

ORACLE®



Oracle Data Mining 11g Release 2

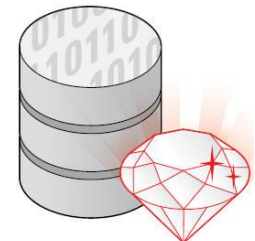
Charlie Berger

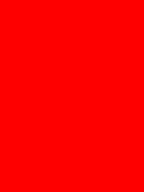
Sr. Director Product Management, Data Mining Technologies

Oracle Corporation

charlie.berger@oracle.com

ORACLE®



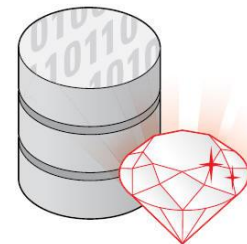


The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described for Oracle's products remains at the sole discretion of Oracle.



Outline

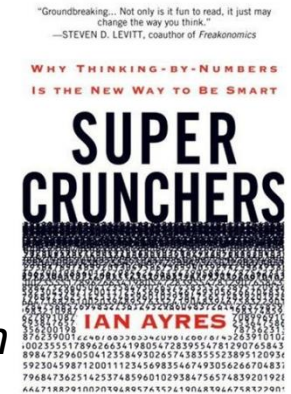
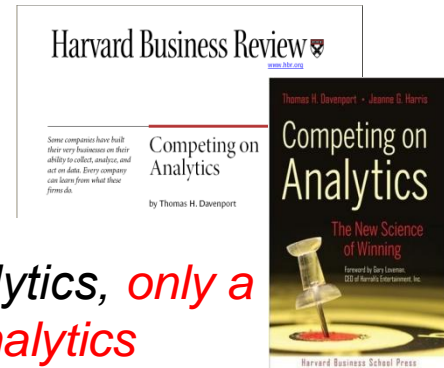
- Market Drivers
- Oracle Data Mining Option
- Positioning & Value Proposition
- Server APIs
 - Oracle Data Mining APIs (SQL & Java)
 - SQL Statistical Functions
- Graphical User Interfaces
 - Oracle Data Miner 11gR1 GUI
 - Oracle Data Miner 11gR2 GUI Preview
- Applications *Powered by* Oracle Data Mining
- Strategic Vision



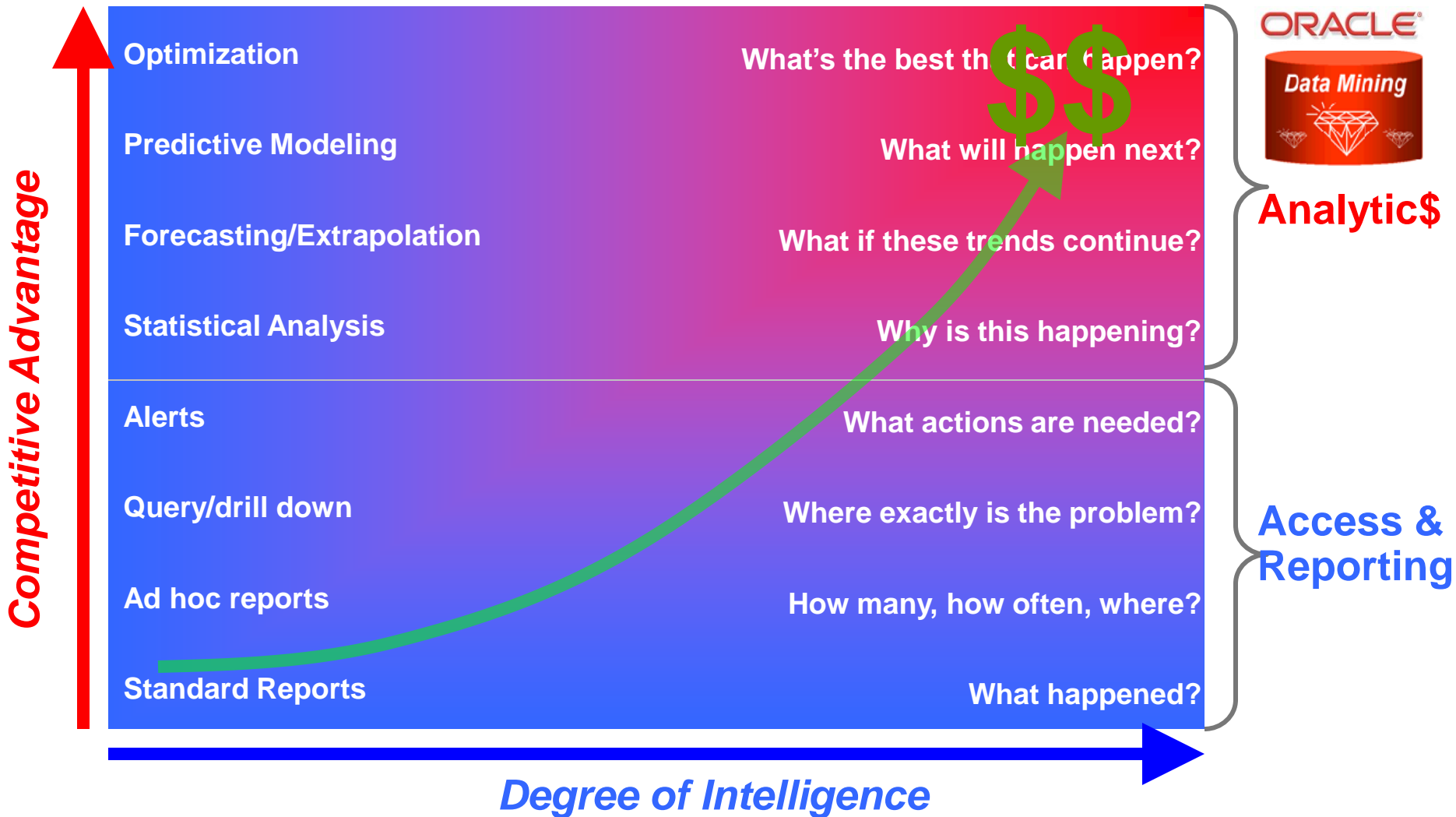
Market Drivers

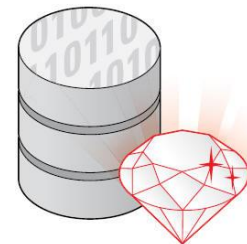
Analytics: Strategic and Mission Critical

- *Competing on Analytics*, by Tom Davenport
 - “Some companies have built their very businesses on their ability to collect, analyze, and act on data.”
 - “Although numerous organizations are embracing analytics, *only a handful have achieved this level of proficiency. But analytics competitors are the leaders in their varied fields—consumer products finance, retail, and travel and entertainment among them.*”
 - “Organizations are moving beyond query and reporting” - IDC 2006
- *Super Crunchers*, by Ian Ayers
 - “In the past, one could get by on intuition and experience. Times have changed. *Today, the name of the game is data.*”
—Steven D. Levitt, author of *Freakonomics*
 - “*Data-mining and statistical analysis have suddenly become cool.... Dissecting marketing, politics, and even sports, stuff th complex and important shouldn't be this much fun to read.*” —Wired



Competitive Advantage

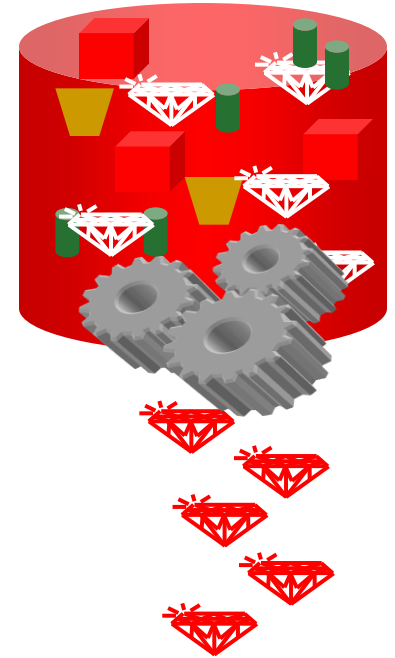




Oracle Data Mining Option

What is Data Mining?

- Automatically sifts through data to find hidden patterns, discover new insights, and make predictions
- Data Mining can provide valuable results:
 - Predict customer behavior (*Classification*)
 - Predict or estimate a value (*Regression*)
 - Segment a population (*Clustering*)
 - Identify factors more associated with a business problem (*Attribute Importance*)
 - Find profiles of targeted people or items (*Decision Trees*)
 - Determine important relationships and “market baskets” within the population (*Associations*)
 - Find fraudulent or “rare events” (*Anomaly Detection*)



Oracle Data Mining Example Use Cases

- **Retail**

- Customer segmentation
- Response modeling
- Recommend next likely product
- Profile high value customers

- **Banking**

- Credit scoring
- Probability of default
- Customer profitability
- Customer targeting

- **Insurance**

- Risk factor identification
- Claims fraud
- Policy bundling
- Employee retention

- **Higher Education**

- Alumni donations
- Student acquisition
- Student retention
- At-risk student identification

- **Healthcare**

- Patient procedure recommendation
- Patient outcome prediction
- Fraud detection
- Doctor & nurse note analysis

- **Life Sciences**

- Drug discovery & interaction
- Common factors in (un)healthy patients
- Cancer cell classification
- Drug safety surveillance

- **Telecommunications**

- Customer churn
- Identify cross-sell opportunities
- Network intrusion detection

- **Public Sector**

- Taxation fraud & anomalies
- Crime analysis
- Pattern recognition in military surveillance

- **Manufacturing**

- Root cause analysis of defects
- Warranty analysis
- Reliability analysis
- Yield analysis

- **Automotive**

- Feature bundling for customer segments
- Supplier quality analysis
- Problem diagnosis

- **Chemical**

- New compound discovery
- Molecule clustering
- Product yield analysis

- **Utilities**

- Predict power line / equipment failure
- Product bundling
- Consumer fraud detection

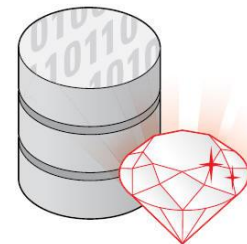


- Oracle Database #1
- Oracle Relational Database #1 in Revenue
- June 1999: acquires Thinking Machines Corporation's Darwin data mining technology and development team
- 10 years “stem celling analytics” into the Oracle Database
 - Designed advanced analytics into database kernel to leverage relational database strengths
 - Naïve Bayes and Association Rules—1st algorithms added
 - Leverages counting, conditional probabilities, and much more
- Now, analytical database platform
 - 12 cutting edge machine learning algorithms and 50+ statistical functions



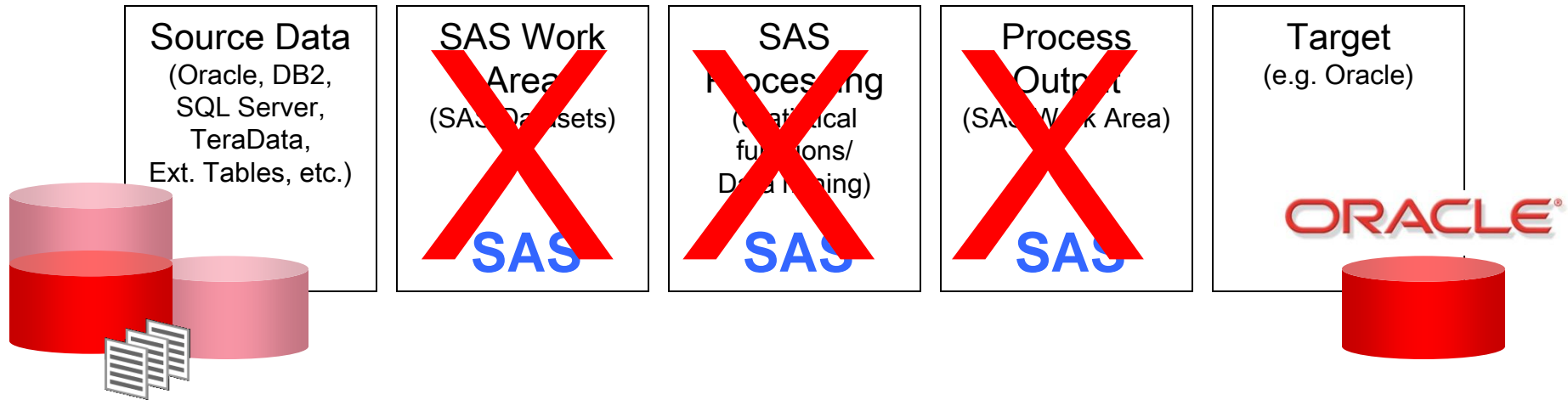
- Rather than add data mining as a bolt-on process outside the database kernel, DMT Dev. team, in collaboration with other ST Dev. teams, has embedded data mining functionality within the Oracle Database.
 - A data mining model is a schema object in the database, built via a PL/SQL API and scored via built-in SQL functions.
 - When building models, leverage existing scalable technology (e.g., parallel execution, bitmap indexes, aggregation techniques) and add new core database technology (e.g., recursion within the parallel infrastructure, IEEE float, etc.)
- True power of embedding within the database is evident when scoring models using built-in SQL functions (incl. Exadata)

```
select cust_id
from customers
where region = 'US'
and prediction_probability(churnmod, 'Y' using *) > 0.8;
```



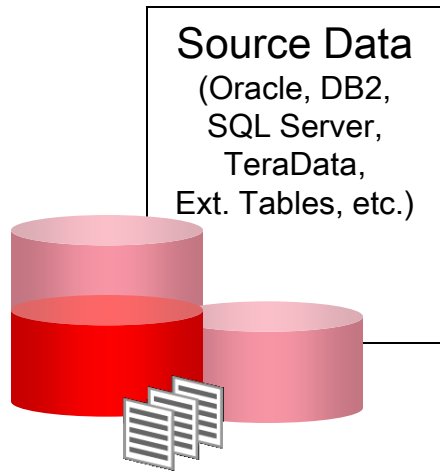
Positioning & Value Proposition

Traditional Analytics (SAS) Environment



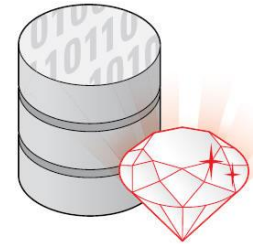
- SAS environment requires:
 - Data movement
 - Data duplication
 - Loss of security

Oracle Architecture

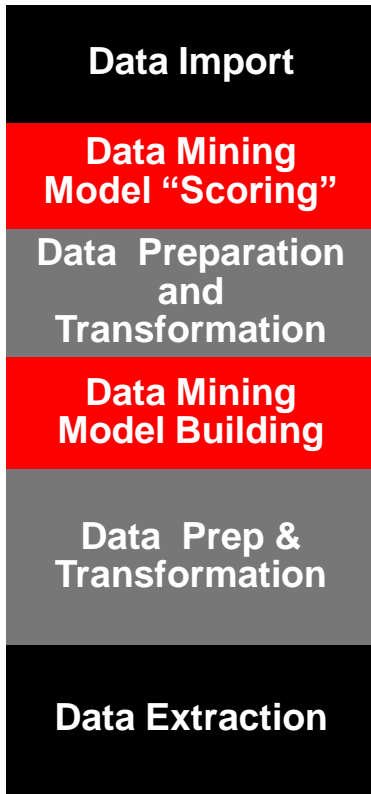


- Oracle environment:
 - Eliminates data movement
 - Eliminates data duplication
 - Preserves security

In-Database Data Mining



Traditional Analytics



Oracle Data Mining



Savings

Results

- Faster time for “Data” to “Insights”
- Lower TCO—Eliminates
 - Data Movement
 - Data Duplication
- Maintains Security

- Model “Scoring”
Data remains in the Database
- Embedded data preparation
- Cutting edge machine learning algorithms inside the SQL kernel of Database
- SQL—Most powerful language for data preparation and transformation
- Data remains in the Database

Hours, Days or Weeks

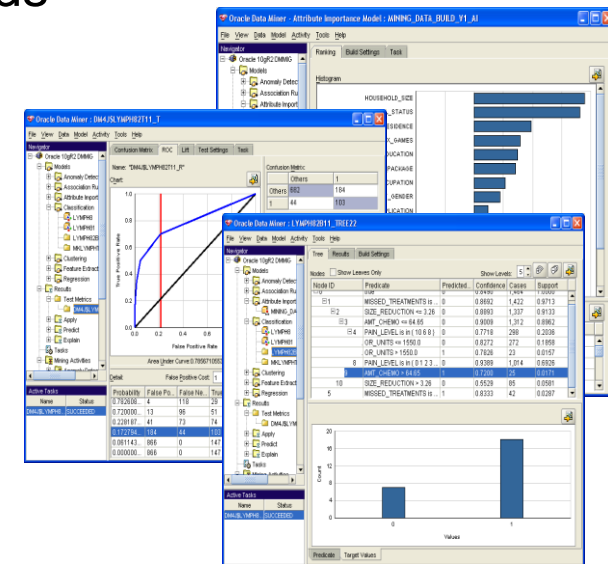
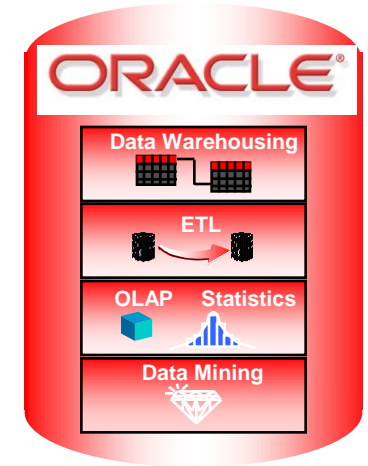


Secs. Mins or Hours

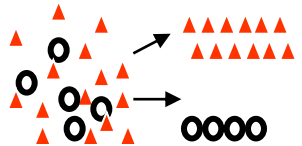
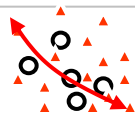
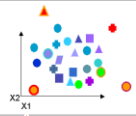
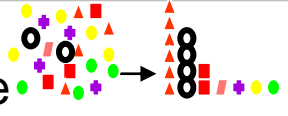
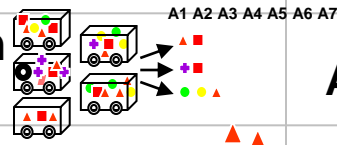
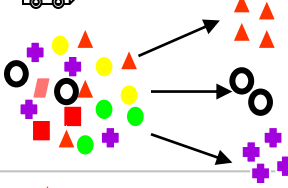
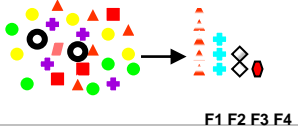


