CA	1		24,554	68,616	4,000	No		11,098	
32	CU3117	4	13,225 SINGLE	M	1 MI	0		16,890	65,560	1,606	No		0	
33	CU3118	4	7,543 SINGLE	M	3 CA	1		28,260...	71,843	700	No		0	
34	CU3119	4	25 DIVORCED	F	3 MI	1		27,410	58,840	0	No		0	
35	CU3120	2	25 SINGLE	M	3 CA	0		1	63,095	0	No		0	
36	CU3121	4	543 DIVORCED	M	3 CA	1		28,178...	60,315	5,700	No		0	
37	CU3123	1	25 SINGLE	M	3 NY	0		24,367.5	72,670	0	No		0	
38	CU3124	4	12,988 SINGLE	M	3 MN	0		17,477.5	70,310	10,100	No		0	
39	CU3125	6	25 DIVORCED	M	3 CA	1		22,995	68,380	5,950	No		66,417	
40	CU3126	2	108 DIVORCED	F	3 MI	1		24,825	60,900	7,200	No		0	

Components

No Components

File Edit View Navigate Run Diagram Team Tools Window Help

Connections Data Miner Start Page BUY\_INSURANCE\_WORKFLOW CUST\_INSUR\_LTV1 100% Parallel Query Off

Connections

- dmuser
  - ACME Mfg Paint Project
    - BERGERS R US
      - Fun with Gov
      - Predictive Analytics WF
      - Chicago Crime
    - Customers R Us Project
      - A+ Students OAA analytics
      - Big Data Analytics w JSON
      - Big Data Analytics
      - BUY\_INSURANCE\_WORKFLOW
      - CARS\_DATA\_MINING
      - Churners01 work flow
      - Claims Fraud Clustering + SVM2
      - Customer Analytics ind RFM
      - Employees\_attrition
      - Insurance Customer Analytics
      - Manufacturing Painting Ops
      - Market Basket Analysis
      - OOW14\_CUSTOMER\_ANALYTICS\_360\_I

Thumbnail

BUY\_INSURANCE\_WORKFLOW - Reports

- Class Build
- Predictive models ind TEXT
- Filter Columns Details
- Scatter Box plots etc.
- Graph
- Anomaly Detection Query
- CUST\_INSUR\_LTV
- CUST\_INSUR\_LTV\_APPLY
- MINING\_DATA\_TEXT\_BUILD\_V
- Model Details
- GLM std coefficients
- PREDICTIONS
- OUTPUT\_12\_15
- Sample
- Likely Buyers
- Customer Segments
- CUST\_INSUR\_LTV1
- Explore Data

Explore Data - Properties

Input

Group By: <Select Group By>

☒ Auto Input Columns Selection

Name	Data Type
AGE	NUMBER
BANK_FUNDS	NUMBER
BUY_INSURANCE	VARCHAR2

Components

Workflow Editor

Data

- Create Table or View
- Data Source
- Explore Data
- Graph
- SQL Query
- Update Table

Transforms

- Aggregate
- Filter Columns
- Filter Columns Details
- Filter Rows
- Join
- JSON Query

Text

Models

- Anomaly Detection
- Association
- Classification
- Clustering
- Feature Extraction
- Model
- Model Details
- Regression

