Learn Predictive Analytics in 2 Hours!

Oracle Advanced Analytics/ Oracle Data Miner Hands on Lab

Charlie Berger, MS Engineering, MBA, Sr. Director Product Management, Advanced Analytics and Machine Learning
charlie.berger@oracle.com
www.twitter.com/CharlieDataMine

Tim Vlamis, Consultant, Vlamis Software Solutions, Inc.

Karl Rexer, President, Rexer Analytics

ORACLE

Copyright © 2016, Oracle and/or its affiliates. All rights reserved. |

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 and discuss use cases!**

- **Take off!—Intermediate/Experts**
  1. **Environment**
     - Oracle 12c on the **Oracle Database Cloud**
     - Will be using SQL Developer 4.2 EA
  2. **Do 3-5 Tutorials**
     - Instructors will walk around helping

Copyright © 2016 Oracle and/or its affiliates. All rights reserved. |

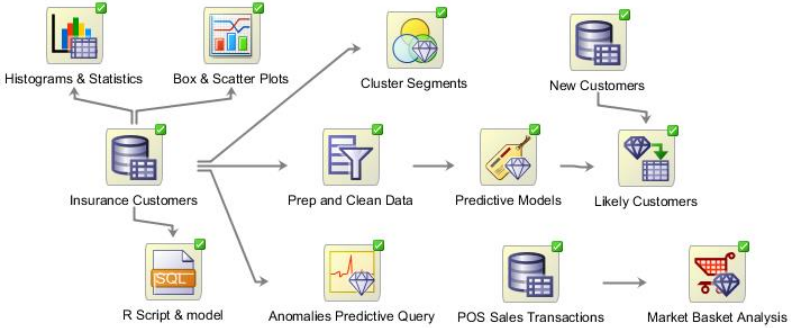
Oracle's Advanced Analytics

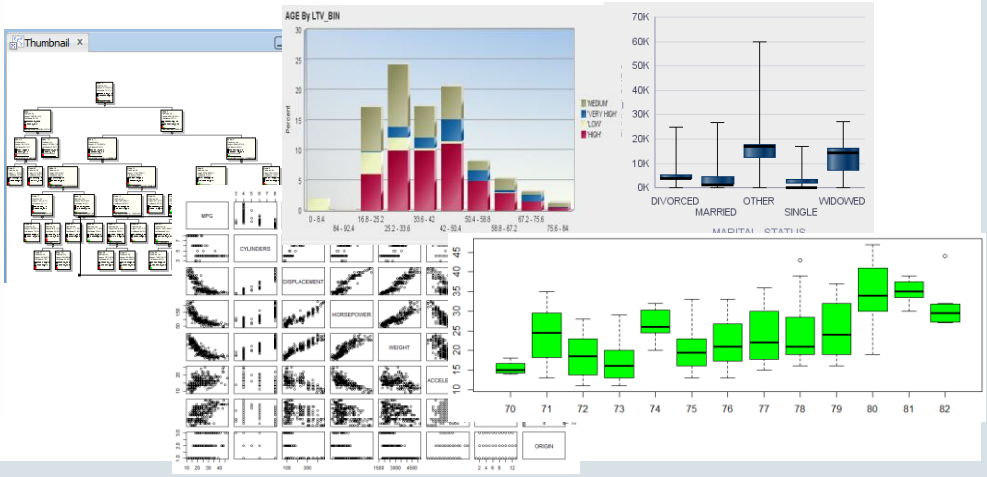
Fastest Way to Deliver Scalable Enterprise-wide Predictive Analytics

Advanced Analytics

Key Features

- Parallel, scalable data mining algorithms and R integration
- In-Database + Hadoop—Don't move the data
- Data analysts, data scientists & developers
- Drag and drop workflow, R and SQL APIs
- Extends data management into powerful advanced/predictive analytics platform
- Enables enterprise predictive analytics deployment + applications

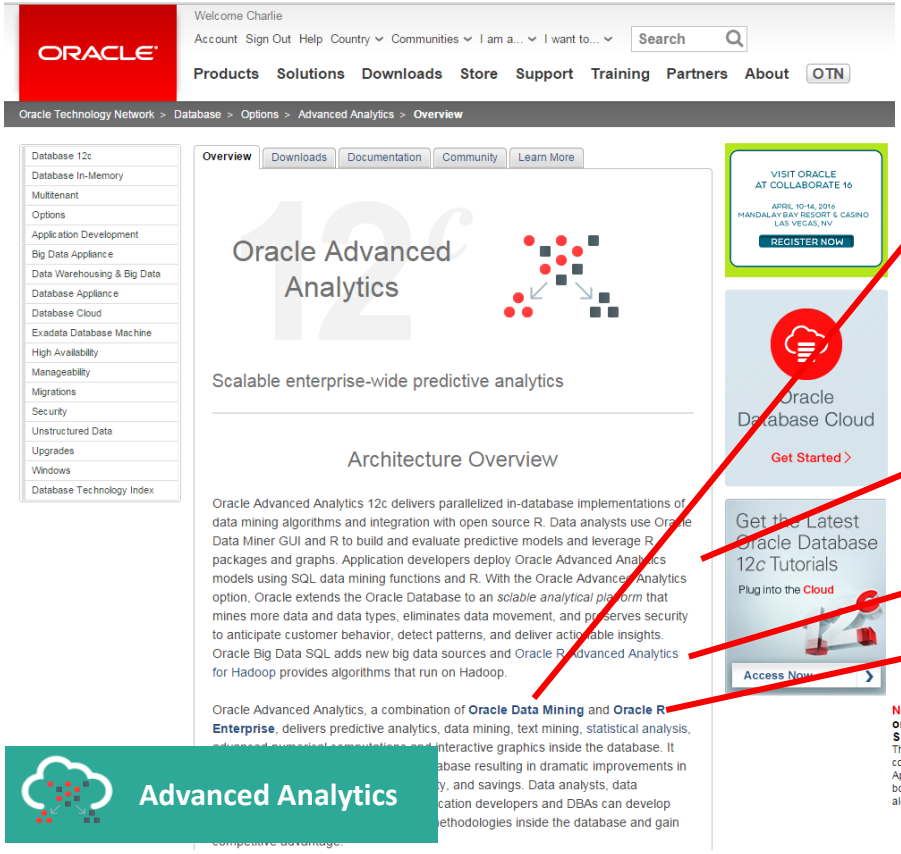


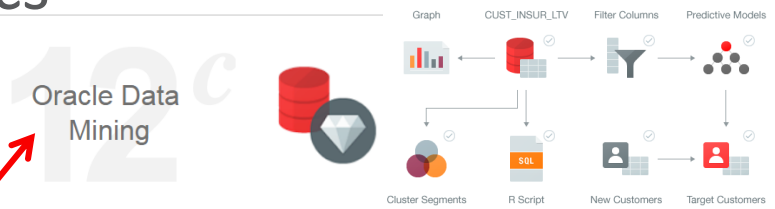


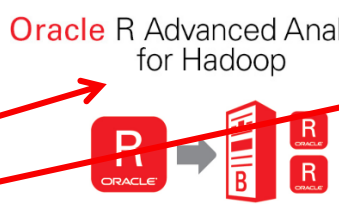
ORACLE

Copyright © 2016, Oracle and/or its affiliates. All rights reserved. |

Google “Oracle Advanced Analytics”





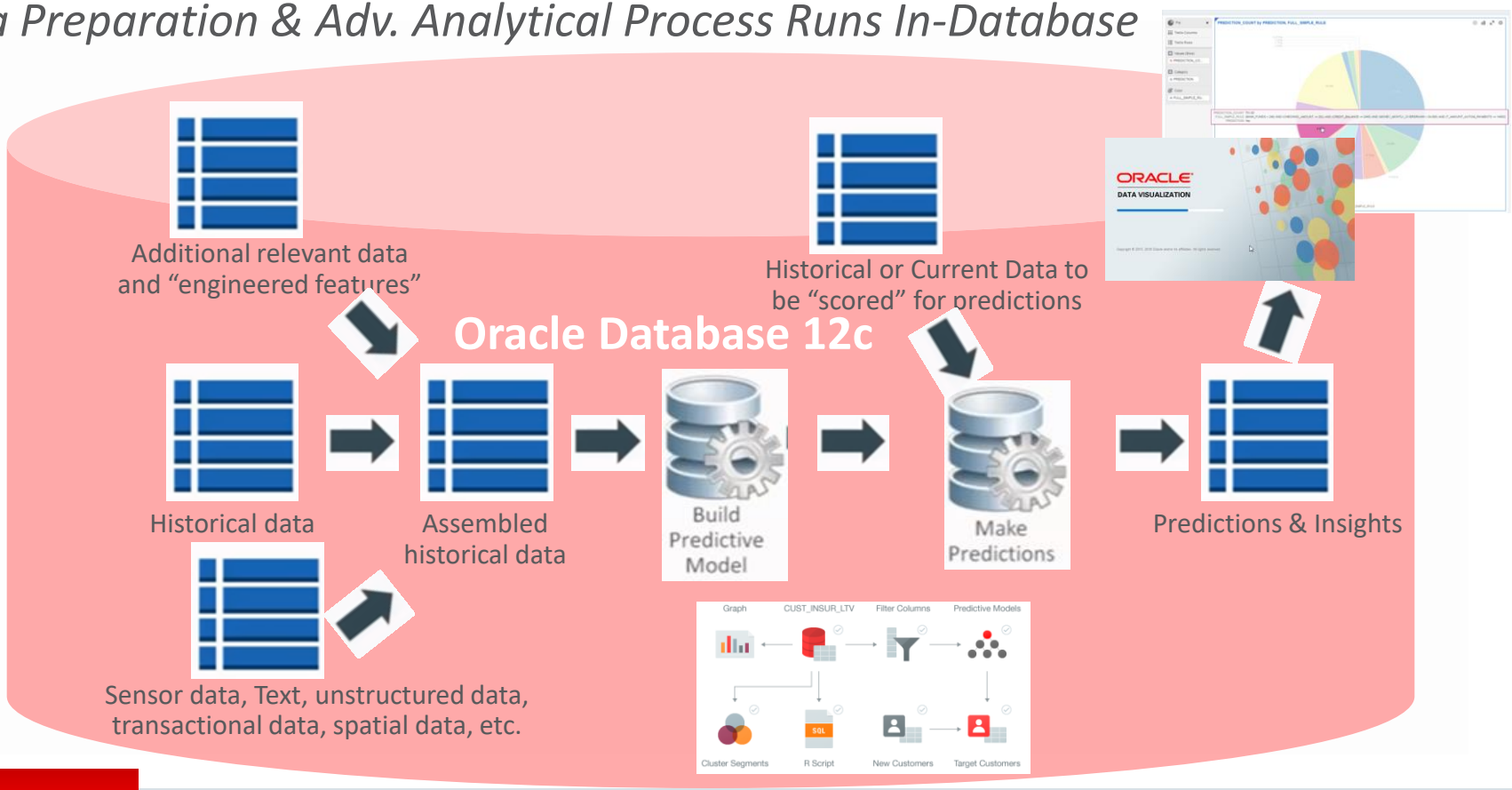


NEW ORAAH RELEASE 2.7.0: Introducing the fastest GLM and LM algorithms on Spark with full summary, enhanced Deep Neural Networks and support for Spark MLlib Gaussian Mixture Models.
The latest release of Oracle R Advanced Analytics for Hadoop (ORAAH), release 2.7.0, is one of the components of the Oracle Big Data Connectors software suite, an option to the Oracle Big Data Appliance. At its core, ORAAH provides an R interface for manipulating data stored in HDFS, using both HIVE transparency capabilities and mapping HDFS as direct input into Machine Learning algorithms that can run as Map Reduce jobs or inside an Apache Spark container.

Copyright © 2016, Oracle and/or its affiliates. All rights reserved. |

Oracle Data Mining/ Machine Learning/Predictive Analytics

Data Preparation & Adv. Analytical Process Runs In-Database



Copyright © 2016, Oracle and/or its affiliates. All rights reserved. | Confidential – Oracle Internal/Restricted/Highly Restricted

Oracle Advanced Analytics 12.2

Model Build Time Performance

Prelim/Unofficial

NEW IN 12.2

OAA 12.2 Algorithms	Rows (Ms)	T7-4 (Sparc & Solaris) Model Build Time (Secs / Degree of Parallelism)	X5-4 (Intel and Linux) Model Build Time (Secs / Degree of Parallelism)
Attributes Importance	640	28s / 512	44s / 72
K Means Clustering	640	161s / 256	268s / 144
Expectation Maximization	159	455s / 512	588s / 144
Naive Bayes Classification	320	17s / 256	23s / 72
GLM Classification	640	154s / 512	363s / 144
GLM Regression	640	55s / 512	93s / 144
Support Vector Machine (IPM solver)	640	404s / 512	1411s / 144
Support Vector Machine (SGD solver)	640	84s / 256	188s / 72

Wow! That's *Fast!*

ORACLE

The way to read their results is that they compare 2 chips: X5 (Intel and Linux) and T7 (Sparc and Solaris). They are measuring scalability (time in seconds) with increase degree of parallelism (dop). The data also has high cardinality categorical columns which translates in 9K mining attributes (when algorithms require explosion). There are no comparisons to 12.1 and it is fair to say that the 12.1 algorithms could not run on data of this size.

Copyright © 2016, Oracle and/or its affiliates. All rights reserved.

Oracle's Advanced Analytics and Machine Learning Platform

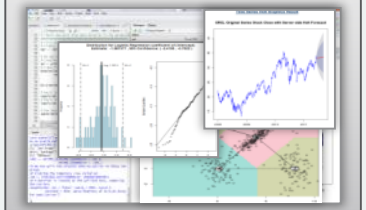
Multiple interfaces across platforms — SQL, R, GUI, Dashboards, Apps

Users

Information Producers

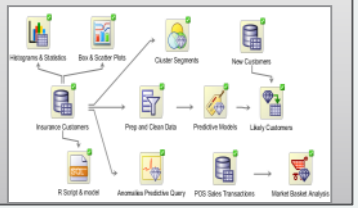
R programmers

R Client



Data & Business Analysts

SQLDEV/ Oracle Data Miner



Information Consumers

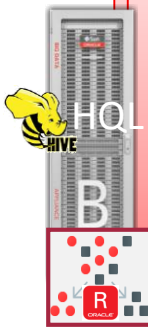
Business Analysts/Mgrs

OBIEE/DV

Domain End Users (HCM, CRM)

Applications

Platform



Hadoop

ORAAH

Parallel, distributed algorithms

Oracle Database Enterprise Edition

Oracle Advanced Analytics - Database Option

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

Oracle Database 12c

Oracle Cloud

ORACLE

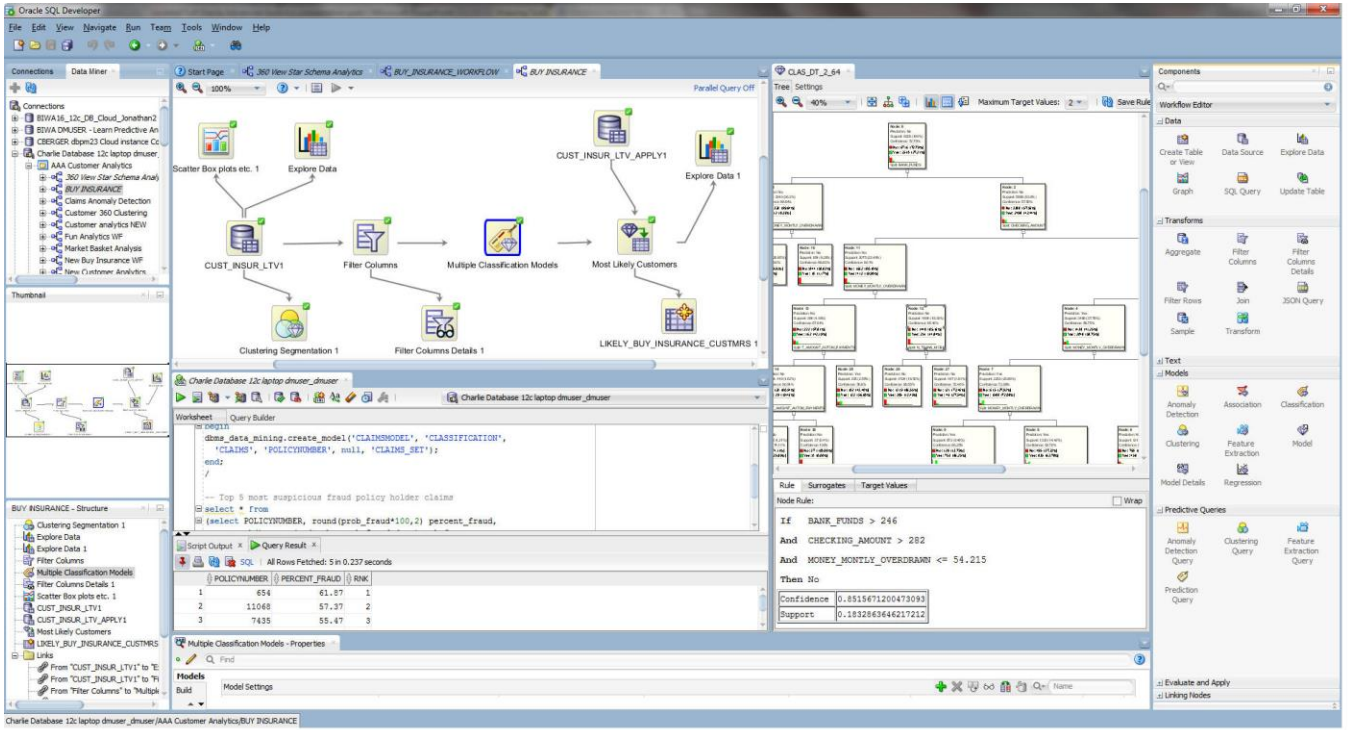
Copyright © 2017, Oracle and/or its affiliates. All rights reserved. |

Oracle Data Miner GUI

Easy to Use for "Citizen Data Scientist"

Advanced Analytics

- Easy to use to define analytical methodologies that can be shared
- SQL Developer Extension
- Workflow API and generates SQL code for immediate deployment



Copyright © 2016 Oracle and/or its affiliates. All rights reserved. |

Take off!—Intermediate/Experts

Quick Set up Overview

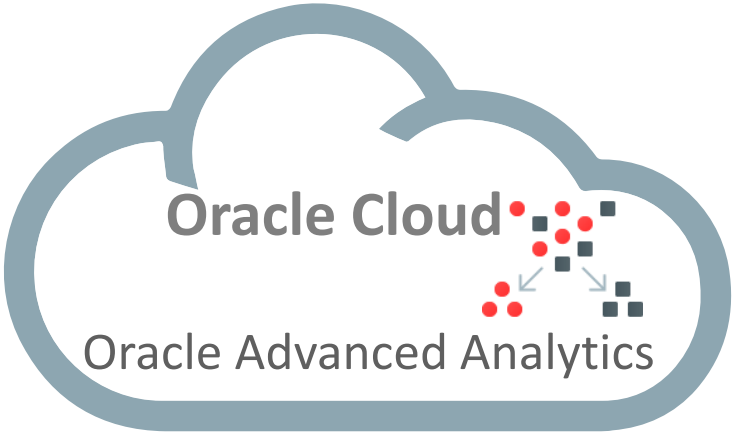
ORACLE

Copyright © 2016 Oracle and/or its affiliates. All rights reserved. |

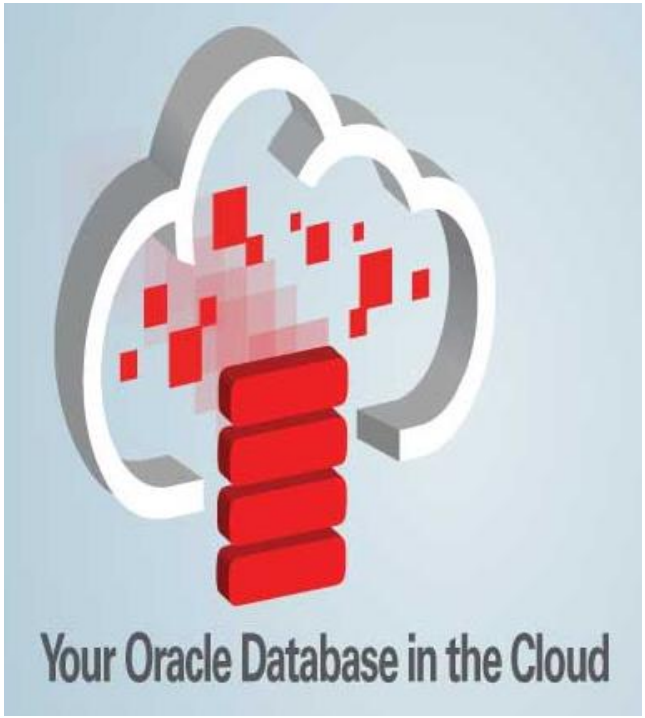
Learn Predictive Analytics in 2 Hours!

Oracle Advanced Analytics Hands on Lab

- We're using the Oracle Database Cloud!



=

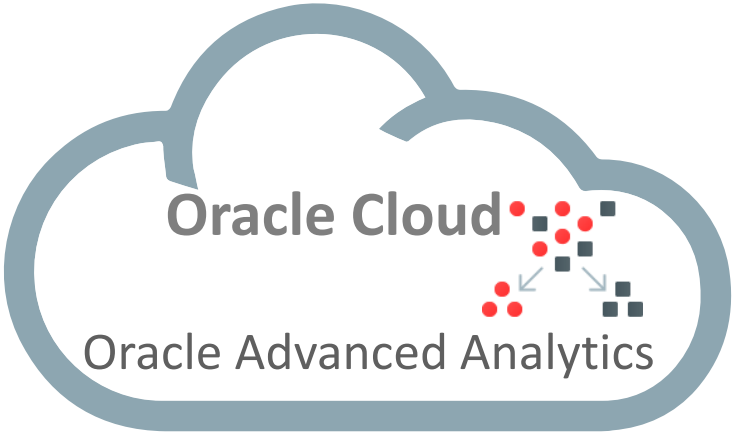


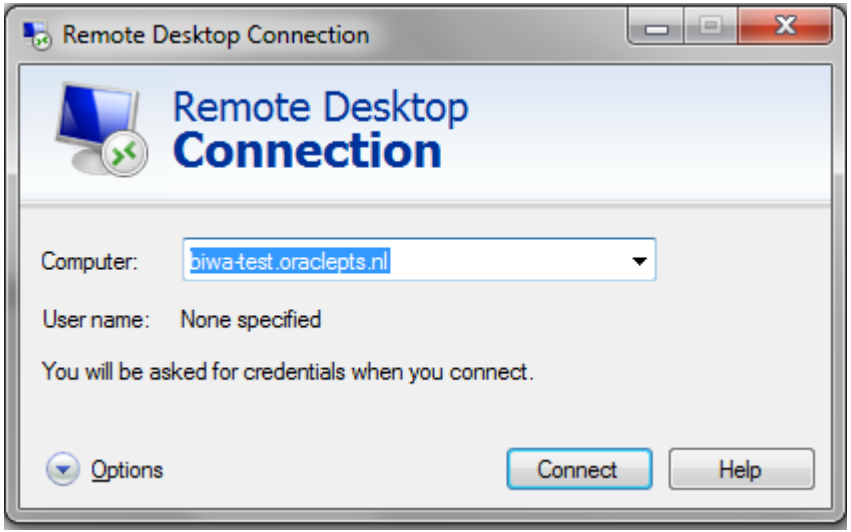
Copyright © 2016 Oracle and/or its affiliates. All rights reserved. |

Learn Predictive Analytics in 2 Hours!