Oracle Data Mining 11g

- Data Mining API Functions (Server)
 - PL/SQL
 - Java
- Oracle Data Miner (GUI)
 - Simplified, guided data mining using wizards
- Wide range of DM algorithms (12)
 - Anomaly detection
 - Association rules (Market Basket analysis)
 - Attribute importance
 - Classification & regression
 - Clustering
 - Feature extraction (NMF)
 - Structured & unstructured data (text mining)
- Predictive Analytics
 - “1-click/automated data mining” (EXPLAIN, PREDICT, PROFILE)



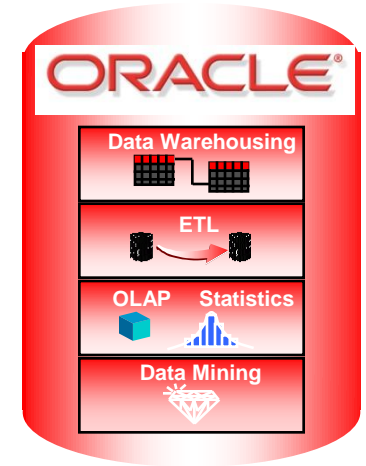
Oracle Data Mining Algorithms

Problem	Algorithm	Applicability
Classification 	Logistic Regression (GLM) Decision Trees Naïve Bayes Support Vector Machine	Classical statistical technique Popular / Rules / transparency Embedded app Wide / narrow data / text
Regression 	Multiple Regression (GLM) Support Vector Machine	Classical statistical technique Wide / narrow data / text
Anomaly Detection 	One Class SVM	Lack examples
Attribute Importance 	Minimum Description Length (MDL)	Attribute reduction Identify useful data Reduce data noise
Association Rules 	Apriori	Market basket analysis Link analysis
Clustering 	Hierarchical K-Means Hierarchical O-Cluster	Product grouping Text mining Gene and protein analysis
Feature Extraction 	NMF	Text analysis Feature reduction

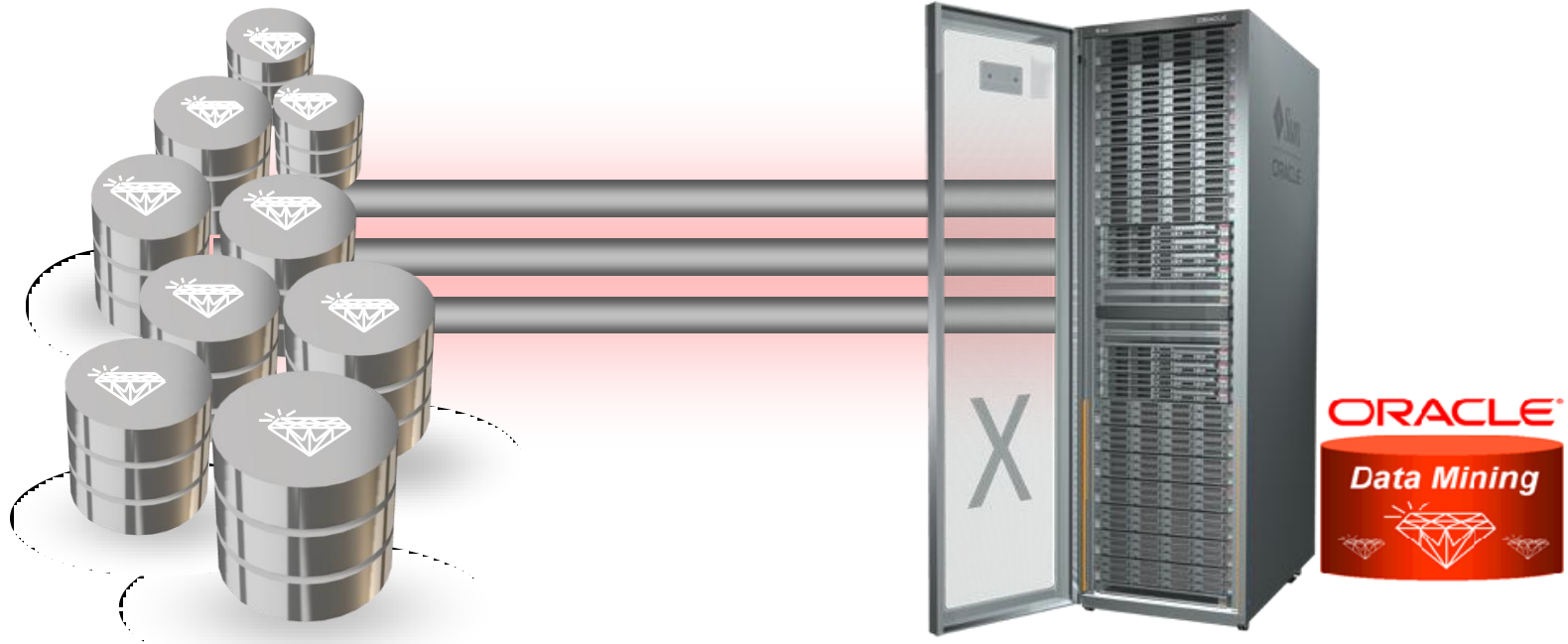
In-Database Data Mining

Advantages

- Data remains in the database
 - Fewer moving parts; shorter information latency
- ODM architecture provides greater
 - Performance, scalability, and security
- Best platform for developing PA/DM Applications
 - Straightforward inclusion within interesting and arbitrarily complex queries
 - “SELECT Customers WHERE Income > 100K, AND **PREDICTION_PROBABILITY(Buy Product A) > .85;**”
 - Enables pipelining of results without costly materialization
- Real-world scalability—available for mission critical appls
 - Fast scoring: 2.5 million records scored in 6 seconds on a single CPU system
 - Real-time scoring: 100 models on a single CPU: 0.085 seconds



Oracle Data Mining + Exadata



- In 11gR2, SQL predicates and Oracle Data Mining models are pushed to storage level for execution

For example, find the US customers likely to churn:

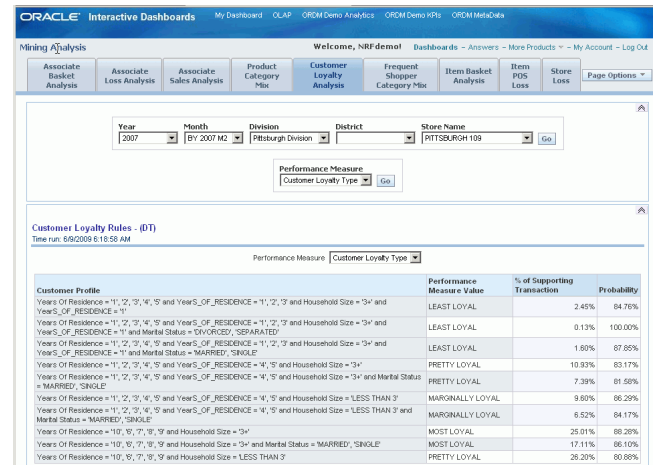
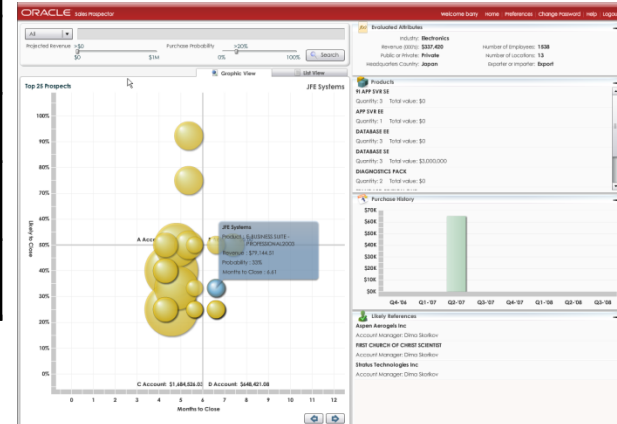
```
select cust_id
from customers
where region = 'US'
and prediction_probability(churnmod, 'Y' using *) > 0.8;
```

Scoring function executed in Exadata

Applications Powered by Oracle Data Mining

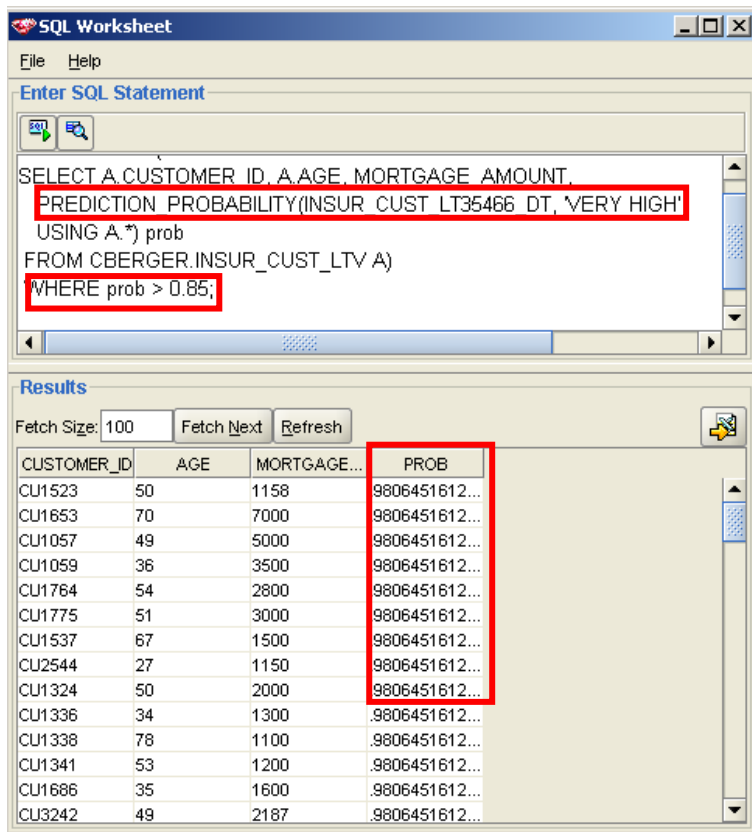
(Partial List as of September, 2009)

Application Name	Status
CRM OnDemand—Sales Prospector	GA—June '08
Oracle Retail Data Model	2Q09
Oracle Open World - Schedule Builder	OOW 2008 & 2009
Applications N...	TBD



Example: Simple, Predictive SQL

Select customers who are **more than 85% likely to be HIGH VALUE customers** & display their AGE & MORTGAGE_AMOUNT



The screenshot shows an Oracle SQL Worksheet window. The query editor contains the following SQL statement:

```
SELECT A.CUSTOMER_ID, A.AGE, MORTGAGE_AMOUNT,  
PREDICTION_PROBABILITY(INSUR_CUST_LT35468_DT, 'VERY HIGH'  
USING A.*) prob  
FROM CBERGER.INSUR_CUST_LTV A)  
WHERE prob > 0.85;
```

The results pane shows a table with the following columns: CUSTOMER_ID, AGE, MORTGAGE_AMOUNT, and PROB. The PROB column is highlighted in red, and the first few rows are also highlighted in red.

CUSTOMER_ID	AGE	MORTGAGE...	PROB
CU1523	50	1158	.9806451612...
CU1653	70	7000	.9806451612...
CU1057	49	5000	.9806451612...
CU1059	36	3500	.9806451612...
CU1764	54	2800	.9806451612...
CU1775	51	3000	.9806451612...
CU1537	67	1500	.9806451612...
CU2544	27	1150	.9806451612...
CU1324	50	2000	.9806451612...
CU1336	34	1300	.9806451612...
CU1338	78	1100	.9806451612...
CU1341	53	1200	.9806451612...
CU1686	35	1600	.9806451612...
CU3242	49	2187	.9806451612...

```
SELECT * from(  
SELECT A.CUSTOMER_ID, A.AGE,  
MORTGAGE_AMOUNT, PREDICTION_PROBABILITY  
(INSUR_CUST_LT4960_DT, 'VERY HIGH'  
USING A.*) prob  
FROM CBERGER.INSUR_CUST_LTV A)  
WHERE prob > 0.85;
```

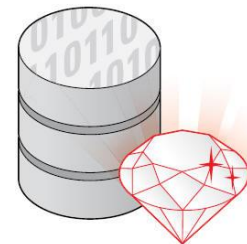
Fraud Prediction Demo

```
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 ('2 to 4', 'more than 4')))
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



Oracle Data Mining APIs

(SQL & Java)

More Interesting SQL

(Missing Value Imputation Example)

Select the 10 customers who are **most likely to attrite** based solely on: age, gender, annual_income, and zipcode. In addition, since annual_income is often missing, perform **null/missing value imputation** for the annual_income attribute using all of the customer demographics.

```
SELECT * FROM (
  SELECT cust_name, cust_contact_info,
         rank() over (ORDER BY
           PREDICTION_PROBABILITY(attrition_model, 'attrite'
            USING age, gender, zipcode,
              NVL(annual_income,
                PREDICTION(estim_income USING *)))
           as annual_income) DESC) as cust_rank
  FROM customers)
WHERE cust_rank < 11;
```

Example of Embedded Predictive SQL

Powers Next Generation Predictive Marketing Tools

Adobe Reader - [1166547846129out.pdf]

File Edit View Document Tools Window Help

Pages

UNIVERSITY.edu

Promoting educational excellence

University.edu
Management Administration
1234 State Street
Bloomington, IN 88888

December 19, 2006

Dear 101504,

Over the past year and a half the University.edu Alumni Association has been busy lending its expertise to the U.edu staff, orchestrating equipment donations, coordinating and participating in a major station clean-up, and establishing an alumni-student mentoring program.

E.EAA wants to raise funds to provide a stipend for the U.Edu general manager and to pay an engineer to repair vital equipment.

Because you have been so generous in the past we are first asking you, our alumni, for financial assistance. We suggest a donation of \$25, \$50, or whatever feels right. (If you or your company would like information on underwriting a program, please contact us at 303.575.3548.)

Your donation is fully tax deductible. We urge you to take your donation this tax year by sending in your donation by December 31, 2006.

In the coming months look for fundraising effort updates in our newsletter. And, be sure to mark your calendar now for our annual meeting and awards dinner -- Saturday evening, April 27, at U.ECP. You'll soon receive more information on this special fund-raiser.

Sincerely,

University.edu Fund Raising Staff

Attachments

Notice that the id of the likely responder is included.

Letter personalized with embedded predictive analytics

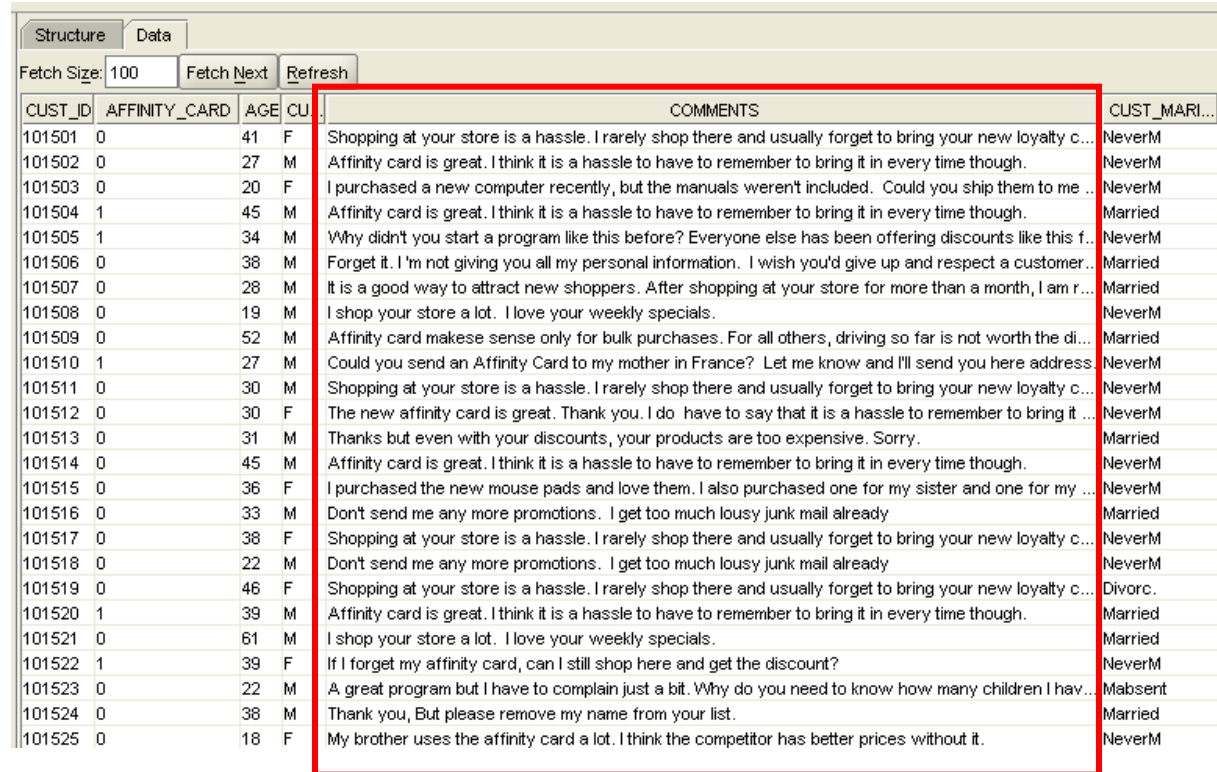
Embedded Data Preparation

Automatically applied when scoring

Attribute	Expression
income	salary + bonus
value	case when revenue < 100 then 'low' when revenue < 500 then 'med' else 'high' end
age	age / 100

Oracle Data Mining and Unstructured Data

- Oracle Data Mining mines unstructured i.e. “text” data
- Include free text and comments in ODM models
- Cluster and Classify documents
- Oracle Text used to preprocess unstructured text



The screenshot shows the Oracle Data Mining interface with a table of customer comments. The table has columns for CUST_ID, AFFINITY_CARD, AGE, GENDER, COMMENTS, and CUST_MARITAL_STATUS. The COMMENTS column contains various customer feedback messages, such as "Shopping at your store is a hassle. I rarely shop there and usually forget to bring your new loyalty c..." and "Affinity card is great. I think it is a hassle to have to remember to bring it in every time though." The table is highlighted with a red border.

CUST_ID	AFFINITY_CARD	AGE	CU	COMMENTS	CUST_MARI...
101501	0	41	F	Shopping at your store is a hassle. I rarely shop there and usually forget to bring your new loyalty c...	NeverM
101502	0	27	M	Affinity card is great. I think it is a hassle to have to remember to bring it in every time though.	NeverM
101503	0	20	F	I purchased a new computer recently, but the manuals weren't included. Could you ship them to me ...	NeverM
101504	1	45	M	Affinity card is great. I think it is a hassle to have to remember to bring it in every time though.	Married
101505	1	34	M	Why didn't you start a program like this before? Everyone else has been offering discounts like this f...	NeverM
101506	0	38	M	Forget it. I'm not giving you all my personal information. I wish you'd give up and respect a customer...	Married
101507	0	28	M	It is a good way to attract new shoppers. After shopping at your store for more than a month, I am r...	Married
101508	0	19	M	I shop your store a lot. I love your weekly specials.	NeverM
101509	0	52	M	Affinity card make sense only for bulk purchases. For all others, driving so far is not worth the di...	Married
101510	1	27	M	Could you send an Affinity Card to my mother in France? Let me know and I'll send you here address	NeverM
101511	0	30	M	Shopping at your store is a hassle. I rarely shop there and usually forget to bring your new loyalty c...	NeverM
101512	0	30	F	The new affinity card is great. Thank you. I do have to say that it is a hassle to remember to bring it ...	NeverM
101513	0	31	M	Thanks but even with your discounts, your products are too expensive. Sorry.	Married
101514	0	45	M	Affinity card is great. I think it is a hassle to have to remember to bring it in every time though.	NeverM
101515	0	36	F	I purchased the new mouse pads and love them. I also purchased one for my sister and one for my ...	NeverM
101516	0	33	M	Don't send me any more promotions. I get too much lousy junk mail already	Married
101517	0	38	F	Shopping at your store is a hassle. I rarely shop there and usually forget to bring your new loyalty c...	NeverM
101518	0	22	M	Don't send me any more promotions. I get too much lousy junk mail already	NeverM
101519	0	46	F	Shopping at your store is a hassle. I rarely shop there and usually forget to bring your new loyalty c...	Divorc.
101520	1	39	M	Affinity card is great. I think it is a hassle to have to remember to bring it in every time though.	Married
101521	0	61	M	I shop your store a lot. I love your weekly specials.	Married
101522	1	39	F	If I forget my affinity card, can I still shop here and get the discount?	NeverM
101523	0	22	M	A great program but I have to complain just a bit. Why do you need to know how many children I hav...	Mabsent
101524	0	38	M	Thank you, But please remove my name from your list.	Married
101525	0	18	F	My brother uses the affinity card a lot. I think the competitor has better prices without it.	NeverM

Performing a Moving Average

The following query computes the moving average of the sales amount between the current month and the previous three months:

```
SQL> --SQL>
SQL> SELECT
  month, SUM(amount) AS month_amount,
  AVG(SUM(amount)) OVER
  (ORDER BY month ROWS BETWEEN 3
  PRECEDING AND CURRENT ROW)
  AS moving_average
FROM all_sales
GROUP BY month
ORDER BY month;
```

MONTH	MONTH_AMOUNT	MOVING_AVERAGE
1	58704.52	58704.52
2	28289.3	43496.91
3	20167.83	35720.55
4	50082.9	39311.1375
5	17212.66	28938.1725
6	31128.92	29648.0775
7	78299.47	44180.9875
8	42869.64	42377.6725
9	35299.22	46899.3125
10	43028.38	49874.1775
11	26053.46	36812.675
12	20067.28	31112.085

12 rows selected.

Complex SQL Transform

- For each customer, compute the amount sold to customer in the past three months and three months prior to that.
- If the increase is greater than 25%, mark the customer as **G**(rowing).
- If the decrease is greater than 25%, mark the customer as **S**(hrinking).
- Otherwise, mark the customer as **U**(nchanged).
- Add special handling for old_sales of 0 by replacing the denominator with new_sales/2, which will yield an increase of more than 25% in the calculation, which is the desired result.

#2

```
select
cust_id,
case when changed_sales > 0.25 then 'G'
      when changed_sales < -0.25 then 'S'
      else 'U' end as cust_value
from (
select
cust_id,
(new_sales - old_sales) /
  decode(old_sales, 0,
         decode(new_sales, 0, 1, new_sales/2), old_sales)
  as changed_sales
from (
select
cust_id,
sum(case when time_id < add_months((select max(time_id) from sh.sales),-3)
      then amount_sold else 0 end) as old_sales,
sum(case when time_id >= add_months((select max(time_id) from sh.sales),-3)
      then amount_sold else 0 end) as new_sales
from sh.sales
where time_id >= add_months((select max(time_id) from sh.sales),-6)
group by cust_id
)
);
```

In-Database Analytics Example

Launch & Evaluate a Marketing Campaign

1. Given a previously built response model, ... **predict who will respond to a campaign, ...and why**

2.find out **how much each customer spent 3 months before and after the campaign**

3.how much for **just DVDs?**

4. Is the success **statistically significant?**

```
select responder, cust_region, count(*) as cnt,
       sum(post_purchase - pre_purchase) as tot_increase,
       avg(post_purchase - pre_purchase) as avg_increase,
       stats_t_test_paired(pre_purchase, post_purchase) as
       significance
from (
  select cust_name,
         prediction(campaign_model using *) as responder,
         sum(case when purchase_date < 15-Apr-2005 then
                purchase_amt else 0 end) as pre_purchase,
         sum(case when purchase_date >= 15-Apr-2005 then
                purchase_amt else 0 end) as post_purchase
  from customers, sales, products@PRODDB
  where sales.cust_id = customers.cust_id
        and purchase_date between 15-Jan-2005 and 14-Jul-2005
        and sales.prod_id = products.prod_id
        and contains(prod_description, 'DVD') > 0
  group by cust_id, prediction(campaign_model using *) )
group by rollup responder, cust_region order by 4 desc;
```

Real-time Prediction

with

```
records as (select
  78000 SALARY,
  250000 MORTGAGE_AMOUNT,
  6 TIME_AS_CUSTOMER,
  12 MONTHLY_CHECKS_WRITTEN,
  55 AGE,
  423 BANK_FUNDS,
  'Married' MARITAL_STATUS,
  'Nurse' PROFESSION,
  'M' SEX,
  4000 CREDIT_CARD_LIMITS,
  2 N_OF_DEPENDENTS,
  1 HOUSE_OWNERSHIP from dual)
```

```
select s.prediction prediction, s.probability probability
```

```
from (
```

```
  select PREDICTION_SET(INSUR_CUST_LT48172_DT, 1 USING *) pset
  from records) t, TABLE(t.pset) s;
```

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

PREDICTION	PROBABILITY
HIGH	.65123504738232096

Prediction Multiple Models/Optimization

```
➤ with records as (select
  178255 ANNUAL_INCOME,
  30 AGE,
  'Bach.' EDUCATION,
  'Married' MARITAL_STATUS,
  'Male' SEX,
  70 HOURS_PER_WEEK,
  98 PAYROLL_DEDUCTION from dual)
select t.*
from (
  select 'CAR_MODEL' MODEL, s1.prediction prediction, s1.probability probability, s1.probability*25000 as
  expected_revenue from (
    select PREDICTION_SET(NBMODEL_JDM, 1 USING *) pset
    from records ) t1, TABLE(t1.pset) s1
  UNION
  select 'MOTOCYCLE_MODEL' MODEL, s2.prediction prediction, s2.probability probability, s1.probability*2000 as
  expected_revenue from (
    select PREDICTION_SET(ABNMODEL_JDM, 1 USING *) pset
    from records ) t2, TABLE(t2.pset) s2
  UNION
  select 'TRICYCLE_MODEL' MODEL, s3.prediction prediction, s3.probability probability, s1.probability*50 as
  expected_revenue from (
    select PREDICTION_SET(TREEMODEL_JDM, 1 USING *) pset
    from records ) t3, TABLE(t3.pset) s3
  UNION
  select 'BICYCLE_MODEL' MODEL, s4.prediction prediction, s4.probability probability, s1.probability*200 as
  expected_revenue from (
    select PREDICTION_SET(SVMCMODEL_JDM, 1 USING *) pset
    from records ) t4, TABLE(t4.pset) s4
  ) t
order by t.expected_revenue desc;
```

**On-the-fly, multiple models;
then sort by expected revenues**

Integration with Oracle BI EE

The screenshot displays the Siebel Analytics Administration Tool interface, divided into three main panels: Presentation, Business Model and Mapping, and Physical. The Presentation panel shows a hierarchy for 'CD_BUYERS' with 'DIM' and 'FACT' folders. 'KEY_FACTOR' and 'IMPORTANCE' are circled in red. The Business Model and Mapping panel shows a similar hierarchy with 'Sources' and 'Sales Facts' folders. The Physical panel shows a database schema for 'Oracle_10gR2' with 'Oracle_cberger' and 'CBERGER' folders. 'KEY_CD_BUYER_ATTRIBUTES' is circled in red. A blue callout box points to the 'KEY_CD_BUYER_ATTRIBUTES' table in the Physical panel, containing the text: 'Oracle Data Mining results available to Oracle BI EE administrators'. Another blue callout box points to the 'AFFINITY_CARD' table in the Presentation panel, containing the text: 'Oracle BI EE defines results for end user presentation'. The top of the window shows the title bar '(Online) Siebel Analytics Administration Tool - AnalyticsWeb' and a menu bar with 'File', 'Edit', 'View', 'Manage', 'Tools', 'Window', and 'Help'. The bottom left corner has the text 'For Help, press F1'.

Presentation

- CD_BUYERS
 - DIM
 - KEY_FACTOR
 - IMPORTANCE
 - FACT
 - RANK
 - Paint
 - Paint Exec
 - Sales_History
 - CUSTOMERS_SH_LIKELY_TO_RESPOND
 - SUPPLEMENTARY_DEMOGRAPHICS
 - CUST_ID
 - EDUCATION
 - OCCUPATION
 - HOUSEHOLD_SIZE
 - YRS_RESIDENCE
 - AFFINITY_CARD
 - BULK_PACK_DISKETTE
 - FLAT_PANEL_MONITOR
 - HOME_THEATER_PACKAGE
 - BOOKKEEPING_APPLICATION
 - PRINTER_SUPPLIES
 - Y_BOX_GAMES
 - OS_DOC_SET_KANJI
 - COMMENTS

Business Model and Mapping

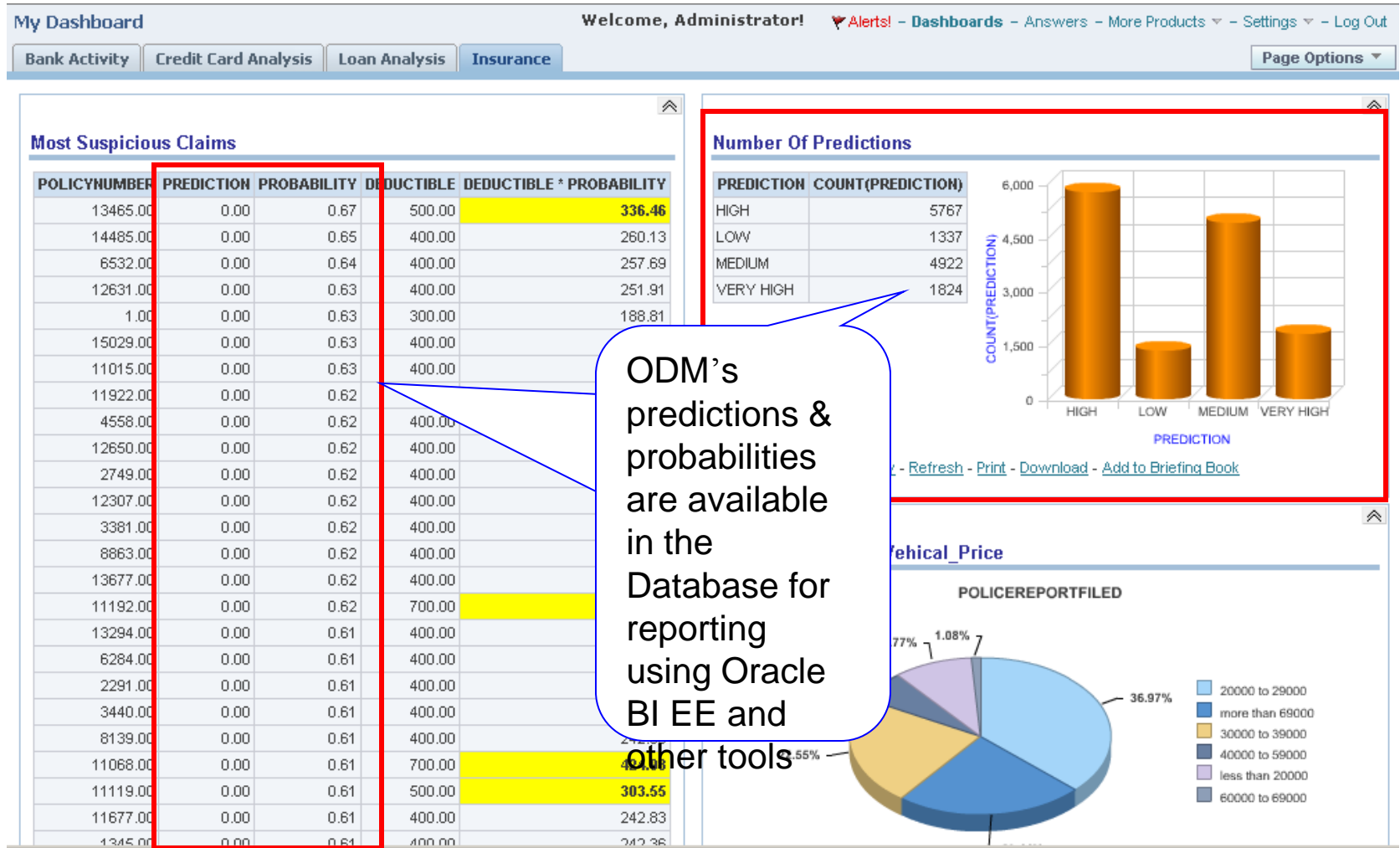
- CD_BUYERS
 - DIM
 - Sources
 - KEY_FACTOR
 - IMPORTANCE
 - FACT
 - RANK
 - Paint
 - MarketDim
 - PeriodDim
 - ProductDim
 - Markets
 - Periods
 - Products
 - Sales Facts
 - Sales_History
 - CUSTOMERS_SH_LIKELY_TO_RESPOND
 - SUPPLEMENTARY_DEMOGRAPHICS
 - Sources
 - CUST_ID
 - EDUCATION
 - OCCUPATION
 - HOUSEHOLD_SIZE
 - YRS_RESIDENCE
 - AFFINITY_CARD
 - BULK_PACK_DISKETTE
 - FLAT_PANEL_MONITOR
 - HOME_THEATER_PACKAGE
 - BOOKKEEPING_APPLICATION
 - PRINTER_SUPPLIES
 - Y_BOX_GAMES
 - OS_DOC_SET_KANJI
 - COMMENTS

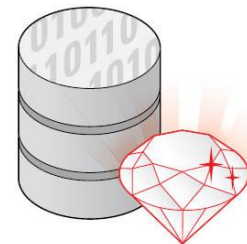
Physical

- Oracle_10gR2
 - Oracle_cberger
 - cberger_pool
 - CBERGER
 - CD_BUYERS44318_SIEBEL_A
 - CD_BUYERS_APPLY394639710_A
 - CD_BUYERS_PREDICT_A
 - CDBUYER_SEGMENT_PROFILES
 - CDBUYER_SEGMENT_STATISTICS
 - CDBUYER_SEGMENTS
 - CUSTOMERS546911500_A
 - KEY_CD_BUYER_ATTRIBUTES
 - SQL_Pair
 - XLS_Fore

Example

Better Information for OBI EE Reports and Dashboards

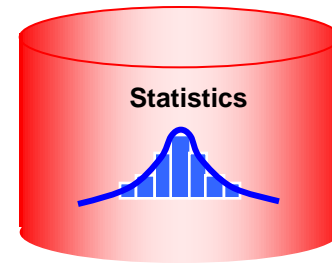




Oracle SQL Statistical Functions

(Free in Every Oracle Database)

11g Statistics & SQL Analytics

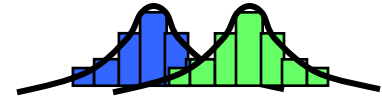


- Ranking functions
 - rank, dense_rank, cume_dist, percent_rank, ntile
- Window Aggregate functions (moving and cumulative)
 - Avg, sum, min, max, count, variance, stddev, first_value, last_value
- LAG/LEAD functions
 - Direct inter-row reference using offsets
- Reporting Aggregate functions
 - Sum, avg, min, max, variance, stddev, count, ratio_to_report
- Statistical Aggregates
 - Correlation, linear regression family, covariance
- Linear regression
 - Fitting of an ordinary-least-squares regression line to a set of number pairs.
 - Frequently combined with the COVAR_POP, COVAR_SAMP, and CORR functions

Descriptive Statistics

- DBMS_STAT_FUNCS: summarizes numerical columns of a table and returns count, min, max, range, mean, stats_mode, variance, standard deviation, median, quantile values, +/- n sigma values, top/bottom 5 values
- Correlations
 - Pearson's correlation coefficients, Spearman's and Kendall's (both nonparametric).
- Cross Tabs
 - Enhanced with % statistics: chi squared, phi coefficient, Cramer's V, contingency coefficient, Cohen's kappa
- Hypothesis Testing
 - Student t-test, F-test, Binomial test, Wilcoxon Signed Ranks test, Chi-square, Mann Whitney test, Kolmogorov-Smirnov test, One-way ANOVA
- Distribution Fitting
 - Kolmogorov-Smirnov Test, Anderson-Darling Test, Chi-Squared Test, Normal, Uniform, Weibull, Exponential

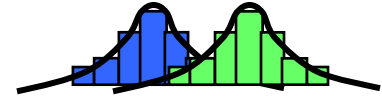
Split Lot A/B Offer testing



- Offer “A” to one population and “B” to another
- Over time period “t” calculate **median** purchase amounts of customers receiving offer A & B
- Perform **t-test** to compare
- If statistically significantly better results achieved from one offer over another, offer everyone higher performing offer



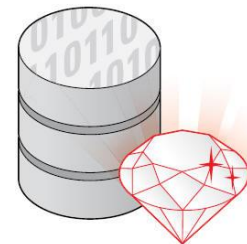
Independent Samples T-Test (Pooled Variances)



- Query compares the mean of AMOUNT_SOLD between MEN and WOMEN within CUST_INCOME_LEVEL ranges