Predictive Queries

Evaluate and Apply

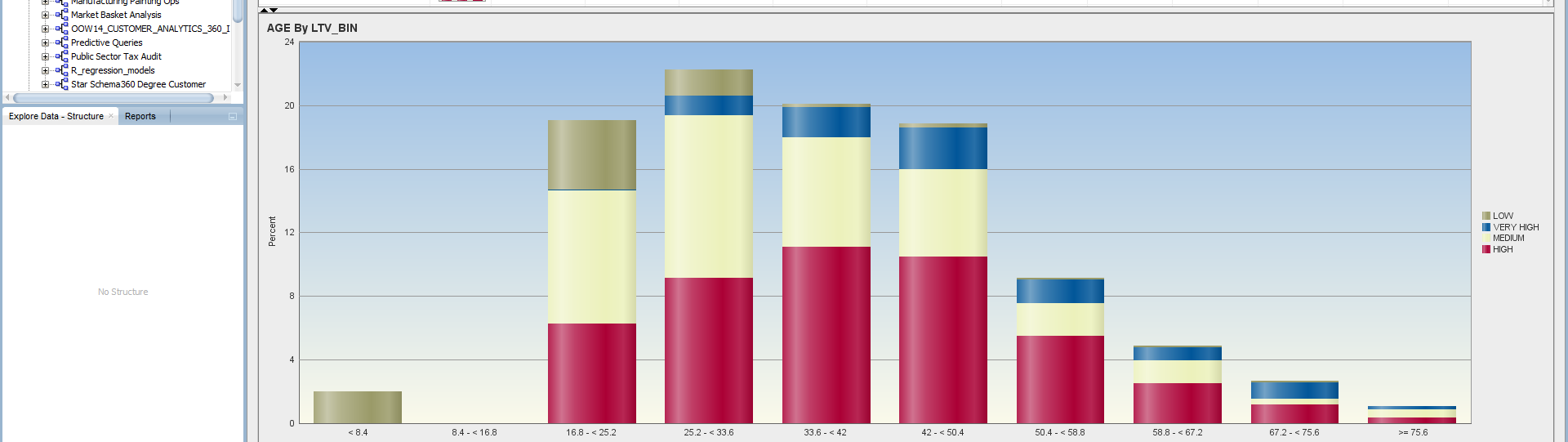
Linking Nodes



Statistics Data | Columns | SQL

Statistics: 10 Columns from 2,005 Rows(Sampled)

Name	Histogram	Data Type	Percent NULLs	Distinct Values	Distinct Percent	Mode	Average	Median	Min Value	Max Value	Standard Deviation	Variance
AGE		NUMBER	0	69	3.4414		37.6823	36	0	84	14.4737	209.4873
BANK_FUNDS		NUMBER	0	425	21.197		2,585.3736	551	0	30,200	4,739.2687	22,460,667.7132
BUY_INSURANCE		VARCHAR2	0	2	0.0998	No						
CAR_OWNERSHIP		NUMBER	0	2	0.0998		0.9247	1	0	1	0.254	0.0697
CHECKING_AMOUNT		NUMBER	0	626	31.2219		1,032.7312	25	25	24,471	3,103.6812	9,632,836.6997
CREDIT_BALANCE		NUMBER	0	187	9.3267		2,369.3825	0	0	185,775	12,753.4182	162,649,675.2952
CREDIT_CARD_LIMITS		NUMBER	0	26	1.2968		1,266.384	1,000	500	5,000	814.3591	663,180.7806
CUST_ID		VARCHAR2	0	2,005	100	<Other>						
FIRST		VARCHAR2	0	1,379	68.7781	<Other>						
HAS_CHILDREN		NUMBER	0	2	0.0998		0.5172	1	0	1	0.4998	0.2498
HOUSE_OWNERSHIP		NUMBER	0	3	0.1496		0.79	1	0	2	0.5116	0.2618



dmuser/Customers R Us Project/BUY\_INSURANCE\_WORKFLOW

FileEditViewNavigateRunDiagramTeamToolsWindowHelp

ConnectionsData MinerStart PageBUY\_INSURANCE\_WORKFLOWCUST\_INSUR\_LTV1Graph 1

Connections  
dmuser  
ACME Mfg Paint Project  
BERGERS R US  
Fun with Gov  
Predictive Analytics WF  
Chicago Crime  
Customers R Us Project  
A+ Students OAA analytics  
Big Data Analytics w JSON  
Big Data Analytics  
BUY\_INSURANCE\_WORKFLOW  
CARS\_DATA\_MINING  
Churners01 work flow  
Claims Fraud Clustering + SVM2  
Customer Analytics ind RFM  
Employees\_attrition  
Insurance Customer Analytics  
Manufacturing Painting Ops  
Market Basket Analysis

Thumbnail  
BUY\_INSURANCE\_WORKFLOW - ...  
Reports  
Predictive Models for Student  
Class Build  
Predictive models ind TEXT  
Filter Columns Details  
Scatter Box plots etc.  
Graph  
Graph 1  
Anomaly Detection Query  
CUST\_INSUR\_LTV  
CUST\_INSUR\_LTV\_APPLY  
MINING\_DATA\_TEXT\_BUILD\_V  
CUST\_INSUR\_LTV1  
Model Details  
GLM std coefficients  
PREDICTIONS  
OUTPUT\_12\_15  
Likely Buyers  
Customer Segments  
Filter Columns

100%  
Parallel Query Off

```
graph LR; CUST_INSUR_LTV1 --> Graph_1[Graph 1]; CUST_INSUR_LTV1 --> Explore_Data[Explore Data]; CUST_INSUR_LTV1 --> Filter[Filter]
```

Filter Columns - Properties  
Columns  
NameTypeOutputHint  
AGENUMBER→  
BANK\_FUNDSNUMBER→  
BUY\_INSURANCEVARCHAR2→  
CAR\_OWNERSHIPNUMBER→  
CHECKING\_AMOUNTNUMBER→

Filter  
Cache  
Details

Connect  
Run  
Force Run  
Edit...  
View Data  
View Attribute Importance  
Generate Apply Chain  
Show Event Log  
Validate Parents  
Deploy  
Save SQL  
Cut  
Copy  
Paste  
Extended Paste...  
Select All  
Parallel Query ...  
Copy Image to Clipboard  
Save Image As...  
Go to Properties  
Navigate

Selected Node  
Selected Node and Children  
Selected Node and Parents  
Child Nodes Only