Oracle Advanced Analytics Hands on Lab

- We're using the Oracle Database Cloud!

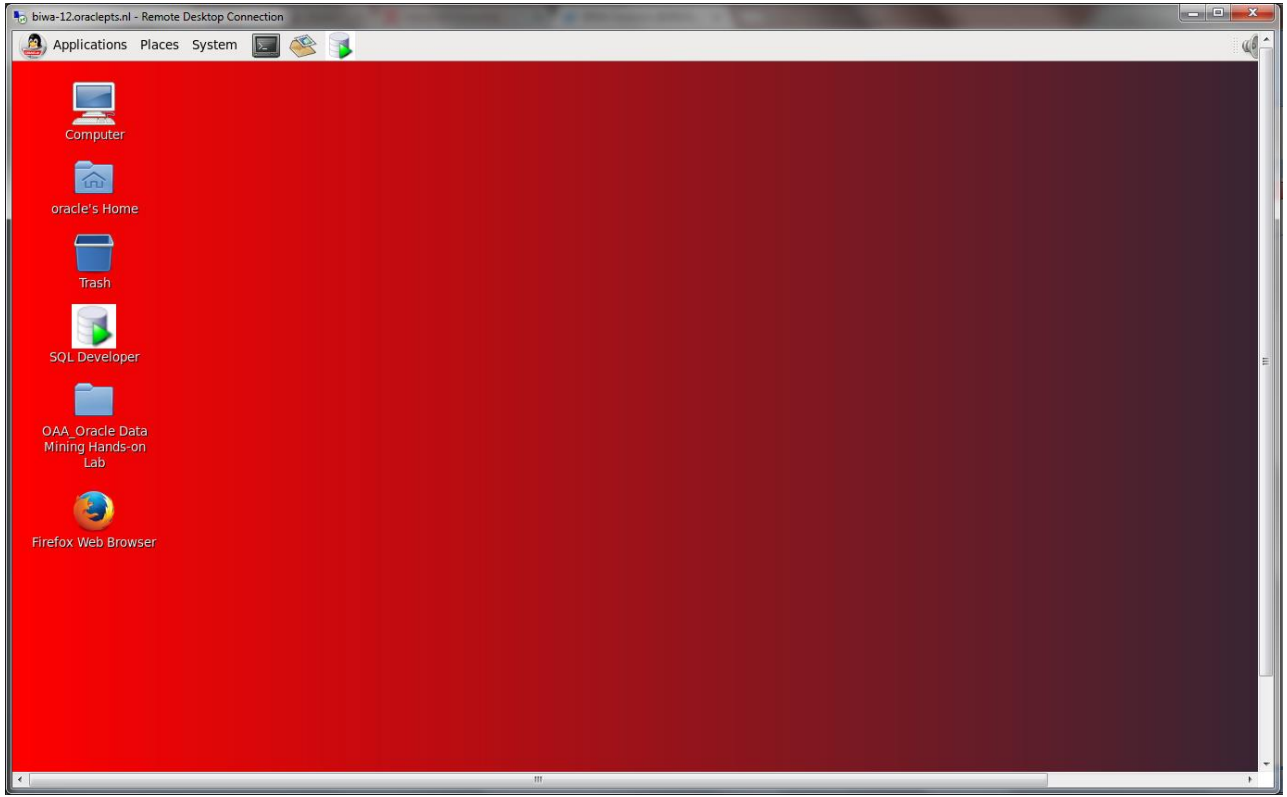




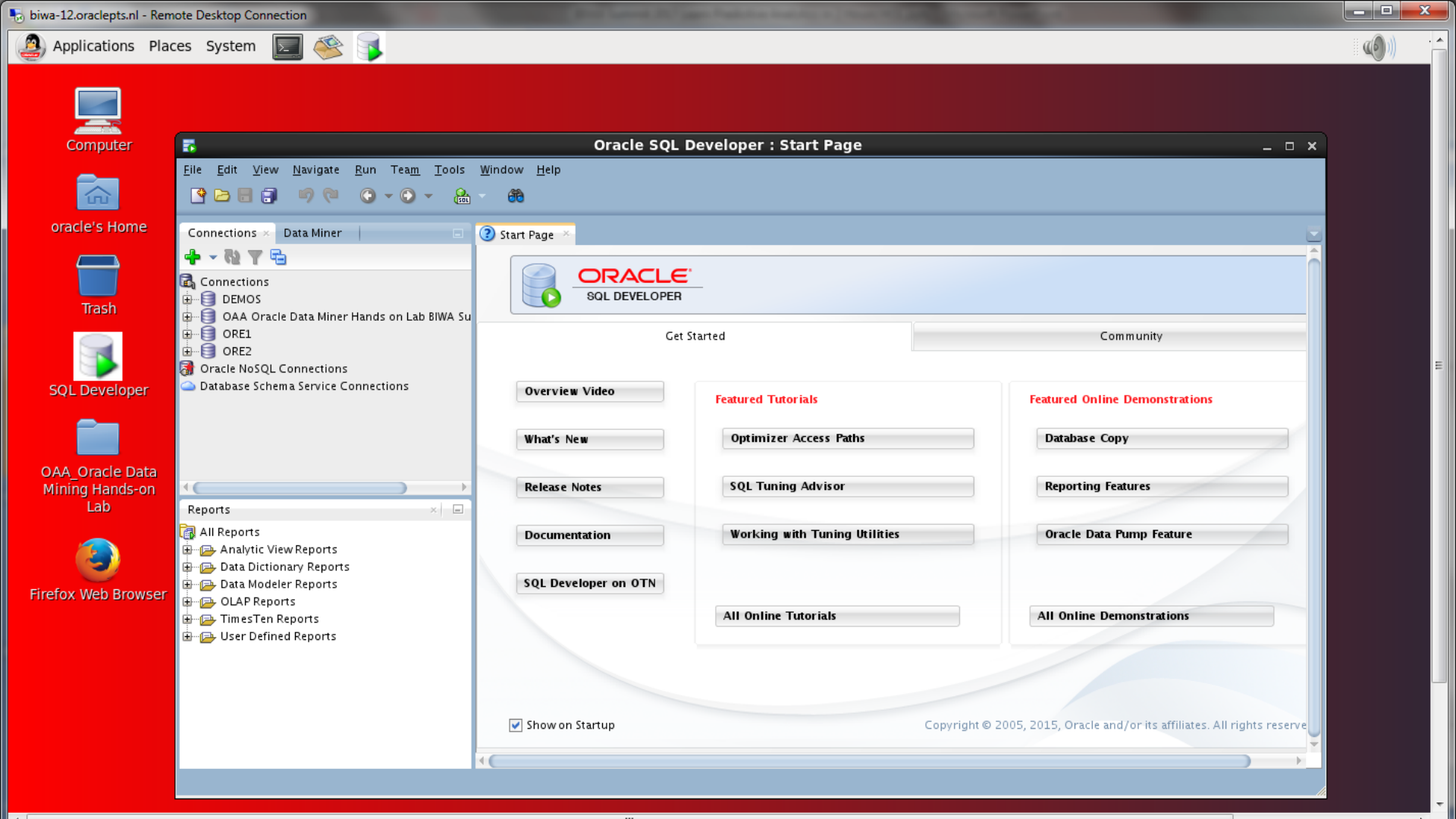
Copyright © 2016 Oracle and/or its affiliates. All rights reserved. |

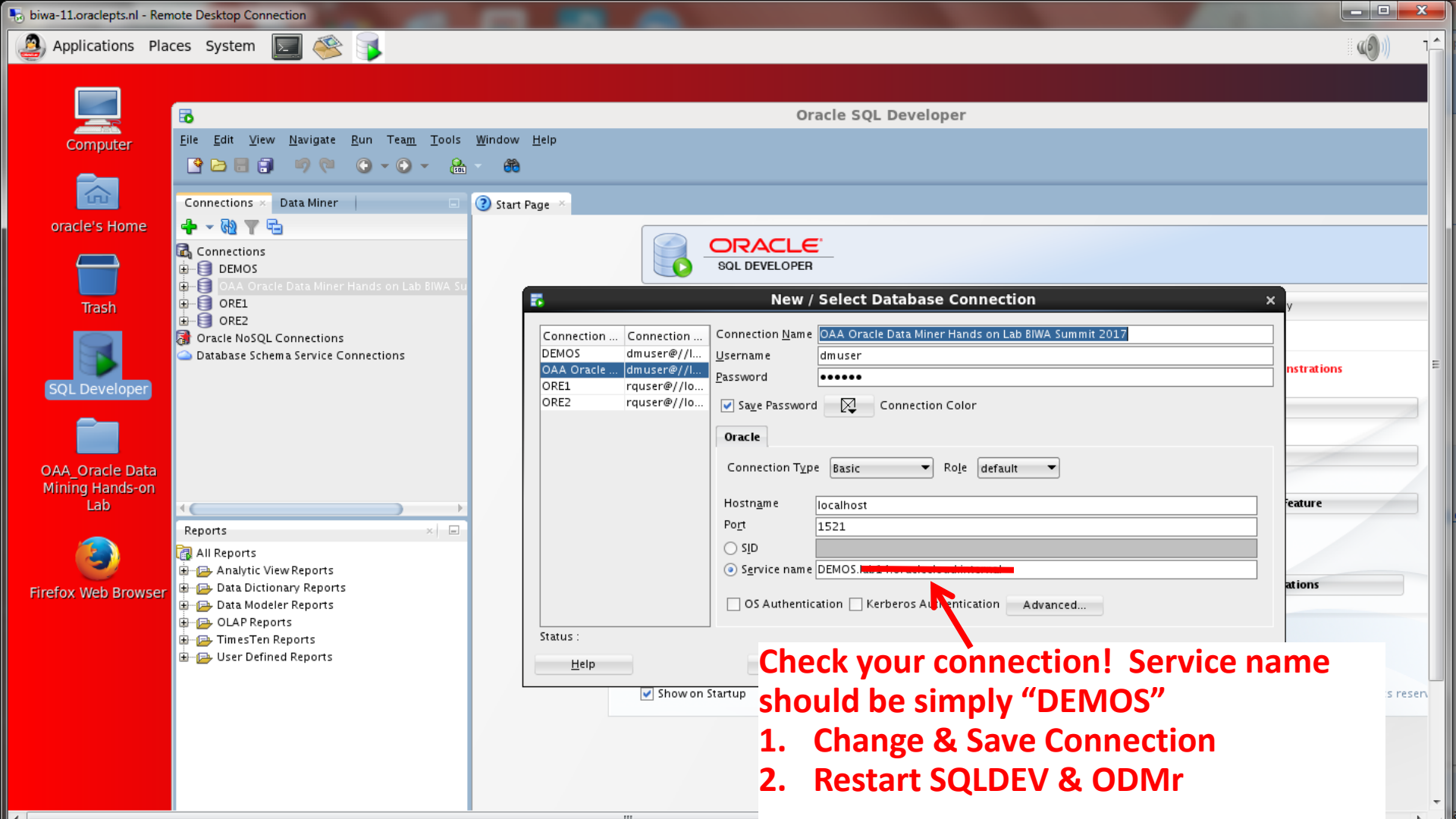
Learn Predictive Analytics in 2 Hours!

Oracle Advanced Analytics Hands on Lab



Copyright © 2016 Oracle and/or its affiliates. All rights reserved. |

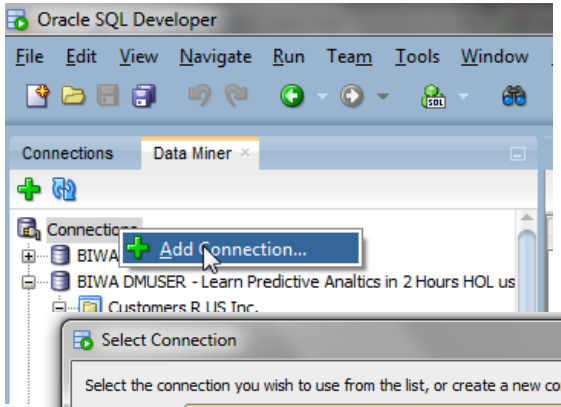

biwa-12.oraclepts.nl - Remote Desktop Connection
Applications
Places
System
Computer
oracle's Home
Trash
SQL Developer
OAA_Oracle Data Mining Hands-on Lab
Firefox Web Browser
Oracle SQL Developer : Start Page
File
Edit
View
Navigate
Run
Team
Tools
Window
Help
Connections
Data Miner
Connections
DEMOS
OAA Oracle Data Miner Hands on Lab BIWA Su
ORE1
ORE2
Oracle NoSQL Connections
Database Schema Service Connections
Reports
All Reports
Analytic View Reports
Data Dictionary Reports
Data Modeler Reports
OLAP Reports
TimesTen Reports
User Defined Reports
Start Page
ORACLE
SQL DEVELOPER
Get Started
Community
Overview Video
What's New
Release Notes
Documentation
SQL Developer on OTN
Featured Tutorials
Optimizer Access Paths
SQL Tuning Advisor
Working with Tuning Utilities
All Online Tutorials
Featured Online Demonstrations
Database Copy
Reporting Features
Oracle Data Pump Feature
All Online Demonstrations
Show on Startup
Copyright © 2005, 2015, Oracle and/or its affiliates. All rights reserve

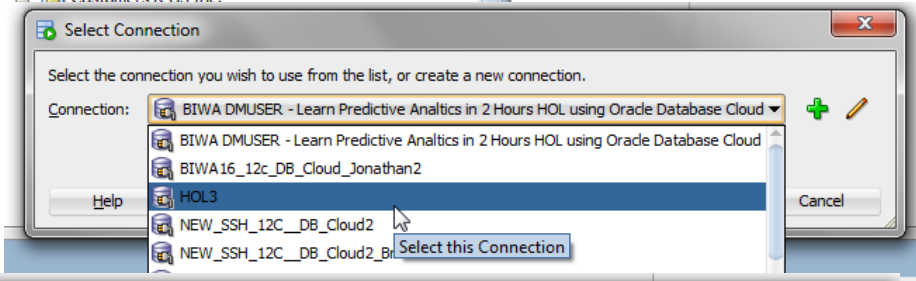

biwa-11.oraclepts.nl - Remote Desktop Connection
Applications Places System
Computer
oracle's Home
Trash
SQL Developer
OAA_Oracle Data Mining Hands-on Lab
Firefox Web Browser
Oracle SQL Developer
File Edit View Navigate Run Team Tools Window Help
Connections Data Miner
Start Page
Connections
DEMOS
OAA Oracle Data Miner Hands on Lab BIWA Su
ORE1
ORE2
Oracle NoSQL Connections
Database Schema Service Connections
Reports
All Reports
Analytic View Reports
Data Dictionary Reports
Data Modeler Reports
OLAP Reports
TimesTen Reports
User Defined Reports
ORACLE SQL DEVELOPER
New / Select Database Connection
Connection ... Connection ...
DEMOS dmuser@//l...
OAA Oracle ... dmuser@//l...
ORE1 rquser@//lo...
ORE2 rquser@//lo...
Connection Name OAA Oracle Data Miner Hands on Lab BIWA Summit 2017
Username dmuser
Password ••••••
Save Password
Connection Color
Oracle
Connection Type Basic
Role default
Hostname localhost
Port 1521
SID
Service name DEMOS
OS Authentication
Kerberos Authentication
Advanced...
Status :
Help
Show on Startup
Check your connection! Service name should be simply "DEMOS"
1. Change & Save Connection
2. Restart SQLDEV & ODMr

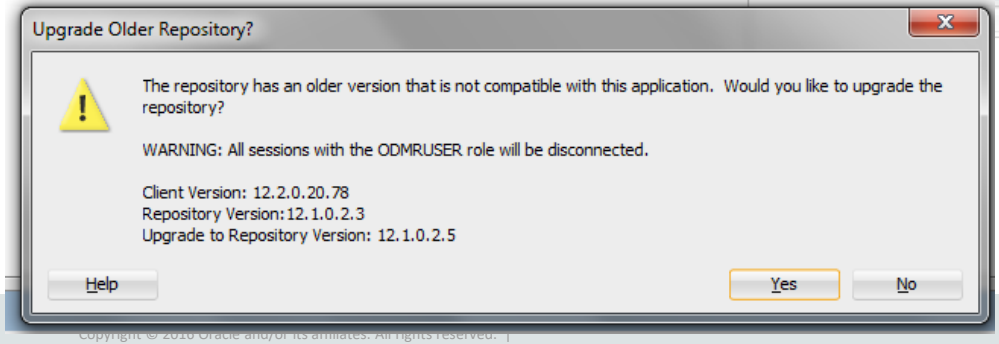
Learn Predictive Analytics in 2 Hours!

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!



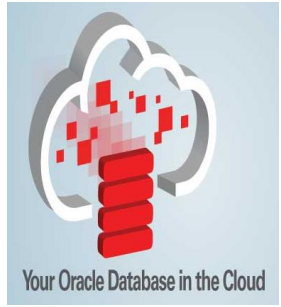




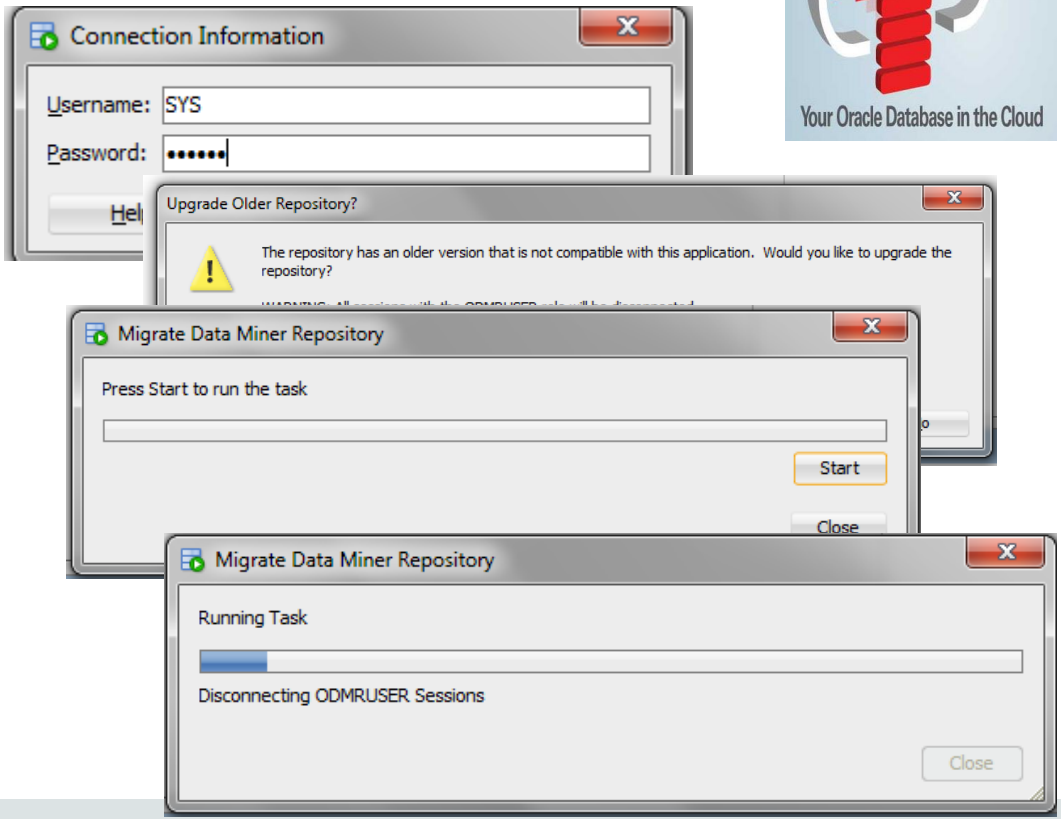
Copyright © 2016 Oracle and/or its affiliates. All rights reserved.