```
SELECT substr(cust_income_level,1,22) income_level,  
       avg(decode(cust_gender, 'M', amount_sold, null)) sold_to_men,  
       avg(decode(cust_gender, 'F', amount_sold, null)) sold_to_women,  
       stats_t_test_indep(cust_gender, amount_sold, 'STATISTIC', 'F')  
       t_observed,  
       stats_t_test_indep(cust_gender, amount_sold) two_sided_p_value  
FROM sh.customers c, sh.sales s  
WHERE c.cust_id=s.cust_id  
GROUP BY rollup(cust_income_level)  
ORDER BY 1;
```

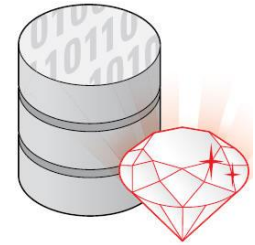
SQL Worksheet



Oracle Data Miner 11gR1 (GUI)

[ODM'r "Classic"]

Oracle Data Miner 11gR1 GUI



Oracle Data Miner - Table : INSUR_CUST_LTY

File View Data Activity Tools Help

Navigator

- Oracle_CB
 - Mining Activities
 - Anomaly Detection
 - Association Rules
 - Attribute Importance
 - Classification
 - Clustering
 - Feature Extraction
 - Regression
 - Data Sources
 - CBERGER
 - Views
 - Tables
 - AAA_INSUR_CLAIMS
 - AAA_INSUR_CLAIMS_VKS20_52
 - AAA_INSUR_CLAIMS171536183...
 - AAA_INSUR_CLAIMS765132333...
 - AAA_INSUR_CLAIMS847455013...
 - AAA_INSUR_DOSG_FEAUM
 - AAA_INSUR...
 - AGE_GENI
 - BANK_DA
 - BANK1
 - BLADDER
 - BOSTON
 - BOSTON_1
 - BPRESS
 - BPRESS_2
 - BPRESS_C
 - BPRESS_C
 - bpress128
 - BRAIN_TUI
 - BRAIN_TUI
 - BRAIN_TUI
 - CAR_STA
 - CD_BUYER
 - CD_BUYER
 - CD_BUYER
 - CD_BUYER

Structure Data

Fetch Size: 100 Fetch Next Refresh

CUSTOMER_ID	LAST	STA...	REGION	SEX	PROFESSION	BUY_INS...	AGE	HAS_CHILD...	LTV_BIN	SALARY	LTV	CAR...	MORTGAGE...	MARITAL_ST...	N_OF_DEPE...
CU2409	MARGIT	CA	West	F	Not specified	No	57	1	LOW	5887	13921.75	1	0	DIVORCED	1
CU2411	GRAHAM	NY	NorthEast	M	Construction...	No	60	1	HIGH	69335	23633.75	1	800	MARRIED	3
CU2412	KALEIGH	FL	South	F	Medical Doctor	Yes	44	1	HIGH	58342	22485.5	1	1584	MARRIED	1
CU2413	MITCH	NY	NorthEast	M	Accountant	Yes	61	1	MEDIUM	67061	16965.25	1	0	SINGLE	0
CU2414	ELDEN	CA	West	M	Software En...	No	68	0	HIGH	67489	29672.25	1	4000	DIVORCED	3
CU2416	SANDY	NY	NorthEast	M	PROF-22	No	32	0	MEDIUM	63963	20190.75	1	0	SINGLE	0
CU2417	ALEX	NY	NorthEast	M	Sales Repre...	No	27	0	MEDIUM	64964	19941	1	0	SINGLE	0
CU2418	ALI	MI	Midwest	F	Construction...	No	47	0	HIGH	58771	27892.75	1	3149	DIVORCED	1
CU2420	MERLE	NY	NorthEast	M	Clerical	No	38	1	MEDIUM	60946	19536.5	1	2000	DIVORCED	3
CU2421	ASLEY	CA	West	F	Clerical	No	52	0	HIGH	67694	29123.5	1	4000	DIVORCED	2
CU2422	GUADAL...	MI	Midwest	F	PROF-40	Yes	52	0	HIGH	61781	29145.25	1	1645	MARRIED	1
CU2423	LIVIA	NY	NorthEast	F	IT Staff	No	64	1	HIGH	63875	25868.75	1	6000	WIDOWED	1
CU3100	SIOBHAN	MI	Midwest	F	PROF-15	No	62	1	MEDIUM	65359	17539.75	1	11000	WIDOWED	6
CU3101	REVA	MI	Midwest	F	Software En...	Yes	43	1	HIGH	61513	23178.25	1	1500	MARRIED	1
CU3102	WELDON	NV	Southwest	M	Bank Teller	No	43	0	MEDIUM	65379	21644.75	1	0	SINGLE	0
CU3104	MONA	NY	NorthEast	F	Truck Driver	No	23	1	MEDIUM	66902	16525.5	1	5000	MARRIED	5
														OTHER	1

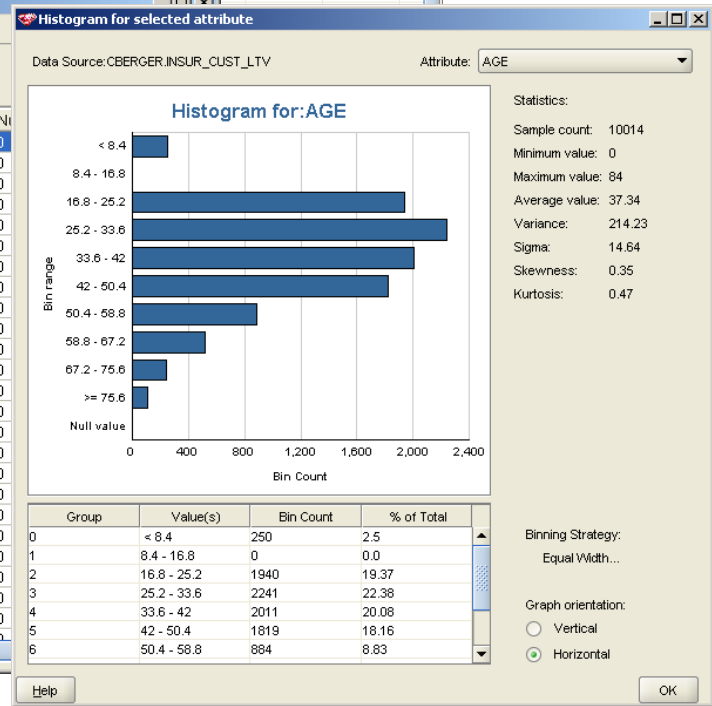
Data Summarization Viewer: CBERGER.INSUR_CUST_LTY

File Help

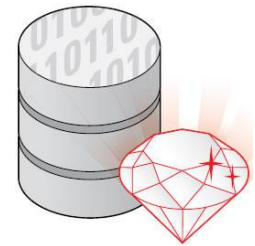
Sample Count: 10014

Attribute Count: 31

Name	Mining ...	Attrib...	Average	Max	Min	Variance	Ni
AGE	numerical	NUMBER	37.34	84	0	214.23	0
BANK_FUNDS	numerical	NUMBER	2,562.97	105,000	0	23,068,...	0
BUY_INSURANCE	categori...	VARCH...					0
CAR_OWNERSHIP	categori...	NUMBER	0.93	1	0	0.06	0
CHECKING_AMOUNT	numerical	NUMBER	1,068.71	24,760	25	10,100,...	0
CREDIT_BALANCE	numerical	NUMBER	2,793.31	571,088	0	341,598,...	0
CREDIT_CARD_LIMITS	numerical	NUMBER	1,266.43	5,000	500	673,981.1	0
CUSTOMER_ID	categori...	VARCH...					0
FIRST	categori...	VARCH...					0
HAS_CHILDREN	categori...	NUMBER	0.49	1	0	0.25	0
HOUSE_OWNERSHIP	categori...	NUMBER	0.8	2	0	0.26	0
LAST	categori...	VARCH...					0
LTV	numerical	FLOAT	22,260.43	47,501.75	0	46,480,...	0
LTV_BIN	categori...	VARCH...					0
MARITAL_STATUS	categori...	VARCH...					0
MONEY_MONTHLY_OVERDRA...	numerical	FLOAT	53.63	93.64	-77.34	9.42	0
MONTHLY_CHECKS_WRITTEN	numerical	NUMBER	4.55	18	0	23.59	0
MORTGAGE_AMOUNT	numerical	NUMBER	2,005.7	90,000	0	10,274,...	0
N_MORTGAGES	categori...	NUMBER	0.8	2	0	0.26	0
N_OF_DEPENDENTS	numerical	NUMBER	2.06	6	0	2.4	0
N_TRANS_ATM	numerical	NUMBER	2.87	8	0	3.48	0
N_TRANS_KIOSK	numerical	NUMBER	1.95	10	0	3.66	0
N_TRANS_TELLER	numerical	NUMBER	1.8	9	0	2.17	0
N_TRANS_WEB_BANK	numerical	NUMBER	1,331.31	45,000	0	3,873,1...	0
PROFESSION	categori...	VARCH...					0



Oracle Data Miner 11gR1 GUI



New Activity Wizard - Step 3 of 4: Data Usage

Review Data Usage Settings

Select the target column, and review the column settings. You can change the column settings to better match your understanding of the data. The default settings have been determined for each column based on the activity type and the characteristics of the data. The options of changing input and mining type vary based on the algorithm chosen. Click Help for more details.

[Data Summary](#)

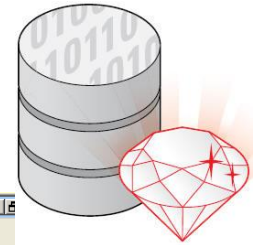
Name	Alias	Target	Input	Data Type	Mining Type	Sparsity
CBERGER.INSUR_C...		<input type="radio"/>	<input checked="" type="checkbox"/>			
AGE	AGE	<input type="radio"/>	<input checked="" type="checkbox"/>	NUMBER	numerical	<input type="checkbox"/>
BANK_FUNDS	BANK_FUNDS	<input type="radio"/>	<input checked="" type="checkbox"/>	NUMBER	numerical	<input type="checkbox"/>
BUY_INSURANCE	BUY_INSURANCE	<input checked="" type="radio"/>	<input checked="" type="checkbox"/>	VARCHAR2	categorical	<input type="checkbox"/>
CAR_OWNERSHIP	CAR_OWNERSHIP	<input type="radio"/>	<input checked="" type="checkbox"/>	NUMBER	categorical	<input type="checkbox"/>
CHECKING_AMOU...	CHECKING_AMOU...	<input type="radio"/>	<input checked="" type="checkbox"/>	NUMBER	numerical	<input type="checkbox"/>
CREDIT_BALANCE	CREDIT_BALANCE	<input type="radio"/>	<input checked="" type="checkbox"/>	NUMBER	numerical	<input type="checkbox"/>
CREDIT_CARD_LI...	CREDIT_CARD_LI...	<input type="radio"/>	<input checked="" type="checkbox"/>	NUMBER	numerical	<input type="checkbox"/>
CUSTOMER_ID	CUSTOMER_ID	<input type="radio"/>	<input type="checkbox"/>	VARCHAR2	categorical	<input type="checkbox"/>
FIRST	FIRST	<input type="radio"/>	<input type="checkbox"/>	VARCHAR2	categorical	<input type="checkbox"/>
HAS_CHILDREN	HAS_CHILDREN	<input type="radio"/>	<input checked="" type="checkbox"/>	NUMBER	categorical	<input type="checkbox"/>
HOUSE_OWNERS...	HOUSE_OWNERS...	<input type="radio"/>	<input checked="" type="checkbox"/>	NUMBER	categorical	<input type="checkbox"/>
LAST	LAST	<input type="radio"/>	<input type="checkbox"/>	VARCHAR2	categorical	<input type="checkbox"/>
LTV	LTV	<input type="radio"/>	<input checked="" type="checkbox"/>	NUMBER	numerical	<input type="checkbox"/>
LTV_BIN	LTV_BIN	<input type="radio"/>	<input checked="" type="checkbox"/>	VARCHAR2	categorical	<input type="checkbox"/>
MARITAL_STATUS	MARITAL_STATUS	<input type="radio"/>	<input checked="" type="checkbox"/>	VARCHAR2	categorical	<input type="checkbox"/>
MONEY_MONTHLY_...	MONEY_MONTHLY_...	<input type="radio"/>	<input checked="" type="checkbox"/>	NUMBER	numerical	<input type="checkbox"/>
MONTHLY CHEC...	MONTHLY CHEC...	<input type="radio"/>	<input checked="" type="checkbox"/>	NUMBER	numerical	<input type="checkbox"/>

Include All Exclude All

Help < Back Next > Finish Cancel

Oracle Data Miner guides the analyst through the data mining process

Oracle Data Miner 11gR1 GUI



Oracle Data Miner - Mining Activity : INSUR_CUST_LTV833819485_BA

Name: INSUR_CUST_LTV833819485_BA
 Type: Decision Tree Mining Activity
 Case Table: CBERGER.INSUR_CUST_LTV
 Unique Identifier: CUSTOMER_ID
 Target: CBERGER.INSUR_CUST_LTV.LTV_BIN
 Comment:

Activity Steps:

- Sample
This step samples the mining data. Although not normally required, you can use this step to sample the data.