Components  
Workflow Editor  
Data  
Create Table or View  
Data Source  
Explore Data  
Graph  
SQL Query  
Update Table  
Transforms  
Aggregate  
Filter Columns  
Filter Columns Details  
Filter Rows  
Join  
JSON Query  
Text  
Models  
Anomaly Detection  
Association  
Classification  
Clustering  
Feature Extraction  
Model  
Model Details  
Regression  
Predictive Queries  
Evaluate and Apply  
Linking Nodes

dmuser/Customers R Us Project/BUY\_INSURANCE\_WORKFLOW

Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

45

FileEditViewNavigateRunDiagramTeamToolsWindowHelp

ConnectionsData MinerStart PageBUY\_INSURANCE\_WORKFLOWCUST\_INSUR\_LTV1Graph 1

ConnectionsdmuserACME Mfg Paint ProjectBERGERS R USFun with GovPredictive Analytics WFChicago CrimeCustomers R Us ProjectA+ Students OAA analytBig Data Analytics w JSONBig Data AnalyticsBUY\_INSURANCE\_WORKFLOWCARS\_DATA\_MININGChurners01 work flowClaims Fraud ClusteringCustomer Analytics ind REmployees\_attritionInsurance Customer AnalManufacturing Painting OMarket Basket Analysis

Thumbnail

BUY\_INSURANCE\_WORKFLOW - ...

Clustering SegmentationExplore/Profile DataExplore DataPrediction Queries by RegionPrediction QueryFilter Cols\_Attrib ImportancePredictive Models for StudentClass BuildPredictive models ind TEXTFilter Columns DetailsScatter Box plots etc.GraphGraph 1Anomaly Detection QueryCUST\_INSUR\_LTV1CUST\_INSUR\_LTV\_APPLYMINING\_DATA\_TEXT\_BUILD\_VCUST\_INSUR\_LTV1Model Details

dmuser/Customers R Us Project/BUY\_INSURANCE\_WORKFLOW

100%

Parallel Query Off

Edit Filter Columns Node

Show Attribute Importance

Show Data Quality

Columns

AllNone

QName

Name	Type	Output	Rank	Importance	% Null	% Unique	% Constant	Hints
1X2 BANK_FUNDS	NUMBER	→	1	0.2039	0	21.3511	35.6016	
1X2 N_TRANS_ATM	NUMBER	→	2	0.1217	0	0.4438	21.499	
1X2 N_TRANS_TELLER	NUMBER	→	3	0.1213	0	0.4931	32.3471	
1X2 MONEY_MONTHLY_OVERDRAWN	NUMBER	→	4	0.1207	0	18.787	16.3215	
1X2 T_AMOUNT_AUTOM_PAYMENTS	NUMBER	→	5	0.1054	0	59.9606	21.5483	
1X2 MONTHLY_CHECKS_WRITTEN	NUMBER	→	6	0.0848	0	0.9369	19.428	
1X2 N_OF_DEPENDENTS	NUMBER	→	7	0.0315	0	0.3452	34.4181	
1X2 TIME_AS_CUSTOMER	NUMBER	→	8	0.0221	0	0.2465	31.3116	
1X2 CHECKING_AMOUNT	NUMBER	→	9	0.0183	0	30.3254	62.7712	
1X2 CREDIT_BALANCE	NUMBER	→	10	0.0135	0	9.4181	90.5819	
1X2 N_TRANS_KIOSK	NUMBER	→	11	0.0084	0	0.5424	43.6884	
1X2 MORTGAGE_AMOUNT	NUMBER	→	12	0.0073	0	21.2525	23.1262	
1X2 SEX	VARCHAR2	→	13	0.0064	0	0.0986	66.075	
1X2 MARITAL_STATUS	VARCHAR2	→	14	0.0056	0	0.2465	34.5168	
1X2 HOUSE_OWNERSHIP	NUMBER	→	15	0.004	0	0.1479	71.1538	
1X2 N_MORTGAGES	NUMBER	→	16	0.004	0	0.1479	71.1538	

HelpOKCancel

Components

Workflow Editor

Data

Data SourceExplore DataSQL QueryUpdate Table

Filter ColumnsFilter Columns DetailsJoinJSON Query

AssociationClassificationFeature ExtractionModelRegression

Predictive QueriesEvaluate and ApplyLinking Nodes

1X2 BANK_FUNDS	NUMBER	→	1	0.2039	0	21.3511	35.6016	
1X2 BUY_INSURANCE	VARCHAR2	→			0	0.0986	71.8442	

ORACLE

Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

46

FileEditViewNavigateRunDiagramTeamToolsWindowHelp

ConnectionsData Miner

Connections

dmuser

ACME Mfg Paint Project

BERGERS R US

Fun with Gov

Predictive Analytics WF

Chicago Crime

Customers R Us Project

A+ Students OAA analytics

Big Data Analytics w JSON

Big Data Analytics

BUY\_INSURANCE\_WORKFLOW

CARS\_DATA\_MINING

Churners01 work flow

Claims Fraud Clustering + SVM2

Customer Analytics ind RFM

Employees\_attrition

Insurance Customer Analytics

Manufacturing Painting Ops

Market Basket Analysis

Thumbnail

BUY\_INSURANCE\_WORKFLOW - ...