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)

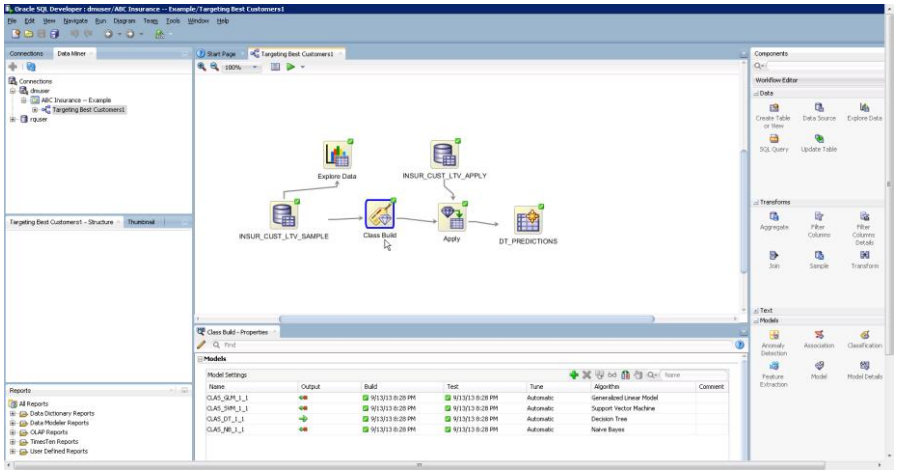


Copyright © 2016 Oracle and/or its affiliates. All rights reserved. |

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*

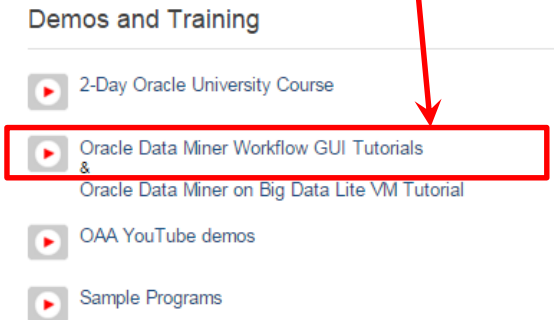


Copyright © 2016 Oracle and/or its affiliates. All rights reserved. |

OAA/Oracle Data Miner 4.2 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

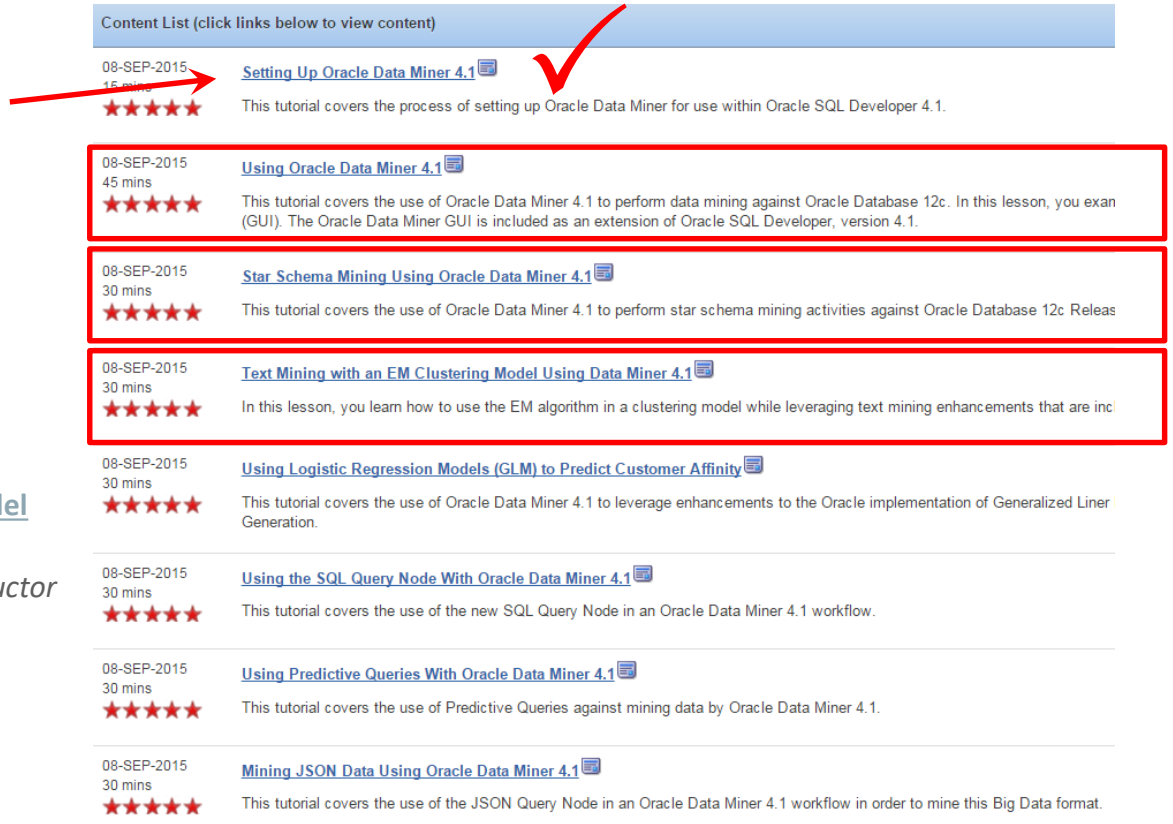


Copyright © 2016 Oracle and/or its affiliates. All rights reserved. |

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*

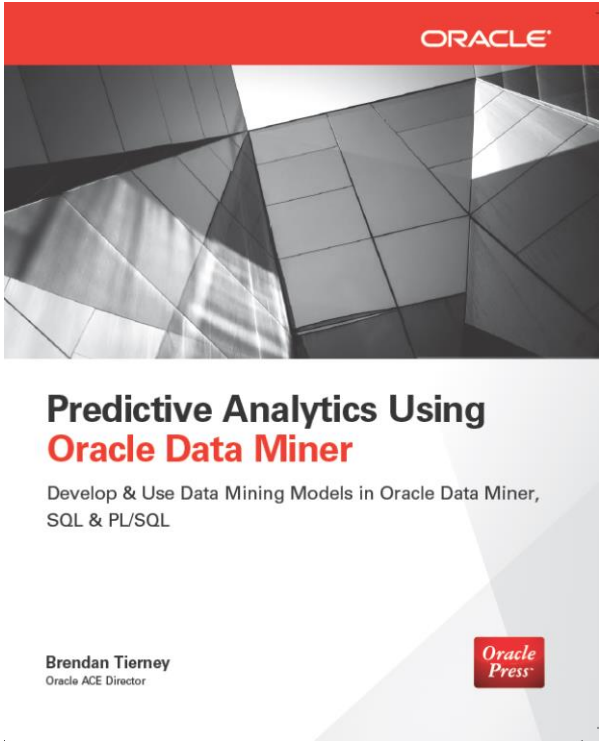


Copyright © 2016 Oracle and/or its affiliates. All rights reserved. |

Great book on Oracle Advanced Analytics

Available on Amazon or from Author

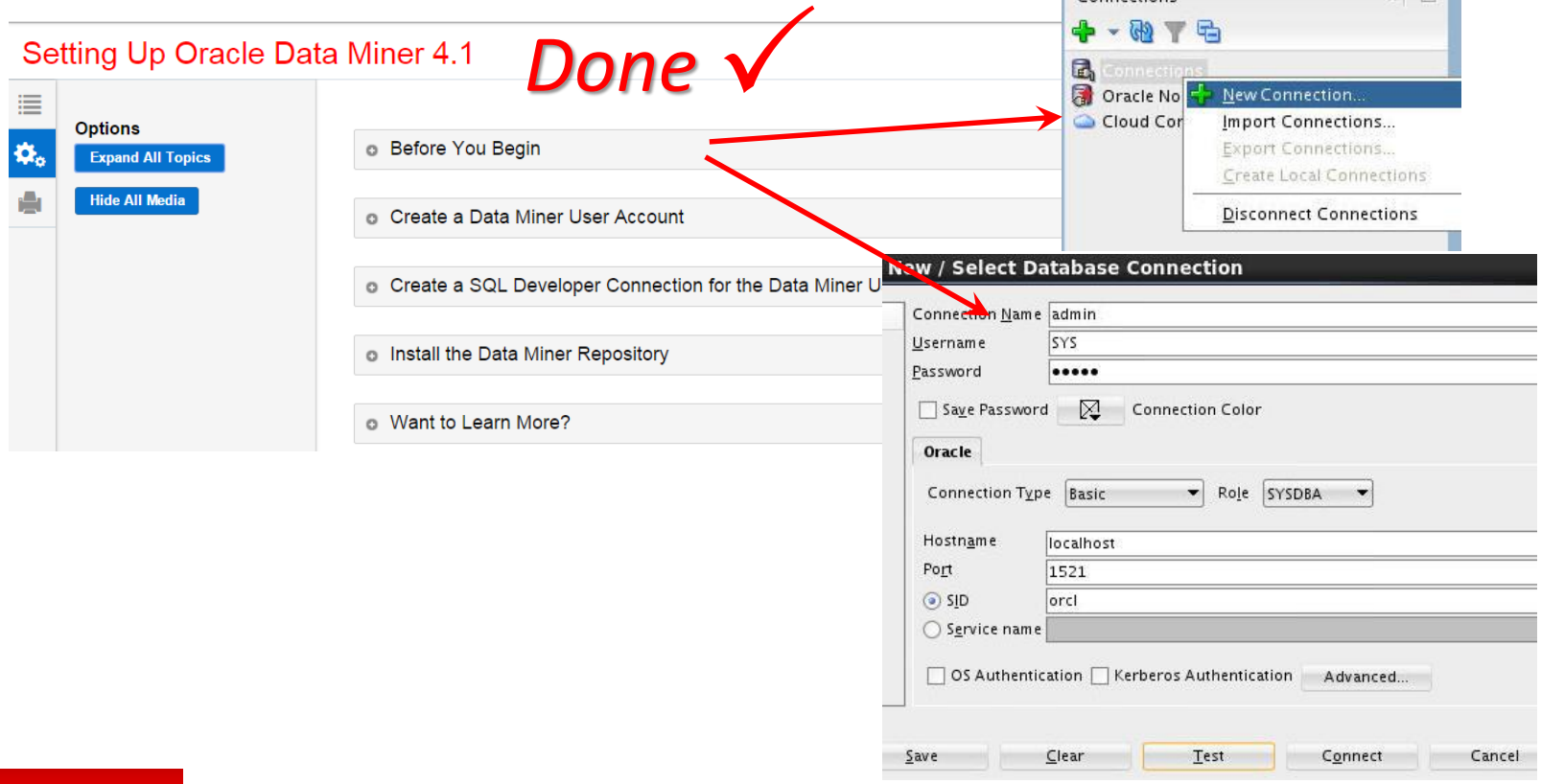
- Predictive Analytics Using Oracle Data Miner: Develop for ODM in SQL & PL/SQL



Copyright © 2016 Oracle and/or its affiliates. All rights reserved. |

OAA/Oracle Data Miner 4.1 HOL

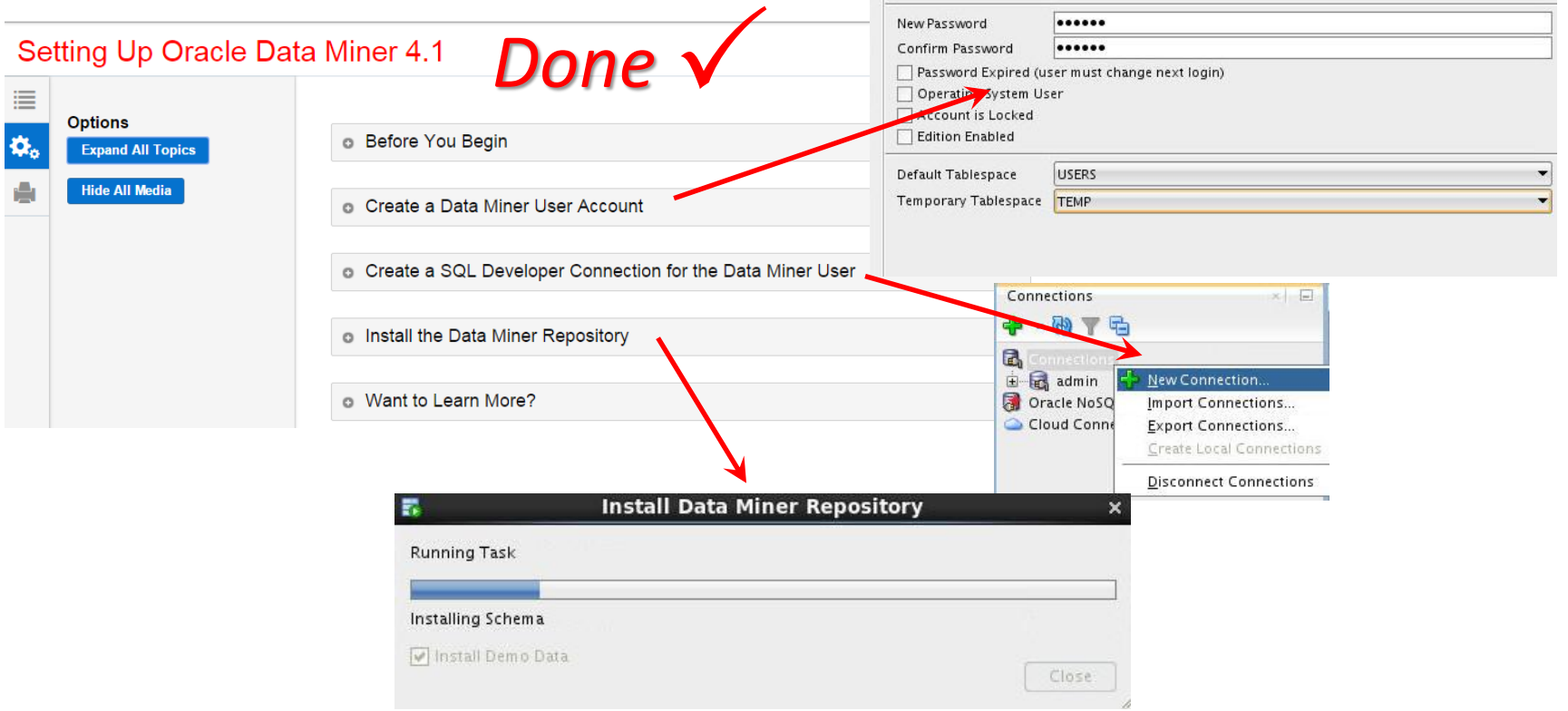
Setting Up Oracle Data Miner



Copyright © 2016 Oracle and/or its affiliates. All rights reserved. |

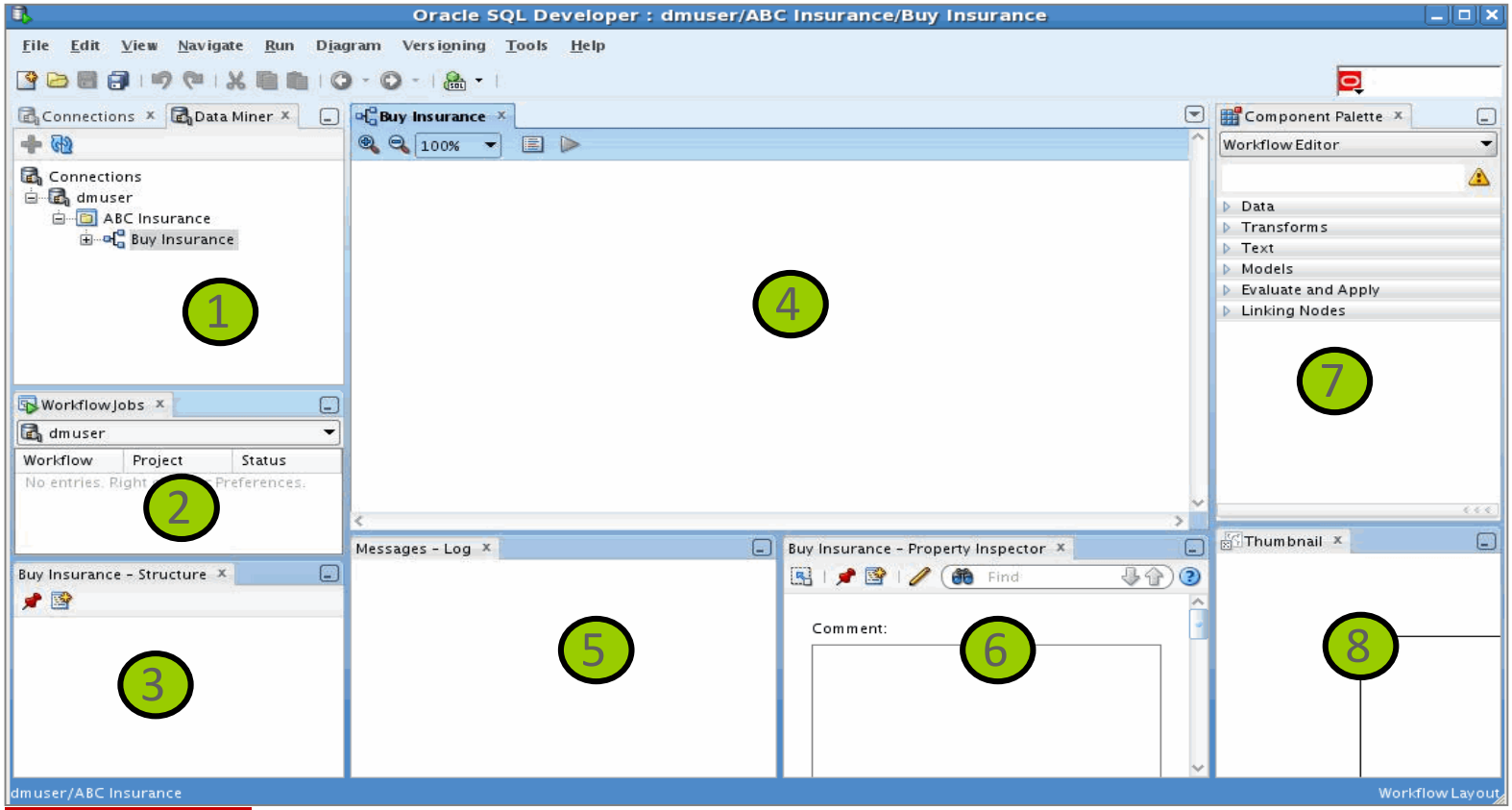
OAA/Oracle Data Miner 4.1 HOL

Setting Up Oracle Data Miner



Copyright © 2016 Oracle and/or its affiliates. All rights reserved. |

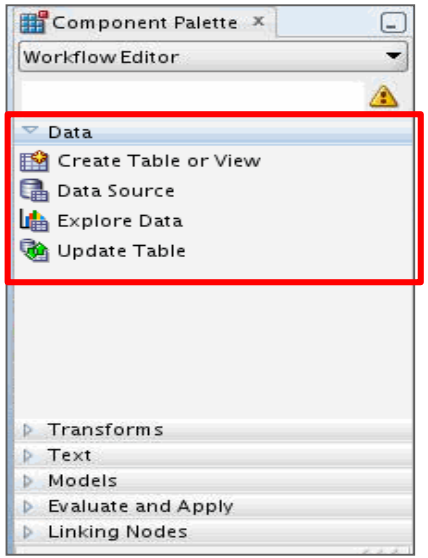
Introducing the Data Miner Interface



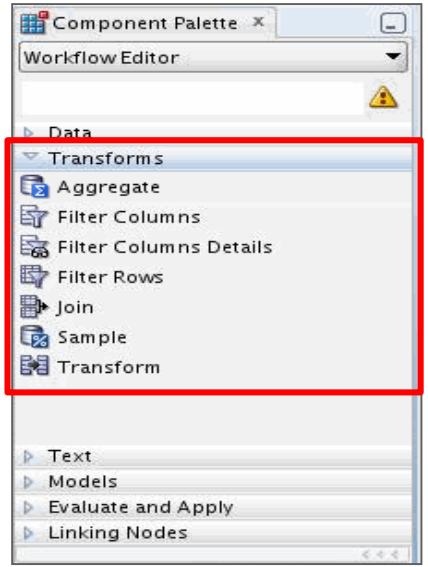
Copyright © 2016 Oracle and/or its affiliates. All rights reserved. |

Examining Oracle Data Miner Nodes

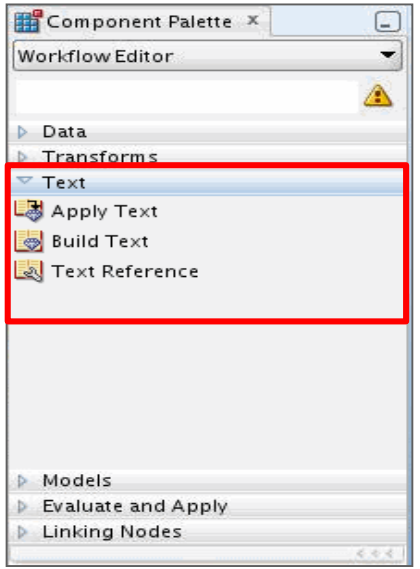
Data



Transforms



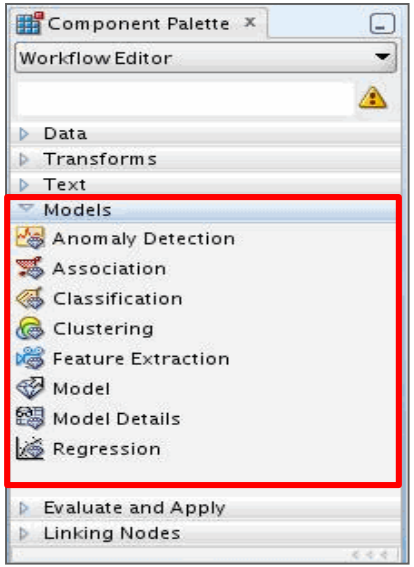
Text



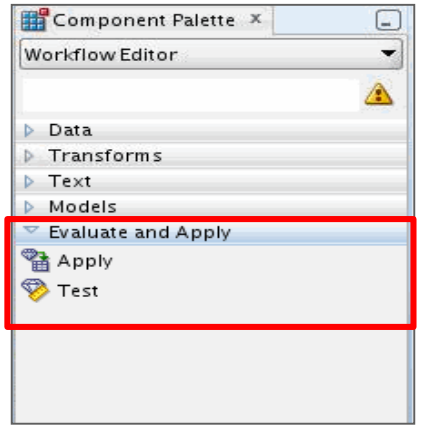
Copyright © 2016 Oracle and/or its affiliates. All rights reserved. |

Examining Oracle Data Miner Nodes

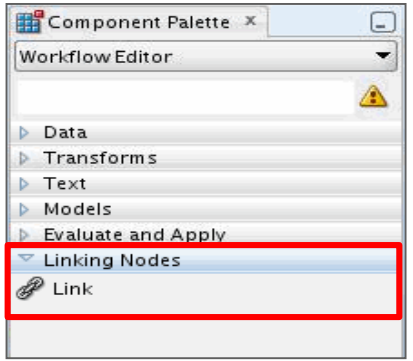
Models



Evaluate and Apply

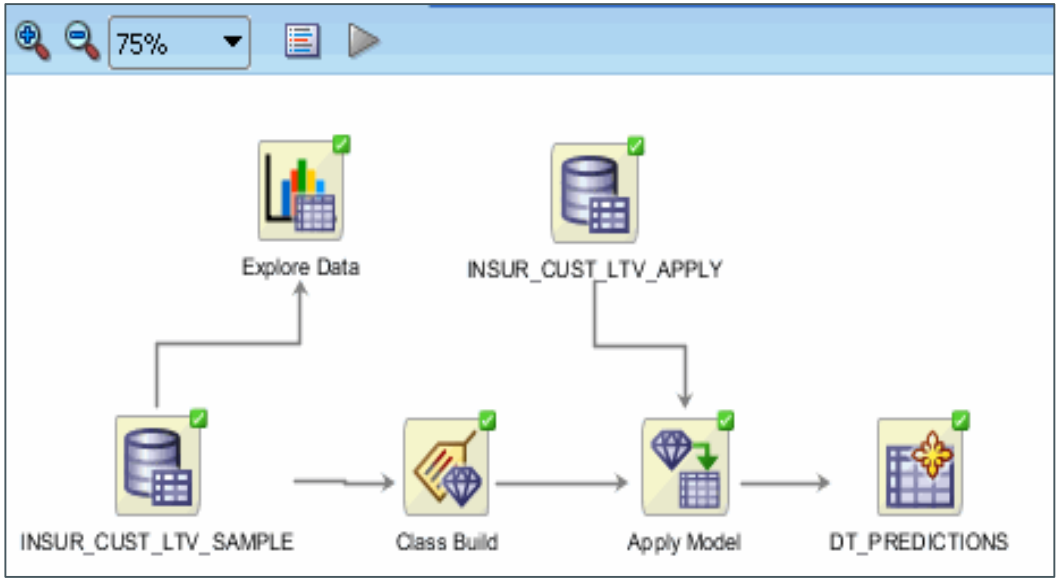


Linking



Copyright © 2016 Oracle and/or its affiliates. All rights reserved. |

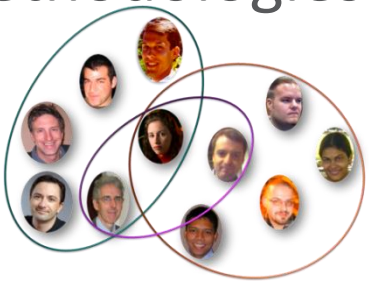
Previewing a Data Miner Workflow



Copyright © 2016 Oracle and/or its affiliates. All rights reserved. |

Rapidly Build, Evaluate & Deploy Analytical Methodologies

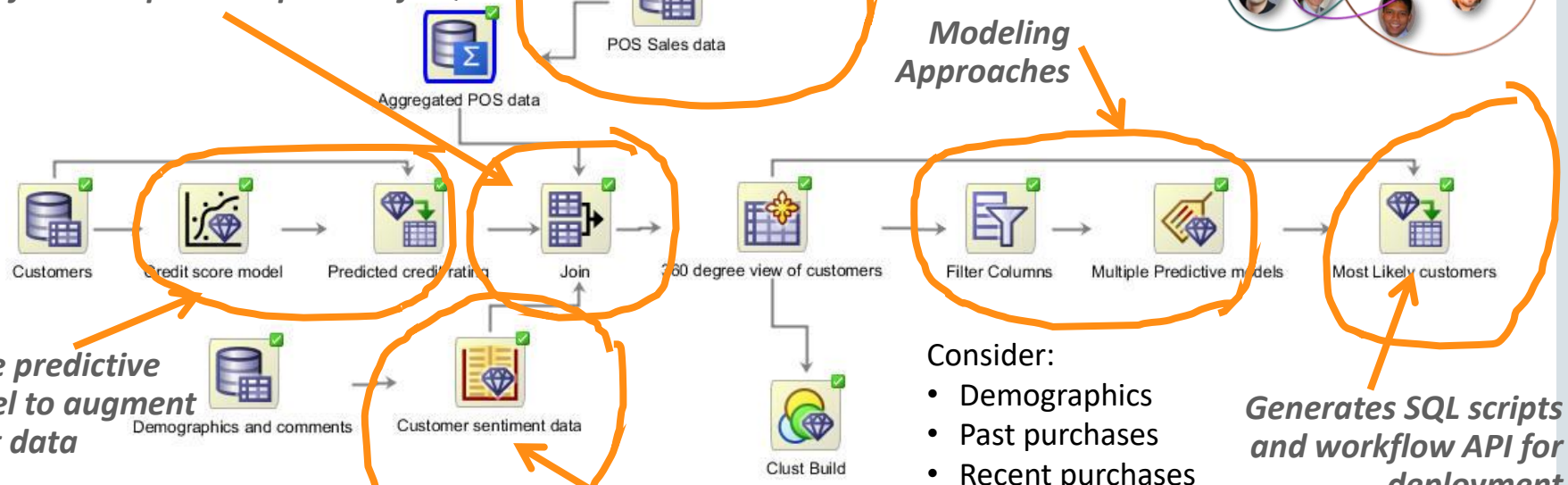
Leveraging a Variety of Data Sources and Types