- Split
This transformation step splits the mining data into build and test data.
- Build
This step builds the mining model. To complete this step manually, click the Build Data button.
- Test Metrics
This step creates a test metric result. To complete this step manually, click the Test Data button.

Activity: INSUR_CUST_LTV833819485_BA

Target Attribute: LTV_BIN

Node ID	Predicate	Confidence	Cases	Support
0	true	0.4849	8,722	1.0000
1	MORTGAGE_AMOUNT <= 0.5	0.5780	2,154	0.2470
21	MORTGAGE_AMOUNT > 0.5	0.6200	6,568	0.7530
22	HOUSE_OWNERSHIP is in 2	0.7921	433	0.0496
60	AGE <= 20.5	0.5000	12	0.0014
23	AGE > 20.5	0.8052	421	0.0483
24	N_OF_DEPENDENTS <= 3.5	0.8991	347	0.0398
25	AGE <= 26.5	0.5263	38	0.0044
61	HAS_CHILDREN is in 0	1.0000	12	0.0014
62	HAS_CHILDREN is in 1	0.7692	26	0.0030
63	AGE > 26.5	0.9515	309	0.0354
64	N_OF_DEPENDENTS > 3.5	0.6351	74	0.0085
26	HOUSE_OWNERSHIP is in 1	0.6500	6,135	0.7034
27	N_OF_DEPENDENTS <= 1.5	0.7774	2,879	0.3301
28	HAS_CHILDREN is in 0	0.6519	1,086	0.1245
29	SALARY <= 64588.0	0.8759	548	0.0628
65	TIME_AS_CUSTOMER is in { 2 4 5 }	0.5875	80	0.0092
66	TIME_AS_CUSTOMER is in 1	0.9615	468	0.0537
80	SALARY > 64588.0	0.5661	538	0.0617
67	AGE <= 23.5	0.6250	16	0.0018
68	AGE > 23.5	0.5824	522	0.0598
31	HAS_CHILDREN is in 1	0.8533	1,793	0.2056
32	SALARY <= 79990.0	0.8778	1,711	0.1962
69	SALARY <= 59841.0	0.5347	346	0.0397
70	SALARY > 59841.0	0.9648	1,365	0.1565
71	SALARY > 79990.0	0.6585	82	0.0094
33	N OF DEPENDENTS > 1.5	0.5375	3,256	0.3733

Predicted Target Value: VERY HIGH
 Support (%): 6.17
 Confidence (%): 56.51
 Cases: 538
 Level: 5

Split Rules: Full Rule Surrogate

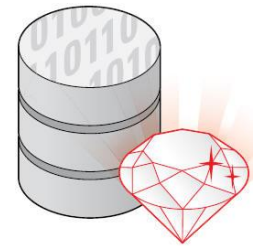
MORTGAGE_AMOUNT > 0.5 AND
 HOUSE_OWNERSHIP is in 1 AND
 N_OF_DEPENDENTS <= 1.5 AND
 HAS_CHILDREN is in 0 AND
 SALARY > 64588.0

Predicate Target Values

Oracle Data Mining builds a model that differentiates HI_VALUE_CUSTOMERS from others

Oracle Data Mining + OBI EE

Targeting High Value Customers



Oracle Data Mining creates a prioritized list of customer who likely to be high value

Activity: INSUR_CUST_LTV1330924475007_AA: Result

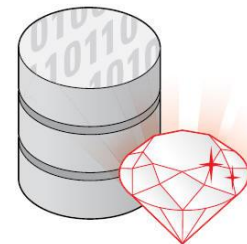
File Publish Help

Apply Output Apply Settings Task

Apply Output Table: INSUR_CUST_LTV_A15728619_A

Fetch Size: 100 Refresh

DMR\$CAS...	PREDICTION	PROBABILITY	COST	RANK	NODE	LAST	AGE1	MARITAL_STATUS1	N_MORTGAGE...
CU3111	HIGH	0.9933	0.0067	1	66	CLIFFORD	47	DIVORCED	1
CU3113	MEDIUM	0.9933	0.0067	1	48	HUMBERTO	38	SINGLE	0
CU3116	HIGH	0.9648	0.0352	1	70	EUNA	39	DIVORCED	1
CU3117	MEDIUM	0.9933	0.0067	1	48	HOYT	45	SINGLE	0
CU3119	HIGH	0.9615	0.0385	1	66	LIZBETH	42	DIVORCED	1
CU3121	HIGH	0.9615	0.0385	1	66	BORIS	46	DIVORCED	1
CU3123	HIGH	1	0	1	52	DANA	52	SINGLE	0
CU3125	MEDIUM	0.8722	0.1278	1	73	TIM	49	DIVORCED	1
CU3126	HIGH	0.9648	0.0352	1	70	LASHAWN	61	DIVORCED	1
CU3127	MEDIUM	0.8127	0.1873	1	49	BUCK	41	SINGLE	0
CU3128	MEDIUM	0.8127	0.1873	1	49	WALTON	46	SINGLE	0
CU3129	VERY HIGH	0.9515	0.0485	1	63	ALDEN	49	MARRIED	2
CU3130	VERY HIGH	0.5824	0.4176	1	68	ANGELICA	41	DIVORCED	1
CU3132	HIGH	0.9648	0.0352	1	70	LIZZETTE	34	DIVORCED	1
CU3133	HIGH	0.9648	0.0352	1	70	ISABELLA	30	DIVORCED	1
CU3134	HIGH	0.9648	0.0352	1	70	DELPHA	46	DIVORCED	1
CU3136	LOW	1	0	1	39	GEORGE	0	SINGLE	0
CU3137	HIGH	0.9648	0.0352	1	70	RAUL	39	MARRIED	1
CU3138	VERY HIGH	0.5875	0.4125	1	65	ANGELO	44	DIVORCED	1
CU3139	MEDIUM	0.9933	0.0067	1	48	GARRET	43	SINGLE	0
CU3141	MEDIUM	0.9933	0.0067	1	48	BRYON	39	SINGLE	0
CU3142	HIGH	0.9648	0.0352	1	70	TAMMI	52	DIVORCED	1
CU3143	HIGH	0.9648	0.0352	1	70	LEEANN	46	DIVORCED	1



Oracle Data Miner 11gR2 (GUI) Preview

[ODM'r "New"]



Connections Navigator

Connections

- Oracle 11gR2 server
 - BERGER_INDUSTRIES
 - Data Mining Fun
 - Loyalty analytics
 - CUSTOMERS
 - Join 4
 - SALES
 - SUPPLEMENTARY_DEMO
 - Marketing
 - Assoc Build 19
 - Class Build 25
 - CLASSIFICATION MODEL
 - CLAS_DT_1_4
 - CLAS_NB_1_4
 - CLAS_SVM_1_4
 - CLAS_SVM_2_4_line
 - CLUSTERING
 - CLUS_KM_1_4
 - CLUS_OC_1_4
 - Column Filter 15
 - CUST_INSUR_LTV
 - EXPLORE
 - SALES_TRANS_CUST
- DW Project
- EJ Project
- MH OOW Project
- MK Project
- MWT Project

Marketing Data Mining Fun

100%

Workflow Editor

Component Palette

Workflow Editor

- Models
 - Anomaly Detection
 - Association
 - Classification
 - Clustering
 - Feature Extraction
 - Link
 - Model
- Evaluate and Apply
 - Apply
 - Link
 - Test
- Data
 - Create Table
 - Data Profile
 - Data Source
 - Link
- Transforms
 - Aggregation
 - Column Filter
 - Column Filter Details
 - Join
 - Link
 - Row Filter
 - Sample
 - Linking Nodes
 - Link

Run Manager

Oracle 11gR2 server

Workflow	Project	Status
Associatio...	MH OOW P...	✓
CUST_INS...	MH OOW P...	✓
workflow3	DW Project	✓
workflow4	DW Project	✓
Loyalty an...	BERGER_I...	✓
workflow	MWT Project	✓

CUST_INSUR_LTV - Property Inspector

Find

Data Use All Data

Sample

Details

Sampling Type: Random

Seed: 12,345

Sampling Size: Number of Rows

2,000



Connections Navigator

- Connections
 - Oracle 11gR2 server
 - BERGER_INDUSTRIES
 - Data Mining Fun
 - Loyalty analytics
 - CUSTOMERS
 - Join 4
 - SALES
 - SUPPLEMENTARY_DEMO
 - Marketing
 - Assoc Build 19
 - Class Build 25
 - CLASSIFICATION MODEL
 - CLAS_DT_1_4
 - CLAS_NB_1_4
 - CLAS_SVM_1_4
 - CLAS_SVM_2_4_line
 - CLUSTERING
 - CLUS_KM_1_4
 - CLUS_OC_1_4
 - Column Filter 15
 - CUST_INSUR_LTV
 - EXPLORE
 - SALES_TRANS_CUST

Marketing | Data Mining Fun | CUST_INSUR_LTV

View: Actual Data | Sort... | Filter: Enter Where Clause

	CUST_ID	LAST	FIRST	STATE	REGION	SEX	PROFESSI...	BUY_INSURANCE	AGE	HAS_CHILDREN	SALARY	N_OF_DEPENDENTS	CAR_C
1	CU1142	MEE	COMER	CA	West	F	Author	Yes	63	1	56,045		1
2	CU1144	LAURI...	ROWLA...	NY	NorthEast	F	PROF-1	Yes	39	1	58,848		3
3	CU1145	ANNETT	MCML...	NY	NorthEast	F	Software ...	Yes	55	1	71,158		1
4	CU1146	THELMA	DELONG	NY	NorthEast	F	Software ...	No	63	1	59,900		1
5	CU1147	CRISE...	HAWKINS	NY	NorthEast	F	PROF-21	Yes	36	1	63,208		1
6	CU1148	DIA	COYLE	NV	Southwest	F	Author	No	82	0	71,288		1
7	CU1150	LILIA	SWANS...	NY	NorthEast	F	IT Staff	No	35	1	64,080		1
8	CU1151	HELLEN	OWEN	NY	NorthEast	F	Clerical	Yes	32	1	63,366		1
9	CU1152	LYNNE	MOSELEY	NV	Southwest	F	IT Staff	No	63	1	57,576		1
10	CU1155	JARRED	CANO	OR	West	M	Accountant	No	53	1	70,384		0
11	CU1156	LYNDSEY	DOZIER	NY	NorthEast	F	IT Staff	No	36	1	68,134		1
12	CU1157	AISHA	CULVER	CA	West	F	PROF-24	No	60	1	63,113		2
13	CU1158	LAURE...	BAUTISTA	CA	West	M	Author	No	81	1	63,113		0
14	CU1159	NAPO...	ELDRIDGE	NY	NorthEast	M	Administra...	No	36	0	57,764		1
15	CU1160	HARL...	PITTS	NY	NorthEast	M	PROF-3	No	33	0	65,416		3
16	CU1161	DARWIN	SEYMOUR	CA	West	M	PROF-6	No	26	1	63,049		0
17	CU1163	LUCIA...	ANDRE...	NY	NorthEast	M	School Te...	No	35	1	73,553		0
18	CU1164	WINF...	ERVIN	MI	Midwest	M	PROF-25	No	64	1	62,029		0
19	CU1166	BERNIE	ENGEL	CA	West	F	Software ...	Yes	53	1	67,124		3
20	CU1167	SPENC...	DIAS	CA	West	M	Cashier	No	70	1	62,685		4
21	CU1168	CORD...	SMART	NY	NorthEast	M	PROF-6	No	37	1	65,272		0
22	CU1169	MARY...	GILLIS	NY	NorthEast	F	PROF-24	No	36	1	72,127		4
23	CU1170	CARO...	RENTE...	FL	South	F	Clerical	Yes	53	1	63,016		1
24	CU1172	LESLIE	HEATH	NY	NorthEast	M	Law Enfor...	No	37	1	65,329		5
25	CU1173	RILEY	PRUETT	CA	West	M	Sales Rep...	Yes	53	1	60,666		1
26	CU1174	KEITH	BURNHAM	NY	NorthEast	M	Clerical	Yes	39	0	66,481		1
27	CU1175	ARNITA	BLUE	NV	Southwest	F	IT Staff	No	54	1	59,830		2
28	CU1176	JOSHUA	CHAST...	CA	West	M	Secretary	No	25	1	56,686		1
29	CU1177	CODY	CHRISTY	CA	West	M	Clerical	Yes	60	1	61,462		1

Run Manager

Oracle 11gR2 server

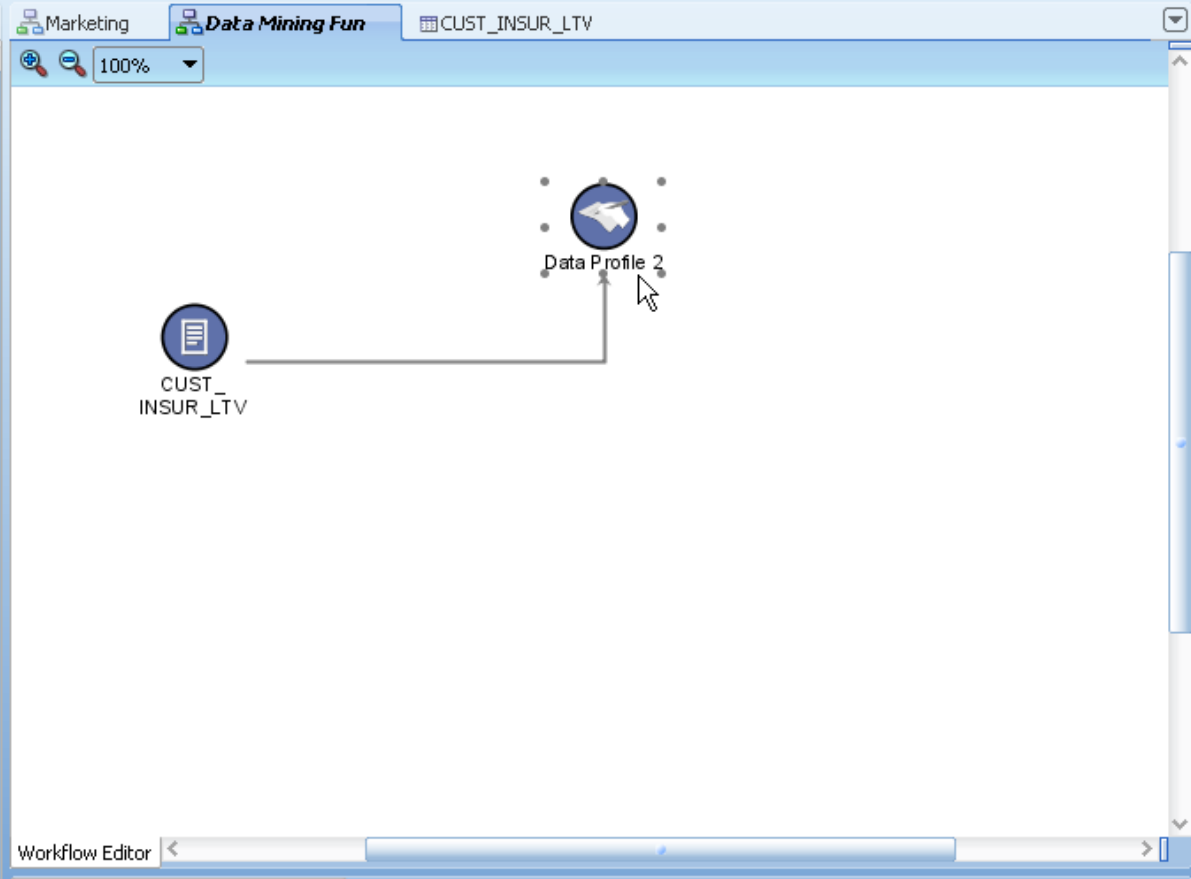
Workflow	Project	Status
Associatio...	MH OOW P...	✓
CUST_INS...	MH OOW P...	✓
workflow3	DW Project	✓