Reports

Clustering Segmentation

Explore/Profile Data

Explore Data

Prediction Queries by Region

Prediction Query

Filter Cols\_Attrb Importance

Filter Columns

Predictive Models for Student

Class Build

Predictive models ind TEXT

Filter Columns Details

Scatter Box plots etc.

Graph

Graph 1

Anomaly Detection Query

CUST\_INSUR\_LTV

CUST\_INSUR\_LTV\_APPLY

MINING\_DATA\_LTV\_BUILD\_V

CUST\_INSUR\_LTV1

Start Page

BUY\_INSURANCE\_WORKFLOW

CUST\_INSUR\_LTV1

Graph 1

100%

Parallel Query Off

Graph 1

CUST\_INSUR\_LTV1

Explore Data

Filter Columns

Class Build

Connect

Run

Force Run

Edit...

Advanced Settings...

View Models

View Test Results

Compare Test Results

Generate Apply Chain

Show Event Log

Validate Parents

Deploy

Cut

Copy

Paste

Extended Paste...

Select All

Parallel Query ...

Copy Image to Clipboard

Save Image As...

Go to Properties

Navigate

Class Build 1 - Properties

Find

Models

Build

Test

Details

Model Settings

Name

Output

Build

Test

CLAS\_GLM\_5\_2

CLAS\_SVM\_5\_2

CLAS\_DT\_4\_2

CLAS\_NB\_4\_2

9/14/15 5:17 PM

9/14/15 5:17 PM

9/14/15 5:17 PM

9/14/15 5:17 PM

Automatic

Automatic

Generalized Linear Model

Support Vector Machine

Decision Tree

Naive Bayes

Components

Workflow Editor

Data

Create Table or View

Data Source

Explore Data

Graph

SQL Query

Update Table

Transforms

Aggregate

Filter Columns

Filter Columns Details

Filter Rows

Join

JSON Query

Text

Models

Anomaly Detection

Association

Classification

Clustering

Feature Extraction

Model

Model Details

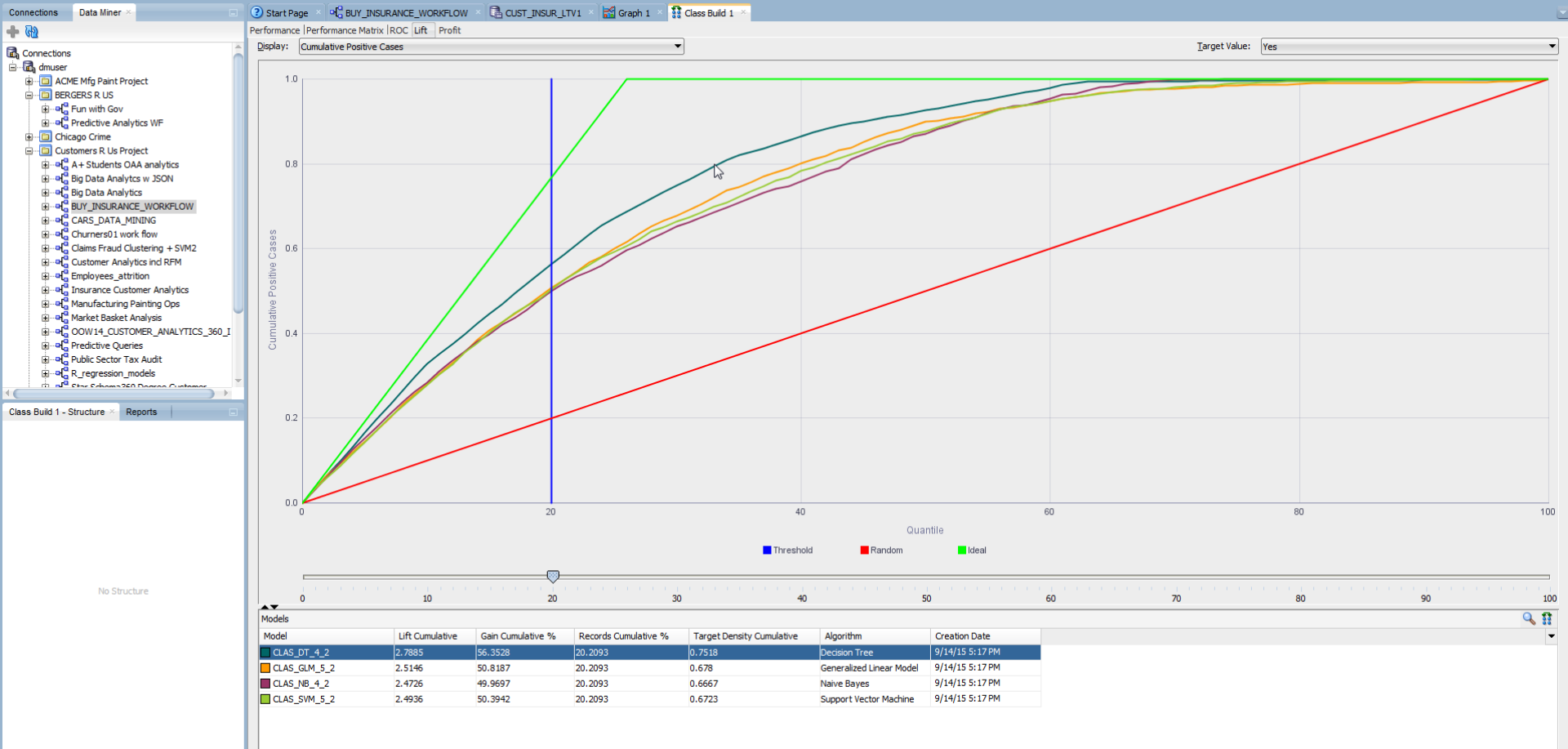
Regression

Predictive Queries

Evaluate and Apply

Linking Nodes

Synchronization complete



FileEditViewNavigateRunDiagramTeamToolsWindowHelp

ConnectionsData Miner

Connections

dmuser

ACME Mfg Paint Project

BERGERS R US

Fun with Gov

Predictive Analytics WF

Chicago Crime

Customers R Us Project

A+ Students OAA analytics

Big Data Analytics w JSON