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

Transactional POS data

Modeling Approaches



Inline predictive model to augment input data

Consider:

- Demographics
- Past purchases
- Recent purchases
- Comments & tweets

Generates SQL scripts and workflow API for deployment

Unstructured data also mined by algorithms

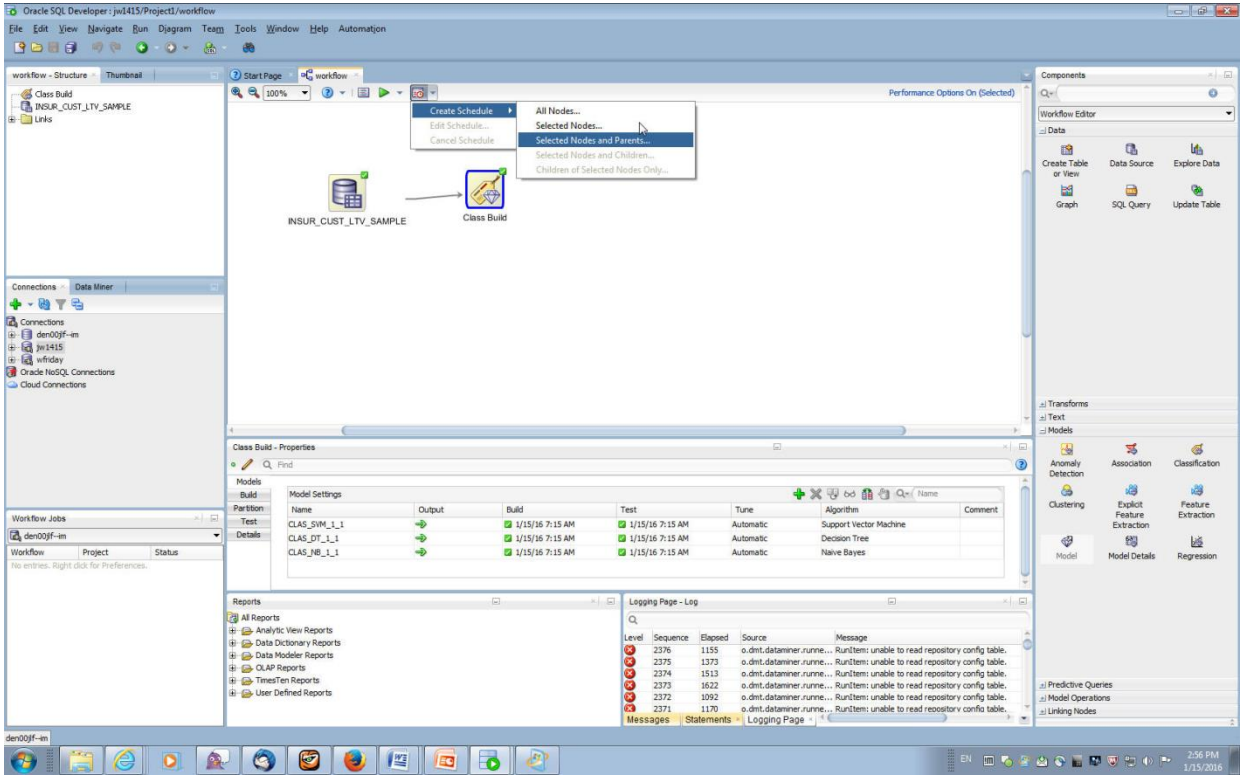
Advanced Analytics

ORACLE

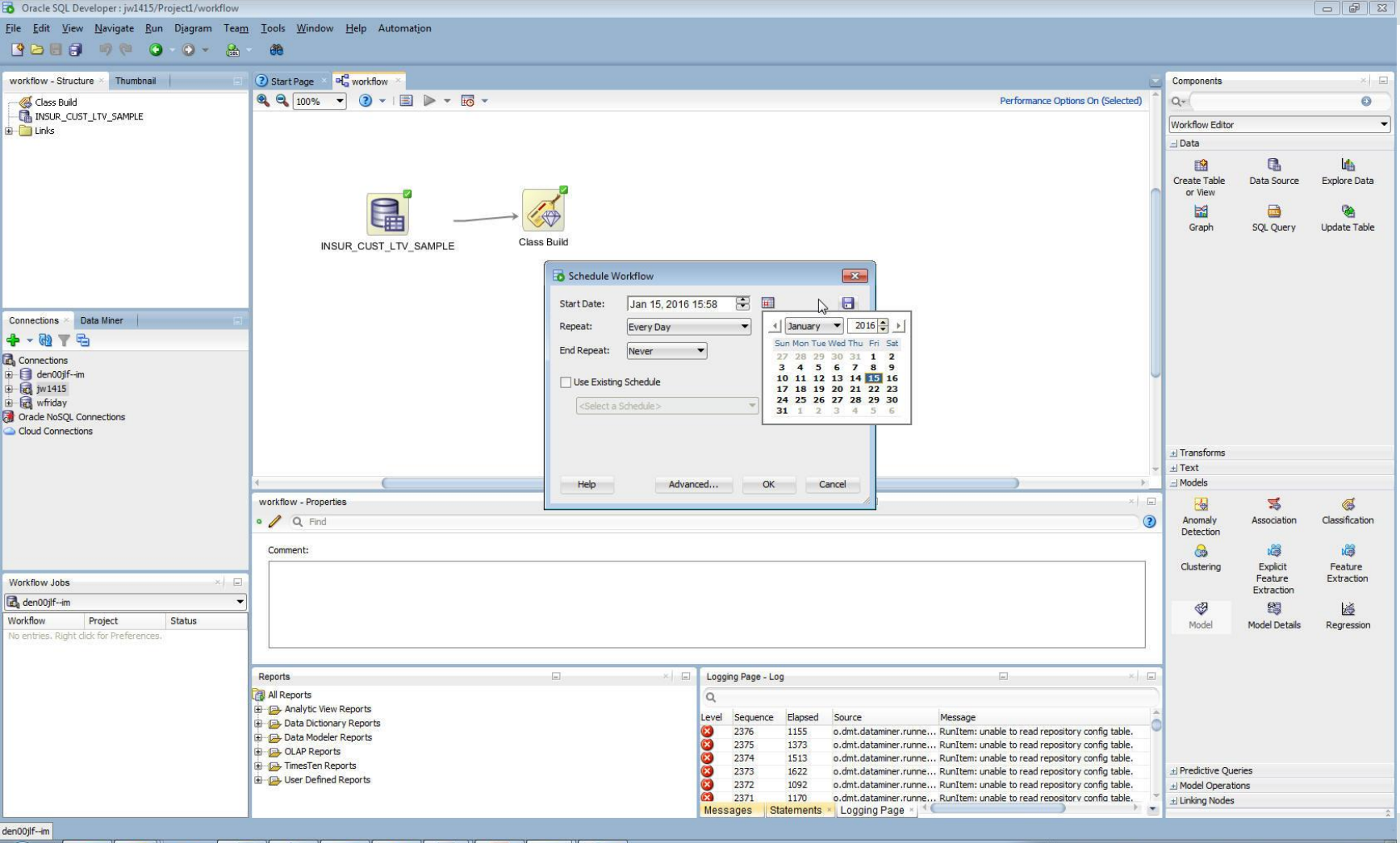
Copyright © 2016, Oracle and/or its affiliates. All rights reserved. |

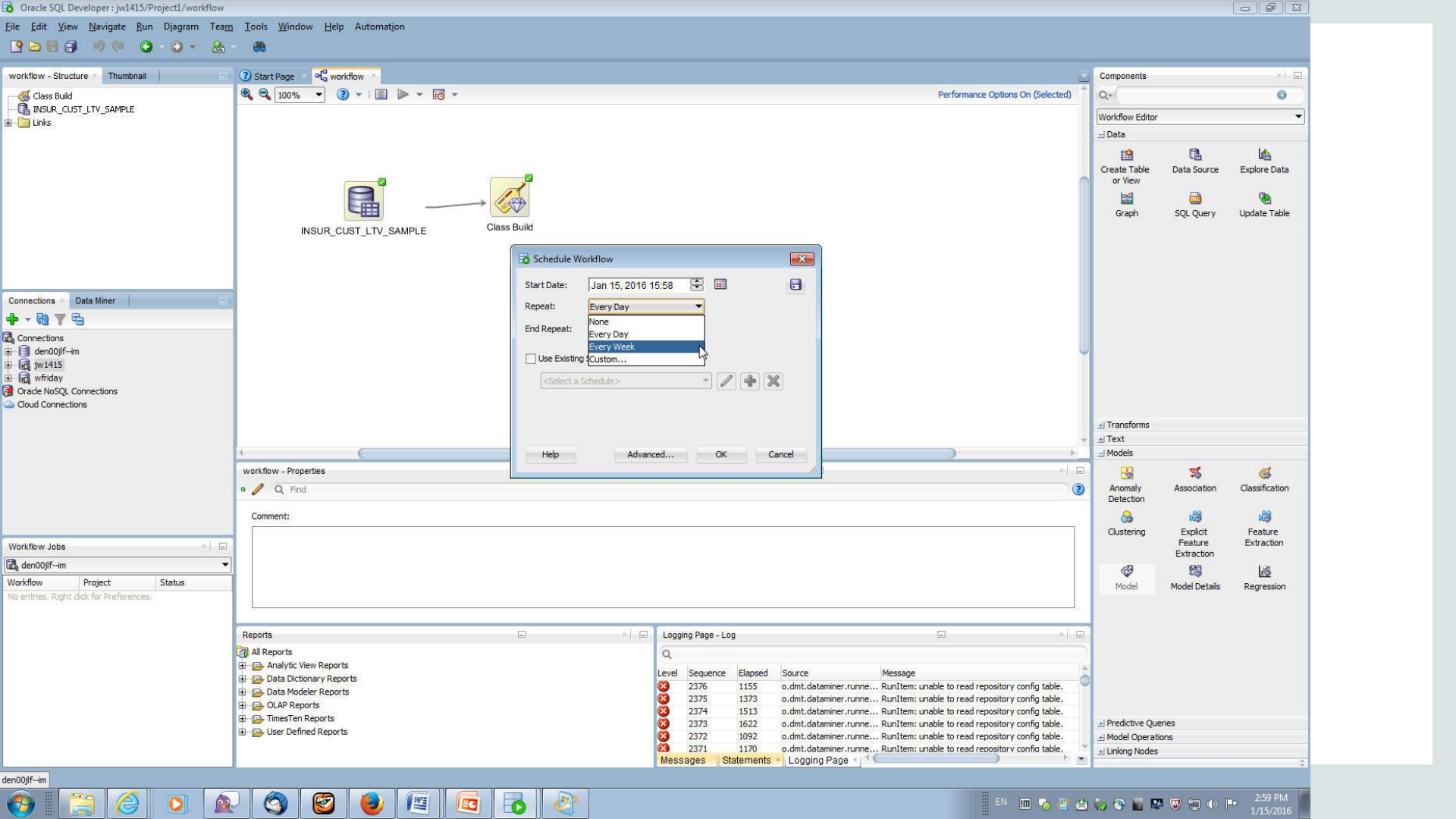
Previewing a 4.2 Feature

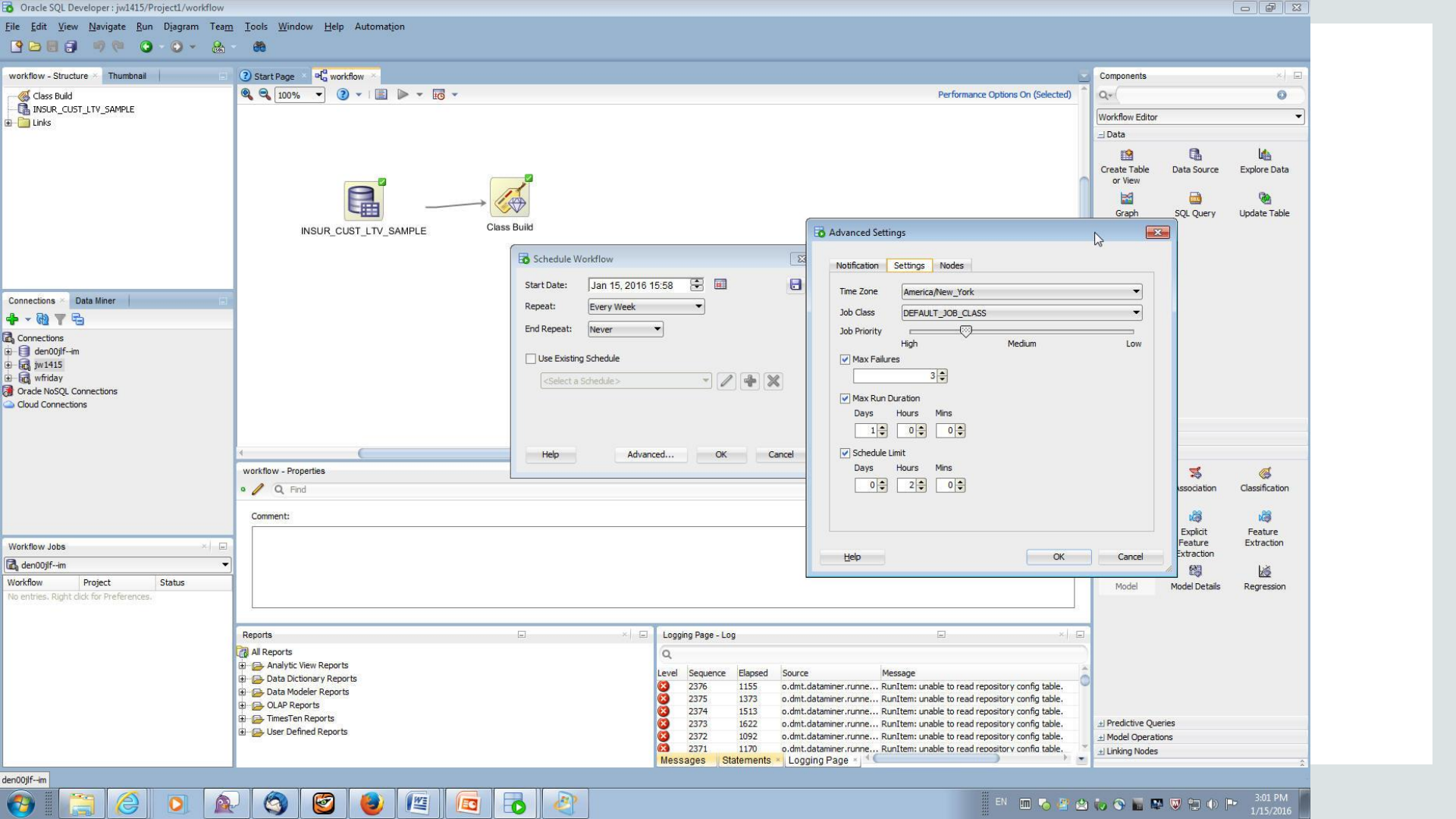
Workflow Scheduler

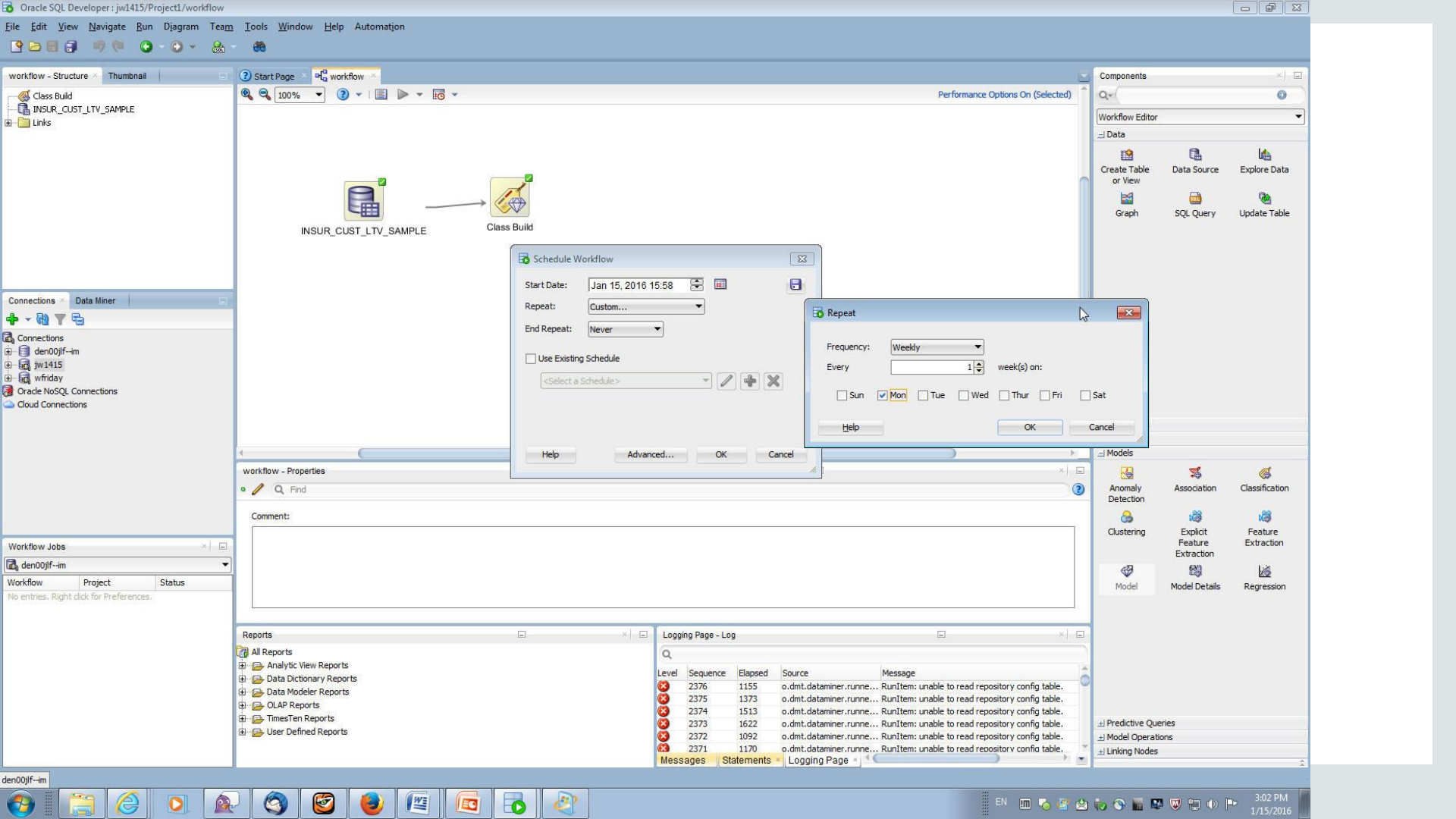


Copyright © 2016 Oracle and/or its affiliates. All rights reserved. |


Oracle SQL Developer : jw1415/Project1/workflow
File Edit View Navigate Run Diagram Team Tools Window Help Automation
workflow - Structure
Thumbnail
Class Build
INSUR_CUST_LTV_SAMPLE
Links
Start Page
workflow
100%
Performance Options On (Selected)
INSUR_CUST_LTV_SAMPLE
Class Build
Schedule Workflow
Start Date: Jan 15, 2016 15:58
Repeat: Every Day
End Repeat: Never
Use Existing Schedule
<Select a Schedule>
January
2016
Sun Mon Tue Wed Thu Fri Sat
Help
Advanced...
OK
Cancel
Components
Workflow Editor
Data
Create Table or View
Data Source
Explore Data
Graph
SQL Query
Update Table
Transforms
Text
Models
Anomaly Detection
Association
Classification
Clustering
Explicit Feature Extraction
Feature Extraction
Model
Model Details
Regression
Predictive Queries
Model Operations
Linking Nodes
Connections
Data Miner
Connections
den00jlf--im
jw1415
wfriday
Oracle NoSQL Connections
Cloud Connections
Workflow Jobs
den00jlf--im
Workflow
Project
Status
No entries. Right click for Preferences.
workflow - Properties
Find
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
den00jlf--im
EN
2:59 PM
1/15/2016