workflow4	DW Project	✓
Loyalty an...	BERGER_I...	✓
workflow	MWT Project	✓



Connections Navigator

Connections

- Oracle 11gR2 server
 - BERGER_INDUSTRIES
 - Data Mining Fun
 - Loyalty analytics
 - CUSTOMERS
 - Join 4
 - SALES
 - SUPPLEMENTARY_DEMO
 - Marketing
 - Assoc Build 19
 - Class Build 25
 - CLASSIFICATION MODEL
 - CLAS_DT_1_4
 - CLAS_NB_1_4
 - CLAS_SVM_1_4
 - CLAS_SVM_2_4_line
 - CLUSTERING
 - CLUS_KM_1_4
 - CLUS_OC_1_4
 - Column Filter 15
 - CUST_INSUR_LTV
 - EXPLORE
 - SALES_TRANS_CUST
- DW Project
- EJ_Project
- MH OOW Project
- MK Project
- MWT Project



Component Palette

Workflow Editor

- Models
 - Anomaly Detection
 - Association
 - Classification
 - Clustering
 - Feature Extraction
 - Link
 - Model
- Evaluate and Apply
 - Apply
 - Link
 - Test
- Data
 - Create Table
 - Data Profile
 - Data Source
 - Link
- Transforms
 - Aggregation
 - Column Filter
 - Column Filter Details
 - Join
 - Link
 - Row Filter
 - Sample
 - Linking Nodes
 - Link

Run Manager

Oracle 11gR2 server

Workflow	Project	Status
Associatio...	MH OOW P...	✓
CUST_INS...	MH OOW P...	✓
workflow3	DW Project	✓
workflow4	DW Project	✓
Loyalty an...	BERGER_I...	✓
workflow	MWT Project	✓

Data Profile 2 - Property Inspector

Find

Profile

Binning: Group By: BUY_INSURANCE

Sample

Details

Name	Data Type
AGE	NUMBER
BANK_FUNDS	NUMBER
BUY_INSURANCE	VARCHAR2
CAR_OWNERSHIP	NUMBER

Connections Navigator

- Oracle 11gR2 server
 - BERGER_INDUSTRIES
 - Data Mining Fun
 - Loyalty analytics
 - CUSTOMERS
 - Join 4
 - SALES
 - SUPPLEMENTARY_DEMO
 - Marketing
 - Assoc Build 19
 - Class Build 25
 - CLASSIFICATION MODEL
 - CLAS_DT_1_4
 - CLAS_NB_1_4
 - CLAS_SVM_1_4
 - CLAS_SVM_2_4_line
 - CLUSTERING
 - CLUS_KM_1_4
 - CLUS_OC_1_4
 - Column Filter 15
 - CUST_INSUR_LTV
 - EXPLORE
 - SALES_TRANS_CUST

Statistics

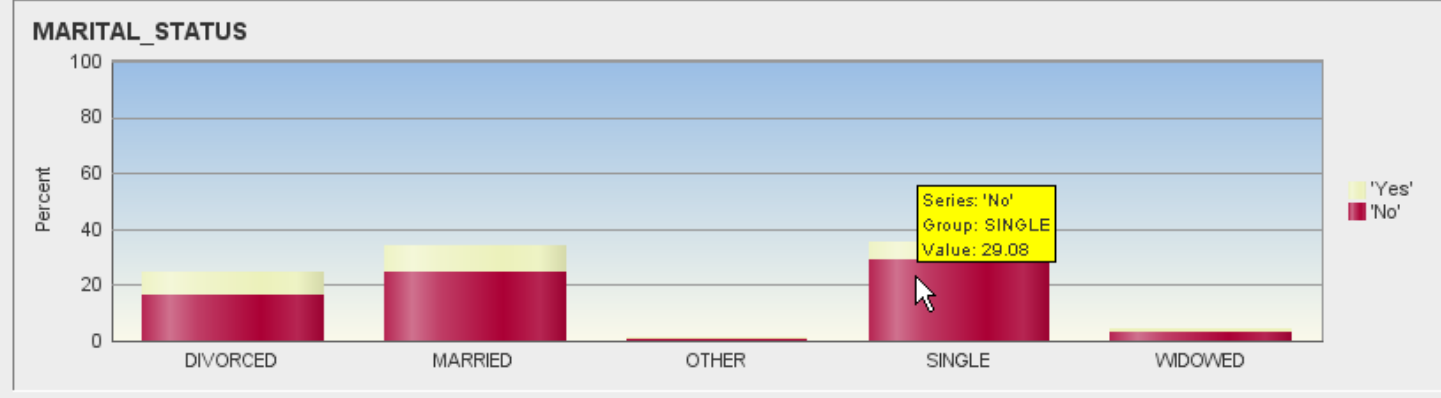
Group by: BUY_INSURANCE Filter: Name

Name	Histogram	Data Type	Percent NULLs	Distinct Values	Mode	Average
MARITAL_STATUS		VARCHAR2	0	5	SINGLE	
STATE		VARCHAR2	0	24	NY	
CREDIT_BALANCE		NUMBER	0	208		2,776.59
TIME_AS_CUSTOMER		NUMBER	0	5		2.4599
MORTGAGE_AMOUNT		NUMBER	0	437		2,009.1137
BANK_FUNDS		NUMBER	0	424		2,554.6334
N_OF_DEPENDENTS		NUMBER	0	7		2.1127
HAS_CHILDREN		NUMBER	0	2		0.4888
SALARY		NUMBER	0	1,904		65,042.007
CUST_ID		VARCHAR2	0	2,005	CU10002	
SEX		VARCHAR2	0	2	M	
PROFESSION		VARCHAR2	0	99	Programmer/Developer	
REGION		VARCHAR2	0	5	NorthEast	
HOUSE_OWNERSHIP		NUMBER	0	3		0.81

Run Manager

Oracle 11gR2 server

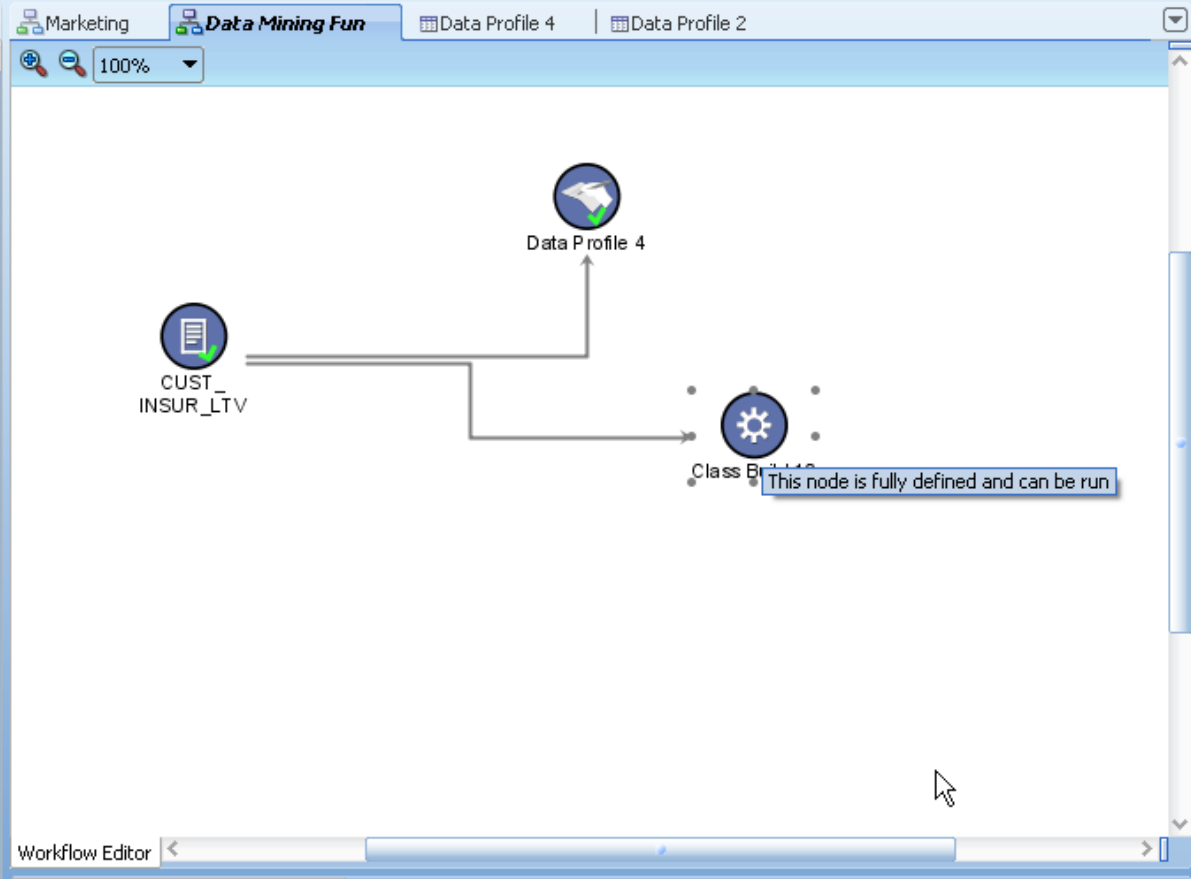
Workflow	Project	Status
Associatio...	MH OOW P...	✓
CUST_INS...	MH OOW P...	✓
workflow3	DW Project	✓
workflow4	DW Project	✓
Loyalty an...	BERGER_I...	✓
workflow	MWT Project	✓





Connections Navigator

- Connections
 - Oracle 11gR2 server
 - BERGER_INDUSTRIES
 - Data Mining Fun
 - Loyalty analytics
 - CUSTOMERS
 - Join 4
 - SALES
 - SUPPLEMENTARY_DEMO
 - Marketing
 - Assoc Build 19
 - Class Build 25
 - CLASSIFICATION MODEL
 - CLAS_DT_1_4
 - CLAS_NB_1_4
 - CLAS_SVM_1_4
 - CLAS_SVM_2_4_line
 - CLUSTERING
 - CLUS_KM_1_4
 - CLUS_OC_1_4
 - Column Filter 15
 - CUST_INSUR_LTV
 - EXPLORE
 - SALES_TRANS_CUST



Component Palette

Workflow Editor

- Models
 - Anomaly Detection
 - Association
 - Classification
 - Clustering
 - Feature Extraction
 - Link
 - Model
 - Evaluate and Apply
 - Apply
 - Link
 - Test
 - Data
 - Create Table
 - Data Profile
 - Data Source
 - Link
 - Transforms
 - Aggregation
 - Column Filter
 - Column Filter Details
 - Join
 - Link
 - Row Filter
 - Sample
 - Linking Nodes
 - Link

Run Manager

Oracle 11gR2 server

Workflow	Project	Status
Associatio...	MH OOW P...	✓
CUST_INS...	MH OOW P...	✓
workflow3	DW Project	✓
workflow4	DW Project	✓
Loyalty an...	BERGER_I...	✓
workflow	MWT Project	✓

Class Build 10 - Property Inspector

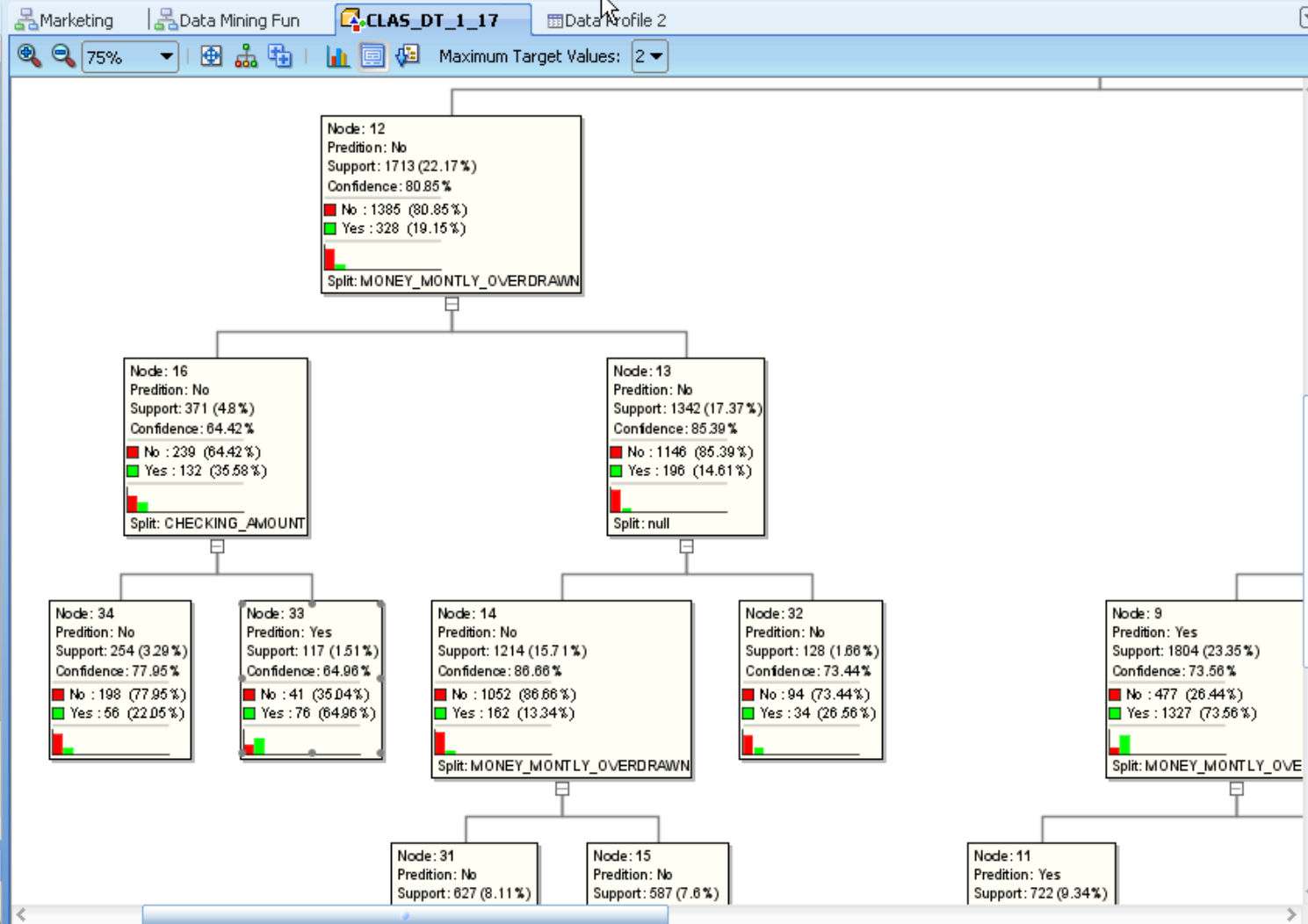
Find

Models

Name	Build	Test	Tune	Algorithm	Comment
CLAS_GLM...	Not built	Not tested	Automatic	Generalized Lin...	
CLAS_SVM...	Not built	Not tested	Automatic	Support Vector...	
CLAS_DT_1...	Not built	Not tested	Automatic	Decision Tree	
CLAS_NB_1...	Not built	Not tested	Automatic	Naive Bayes	

Connections Navigator

- Oracle 11gR2 server
 - BERGER_INDUSTRIES
 - Class Build 10
 - CUST_INSUR_LTV
 - Data Profile 4
 - Loyalty analytics
 - CUSTOMERS
 - Join 4
 - SALES
 - SUPPLEMENTARY_DEMO
 - Marketing
 - Assoc Build 19
 - Class Build 25
 - CLASSIFICATION MODEL
 - CLAS_DT_1_4
 - CLAS_NB_1_4
 - CLAS_SVM_1_4
 - CLAS_SVM_2_4_line
 - CLUSTERING
 - CLUS_KM_1_4
 - CLUS_OC_1_4
 - Column Filter 15
 - CUST_INSUR_LTV
 - EXPLORE
 - SALES_TRANS_CUST



Run Manager

Oracle 11gR2 server

Workflow	Project	Status
Associatio...	MH OOW P...	✓
CUST_INS...	MH OOW P...	✓
workflow3	DW Project	✓
workflow4	DW Project	✓
Loyalty an...	BERGER_I...	✓
workflow	MWT Project	✓

Rule Surrogates Target Values

Value	Record Count	% Percentage
Yes	76	64.95%
No	41	35.04%

Tree Settings



Connections Navigator

- Connections
 - Oracle 11gR2 server
 - BERGER_INDUSTRIES
 - Class Build 10
 - CLAS_DT_1_17
 - CLAS_GLM_1_17
 - CLAS_NB_1_17
 - CLAS_SVM_1_17
 - CUST_INSUR_LTV
 - Data Profile 4
 - Loyalty analytics
 - CUSTOMERS
 - Join 4
 - SALES
 - SUPPLEMENTARY_DEMO
 - Marketing
 - Assoc Build 19
 - Class Build 25
 - CLASSIFICATION MODEL
 - CLAS_DT_1_4
 - CLAS_NB_1_4
 - CLAS_SVM_1_4
 - CLAS_SVM_2_4_line.
 - CLUSTERING
 - CLUS_KM_1_4
 - CLUS_OC_1_4
 - Column Filter 15
 - CUST_INSUR_LTV

Marketing | Data Mining Fun | **CLAS_NB_1_17** | CLAS_DT_1_17 | Data Profile 2

Target Value: Yes Query

Fetch Size: 10,000

Probabilities: 54 out of 54 Attribute

Attribute	Value	Probability(%) for Yes
LTV	(7040.375;)	99.42829919
AGE	(18.5;)	99.28537399
CAR_OWNERSHIP	1	97.33206289
BANK_FUNDS	(271;)	97.28442115
CREDIT_BALANCE	(; 42), [42; 42]	95.75988566
N_OF_DEPENDENTS	(.5;)	95.23582658
MONTHLY_CHECKS_WRITTEN	(1.5;)	92.23439733
TIME_AS_CUSTOMER	1, 3, 4, 5	92.13911386
N_TRANS_KIOSK	(; 4.5), [4.5; 4.5]	89.09004288
CREDIT_CARD_LIMITS	(750;)	86.08861363
HOUSE_OWNERSHIP	1, 2	83.80181039
N_MORTGAGES	1, 2	83.80181039
MARITAL_STATUS	'DIVORCED', 'MARRIED', 'OTHER', '1, 2', 'POWERD'	75.22629824
N_TRANS_WEB_BANK	(.5; 2662]	75.17865650
MORTGAGE_AMOUNT	(906;)	68.31824678
CHECKING_AMOUNT	(; 25.5), [25.5; 25.5]	64.74511672
SEX	M	58.40876608
HAS_CHILDREN	1	58.21819914
T_AMOUNT_AUTOM_PAYMENTS	(556.5; 7287]	55.40733683
LTV_BIN	'HIGH'	52.92996665
<PRIOR>	<PRIOR>	50.00000000
LTV_BIN	'LOW', 'MEDIUM', 'VERY HIGH'	47.07003335
N_TRANS_ATM	(4.5;)	43.44926155
N_TRANS_TELLER	(2.5; 5.5]	43.35397808
HAS_CHILDREN	0	41.78180086
SEX	F	41.59123392
MONEY_MONTHLY_OVERDRAWN	(53.285; 54.145]	40.97189138
MONEY_MONTHLY_OVERDRAWN	(54.145;)	38.06574559
MORTGAGE_AMOUNT	(; 906), [906; 906]	31.68175322
N_TRANS_TELLER	(1.5; 2.5]	29.58551691
N_TRANS_ATM	(3.5; 4.5]	28.20390662
T_AMOUNT_AUTOM_PAYMENTS	(100.5; 489.5]	27.63220581
N_TRANS_TELLER	(.5; 1.5]	25.34540257
MARITAL_STATUS	'SINGLE'	24.77370176
HOUSE_OWNERSHIP	0	16.10810961

Run Manager

Oracle 11gR2 server

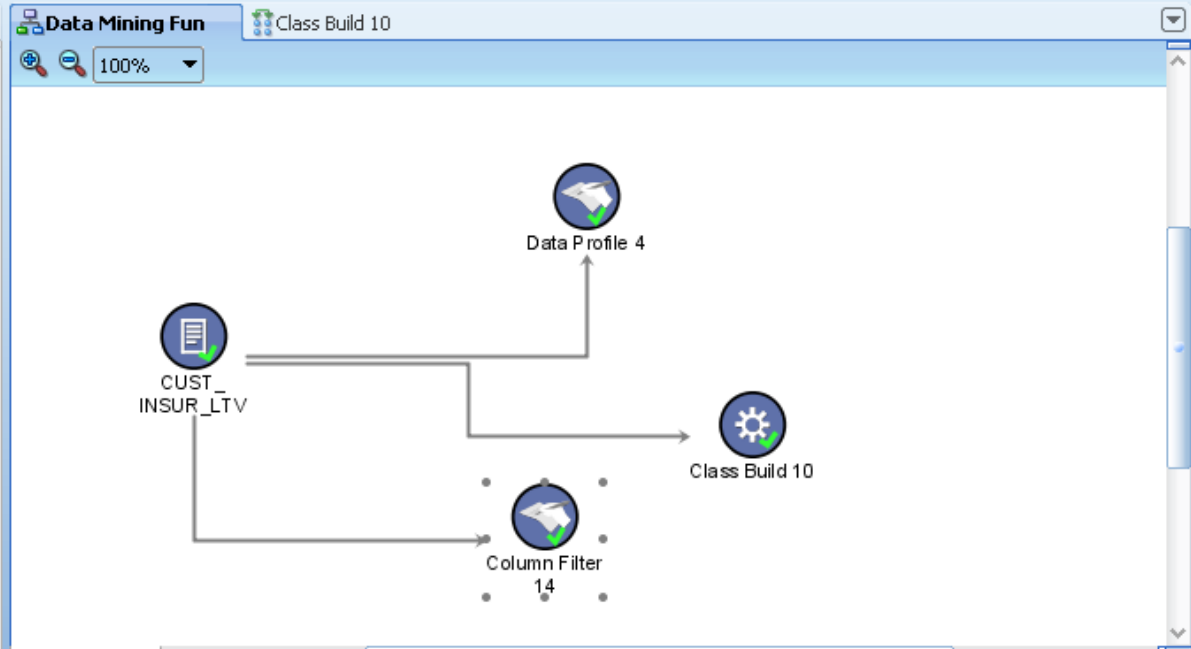
Workflow	Project	Status
Associati...	MH OOW P...	✓
CUST_INS...	MH OOW P...	✓
workflow3	DW Project	✓
workflow4	DW Project	✓
Loyalty an...	BERGER_I...	✓
workflow	MWT Project	✓



Connections Navigator

Connections

- Oracle 11gR2 server
 - BERGER_INDUSTRIES
 - Data Mining Fun
 - Class Build 10
 - CLAS_DT_1_17
 - CLAS_GLM_1_17
 - CLAS_NB_1_17
 - CLAS_SVM_1_17
 - Column Filter 14
 - CUST_INSUR_LTV
 - Data Profile 4
 - Loyalty analytics
 - CUSTOMERS
 - Join 4
 - SALES
 - SUPPLEMENTARY_DEMOGR
 - Marketing
 - DW Project
 - EJ_Project
 - MH OOW Project
 - MK Project
 - MWT Project



Component Palette

Workflow Editor

- Models
 - Anomaly Detection
 - Association
 - Classification
 - Clustering
 - Feature Extraction