Big Data Analytics

BUY\_INSURANCE\_WORKFLOW

CARS\_DATA\_MINING

Churners01 work flow

Claims Fraud Clustering + SVM2

Customer Analytics ind RFM

Employees\_attrition

Insurance Customer Analytics

Manufacturing Painting Ops

Market Basket Analysis

Thumbnail

CLAS\_DT\_4\_2 - StructureReports

Node: 0

Node: 1

Node: 15

Node: 16

Node: 2

Node: 3

Node: 4

Node: 5

Node: 17

Node: 6

Node: 18

Node: 19

Node: 7

Node: 8

Node: 20

Node: 21

Node: 9

Node: 22

Node: 23

Start PageBUY\_INSURANCE\_WORKFLOWCUST\_INSUR\_LTV1Graph 1CLAS\_DT\_4\_2

Tree Settings

100%Maximum Target Values: 2

Confidence: 86.55%  
No : 1319 (86.55%)  
Yes : 205 (13.45%)

Confidence: 72.46%  
No : 121 (72.46%)  
Yes : 46 (27.54%)

Confidence: 72.08%  
No : 615 (27.92%)  
Yes : 1588 (72.08%)  
Split: MONEY\_MONTHLY\_OVERDRAWN

Confidence: 64.15%  
No : 823 (64.15%)  
Yes : 460 (35.85%)  
Split: N\_TRANS\_ATM

Confidence: 82%  
No : 123 (82.00%)  
Yes : 27 (18.00%)

Node: 9  
Prediction: Yes  
Support: 873 (9.46%)  
Confidence: 86.25%  
No : 120 (13.75%)  
Yes : 753 (86.25%)  
Split: T\_AMOUNT\_AUTOM\_PAYMENTS

Node: 8  
Prediction: Yes  
Support: 1330 (14.42%)  
Confidence: 62.78%  
No : 495 (37.22%)  
Yes : 835 (62.78%)  
Split: T\_AMOUNT\_AUTOM\_PAYMENTS

Node: 6  
Prediction: No  
Support: 1218 (13.2%)  
Confidence: 62.4%  
No : 760 (62.40%)  
Yes : 458 (37.60%)  
Split: T\_AMOUNT\_AUTOM\_PAYMENTS

Node: 17  
Prediction: No  
Support: 65 (0.7%)  
Confidence: 96.92%  
No : 63 (96.92%)  
Yes : 2 (3.08%)

Node: 22  
Prediction: Yes  
Support: 851 (9.22%)  
Confidence: 88.25%  
No : 100 (11.75%)  
Yes : 751 (88.25%)

Node: 23  
Prediction: No  
Support: 22 (0.24%)  
Confidence: 90.91%  
No : 20 (90.91%)  
Yes : 2 (9.09%)

Node: 20  
Prediction: Yes  
Support: 1244 (13.48%)  
Confidence: 66%  
No : 423 (34.00%)  
Yes : 821 (66.00%)

Node: 21  
Prediction: No  
Support: 86 (0.93%)  
Confidence: 83.72%  
No : 72 (83.72%)  
Yes : 14 (16.28%)