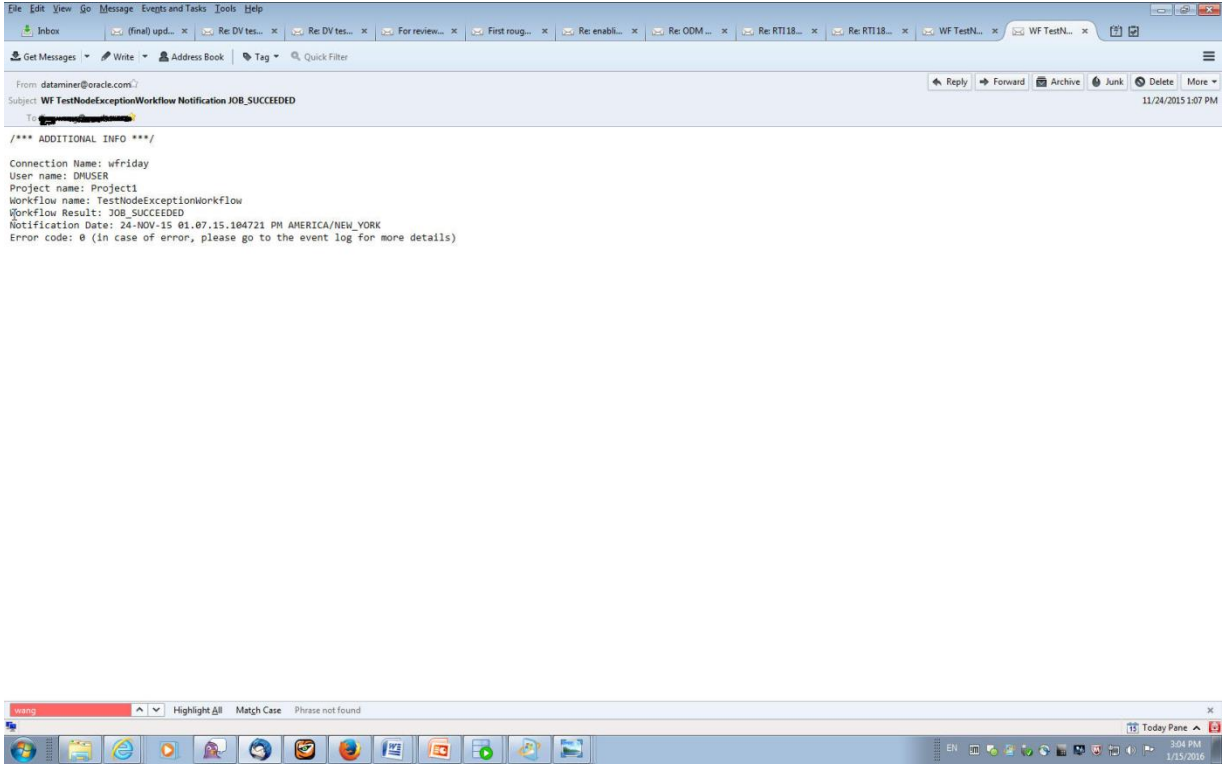

Oracle SQL Developer : jw1415/Project1/workflow
File Edit View Navigate Run Diagram Team Tools Window Help Automation
workflow - Structure
Thumbnail
Class Build
INSUR_CUST_LTV_SAMPLE
Links
Start Page
workflow
100%
Performance Options On (Selected)
INSUR_CUST_LTV_SAMPLE
Class Build
Schedule Workflow
Start Date:
Jan 15, 2016 15:58
Repeat:
Every Day
End Repeat:
None
Every Day
Every Week
Custom...
Use Existing
<Select a Schedule>
Help
Advanced...
OK
Cancel
Components
Workflow Editor
Data
Create Table or View
Data Source
Explore Data
Graph
SQL Query
Update Table
Transforms
Text
Models
Anomaly Detection
Association
Classification
Clustering
Explicit Feature Extraction
Feature Extraction
Model
Model Details
Regression
Predictive Queries
Model Operations
Linking Nodes
Connections
Data Miner
Connections
den00jlf--im
jw1415
wfriday
Oracle NoSQL Connections
Cloud Connections
Workflow Jobs
den00jlf--im
Workflow
Project
Status
No entries. Right click for Preferences.
workflow - Properties
Find
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
den00jlf--im
EN
2:59 PM
1/15/2016


Oracle SQL Developer : jw1415/Project1/workflow
File Edit View Navigate Run Diagram Team Tools Window Help Automation
workflow - Structure
Thumbnail
Class Build
INSUR_CUST_LTV_SAMPLE
Links
Start Page
workflow
100%
Performance Options On (Selected)
INSUR_CUST_LTV_SAMPLE
Class Build
Components
Workflow Editor
Data
Create Table or View
Data Source
Explore Data
Graph
SQL Query
Update Table
Schedule Workflow
Start Date: Jan 15, 2016 15:58
Repeat: Every Week
End Repeat: Never
Use Existing Schedule
<Select a Schedule>
Help
Advanced...
OK
Cancel
Advanced Settings
Notification
Settings
Nodes
Time Zone America/New_York
Job Class DEFAULT_JOB_CLASS
Job Priority
High
Medium
Low
Max Failures
3
Max Run Duration
Days Hours Mins
1 0 0
Schedule Limit
Days Hours Mins
0 2 0
Help
OK
Cancel
workflow - Properties
Find
Comment:
Connections
Data Miner
Connections
den00jlf--im
jw1415
wfriday
Oracle NoSQL Connections
Cloud Connections
Workflow Jobs
den00jlf--im
Workflow
Project
Status
No entries. Right click for Preferences.
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
Association
Classification
Explicit Feature Extraction
Feature Extraction
Model
Model Details
Regression
Predictive Queries
Model Operations
Linking Nodes
den00jlf--im
EN
3:01 PM
1/15/2016


Oracle SQL Developer : jw1415/Project1/workflow
File Edit View Navigate Run Diagram Team Tools Window Help Automation
workflow - Structure
Thumbnail
Class Build
INSUR_CUST_LTV_SAMPLE
Links
Start Page
workflow
100%
Performance Options On (Selected)
INSUR_CUST_LTV_SAMPLE
Class Build
Components
Workflow Editor
Data
Create Table or View
Data Source
Explore Data
Graph
SQL Query
Update Table
Schedule Workflow
Start Date: Jan 15, 2016 15:58
Repeat: Custom...
End Repeat: Never
Use Existing Schedule
<Select a Schedule>
Help
Advanced...
OK
Cancel
Repeat
Frequency: Weekly
Every 1 week(s) on:
Sun Mon Tue Wed Thur Fri Sat
Help
OK
Cancel
Connections
Data Miner
Connections
den00jlf--im
jw1415
wfriday
Oracle NoSQL Connections
Cloud Connections
Workflow Jobs
den00jlf--im
Workflow
Project
Status
No entries. Right click for Preferences.
workflow - Properties
Find
Comment:
Models
Anomaly Detection
Association
Classification
Clustering
Explicit Feature Extraction
Feature Extraction
Model
Model Details
Regression
Predictive Queries
Model Operations
Linking Nodes
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
den00jlf--im
EN
3:02 PM
1/15/2016

Previewing a 4.2 Feature

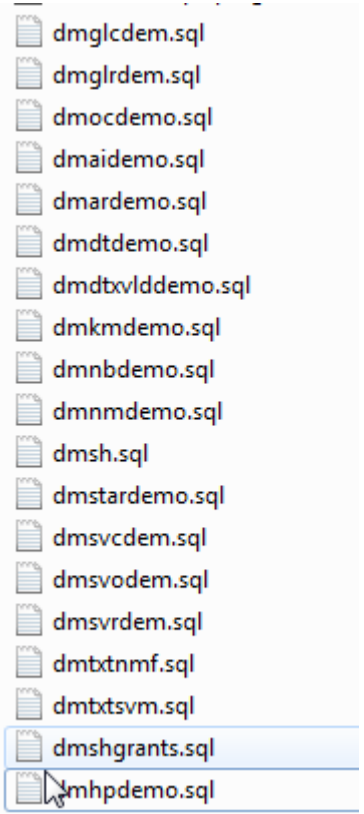
Workflow Scheduler—Email Notifications



Copyright © 2016 Oracle and/or its affiliates. All rights reserved. |

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.



Copyright © 2016 Oracle and/or its affiliates. All rights reserved. |

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 ×

Task completed in 5.096 seconds

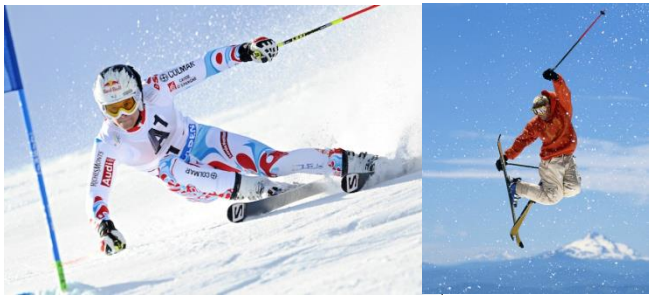
ATTRIBUTE_NAME	EXPLANATORY_VALUE	RANK
HOUSEHOLD_SIZE	.195	1
CUST_MARITAL_STATUS	.194	2
YRS_RESIDENCE	.115	3
EDUCATION	.106	4
AGE	.104	5
OCCUPATION	.092	6
Y_BOX_GAMES	.077	7
HOME_THEATER_PACKAGE	.069	8
CUST_GENDER	.043	9
BOOKKEEPING_APPLICATION	.024	10
BULK_PACK_DISKETTES	.000	11
COUNTRY_NAME	.000	11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

Copyright © 2016 Oracle and/or its affiliates. All rights reserved. |

Learn Predictive Analytics in 2 Hours!

Oracle Advanced Analytics Hands on Lab



- READY, SET, GO!!!!
- 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*

Copyright © 2016 Oracle and/or its affiliates. All rights reserved. |

Novice/Introductory/Overviews

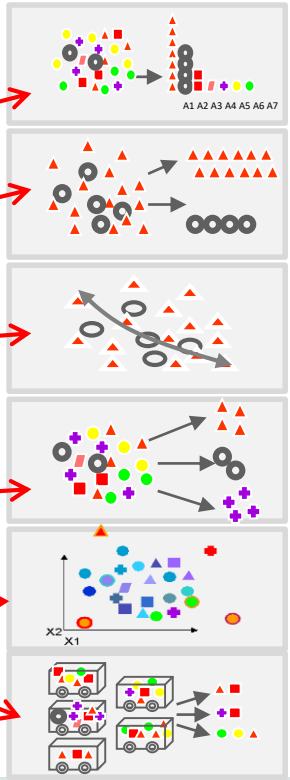
Quick Overview of Concepts, Process and Use Cases

Copyright © 2016 Oracle and/or its affiliates. All rights reserved. |

What is Machine Learning, 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)*

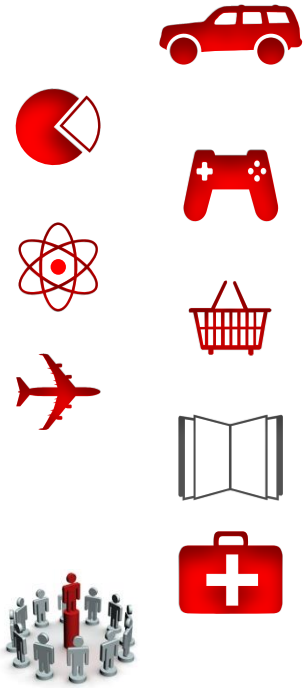


Copyright © 2016 Oracle and/or its affiliates. All rights reserved. |

Predictive Analytics & Data Mining

Typical Use Cases

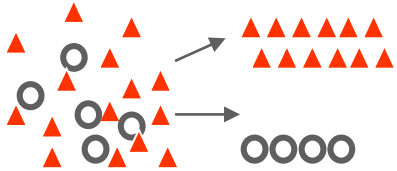
- Targeting the right customer with the right offer
- How is a customer likely to respond to an offer?
- Finding the most profitable growth opportunities
- Finding and preventing customer churn
- Maximizing cross-business impact
- Security and suspicious activity detection
- Understanding sentiments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

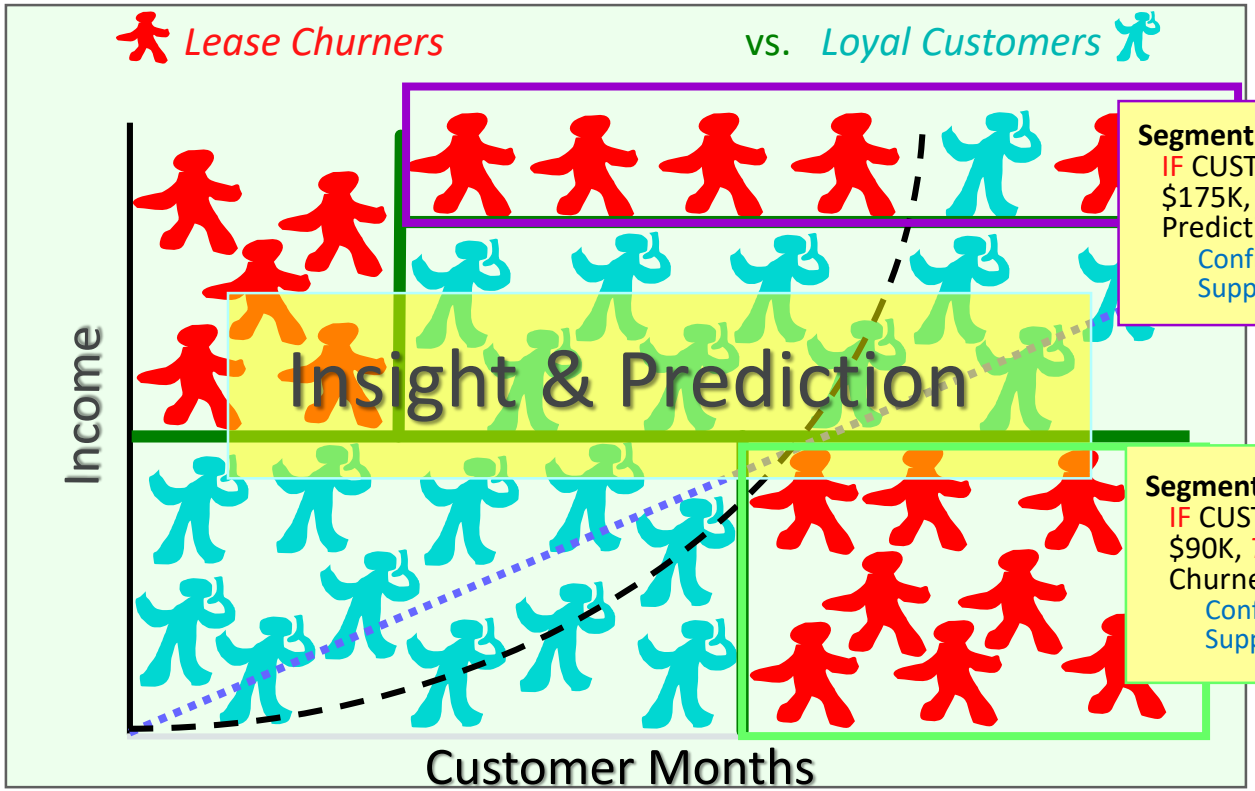


Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

Data Mining Provides

Better Information, Valuable Insights and Predictions





Segment #3
IF CUST_MO > 7 AND INCOME < $175K, *THEN* Prediction = Lease Churner,
Confidence = 83%
Support = 6/39

Segment #1
IF CUST_MO > 14 AND INCOME < $90K, *THEN* Prediction = Lease Churner
Confidence = 100%
Support = 8/39

Source: Inspired from *Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management* by Michael J. A. Berry, Gordon S. Linoff

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

Oracle Advanced Analytics DB Option

Advanced Analytics

In-Database Machine Learning Algorithms*—SQL & R & GUI Access

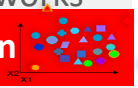
Classification

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

Regression

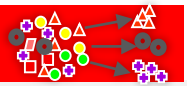
- Multiple Regression (GLM)
- Support Vector Machine (SVM)
- Stepwise Linear Regression
- Linear Model
- Generalized Linear Model
- Multi-Layer Neural Networks

Anomaly Detection



- 1-Class Support Vector Machine

Clustering



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

Attribute Importance

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

Market Basket Analysis

- Apriori – Association Rules

Text Mining

- All OAA/ODM SQL ML support
- Explicit Semantic Analysis

Predictive Queries

- Clustering
- Regression
- Anomaly Detection
- Feature Extraction

Feature Extraction & Creation

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

Time Series

- Single & Double Exp. Smoothing

Open Source R Algorithms

- Ability to run any R package (9,000+)via Embedded R mode

+ Ability to Mine Unstructured, Structured & Transactional data
+ Partitioned Models

ORACLE

Copyright © 2016, Oracle and/or its affiliates. All rights reserved. |

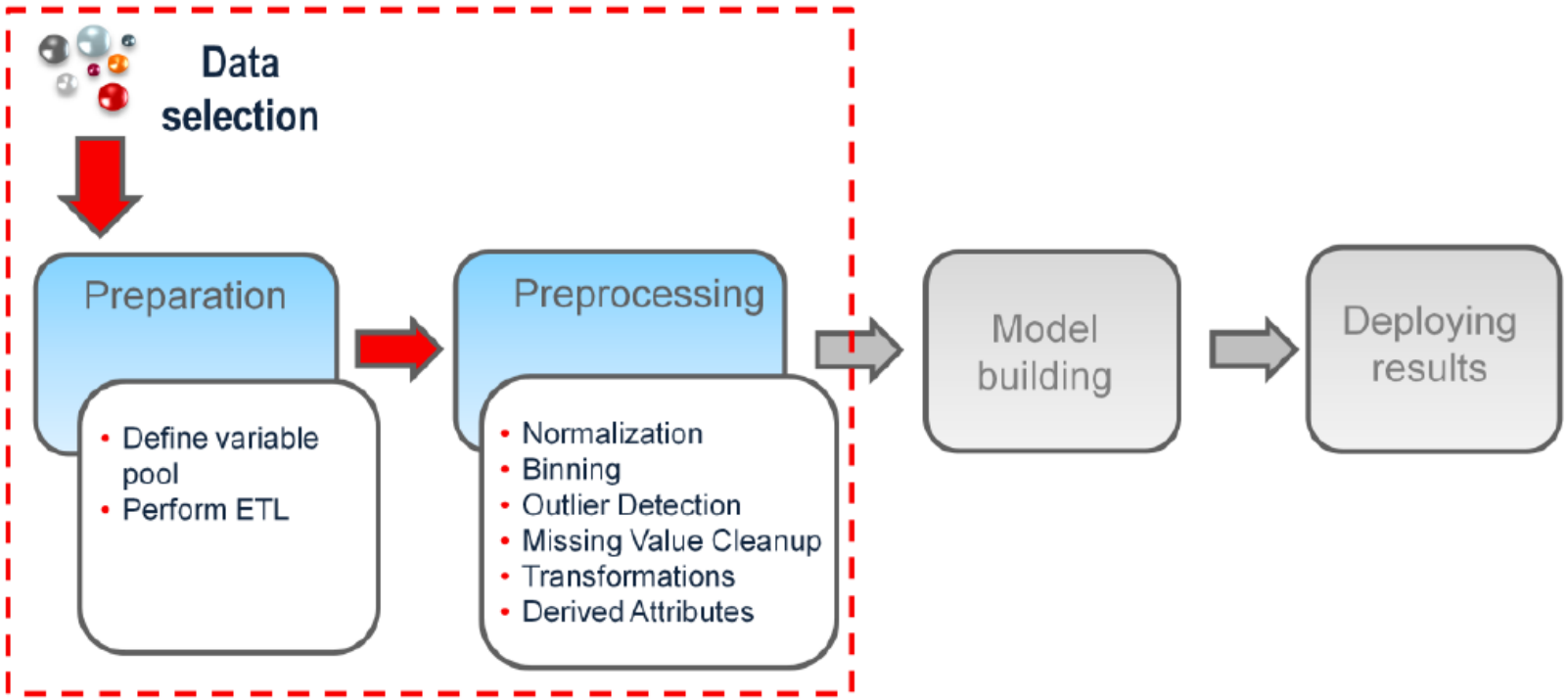
Oracle University's Learn Predictive Analytics Using Oracle Data Mining Course Agenda

Day 1	Day 2
1. Introduction	7. Using Clustering Models
2. Predictive Analytics and Data Mining Concepts	8. Performing Market Basket Analysis
3. The Data Mining Process	9. Performing Anomaly Detection
4. Introducing Oracle Data Miner 4.1	10. Performing Star Schema Mining
5. Using Classification Models	11. Using Predictive Queries
6. Using Regression Models	12. Deploying Data Mining Results

ORACLE

Copyright © 2016 Oracle and/or its affiliates. All rights reserved. |

The Data Mining Process



ORACLE

Copyright © 2016 Oracle and/or its affiliates. All rights reserved. |

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

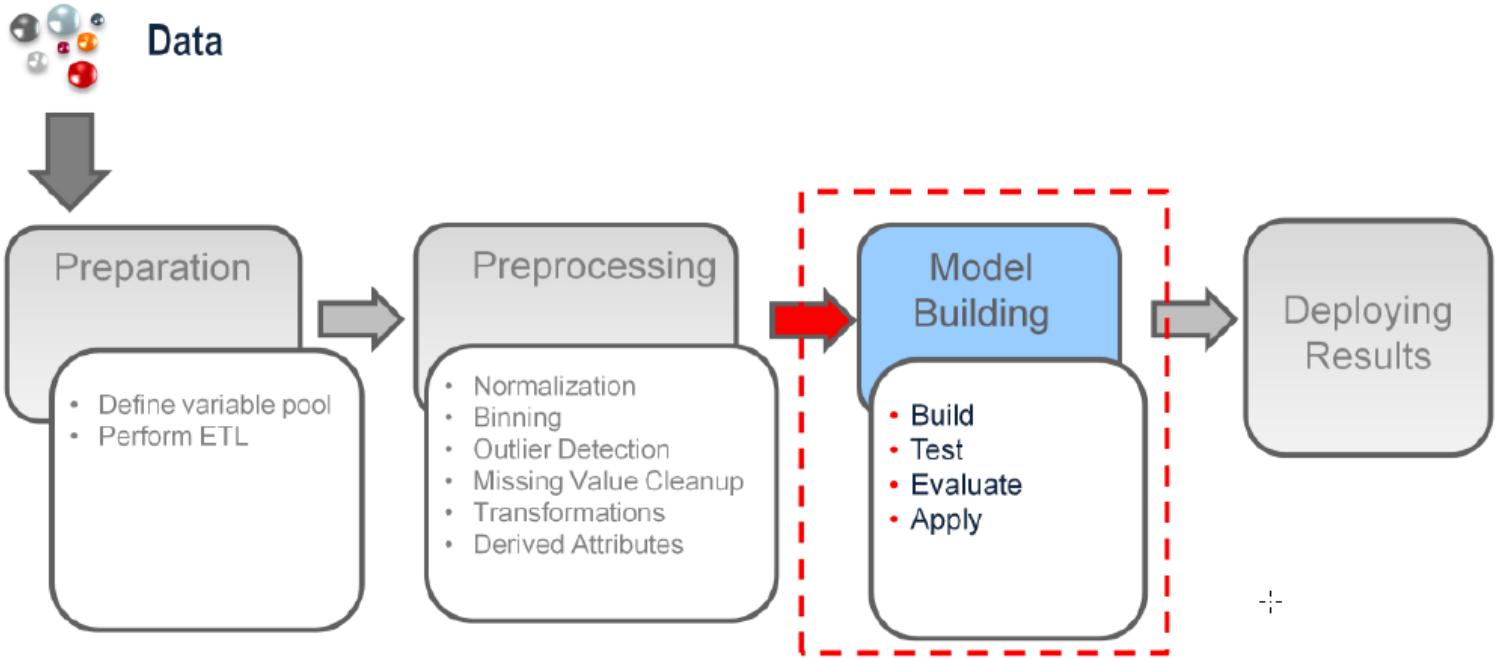
Target

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

Active

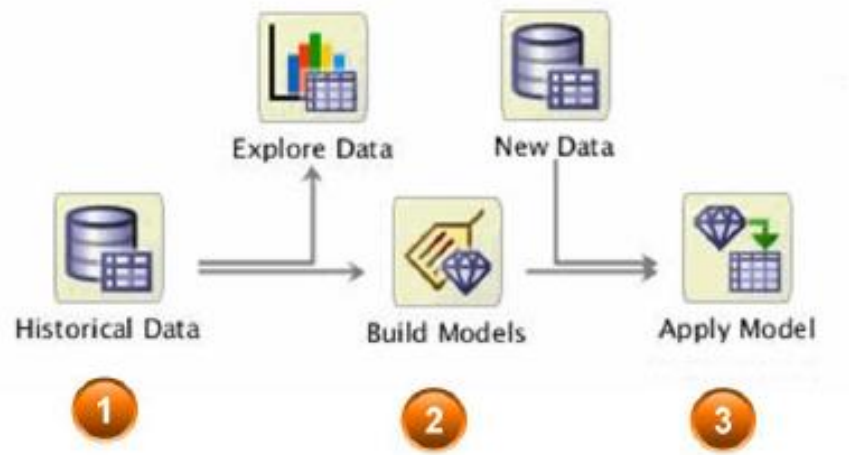
Copyright © 2016 Oracle and/or its affiliates. All rights reserved. |