 - Link
 - Model
- Evaluate and Apply
 - Apply
 - Link
 - Test
- Data
 - Create Table
 - Data Profile
 - Data Source
 - Link
- Transforms
 - Aggregation
 - Column Filter
 - Column Filter Details
 - Join
 - Link
 - Row Filter
 - Sample
 - Linking Nodes
 - Link

Workflow Editor

Column Filter 14 - Property Inspector

Columns

Filters

- Data Quality
 - % Nulls less than or equal
 - % Unique less than or equal
 - % Constant less than or equal
- Attribute Importance
 - Target: BUY_INSURANCE
 - Importance Cutoff
 - Top N
- Sampling (Stratified)
 - Sample Size

Run Manager

Oracle 11gR2 server

Workflow	Project	Status
Associatio...	MH OOW P...	✓
CUST_INS...	MH OOW P...	✓
workflow3	DW Project	✓
workflow4	DW Project	✓
Loyalty an...	BERGER_I...	✓
workflow	MWT Project	✓

Connections Navigator

Connections

- Oracle 11gR2 server
 - BERGER_INDUSTRIES
 - Data Mining Fun
 - Class Build 10
 - CLAS_DT_1_17
 - CLAS_GLM_1_17
 - CLAS_NB_1_17
 - CLAS_SVM_1_17
 - Column Filter 14
 - CUST_INSUR_LTV
 - Data Profile 4
 - Loyalty analytics
 - CUSTOMERS
 - Join 4
 - SALES
 - SUPPLEMENTARY_DEMOGR
 - Marketing
 - DW Project
 - EJ_Project
 - MH OOW Project
 - MK Project
 - MWT Project

Automatic Filterings

Remove Missing Input Automatically

Name	Type	Automatic	Output	Rank	Importance	% Null	% Unique	% Constant	Hints/Reject Reasons
AGE	NUMBER	<input checked="" type="checkbox"/>	→*	22	0	0	0	0	Min importance not reached
BANK_FUNDS	NUMBER	<input checked="" type="checkbox"/>	→	3	0.2037	0	0	0	
BUY_INSURANCE	VARCHAR2	<input checked="" type="checkbox"/>	→	1	0.8412	0	0.0997	72.297	
CAR_OWNERSHIP	NUMBER	<input checked="" type="checkbox"/>	→	16	0.0052	0	0	0	
CHECKING_AMOUNT	NUMBER	<input checked="" type="checkbox"/>	→	11	0.0116	0	0	0	
CREDIT_BALANCE	NUMBER	<input checked="" type="checkbox"/>	→*	22	0	0	0	0	Min importance not reached
CREDIT_CARD_LIMITS	NUMBER	<input checked="" type="checkbox"/>	→	20	0.003	0	0	0	
CUST_ID	VARCHAR2	<input checked="" type="checkbox"/>	→*	22	0	0	100	0.0498	Exceed % unique, Min importa...
FIRST	VARCHAR2	<input checked="" type="checkbox"/>	→*	22	0	0	69.3572	0.2491	Min importance not reached
HAS_CHILDREN	NUMBER	<input checked="" type="checkbox"/>	→	21	0.0004	0	0	0	
HOUSE_OWNERSHIP	NUMBER	<input checked="" type="checkbox"/>	→	18	0.0047	0	0	0	
LAST	VARCHAR2	<input checked="" type="checkbox"/>	→*	22	0	0	68.6099	0.299	Min importance not reached
LTV	NUMBER	<input checked="" type="checkbox"/>	→*	22	0	0	0	0	Min importance not reached
LTV_BIN	VARCHAR2	<input checked="" type="checkbox"/>	→*	22	0	0	0.1993	48.9287	Min importance not reached
MARITAL_STATUS	VARCHAR2	<input checked="" type="checkbox"/>	→	15	0.006	0	0.2491	34.6288	
MONEY_MONTHLY_OVERDRAWN	NUMBER	<input checked="" type="checkbox"/>	→	4	0.1486	0	0	0	
MONTHLY_CHECKS_WRITTEN	NUMBER	<input checked="" type="checkbox"/>	→	8	0.0853	0	0	0	
MORTGAGE_AMOUNT	NUMBER	<input checked="" type="checkbox"/>	→	13	0.0078	0	0	0	
N_MORTGAGES	NUMBER	<input checked="" type="checkbox"/>	→	17	0.0048	0	0	0	
N_OF_DEPENDENTS	NUMBER	<input checked="" type="checkbox"/>	→	9	0.0188	0	0	0	
N_TRANS_ATM	NUMBER	<input checked="" type="checkbox"/>	→	5	0.1389	0	0	0	
N_TRANS_KIOSK	NUMBER	<input checked="" type="checkbox"/>	→	19	0.0036	0	0	0	

Help OK Cancel

Run Manager

Oracle 11gR2 server

Workflow	Project	Status
Associatio...	MH OOW P...	✓
CUST_INS...	MH OOW P...	✓
workflow3	DW Project	✓
workflow4	DW Project	✓
Loyalty an...	BERGER_I...	✓
workflow	MWT Project	✓

Filters

Sample

Details

Data Quality

- % Nulls less than or equal: 95
- % Unique less than or equal: 95
- % Constant less than or equal: 95

Attribute Importance

Target: BUY_INSURANCE

- Importance Cutoff: 0
- Top N: 100

Sampling (Stratified)

Sample Size: 2,000

Transforms

- Aggregation
- Column Filter
- Column Filter Details
- Join
- Link
- Row Filter
- Sample
- Linking Nodes
- Link

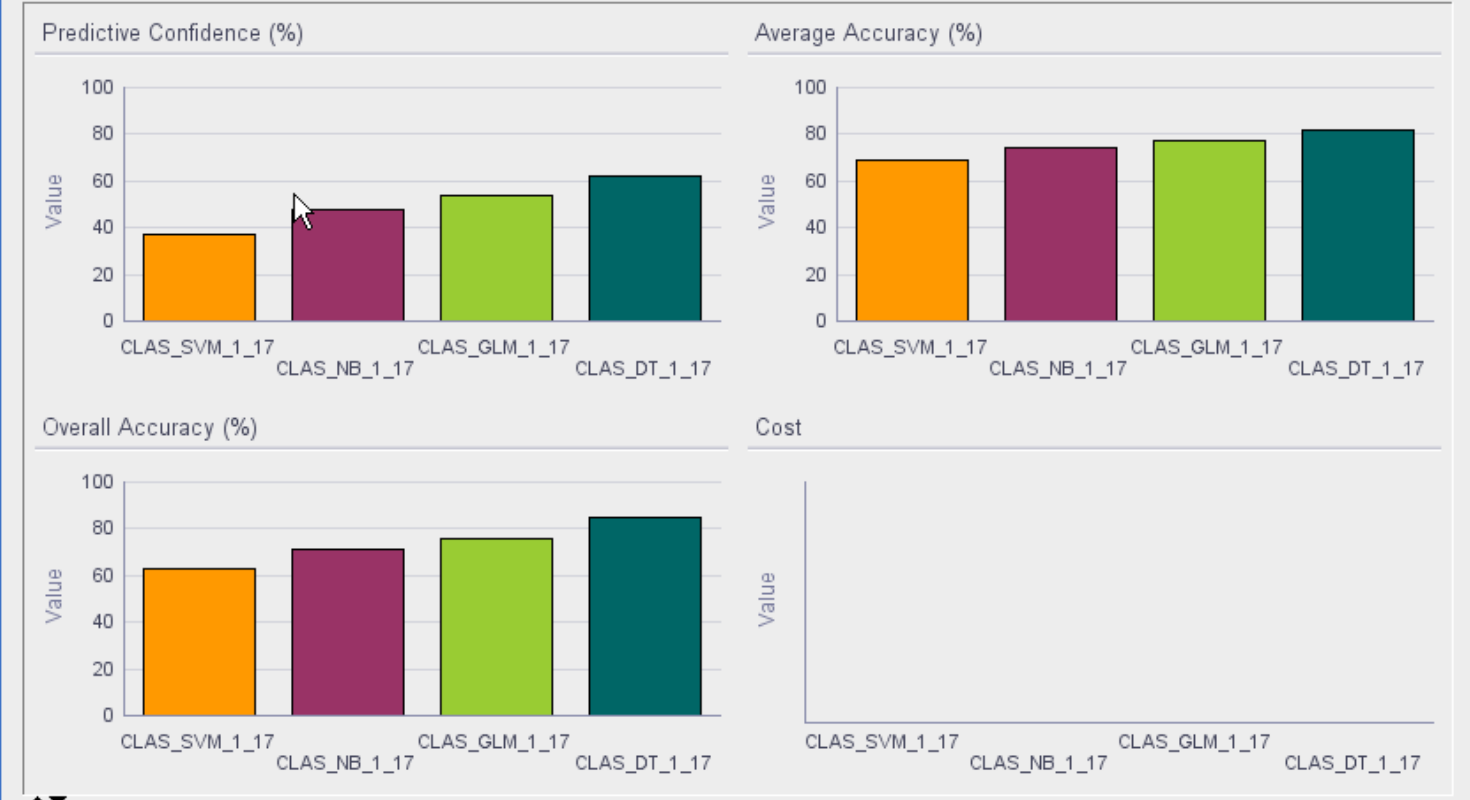


Connections Navigator

- Connections
 - Oracle 11gR2 server
 - BERGER_INDUSTRIES
 - Class Build 10
 - CLAS_DT_1_17
 - CLAS_GLM_1_17
 - CLAS_NB_1_17
 - CLAS_SVM_1_17
 - CUST_INSUR_LTV
 - Data Profile 4
 - Loyalty analytics
 - CUSTOMERS
 - Join 4
 - SALES
 - SUPPLEMENTARY_DEMO
 - Marketing
 - Assoc Build 19
 - Class Build 25
 - CLASSIFICATION MODEL
 - CLAS_DT_1_4
 - CLAS_NB_1_4
 - CLAS_SVM_1_4
 - CLAS_SVM_2_4_line
 - CLUSTERING
 - CLUS_KM_1_4
 - CLUS_OC_1_4
 - Column Filter 15
 - CUST_INSUR_LTV

Marketing | Data Mining Fun | **Class Build 10** | CLAS_NB_1_17 | CLAS_DT_1_17 | Data Profile 2

Measure: All Measures | Sort By: Name | Descending



Models

Name	Predictive Confidence %	Overall Accuracy %	Average Accuracy %	Cost	Algorithm	Created
CLAS_DT_1_17	62.1368	84.6265	81.0684		Decision Tree	Tue C
CLAS_GLM_1_17	53.8844	75.4234	76.9422		Generalized Linear Model	Tue C
CLAS_NB_1_17	47.1474	70.9991	73.5737		Naive Bayes	Tue C
CLAS_SVM_1_17	36.7506	62.6362	68.3753		Support Vector Machine	Tue C

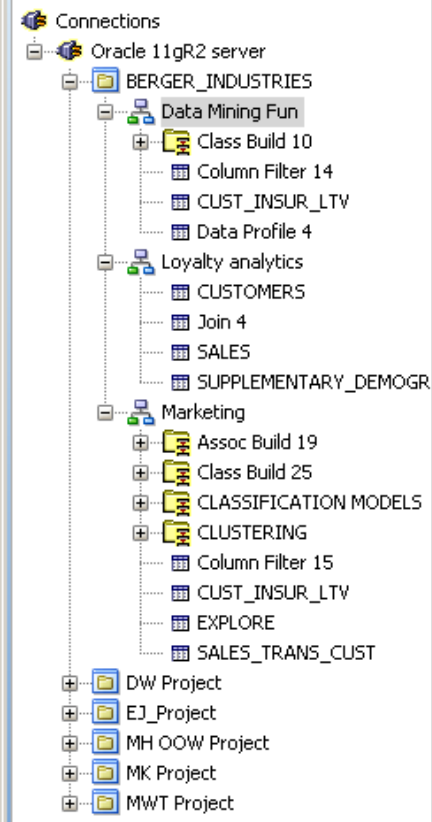
Run Manager

Oracle 11gR2 server

Workflow	Project	Status
Associatio...	MH OOW P...	✓
CUST_INS...	MH OOW P...	✓
workflow3	DW Project	✓
workflow4	DW Project	✓
Loyalty an...	BERGER_I...	✓
workflow	MWT Project	✓

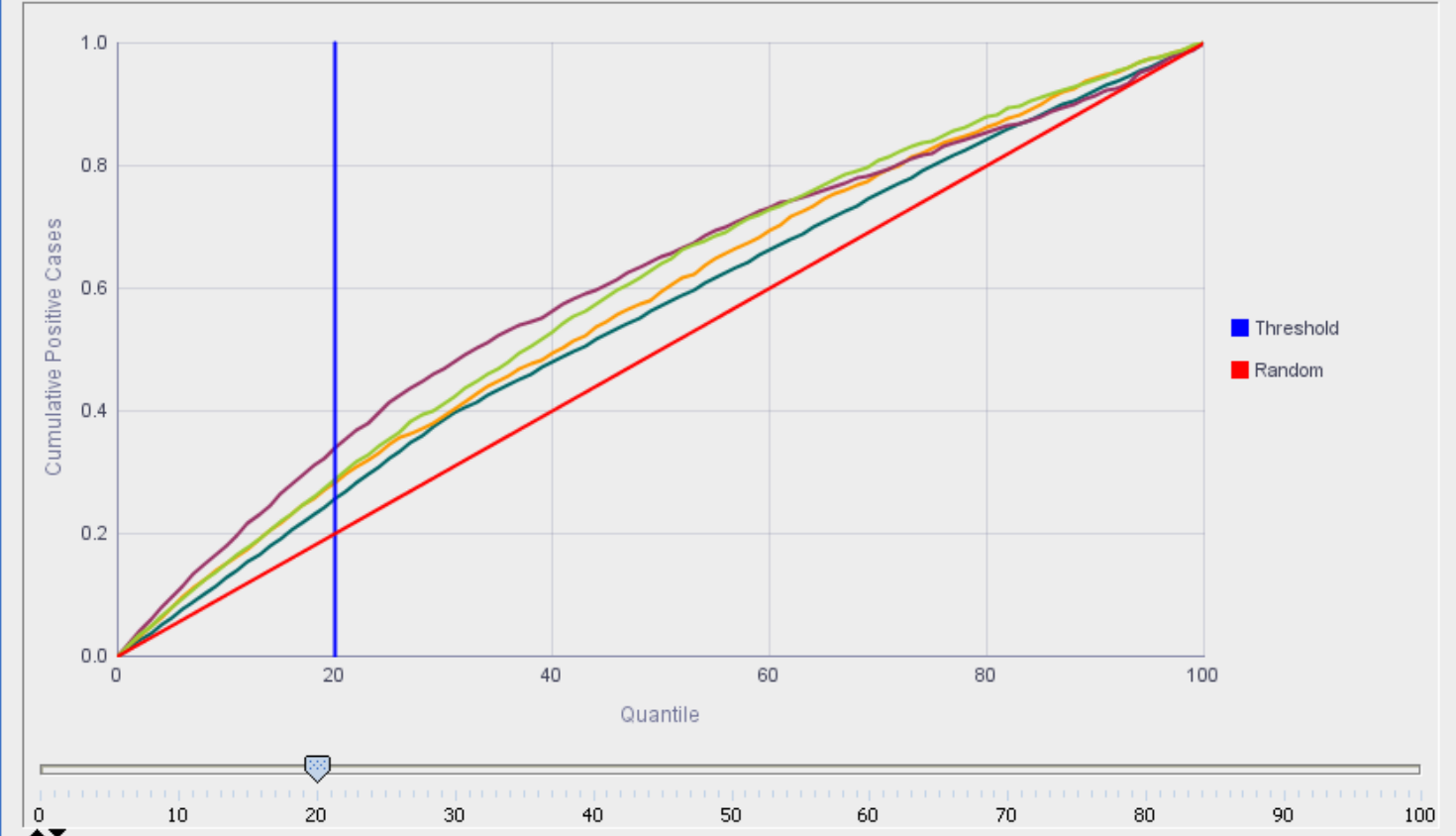


Connections Navigator



Data Mining Fun Class Build 10 Loyalty analytics CLAS_DT_2_4 Marketing Class Build 25

Display: Cumulative Positive Cases Target Value: Yes



Run Manager

Workflow	Project	Status
Associatio...	MH OOW P...	✓
CUST_INS...	MH OOW P...	✓
workflow3	DW Project	✓
workflow4	DW Project	✓
Loyalty an...	BERGER_I...	✓
workflow	MWT Project	✓

Model	Lift Cumulative	Gain Cumulative %	Records Cumulative %	Target Density Cumulative	Algorithm
CLAS_DT_1_17	1.2467	25.7297	20.6379	0.8836	Decision Tree
CLAS_GLM_1_17	1.4077	28.8392	20.4873	0.7465	Generalized Linear Mode
CLAS_NB_1_17	1.6828	34.1299	20.2814	0.8049	Naive Bayes
CLAS_SVM_1_17	1.3842	28.1944	20.369	0.5602	Support Vector Machine

Performance Performance Matrix ROC Lift Profit



Connections Navigator

- Connections
 - Oracle 11gR2 server
 - BERGER_INDUSTRIES
 - Data Mining Fun
 - Assoc Build 23
 - Class Build 10
 - Column Filter 14
 - CUST_INSUR_LTV
 - Data Profile 4
 - SALES_TRANS_CUST
 - Loyalty analytics
 - CUSTOMERS
 - Join 4
 - SALES
 - SUPPLEMENTARY_DEMO
 - Marketing
 - Assoc Build 19
 - Class Build 25
 - CLASSIFICATION MODEL
 - CLUSTERING
 - Column Filter 15
 - CUST_INSUR_LTV
 - EXPLORE
 - SALES_TRANS_CUST
 - DW Project
 - EJ_Project
 - MH OOW Project
 - MK Project
 - MWT Project

Data Mining Fun ASSOC_AP_2_17 Class Build 10

Sort by: Lift Descending

Fetch Size: 1,000

Query More

Rule Content: Subname

Rules: 15 out of 15

ID	Antecedent	Consequent	Lift	Confidence(%)	Support(%)	Length
14	Mouse Pad AND Extension Cable	Standard Mouse	2.7212	87.4251	15.5319	2
15	Mouse Pad AND Standard Mouse	Extension Cable	2.7075	84.3931	15.5319	2
13	Standard Mouse AND Extension Cable	Mouse Pad	2.6643	85.8824	15.5319	2
7	Extension Cable	Standard Mouse	1.8059	58.0205	18.0851	1
8	Standard Mouse	Extension Cable	1.8059	56.2914	18.0851	1
1	Standard Mouse	Mouse Pad	1.7772	57.2848	18.4043	1
2	Mouse Pad	Standard Mouse	1.7772	57.0957	18.4043	1
3	Extension Cable	Mouse Pad	1.7682	56.9966	17.766	1
4	Mouse Pad	Extension Cable	1.7682	55.1155	17.766	1
9	CD-RW, High Speed Pack of 5	External 8X CD-ROM	1.7150	52.3636	15.3191	1
10	External 8X CD-ROM	CD-RW, High Speed Pack ...	1.7150	50.1742	15.3191	1
11	18" Flat Panel Graphics Monitor	SIMM- 16MB PCMCIAII card	1.5611	49.6575	15.4255	1
12	SIMM- 16MB PCMCIAII card	18" Flat Panel Graphics Mo...	1.5611	48.495	15.4255	1

Rule Details:

ID: 14

IF

Mouse Pad AND
Extension Cable

THEN

Standard Mouse

Lift	2.7212
Confidence(%)	87.4251
Support(%)	15.5319

Run Manager

Oracle 11gR2 server