Node: 19  
Prediction: No  
Support: 418 (4.53%)  
Confidence: 74.88%  
No : 313 (74.88%)  
Yes : 105 (25.12%)

Node: 18  
Prediction: No  
Support: 800 (8.67%)  
Confidence: 55.88%  
No : 447 (55.88%)  
Yes : 353 (44.12%)

RuleSurrogatesTarget Values

Node Rule:

IfBANK\_FUNDS > 246

AndCHECKING\_AMOUNT <= 282

AndCREDIT\_BALANCE <= 2445

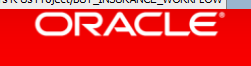
AndMONEY\_MONTHLY\_OVERDRAWN > 54.095

AndT\_AMOUNT\_AUTOM\_PAYMENTS <= 14993

ThenYes

Confidence	0.882491186839013
Support	0.0922393236505279

Wrap



File Edit View Navigate Run Diagram Team Tools Window Help

Connections Data Miner Start Page BUY\_INSURANCE\_WORKFLOW CUST\_INSUR\_LTV1 Graph 1 CLAS\_DT\_4\_2

Connections

- dmuser
- ACME Mfg Paint Project
- BERGERS R US
  - Fun with Gov
  - Predictive Analytics WF
  - Chicago Crime
- Customers R Us Project
  - A+ Students OAA analytics
  - Big Data Analytics w JSON
  - Big Data Analytics
  - BUY\_INSURANCE\_WORKFLOW
  - CARS\_DATA\_MINING
  - Churners01 work flow
  - Claims Fraud Clustering + SVM2
  - Customer Analytics ind RFM
  - Employees\_attrition
  - Insurance Customer Analytics
  - Manufacturing Painting Ops

Thumbnail

BUY\_INSURANCE\_WORKFLOW - ... Reports

- Clustering Segmentation
- Explore/Profile Data
- Explore Data
- Prediction Queries by Region
- Prediction Query
- Filter Cols\_Attrib Importance
- Filter Columns
- Predictive Models for Student
- Class Build
- Predictive models ind TEXT
- Class Build 1
- Filter Columns Details
- Scatter Box plots etc.
- Graph
- Graph 1
- Anomaly Detection Query
- CUST\_INSUR\_LTV
- CUST\_INSUR\_LTV\_APPLY
- MINING\_DATA\_TEXT\_BUILD\_V
- CUST\_INSUR\_LTV1

Workflow Diagram:

Graph 1

CUST\_INSUR\_LTV1

Explore Data

Filter Columns

Class Build 1

CUST\_INSUR\_LTV\_APPLY1

Context Menu:

- Connect
- Run
- Force Run
- Edit...
- View Data
- Generate Apply Chain
- Show Event Log
- Validate Parents
- Deploy
- Save SQL
- Cut
- Copy
- Paste
- Extended Paste...
- Select All
- Parallel Query ...
- Copy Image to Clipboard
- Save Image As...
- Go to Properties
- Navigate

Apply - Properties

Predictions

Additional Output

Cache

Details

Automatic Settings Case ID: CUST\_ID

Output Apply Columns

Column	Function	Parameter(s)	Model	Node
CLAS_NB_4_2_PROB_Yes	Prediction Probability	Prediction: Yes	CLAS_NB_4_2	Class Build 1
CLAS_NB_4_2_PDET	Prediction Details	Prediction: Yes, Sort: Absolute, Length: 5	CLAS_NB_4_2	Class Build 1

Components

Workflow Editor

Data

- Create Table or View
- Data Source
- Explore Data
- Graph
- SQL Query
- Update Table

Transforms

- Aggregate
- Filter Columns
- Filter Columns

Text

Models

- Anomaly Detection
- Association
- Classification
- Clustering
- Feature Extraction
- Model

Predictive Queries

Evaluate and Apply

- Apply
- Test

Linking Nodes

dmuser/Customers R Us Project/BUY\_INSURANCE\_WORKFLOW

FileEditViewRunTeamToolsWindowHelp

ConnectionsData Miner

Start PageBUY\_INSURANCE\_WORKFLOWCUST\_INSUR\_LTV1Graph 1CLAS\_DT\_4\_2Apply

DataColumnsSQL

ViewActual DataSort...Parallel Query Off...FilterEnter Where Clause

	CLAS_NB_4_2_PROB_Yes	CLAS_NB_4_2_PDCT	CUST_ID	LAST	N_MORTGAGES	SALARY
1	0.9998904466629028	<Details algorithm="Naive Bayes" das...	CU3367	LOUISE	1	63,966
2	0.9998904466629028	<Details algorithm="Naive Bayes" das...	CU9432	CLEOTILDE	1	96,573
3	0.9998646974563599	<Details algorithm="Naive Bayes" das...	CU6274	SUDIE	1	69,159
4	0.9998481869697571	<Details algorithm="Naive Bayes" das...	CU2126	DWANA	1	65,581
5	0.9998325705528259	<Details algorithm="Naive Bayes" das...	CU12392	FLOSSIE	1	93,595
6	0.999816358089447	<Details algorithm="Naive Bayes" das...	CU13307	JULIO	1	70,535