Building and Evaluating Models



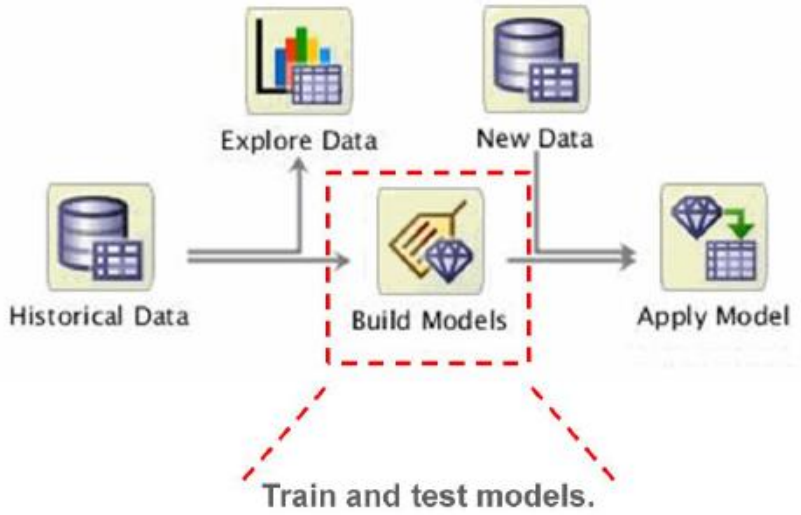
Copyright © 2016 Oracle and/or its affiliates. All rights reserved. |

Model Building Tasks



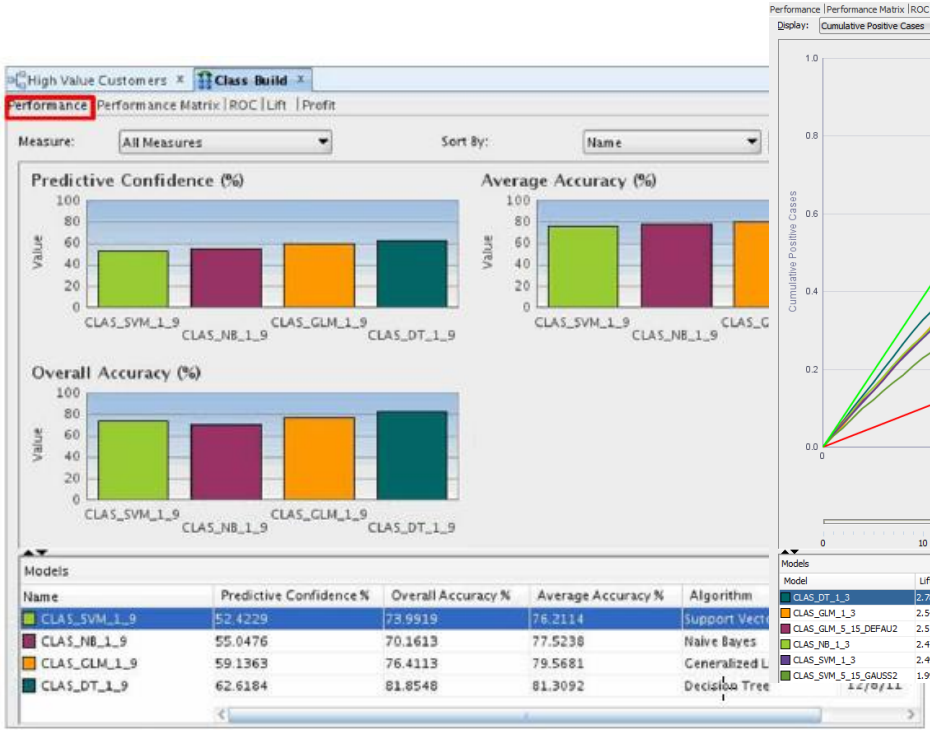
Copyright © 2016 Oracle and/or its affiliates. All rights reserved. |

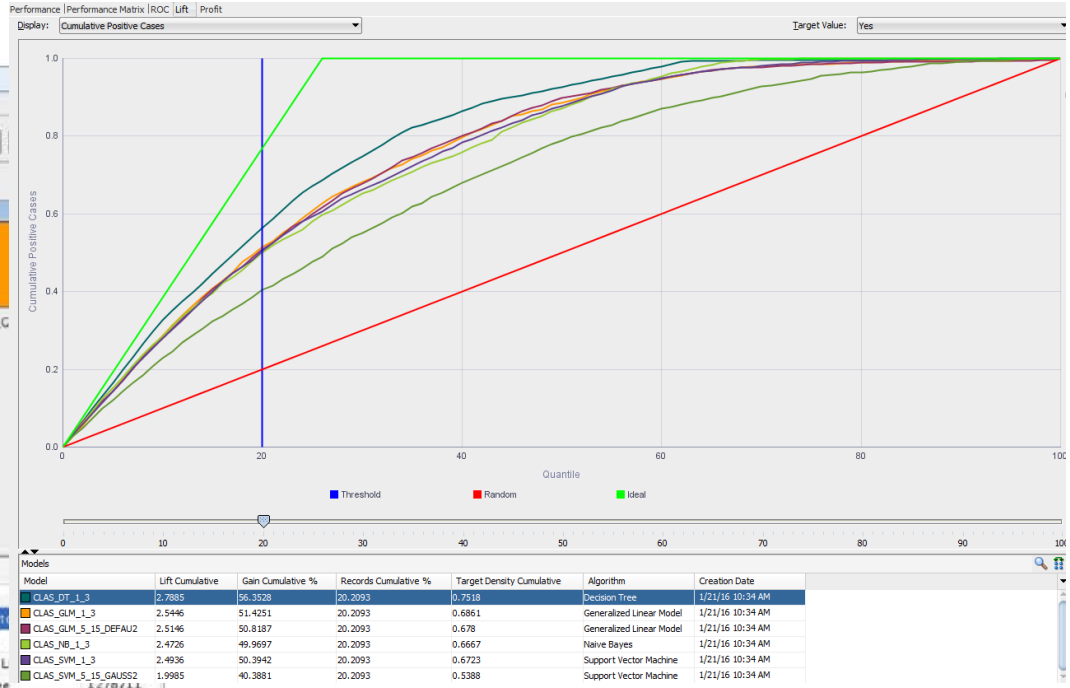
Model Train & Test: Supervised Learning



Copyright © 2016 Oracle and/or its affiliates. All rights reserved. |

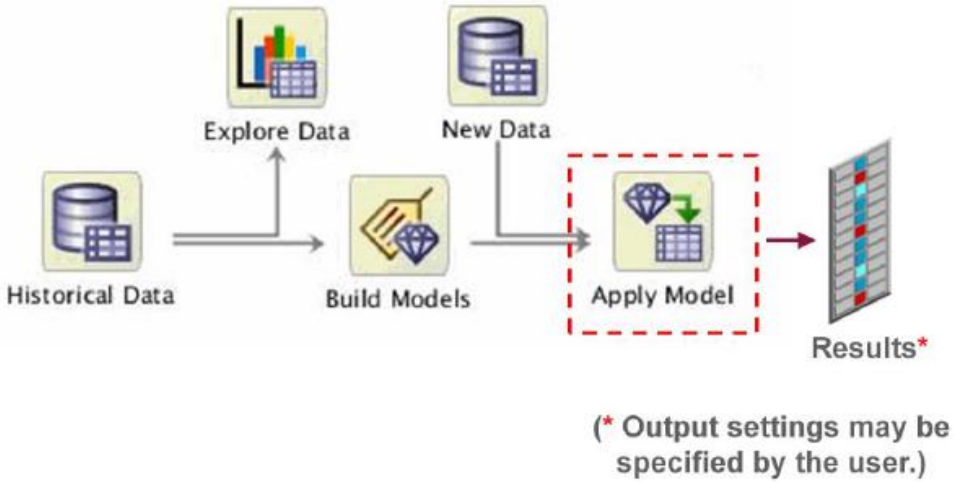
Model Evaluation: Supervised Learning





Copyright © 2016 Oracle and/or its affiliates. All rights reserved. |

Applying the Selected Model(s)



Copyright © 2016 Oracle and/or its affiliates. All rights reserved. |

Oracle Advanced Analytics

Brief Demos

ORACLE

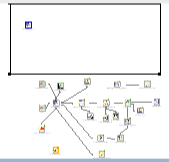
Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

File Edit View Navigate Run Diagram Team Tools Window Help

Connections | Data 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 Analytcs w JSON
    - Big Data Analytics
    - *BUY_INSURANCE_WORKFLOW*
    - CARS_DATA_MINING
    - Churners01 work flow
    - Claims Fraud Clustering + SVM2
    - Customer Analytics incl 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_Attrib Importance
- Predictive Models for Student
- Class Build
- Predictive models incl 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 | *BUY_INSURANCE_WORKFLOW*

100% | Parallel Query Off

CUST_INSUR_LTV1

CUST_INSUR_LTV1 - Properties

Find

Data
Cache
Details

Source Table: DMUSER.CUST_INSUR_LTV

Data

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

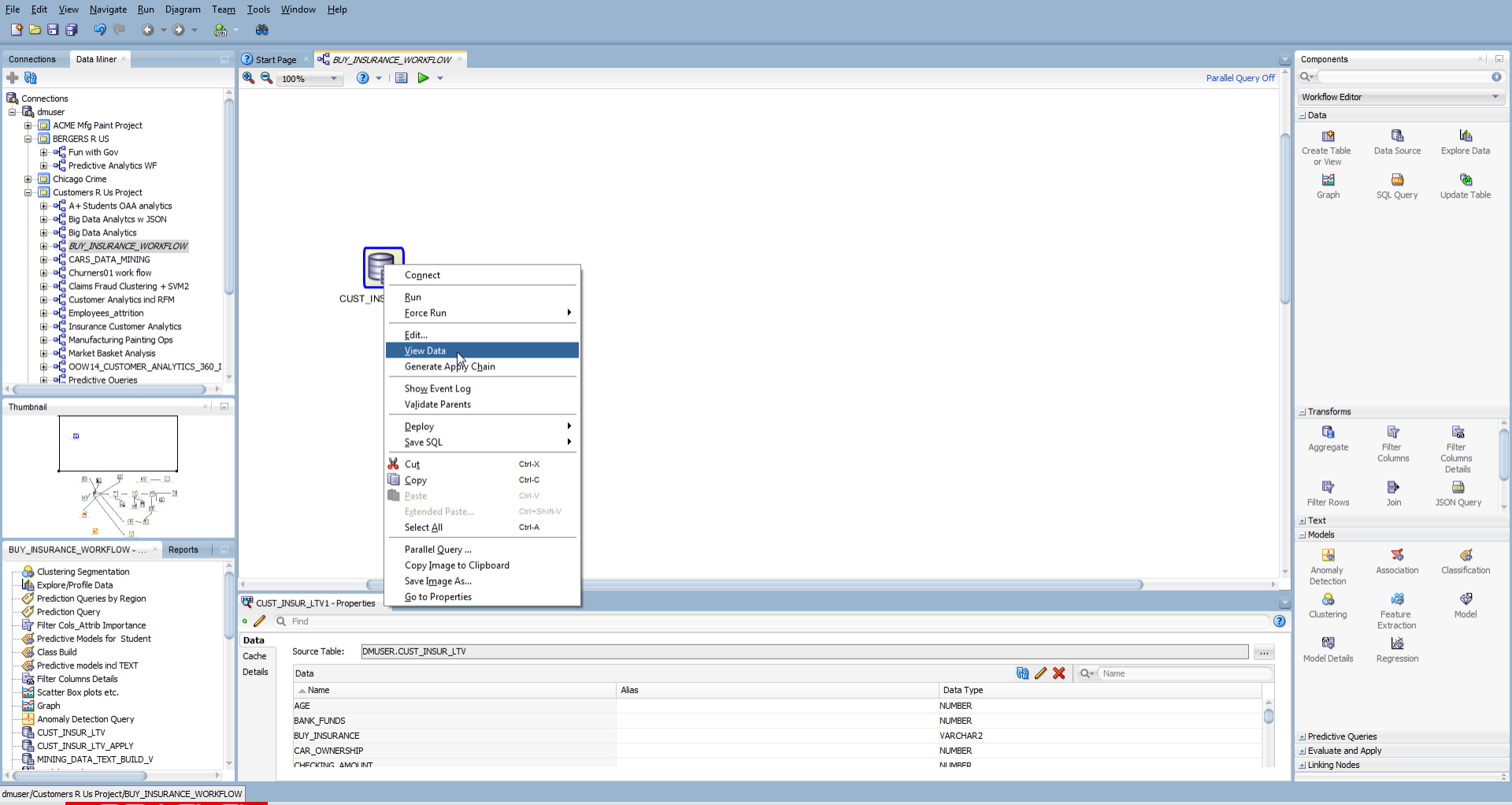
Predictive Queries

Evaluate and Apply

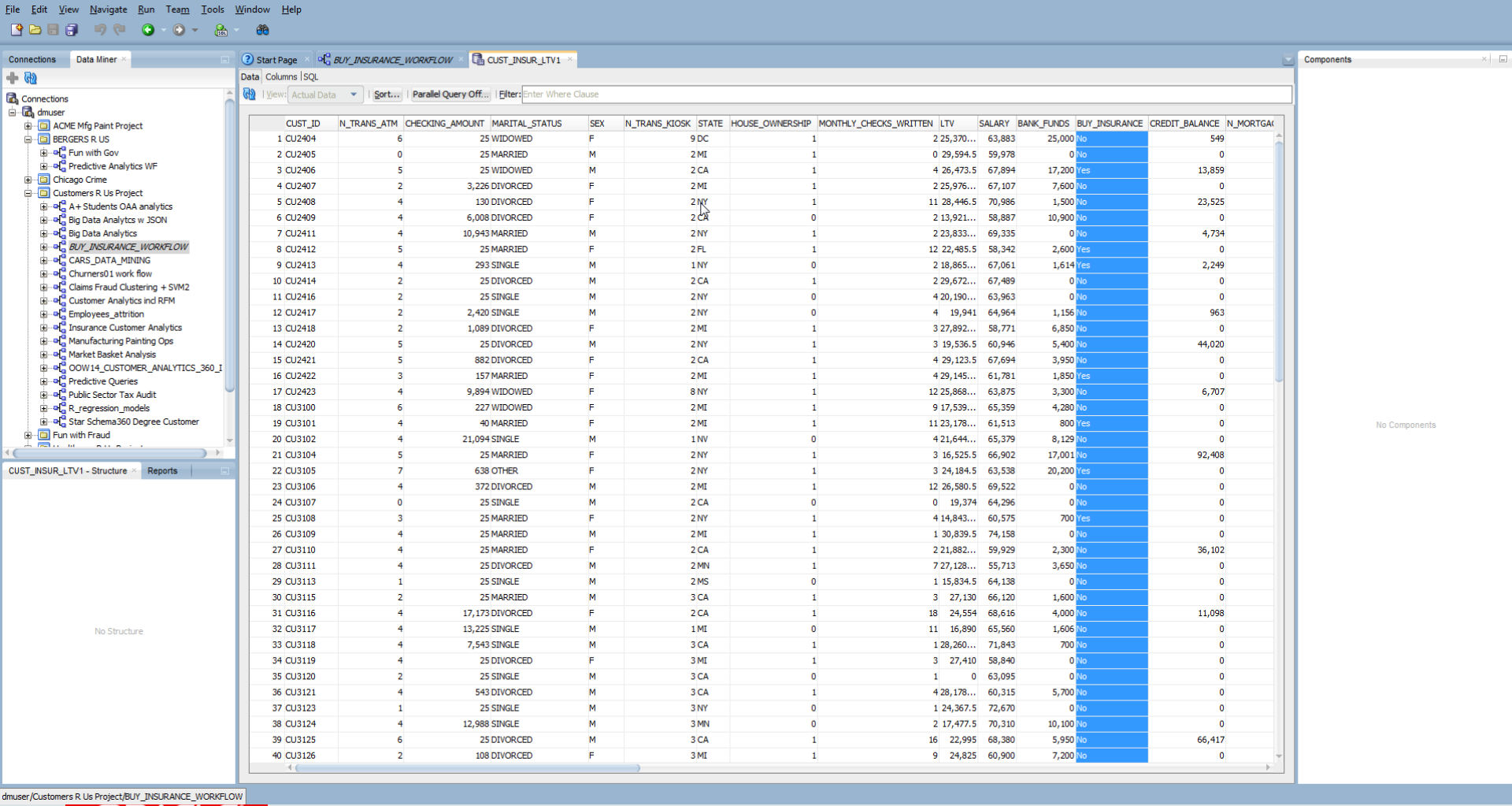
Linking Nodes

dmuser/Customers R Us Project/BUY_INSURANCE_WORKFLOW

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

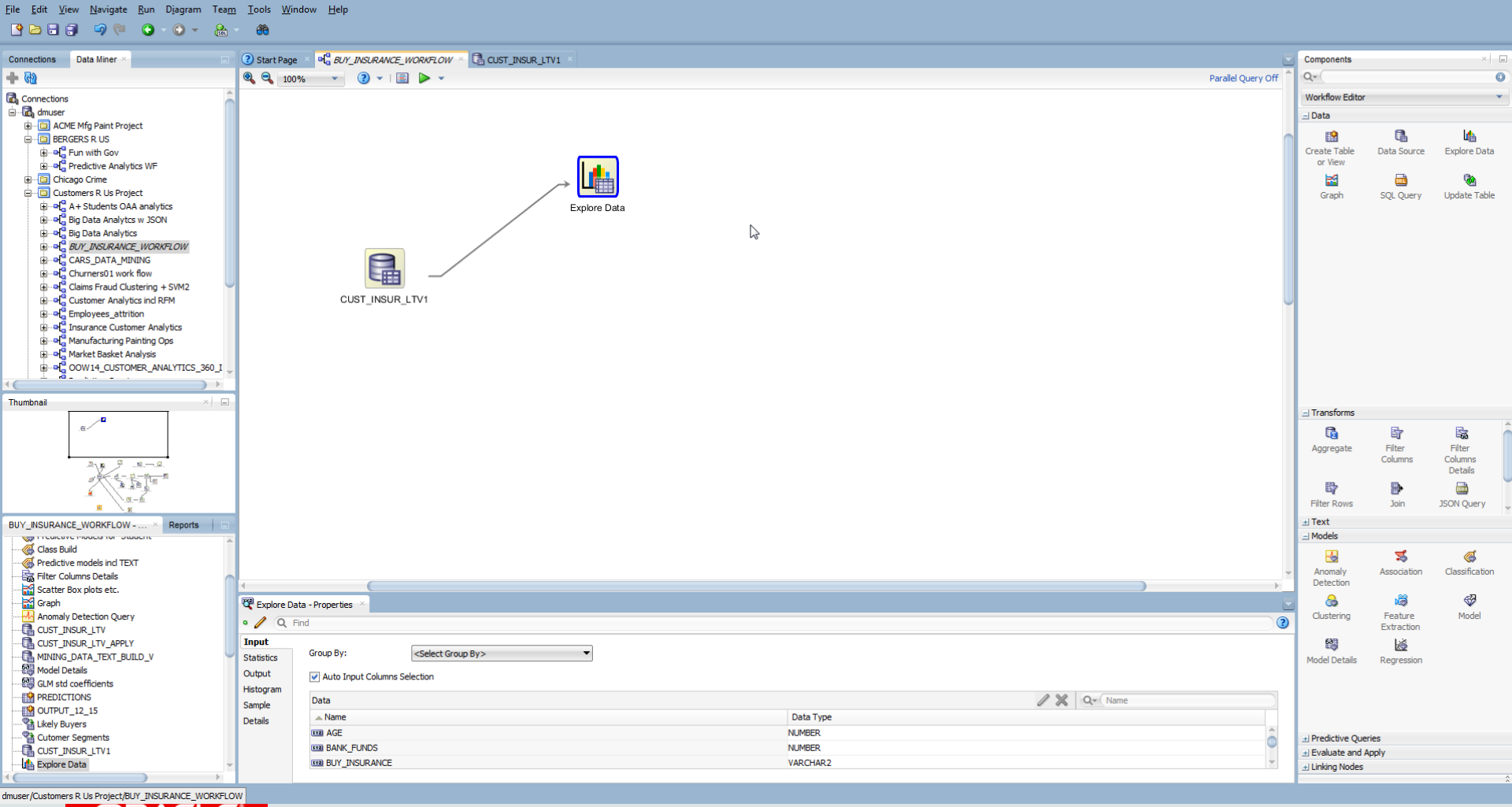


Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |



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

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



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

Oracle SQL Developer : Charlie Database 12c laptop dmuser_dmuser/AAA Customer Analytics/BUY INSURANCE/Scatter Box plots etc.

File Edit View Navigate Run Team Tools Window Help

Connections | Data Miner

Connections
- BIWA16_12c_DB_Cloud_Jonathan2
- BIWA DMUSER - Learn Predictive Analtics in 2 Hours HO
- CBERGER dbpm23 Cloud instance Connection
- Charlie Database 12c laptop dmuser_dmuser
  - AAA Customer Analytics
    - 360 View Star Schema Analytics
    - BUY INSURANCE
    - Claims Anomaly Detection
    - Customer 360 Clustering
    - Customer analytics NEW
    - Fun Analytics WF
    - Market Basket Analysis
    - My First Workflow
    - New Buy Insurance WF
    - New Customer Analytics
    - New WF
    - Predictive Queries
    - RFM input4 Customer Analytics
    - Structured_Unstructured Data

Thumbnail

SSH Hosts
- SSH Hosts
  - CBERGER dbpm23 Cloud instance
  - Express Cloud PDB May 25
  - HOL3
  - NEW_SSH_BIWA_Brian2
  - OracleCloudDatabase1
  - OracleDatabaseCloud2
  - SSH_BIWA_Jonathan2

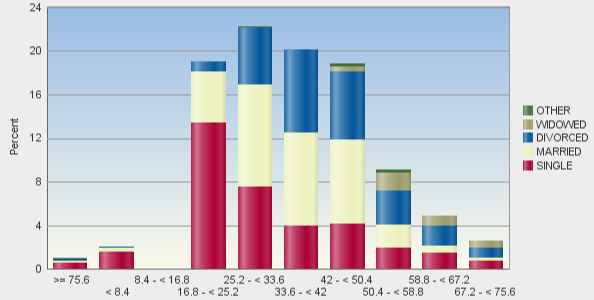
Start Page | BUY INSURANCE | Charlie Database 12c laptop dmuser_dmuser2.sql | Profile Data

Statistics | Data | Columns | SQL

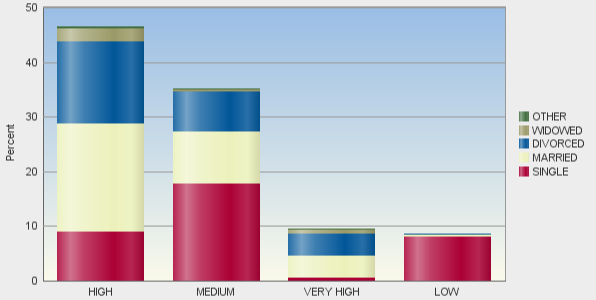
Statistics: 10 Columns from 2,005 Rows(Sampled) — Group by: MARITAL_STATUS — Show Nulls — Filter: Name

Name	Histogram	Data Type	Percent NULLs	Distinct Values	Distinct Per…	Mode	Average	Median	Min Value	Max Value	Standard Devi…	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	22,460,667.7…
BUY_INSURANCE		VARCHAR2	0	2	0.0998	No						
CAR_OWNERSHIP		NUMBER	0	2	0.0998		0.9247	1	0	1	0.264	0.0697
CHECKING_AMOUNT		NUMBER	0	626	31.2219		1,032.7312	25	25	24,471	3,103.6812	9,632,836.6997

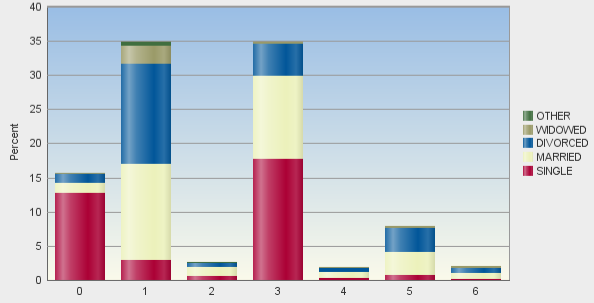
AGE By MARITAL_STATUS



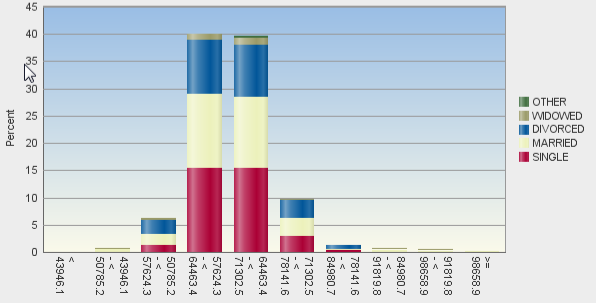
LTV_BIN By MARITAL_STATUS



N_OF_DEPENDENTS By MARITAL_STATUS



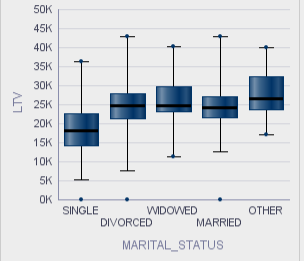
SALARY By MARITAL_STATUS



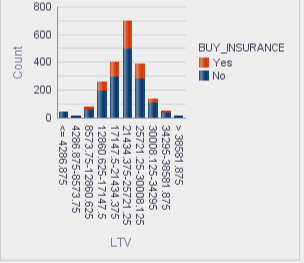
Scatter Box plots etc. | Components

View: Sample Data | Parallel Query Off…

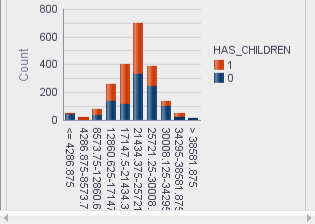
Graph1



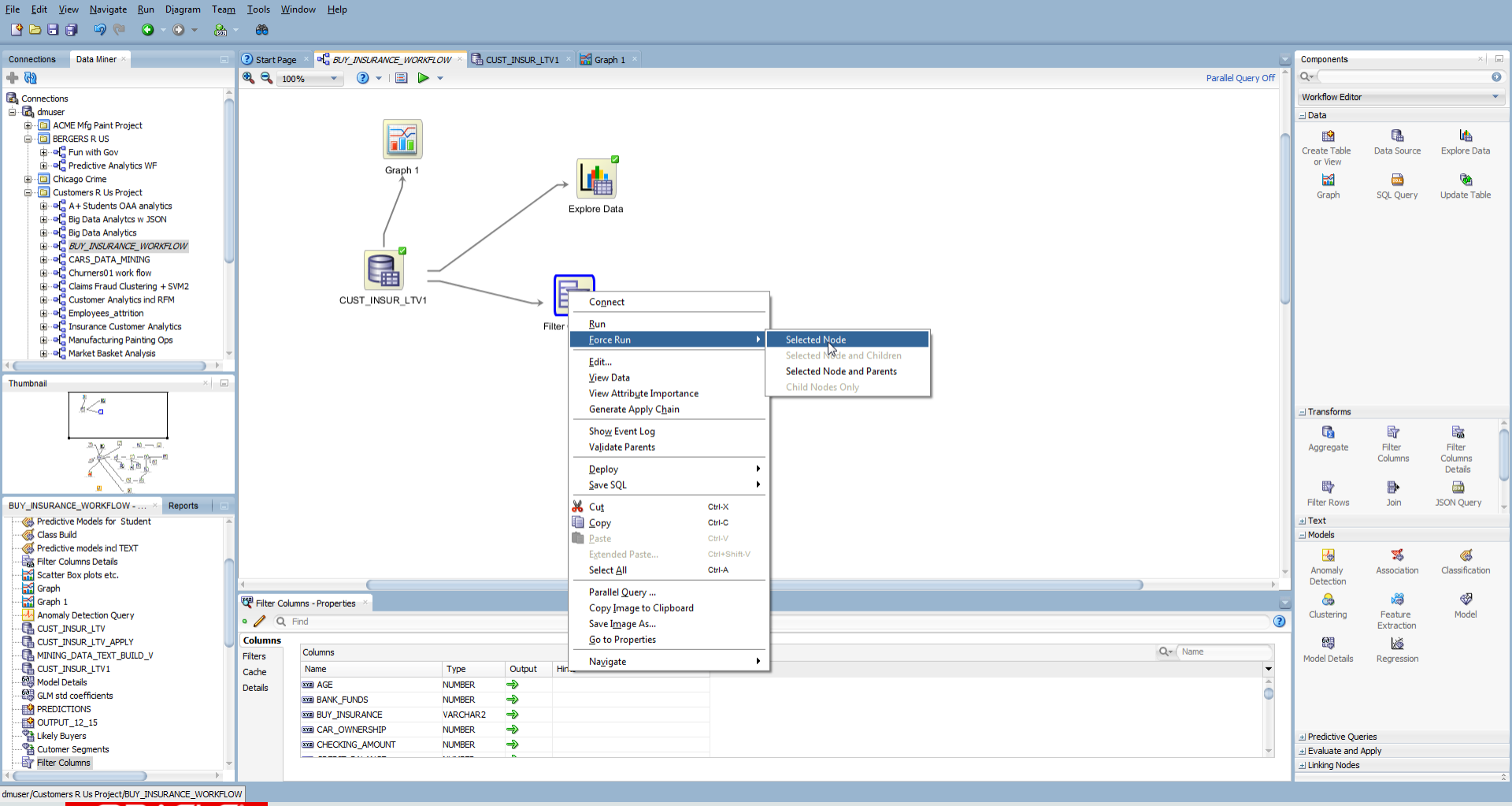
Graph2



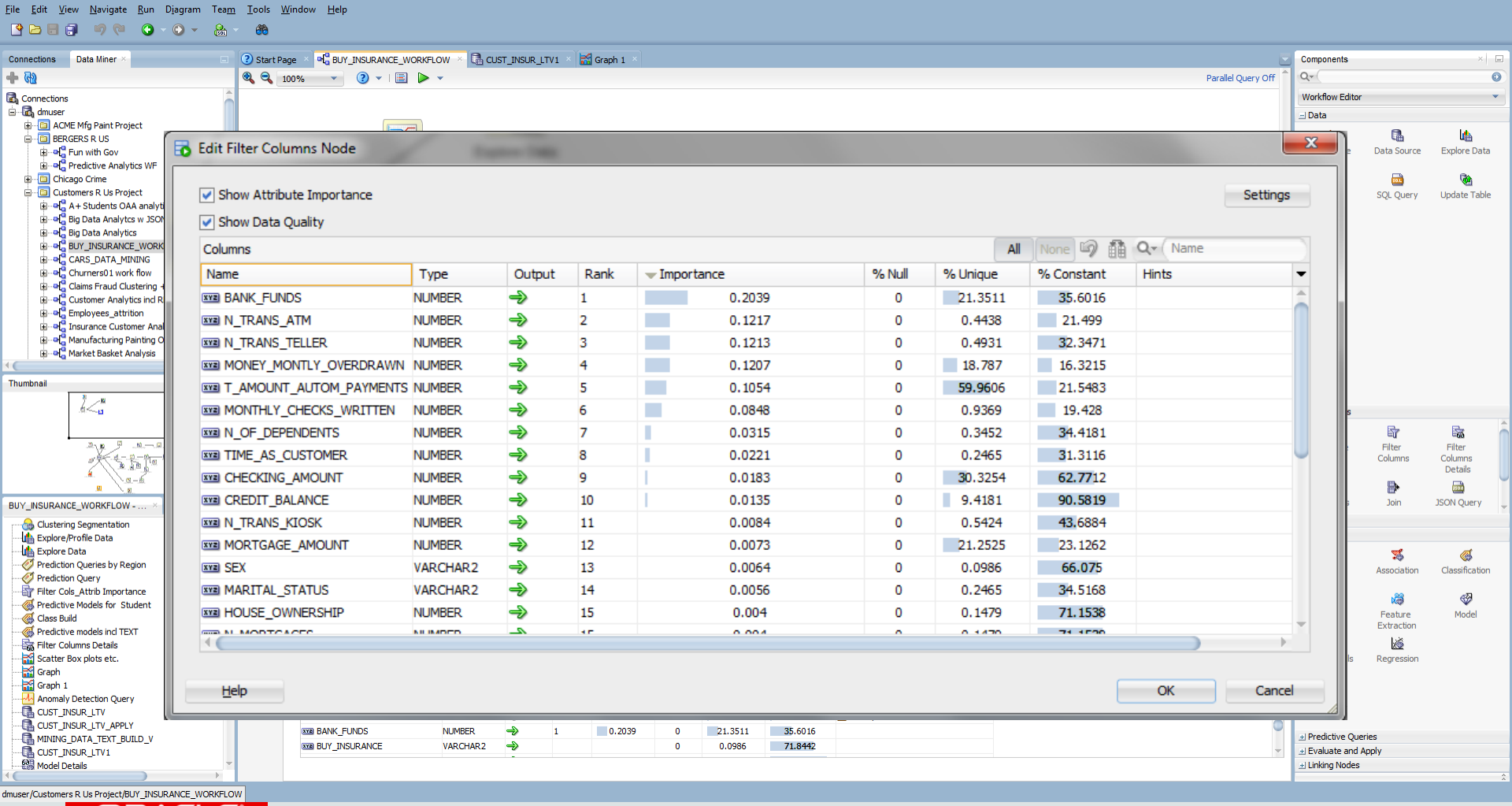
Graph3



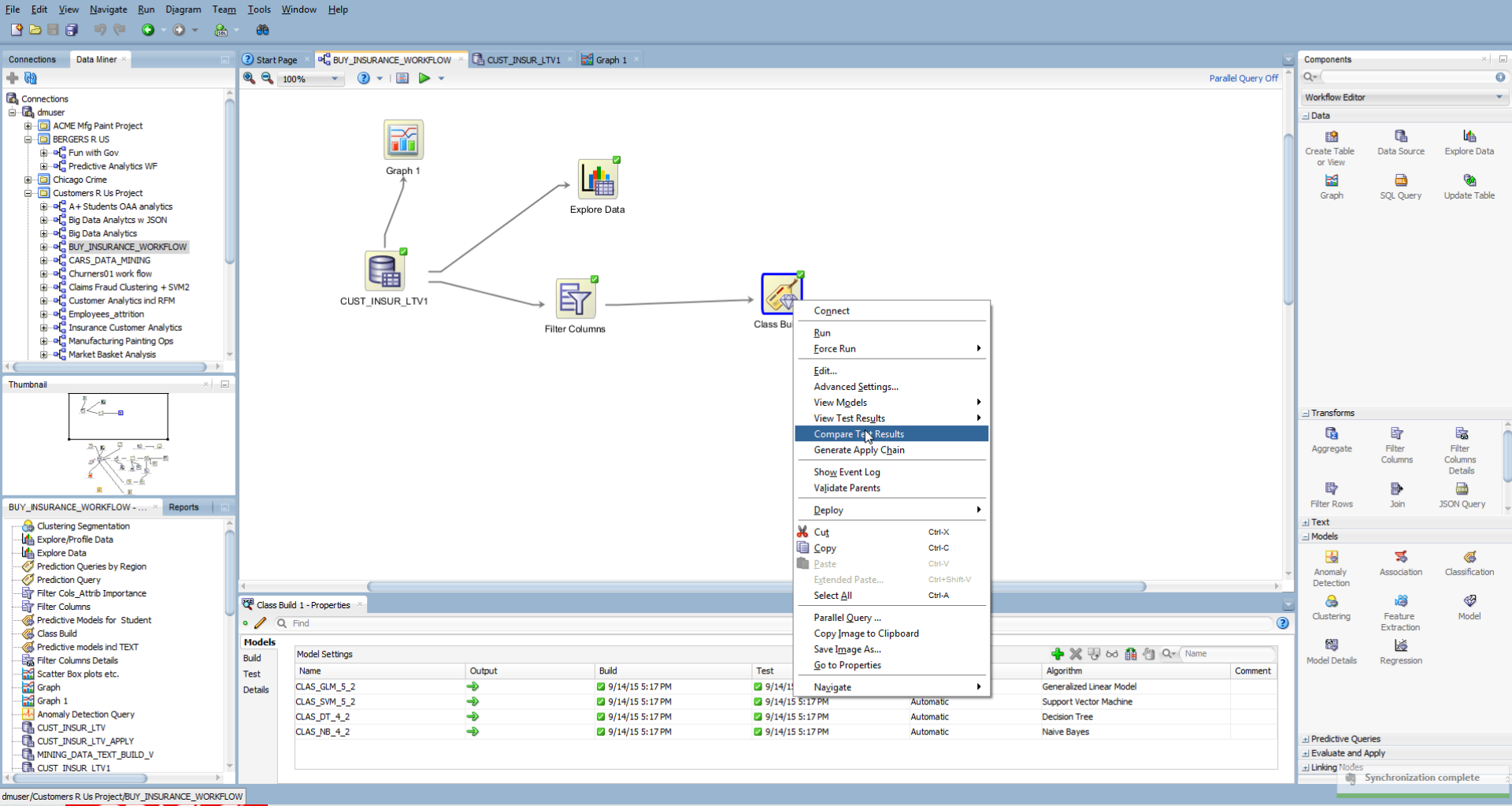
Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |



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



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



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

File Edit View Navigate Run Team Tools Window Help

Connections | Data 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 Analytcs w JSON
    - Big Data Analytics
    - BUY_INSURANCE_WORKFLOW
    - CARS_DATA_MINING
    - Churners01 work flow
    - Claims Fraud Clustering + SVM2
    - Customer Analytics incl RFM
    - Employees_attrition
    - Insurance Customer Analytics
    - Manufacturing Painting Ops
    - Market Basket Analysis
    - OOW14_CUSTOMER_ANALYTICS_360_I
    - Predictive Queries
    - Public Sector Tax Audit
    - R_regression_models

Class Build 1 - Structure | Reports

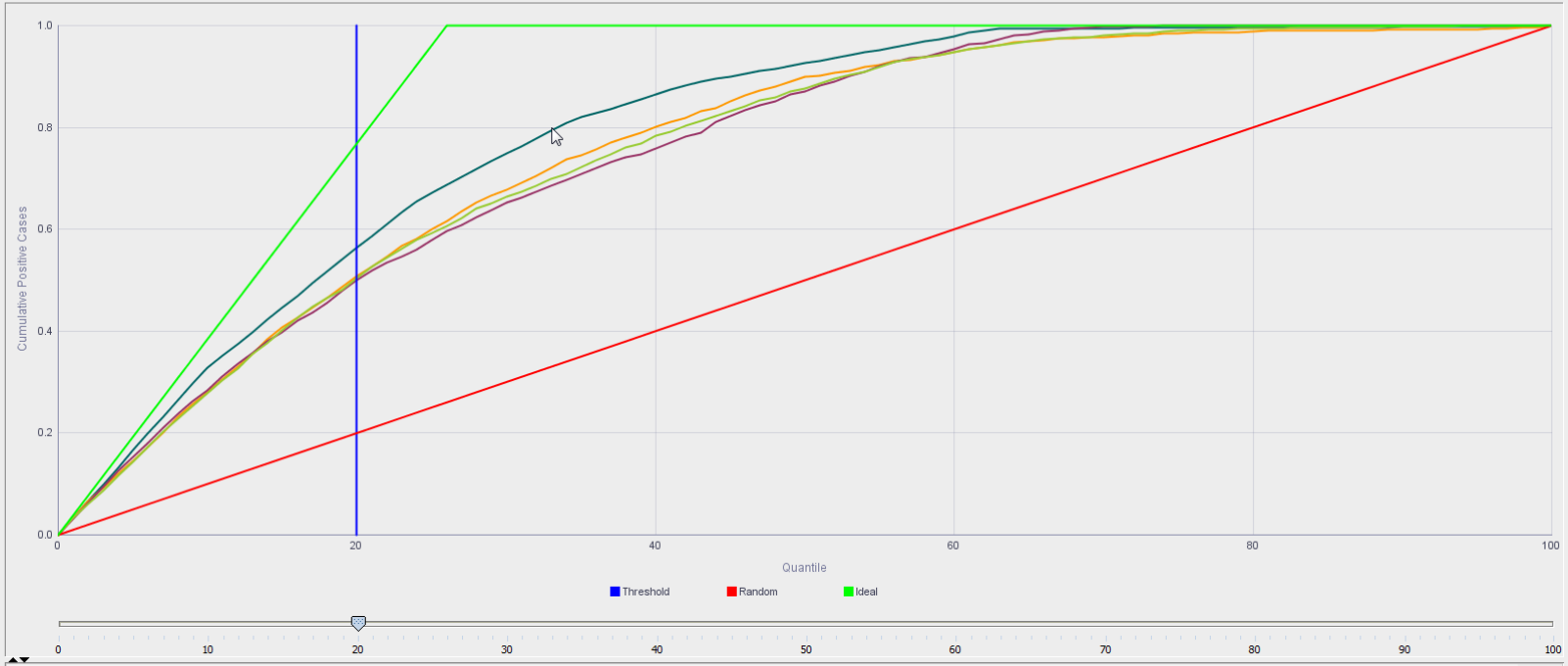
No Structure

Start Page | BUY_INSURANCE_WORKFLOW | CUST_INSUR_LTV1 | Graph 1 | Class Build 1

Performance | Performance Matrix | ROC | Lift | Profit

Display: Cumulative Positive Cases

Target Value: Yes



Models

Model	Lift Cumulative	Gain Cumulative %	Records Cumulative %	Target Density Cumulative	Algorithm	Creation Date
CLAS_DT_4_2	2.7885	56.3528	20.2093	0.7518	Decision Tree	9/14/15 5:17 PM
CLAS_GLM_5_2	2.5146	50.8187	20.2093	0.678	Generalized Linear Model	9/14/15 5:17 PM
CLAS_NB_4_2	2.4726	49.9697	20.2093	0.6667	Naive Bayes	9/14/15 5:17 PM
CLAS_SVM_5_2	2.4936	50.3942	20.2093	0.6723	Support Vector Machine	9/14/15 5:17 PM

dmuser/Customers R Us Project/BUY_INSURANCE_WORKFLOW

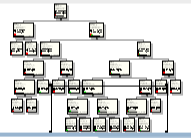
Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

File Edit View Navigate Run Diagram Team Tools Window Help

Connections | Data 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 Analytcs w JSON
      - Big Data Analytics
      - BUY_INSURANCE_WORKFLOW
      - CARS_DATA_MINING
      - Churners01 work flow
      - Claims Fraud Clustering + SVM2
      - Customer Analytics incl RFM
      - Employees_attrition
      - Insurance Customer Analytics
      - Manufacturing Painting Ops
      - Market Basket Analysis

Thumbnail



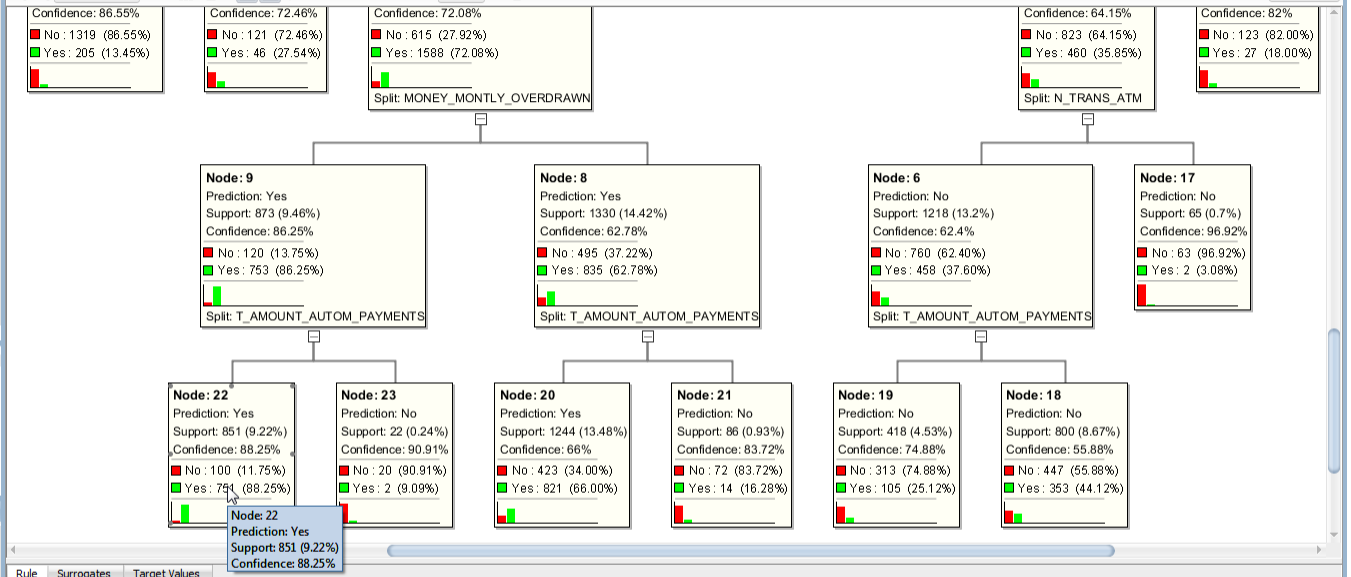
CLAS_DT_4_2 - Structure | Reports

- 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
- Node: 10

Start Page | BUY_INSURANCE_WORKFLOW | CUST_INSUR_LTV1 | Graph 1 | CLAS_DT_4_2

Tree | Settings

100% | Maximum Target Values: 2 | Save Rules...