Connections

dmuser

ACME Mfg Paint Project

BERGERS R US

Fun with Gov

Predictive Analytics WF

Chicago Crime

Customers R Us Project

A+ Students OAA analytics

Big Data Analytics w JSON

Big Data Analytics

BUY\_INSURANCE\_WORKFLOW

CARS\_DATA\_MINING

Churners01 work flow

Claims Fraud Clustering

Customer Analytics ind RF

Employees\_attrition

Insurance Customer Analy

Manufacturing Painting Op

Market Basket Analysis

OOOW14\_CUSTOMR\_ANA

Predictive Queries

Public Sector Tax Audit

R regression models

Components

No Components

View Value

Find

```
<Details algorithm="Naive Bayes" class="Yes">
  <Attribute name="LTV_BIN" actualValue="MEDIUM" operator="in" range="LOW,MEDIUM,VERY HIGH" weight=".987" rank="1"/>
  <Attribute name="N_TRANS_WEB_BANK" actualValue="2300" operator="greaterThan" value="1419.5" weight=".963" rank="2"/>
  <Attribute name="LTV" actualValue="21389.75" operator="greaterThan" value="6861.625" weight=".848" rank="3"/>
  <Attribute name="CAR_OWNERSHIP" actualValue="1" weight=".813" rank="4"/>
  <Attribute name="AGE" actualValue="51" operator="between" range="(17.5:55.5]" weight=".8" rank="5"/>
</Details>
```

Close

dmuser/Customers R Us Project/BUY\_INSURANCE\_WORKFLOW

ORACLE

Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

51



Oracle SQL Developer: jw1415/Project1/workflow

File Edit View Navigate Run Diagram Team Tools Window Help Automation

workflow - Structure Thumbnail

Start Page 100% Performance Options On (Selected)

INSUR\_CUST\_LTV\_SAMPLE Class Build

Connections Data Miner

Connections

- den00jf-im
- jw1415
- wfriday
- Oracle NoSQL Connections
- Cloud Connections

Workflow Jobs

den00jf-im

Workflow Project Status

No entries. Right click for Preferences.

workflows - Properties

Comment:

Reports

- All Reports
- Analytic View Reports
- Data Dictionary Reports
- Data Modeler Reports
- OLAP Reports
- TimesTen Reports
- User Defined Reports

Logging Page - Log

Level	Sequence	Elapsed	Source	Message
✖	2376	1155	o.dmt.dataminer.runne...	RunItem: unable to read repository config table.
✖	2375	1373	o.dmt.dataminer.runne...	RunItem: unable to read repository config table.
✖	2374	1513	o.dmt.dataminer.runne...	RunItem: unable to read repository config table.
✖	2373	1622	o.dmt.dataminer.runne...	RunItem: unable to read repository config table.
✖	2372	1092	o.dmt.dataminer.runne...	RunItem: unable to read repository config table.
✖	2371	1170	o.dmt.dataminer.runne...	RunItem: unable to read repository config table.

Messages Statements Logging Page

Components

Workflow Editor

Data

- Create Table or View
- Data Source
- Explore Data
- Graph
- SQL Query
- Update Table

Models

- Anomaly Detection
- Association
- Classification
- Clustering
- Explicit Feature Extraction
- Feature Extraction
- Model
- Model Details
- Regression

Predictive Queries

Model Operations

Linking Nodes

Repeat

Frequency: Weekly

Every 1 week(s) on:

☐ Sun ☒ Mon ☐ Tue ☐ Wed ☐ Thur ☐ Fri ☐ Sat

Help OK Cancel

Schedule Workflow

Start Date: Jan 15, 2016 15:58

Repeat: Custom...

End Repeat: Never

☐ Use Existing Schedule

<Select a Schedule>

Help Advanced... OK Cancel

3:02 PM 1/15/2016

# Fraud Prediction Demo

## Automated In-DB Analytical Methodology



```
drop table CLAIMS_SET;  
exec dbms_data_mining.drop_model('CLAIMSMODEL');  
create table CLAIMS_SET (setting_name varchar2(30), setting_value varchar2(4000));  
insert into CLAIMS_SET values ('ALGO_NAME','ALGO_SUPPORT_VECTOR_MACHINES');  
insert into CLAIMS_SET values ('PREP_AUTO','ON');  
commit;
```

```
begin  
dbms_data_mining.create_model('CLAIMSMODEL', 'CLASSIFICATION',  
  'CLAIMS', 'POLICYNUMBER', null, 'CLAIMS_SET');  
end;  
/
```

```
-- Top 5 most suspicious fraud policy holder claims  
select * from  
(select POLICYNUMBER, round(prob_fraud*100,2) percent_fraud,  
      rank() over (order by prob_fraud desc) rnk from  
(select POLICYNUMBER, prediction_probability(CLAIMSMODEL, '0' using *) prob_fraud  
 from CLAIMS  
 where PASTNUMBEROFCLAIMS in ('2to4', 'morethan4'))  
 where rnk <= 5  
 order by percent_fraud desc;
```