Rule | Surrogates | Target Values

Node Rule: Wrap

```
If   BANK_FUNDS > 246
And  CHECKING_AMOUNT <= 282
And  CREDIT_BALANCE <= 2445
And  MONEY_MONTLY_OVERDRAWN > 54.095
And  T_AMOUNT_AUTOM_PAYMENTS <= 14993
Then Yes
```

Confidence	0.882491186839013
Support	0.09223932365055279

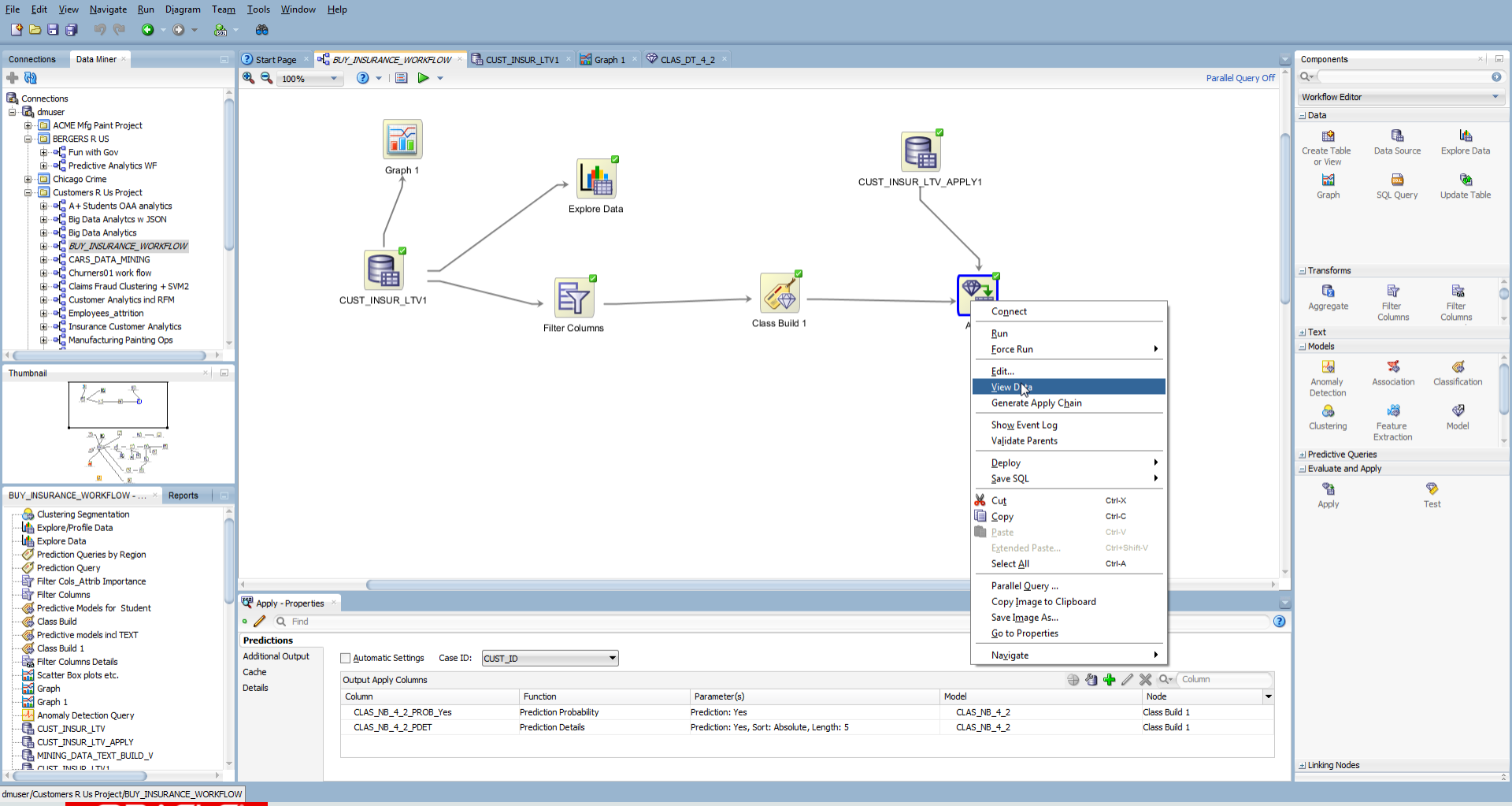
Components

No Components

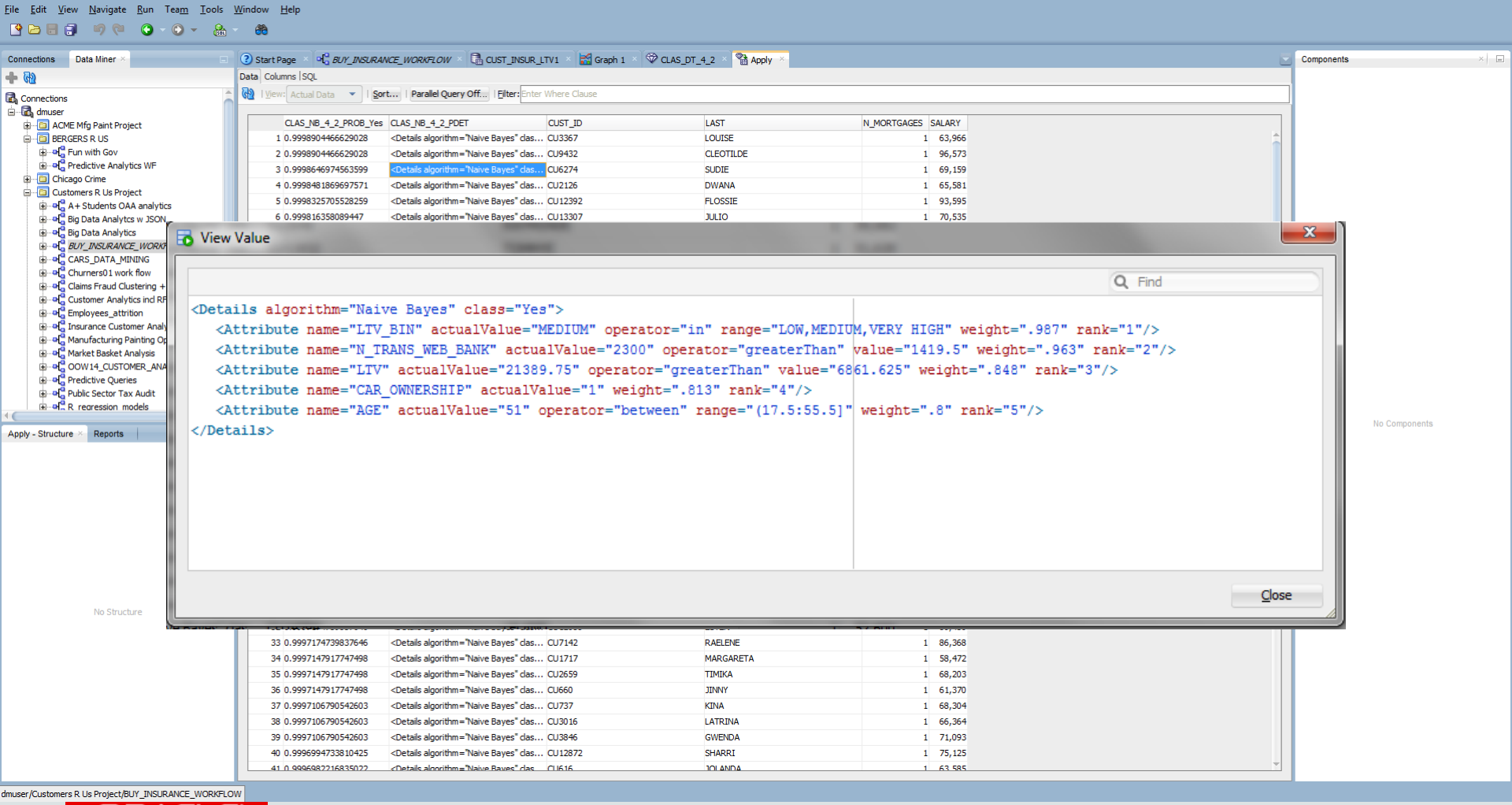
dmuser/Customers R Us Project/BUY_INSURANCE_WORKFLOW

ORACLE

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



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



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

File Edit View Navigate Run Diagram Team Tools Window Help

Connections | Data Miner

Customers R Us Project
- A+ Students OAA analytics
- Big Data Analytcs w JSON
- Big Data Analytics
- BUY_INSURANCE_WORKFLOW
- CARS_DATA_MINING
- Churners01 work flow
- Claims Fraud Clustering + SVM2
- Customer Analytics incl RFM
- *Employees_attrition*
- Insurance Customer Analytics
- Manufacturing Painting Ops
- Market Basket Analysis
- OOW14_CUSTOMER_ANALYTICS_360_
- PAW Cust Analytics
- Predictive Queries
- Public Sector Tax Audit
- R_regression_models
- RFM Customer Analytics
- RFM CUSTOMER ANALYTICS2
- Star Schema360 Degree Customer

Thumbnail

Start Page | *Employees_attrition* | Hist. Employee data | CLAS_DT_1_10

Tree | Settings

64% | Maximum Target Values: 2 | Save Rules...

Node: 29 Prediction: NO Support: 162 (2.1%) Confidence: 56.17% YES : 71 (43.83%) NO : 91 (56.17%)

Node: 28 Prediction: YES Support: 930 (12.07%) Confidence: 82.15% YES : 764 (82.15%) NO : 166 (17.85%)

Node: 27 Prediction: YES Support: 360 (4.67%) Confidence: 91.94% YES : 331 (91.94%) NO : 29 (8.06%)

Node: 8 Prediction: NO Support: 2065 (26.8%) Confidence: 69.39% YES : 632 (30.61%) NO : 1433 (69.39%)

Node: 7 Prediction: YES Support: 745 (9.67%) Confidence: 67.52% YES : 503 (67.52%) NO : 242 (32.48%)

Rule | Surroga

Node Rule:

If BONU
And HEAL
And STOC
And 2.5
And AMT_
Then NO

Confidence	0.6183533447684391
Support	0.15131066701271736

At Risk Employees

Data | Columns | SQL

View: Cache Data | Sort... | Parallel Query Off... | Filter: Enter Where Clause

	CLAS_GLM_1_10_PROB_NO	CLAS_GLM_1_10_PDET
1	0.4790120104773802	<Details algorithm="Generalized Linea...
2	0.03718578019711771	<Details algorithm="Generalized Linea...
3	0.1457710019211804	<Details algorithm="Generalized Linea...
4	0.12768928733510213	<Details algorithm="Generalized Linea...
5	0.5698507390589098	<Details algorithm="Generalized Linea...
6	0.8124155905506835	<Details algorithm="Generalized Linea...
7	0.2038408741589256	<Details algorithm="Generalized Linea...
8	0.8968650734058157	<Details algorithm="Generalized Linea...
9	0.6750795593148387	<Details algorithm="Generalized Linea...
10	0.7915741097722724	<Details algorithm="Generalized Linea...
	70587337939665	<Details algorithm="Generalized Linea...
	148971355681	<Details algorithm="Generalized Linea...
	53665860052554	<Details algorithm="Generalized Linea...
	07470752471	<Details algorithm="Generalized Linea...
	695885974645	<Details algorithm="Generalized Linea...
36	0.4310836245766258 6	<Details algorithm="Generalized Linea...
37	0.2988898709403637	<Details algorithm="Generalized Linea...

ORACLE

DATA VISUALIZATION

Copyright © 2015, 2016 Oracle and/or its affiliates. All rights reserved.

Find

```
Model" class="NO">
ctualValue="5" weight=".246" rank="1"/>
alue="1" weight="-.046" rank="2"/>
e="F" weight=".036" rank="3"/>
actualValue="25" weight=".022" rank="4"/>
27" weight=".021" rank="5"/>
```

Close

Untitled - Oracle Data Visualization Desktop

ORACLE Data Visualization Desktop

Search

CUST_INSUR_PROFILES
- PROFILE_ID
- PREDICTION
- PREDICTION_COUNT
- RECORD_COUNT
- TOTAL_RECORD_COUNT
- IS_LEAF_NODE
- CONFIDENCE
- SUPPORT
- FULL_SIMPLE_RULE

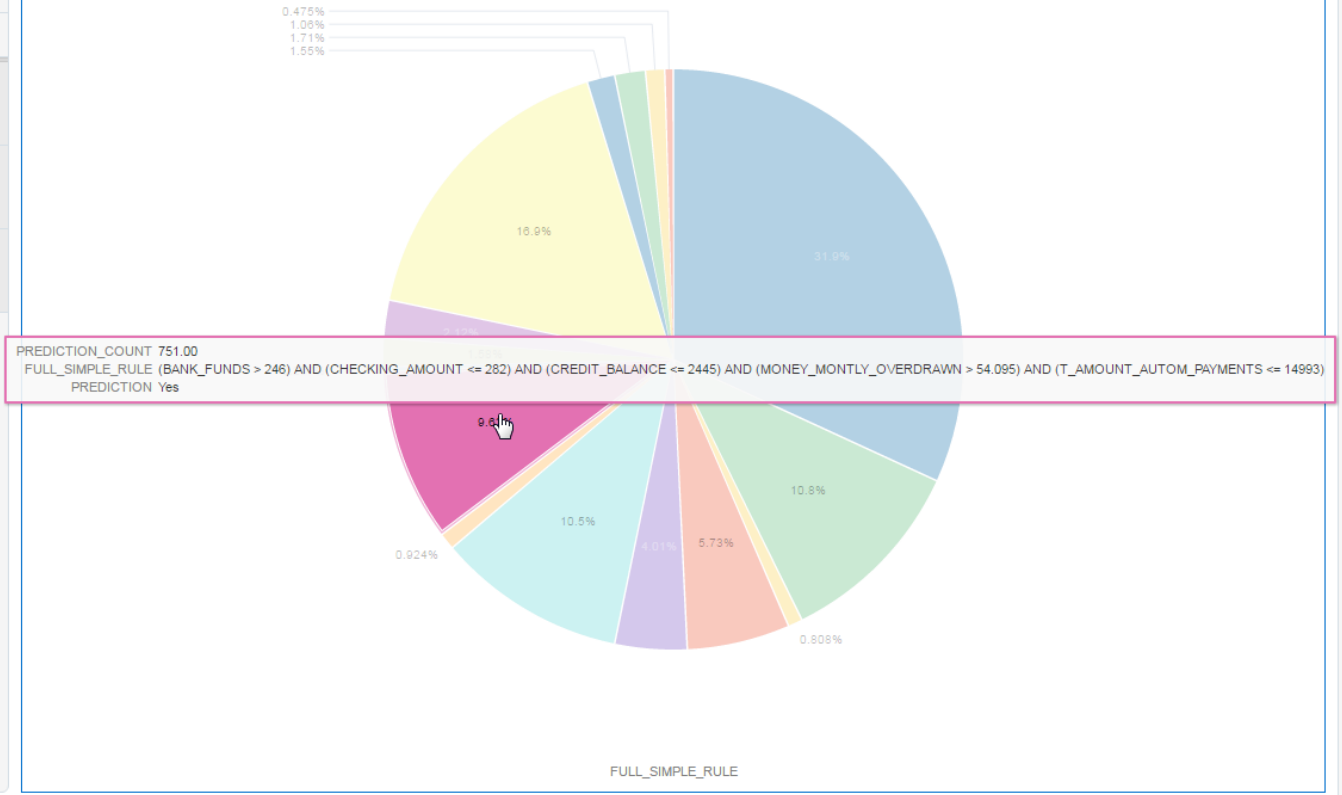
My Calculations
Value Labels

Untitled

Click here or drag data elements to add a filter

Pie
Trellis Columns
Trellis Rows
Values (Slice)
PREDICTION_CO...
Category
PREDICTION
Color
FULL_SIMPLE_RU..

PREDICTION_COUNT by PREDICTION, FULL_SIMPLE_RULE

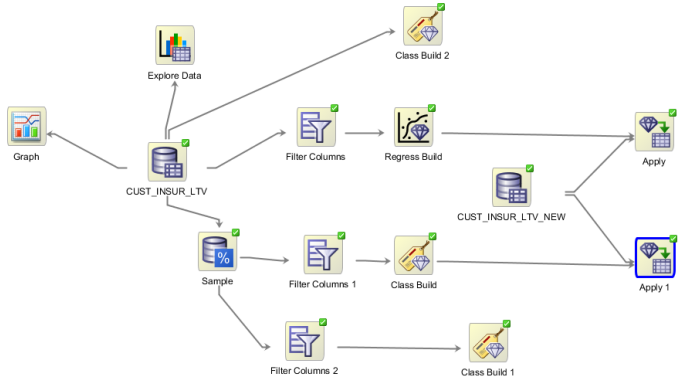


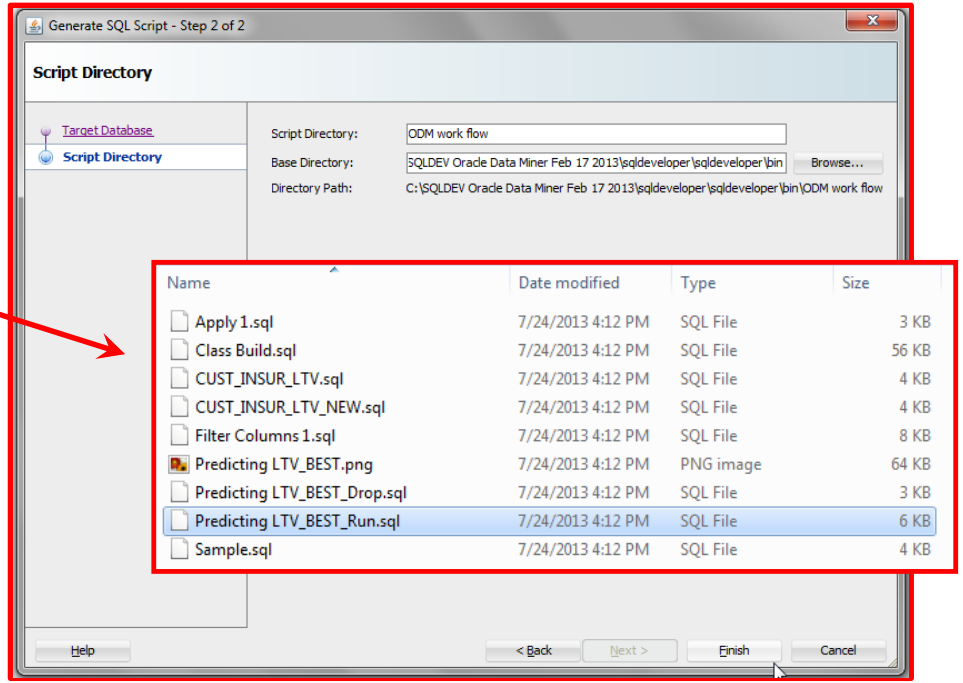
PREDICTION_COUNT 751.00
FULL_SIMPLE_RULE (BANK_FUNDS > 246) AND (CHECKING_AMOUNT <= 282) AND (CREDIT_BALANCE <= 2445) AND (MONEY_MONTLY_OVERDRAWN > 54.095) AND (T_AMOUNT_AUTOM_PAYMENTS <= 14993)
PREDICTION Yes

Sharing, Automation and Deployment

Immediately Go to "Productionization" of Analytical Methodologies

- Share ODMr workflows
- Workflow API for 100% automation
  - Immediate deployment of data analyst's methodologies
- SQL Script Generation
  - Deploy methodology as SQL scripts





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

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;
```

Script Output | Query Result

SQL | All Rows Fetched: 5 in 0.064 seconds

	POLICYNUMBER	PERCENT_FRAUD	RNK
1	654	61.87	1
2	11068	57.37	2
3	7435	55.47	3
4	3599	55.4	4
5	14877	55.37	5

Automated Monthly "Application"!

Just add:

```
Create
View CLAIMS2_30
As
Select * from CLAIMS2
Where mydate > SYSDATE – 30
```

Time measure: set timing on;

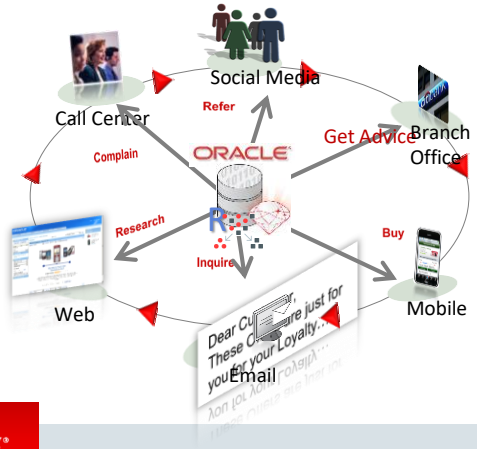
Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

Oracle Advanced Analytics

Real-Time Scoring, Predictions and Recommendations

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

```
Select prediction_probability(CLAS_DT_1_64, 'Yes'
  USING 7800 as bank_funds, 125 as checking_amount, 20 as
  credit_balance, 55 as age, 'Married' as marital_status,
  250 as MONEY_MONTLY_OVERDRAWN, 1 as house_ownership)
from dual;
```



Likelihood to respond:

Query Result

SQL | All Rows Fetched: 1 in 0 seconds

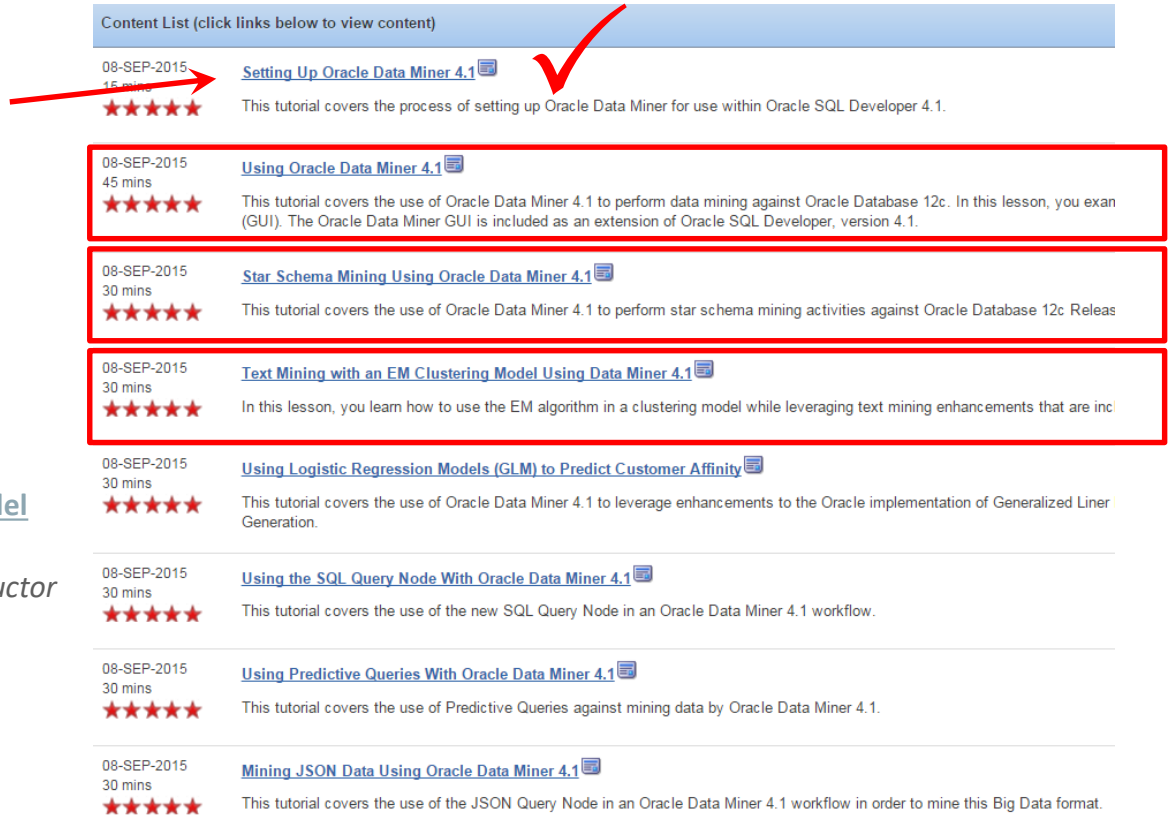
PREDICTION_PROB...
0.8382936507936...

Copyright © 2016, Oracle and/or its affiliates. All rights reserved. |

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*



Copyright © 2016 Oracle and/or its affiliates. All rights reserved. |

Copyright © 2016 Oracle and/or its affiliates. All rights reserved. |

ORACLE®