POLICYNUMBER	PERCENT_FRAUD	RNK	
-----	-----	-----	
6532	64.78	1	
2749	64.17	2	
3440	63.22	3	
654	63.1		4
12650	62.36	5	

**Automated Monthly “Application”!** *Just add:*

Create  
View CLAIMS2\_30  
As  
Select \* from CLAIMS2  
Where mydate > SYSDATE – 30

Time measure: set timing on;

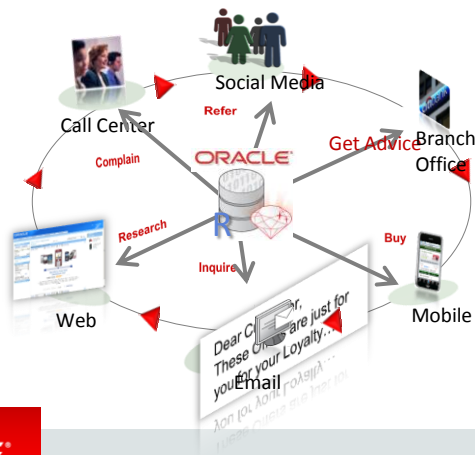
# Oracle Advanced Analytics

## More Details

- On-the-fly, single record apply with new data (e.g. from call center)



```
Select prediction_probability(CLAS_DT_1_2, 'Yes'  
  USING 7800 as bank_funds, 125 as checking_amount, 20 as  
  credit_balance, 55 as age, 'Married' as marital_status,  
  250 as MONEY_MONTHLY_OVERDRAWN, 1 as house_ownership)  
from dual;
```



### Likelihood to respond:

Query Result	
All Rows Fetched: 1 in 0 seconds	
PREDICTION_PROB...	
0.8382936507936...	

# OAA/Oracle Data Miner 4.1 HOL

## Uses Oracle by Example Free Online Tutorials

- There are 6 Tutorials

- The first tutorial is already done for you

- Recommend doing 3-5 Tutorials

1. [Using Oracle Data Miner 4.1](#)

2. [Star Schema Mining Using Oracle Data Miner 4.1](#)

3. [Text Mining with an EM Clustering Model Using Data Miner 4.1](#)

4. **Anomaly Detection (CLAIMS)** *See Instructor for assistance*

5. **Market Basket Analysis (SH.SALES)** *See Instructor for assistance*

### Content List (click links below to view content)

08-SEP-2015  
15 mins  
★★★★★

[Setting Up Oracle Data Miner 4.1](#)

This tutorial covers the process of setting up Oracle Data Miner for use within Oracle SQL Developer 4.1.

08-SEP-2015  
45 mins  
★★★★★

[Using Oracle Data Miner 4.1](#)

This tutorial covers the use of Oracle Data Miner 4.1 to perform data mining against Oracle Database 12c. In this lesson, you examine the Oracle Data Miner GUI. The Oracle Data Miner GUI is included as an extension of Oracle SQL Developer, version 4.1.

08-SEP-2015  
30 mins  
★★★★★

[Star Schema Mining Using Oracle Data Miner 4.1](#)

This tutorial covers the use of Oracle Data Miner 4.1 to perform star schema mining activities against Oracle Database 12c Release 1.

08-SEP-2015  
30 mins  
★★★★★

[Text Mining with an EM Clustering Model Using Data Miner 4.1](#)

In this lesson, you learn how to use the EM algorithm in a clustering model while leveraging text mining enhancements that are included in Oracle Database 12c Release 1.

08-SEP-2015  
30 mins  
★★★★★

[Using Logistic Regression Models \(GLM\) to Predict Customer Affinity](#)

This tutorial covers the use of Oracle Data Miner 4.1 to leverage enhancements to the Oracle implementation of Generalized Linear Models (GLM) to predict customer affinity.

08-SEP-2015  
30 mins  
★★★★★

[Using the SQL Query Node With Oracle Data Miner 4.1](#)

This tutorial covers the use of the new SQL Query Node in an Oracle Data Miner 4.1 workflow.

08-SEP-2015  
30 mins  
★★★★★

[Using Predictive Queries With Oracle Data Miner 4.1](#)

This tutorial covers the use of Predictive Queries against mining data by Oracle Data Miner 4.1.

08-SEP-2015  
30 mins  
★★★★★

[Mining JSON Data Using Oracle Data Miner 4.1](#)

This tutorial covers the use of the JSON Query Node in an Oracle Data Miner 4.1 workflow in order to mine this Big Data format.



ORACLE®