Workflow	Project	Status
Associatio...	MH OOW P...	✓
CUST_INS...	MH OOW P...	✓
workflow3	DW Project	✓
workflow4	DW Project	✓
Loyalty an...	BERGER_I...	✓
workflow	MWT Project	✓



Connections Navigator

- Connections
 - Oracle 11gR2 server
 - BERGER_INDUSTRIES
 - Data Mining Fun
 - Assoc Build 23
 - Class Build 10
 - Clust Build 27
 - Column Filter 14
 - CUST_INSUR_LTV
 - Data Profile 4
 - SALES_TRANS_CUST
 - Loyalty analytics
 - CUSTOMERS
 - Join 4
 - SALES
 - SUPPLEMENTARY_DEMO
 - Marketing
 - Assoc Build 19
 - Class Build 25
 - CLASSIFICATION MODEL
 - CLUSTERING
 - Column Filter 15
 - CUST_INSUR_LTV
 - EXPLORE
 - SALES_TRANS_CUST

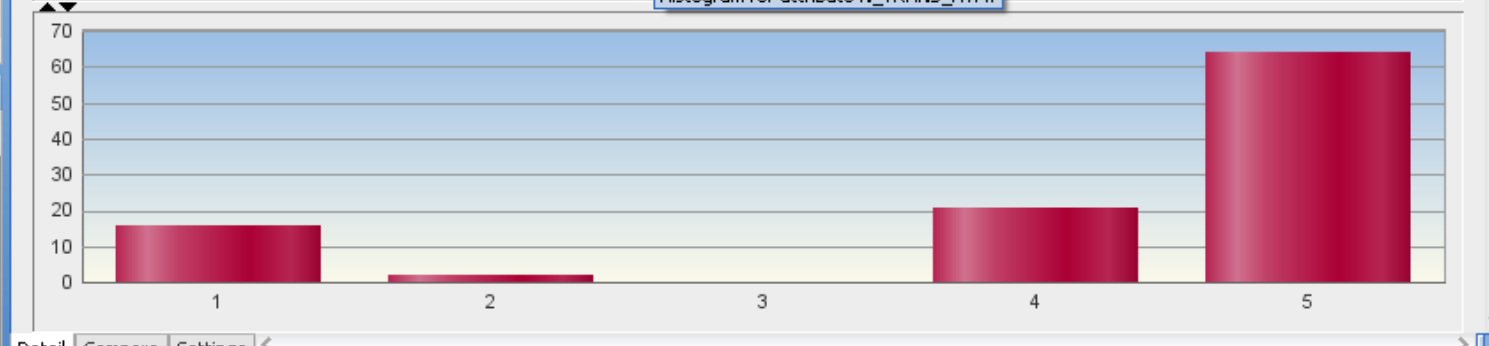
Data Mining Fun | **CLUS_KM_1_17** | ASSOC_AP_2_17 | Class Build 10

Cluster: Leaves Only

Fetch Size:

Attributes: 27 out of 27

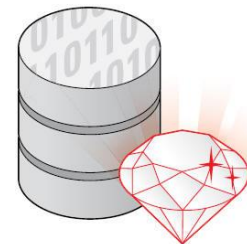
Attribute	Histogram	Confidence(%)	Support	Mode	Mean
TIME_AS_CUSTOMER		60.0000	1,160	5	
LTV		60.0000	1,105		18,913.0989
SEX		50.0000	1,179	F	
HOUSE_OWNERSHIP		50.0000	1,157	1	
N_MORTGAGES		50.0000	1,157	1	
N_OF_DEPENDENTS		50.0000	1,101		4.1555
LTV_BIN		50.0000	1,083	MEDIUM	
HAS_CHILDREN		50.0000	958	1	
MONTHLY_CHECKS_WRITTEN		40.0000	1,144		5.0862
MARITAL_STATUS		33.3333	1,042	DIVORCED	
STATE		22.2222	1,032	NY	
N_TRANS_TELLER		20.0000	1,051		2.2232
AGE		16.6667	1,134		38.8715
N_TRANS_ATM		14.2857	1,098		3.7371



Run Manager

Oracle 11gR2 server

Workflow	Project	Status
Associatio...	MH OOW P...	✓
CUST_INS...	MH OOW P...	✓
workflow3	DW Project	✓
workflow4	DW Project	✓
Loyalty an...	BERGER_I...	✓
workflow	MWT Project	✓



Applications *Powered by* Oracle Data Mining

CRM OnDemand—Sales Prospector

Predictions

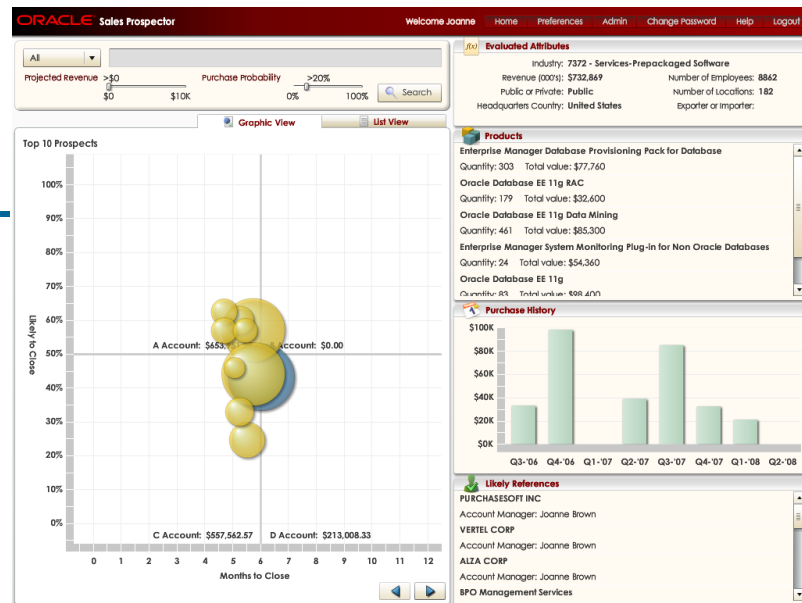
- Revenue
- Probability
- Time to close

Analysis

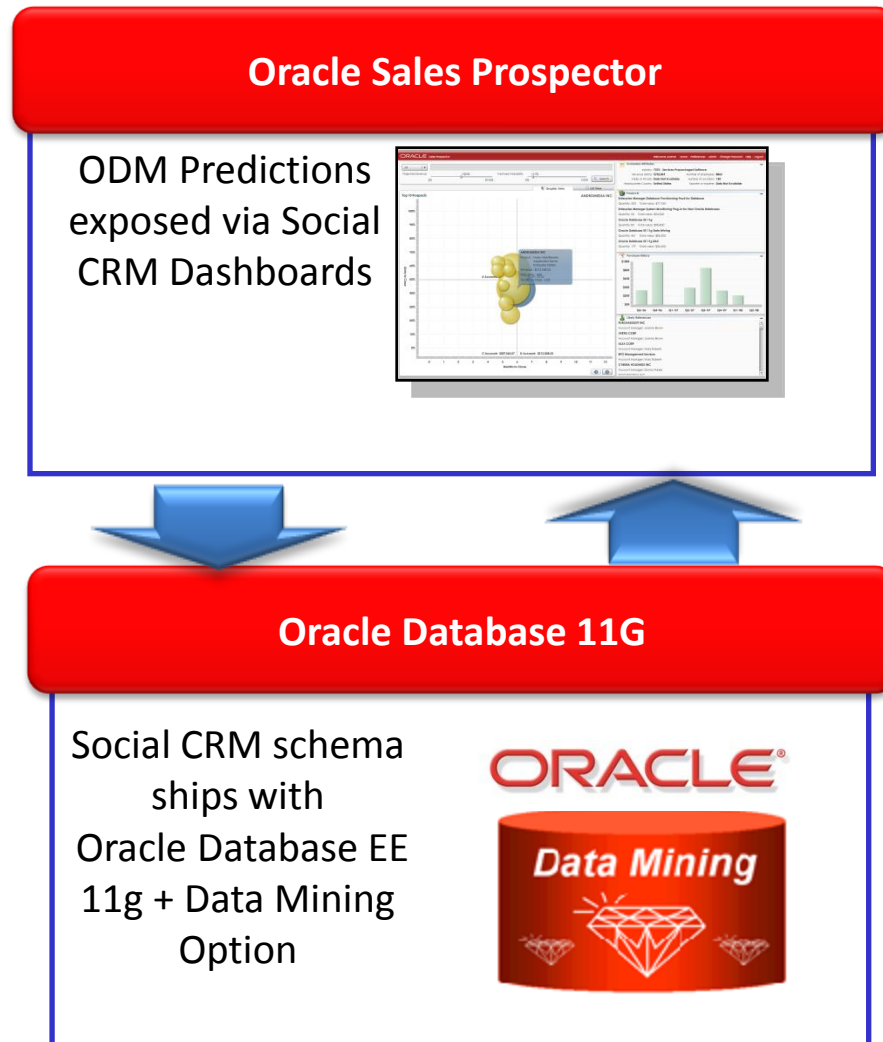
- Customer attributes
- Products owned
- Purchase history

References

- Similar customers
- Similar products



CRM OnDemand—Sales Prospector

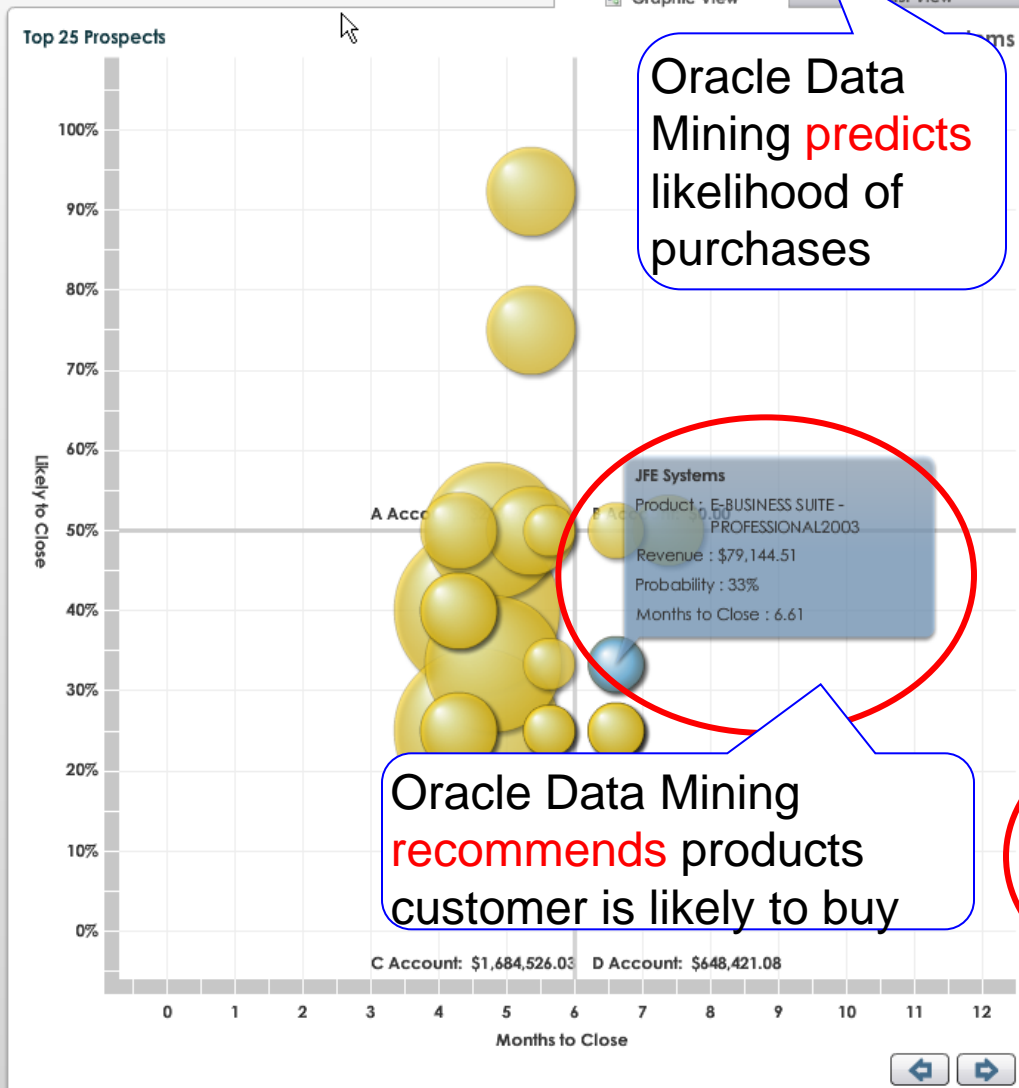


All

Projected Revenue >\$0 Purchase Probability >20%

\$0 \$1M 0% 100%

Search



Oracle Data Mining **predicts** likelihood of purchases

Oracle Data Mining **recommends** products customer is likely to buy

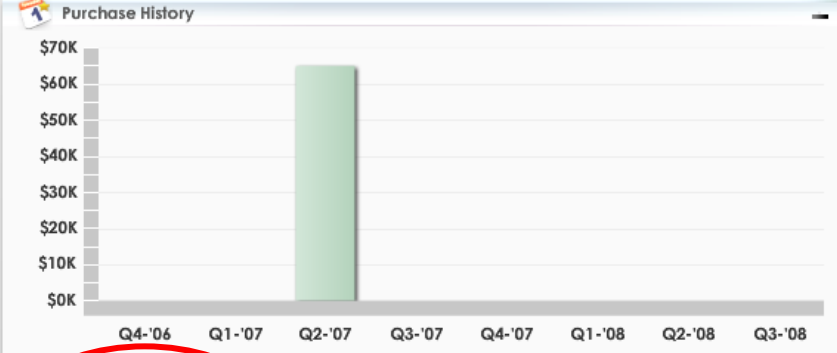
Evaluated Attributes

Industry: **Electronics**
 Revenue (000's): **\$337,420**
 Public or Private: **Private**
 Headquarters Country: **Japan**

Number of Employees: **1538**
 Number of Locations: **13**
 Exporter or Importer: **Export**

Products

91 APP SVR SE	Quantity: 3	Total value: \$0
APP SVR EE	Quantity: 1	Total value: \$0
DATABASE EE	Quantity: 3	Total value: \$0
DATABASE SE	Quantity: 3	Total value: \$3,000,000
DIAGNOSTICS PACK	Quantity: 2	Total value: \$0



Likely References

Aspen Aerogels Inc	Account Manager: Dima Skorikov
FIRST CHURCH OF CHRIST SCIENTIST	Account Manager: Dima Skorikov
Stratus Technologies Inc	Account Manager: Dima Skorikov

Oracle Data Mining **suggests** likely references

Oracle Open World (OOW) Schedule Builder

Session Recommendation Engine

- Build Personal OOW Agendas
 - Recommends sessions, exhibitors and demos based on profile
 - Identify related sessions to selected session
- Get Recommendations
- Status
 - Production use at OOW'08 and OOW'09
 - 40,000+ attendees
- Tech details
 - Solution includes in-database transformations, ODM clustering (text mining) and classification algorithms with code generation from Oracle Data Miner

The screenshot displays the Oracle Open World Schedule Builder interface. At the top, it shows the event details: "OCTOBER 11-15, 2009 MOSCONE CENTER SAN FRANCISCO" and the slogan "Come with questions. Leave with answers." Below this is a navigation bar with "Home", "Build Schedule", "Saved Schedule & Interests", "Session Changes", and "Logout".

The main content area includes a search section with "Basic Search" and "Advanced Search" tabs, and a "Recommended Exhibitors" list featuring companies like SAP AG, Microstrategy, and IBM. Below that is an "Exhibitors" list and a "Recommended Oracle Demos" section.

The central part of the interface is a calendar grid showing sessions from Sunday, Oct 11 to Thursday, Oct 15. The grid is organized by time slots (8:00 to 15:00). Key sessions include:

- Monday, Oct 12, 9:00:** Keynote: Oracle Develop--What Are We Still Doing Wrong?
- Monday, Oct 12, 10:00:** Keynote: Capitalizing on
- Monday, Oct 12, 13:00:** (*R) Oracle Data Mining 11g: Overview, Demos, Oracle Exadata, and
- Tuesday, Oct 13, 9:00:** Keynote: The Art of the Possible--Charles Phillips and Safra Catz, Oracle
- Tuesday, Oct 13, 10:00:** Keynote: The Future of Enterprise
- Tuesday, Oct 13, 11:30:** IT Convergence
- Wednesday, Oct 14, 9:00:** Keynote
- Wednesday, Oct 14, 10:00:** Keynote: A Thorough Program for Modern Customers
- Thursday, Oct 15, 9:00:** Keynote: Primavera Project Portfolio Management Road
- Thursday, Oct 15, 14:30:** Harnessing the Power of Information to Drive
- Thursday, Oct 15, 15:00:** Keynote: Seven

Oracle Retail Data Model

ORACLE Interactive Dashboards My Dashboard OLAP ORDM Demo Analytics ORDM Demo KPIs ORDM MetaData

Mining Analysis Welcome, NRFdemo! Dashboards - Answers - More Products - My Account - Log Out

Associate Basket Analysis Associate Loss Analysis Associate Sales Analysis Product Category Mix Customer Loyalty Analysis Frequent Shopper Category Mix Item Basket Analysis Item ROS Loss Store Loss Page Options

Year: 2007 Month: BY 2007 M2 Division: Pittsburgh Division District: Store Name: PITTSBURGH 109

Performance Measure: Customer Loyalty Type Go

Customer Loyalty Type

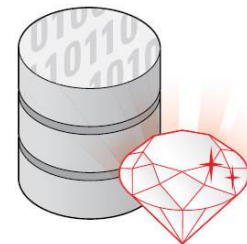
Customer Loyalty Type

Time run: 6/9/2007

Customer Profile	Performance Measure Value	% of Supporting Transaction	Probability
Years Of Residence = '1', '2', '3', '4', '5' and YearS_OF_RESIDENCE = '1', '2', '3' and Household Size = '3+' and YearS_OF_RESIDENCE = '1'	LEAST LOYAL	2.45%	84.76%
Years Of Residence = '1', '2', '3', '4', '5' and YearS_OF_RESIDENCE = '1', '2', '3' and Household Size = '3+' and YearS_OF_RESIDENCE = '1' and Marital Status = 'DIVORCED', 'SEPARATED'	LEAST LOYAL	0.13%	100.00%
Years Of Residence = '1', '2', '3', '4', '5' and YearS_OF_RESIDENCE = '1', '2', '3' and Household Size = '3+' and YearS_OF_RESIDENCE = '1' and Marital Status = 'MARRIED', 'SINGLE'	LEAST LOYAL	1.60%	87.85%
Years Of Residence = '1', '2', '3', '4', '5' and YearS_OF_RESIDENCE = '4', '5' and Household Size = '3+'	PRETTY LOYAL	10.93%	83.17%
Years Of Residence = '1', '2', '3', '4', '5' and YearS_OF_RESIDENCE = '4', '5' and Household Size = '3+' and Marital Status = 'MARRIED', 'SINGLE'	PRETTY LOYAL	7.33%	81.58%
Years Of Residence = '1', '2', '3', '4', '5' and YearS_OF_RESIDENCE = '4', '5' and Household Size = 'LESS THAN 3'	MARGINALLY LOYAL	9.00%	86.29%
Years Of Residence = '1', '2', '3', '4', '5' and YearS_OF_RESIDENCE = '4', '5' and Household Size = 'LESS THAN 3' and Marital Status = 'MARRIED', 'SINGLE'	MARGINALLY LOYAL	6.52%	84.17%
Years Of Residence = '10', '6', '7', '8', '9' and Household Size = '3+'	MOST LOYAL	25.01%	88.28%
Years Of Residence = '10', '6', '7', '8', '9' and Household Size = '3+' and Marital Status = 'MARRIED', 'SINGLE'	MOST LOYAL	17.11%	86.10%
Years Of Residence = '10', '6', '7', '8', '9' and Household Size = 'LESS THAN 3'	PRETTY LOYAL	26.20%	80.88%

Out-of-the box, Oracle Data Mining generates profiles of customers

Oracle Data Mining automatically mines data for analysis reports

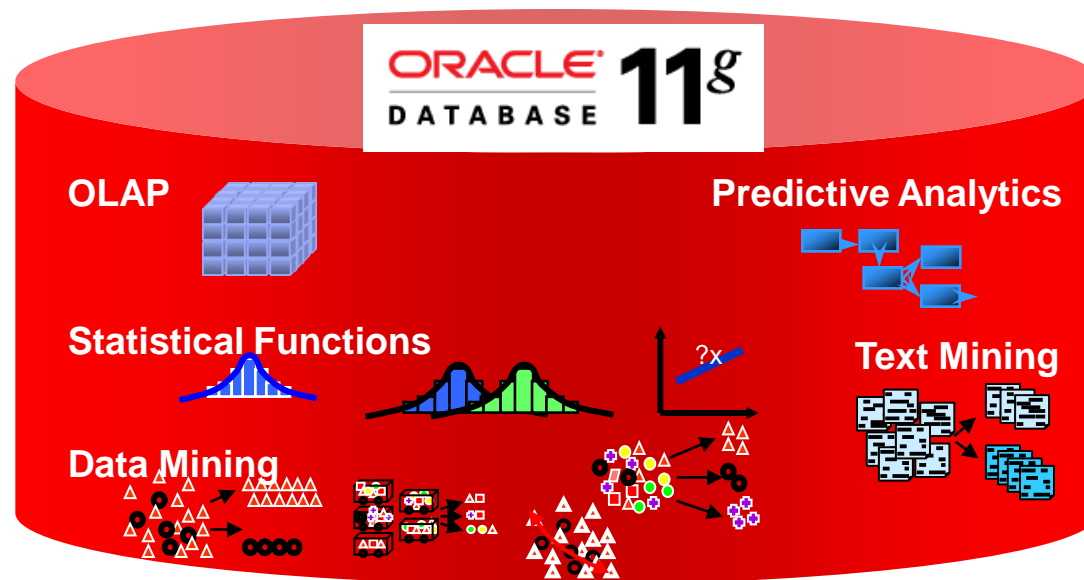


Strategic Vision

An Analytical Database Changes— Everything!



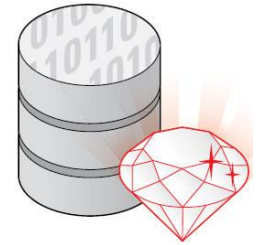
Less data movement = **faster** analytics, ...and
faster analytics = **better** BI throughout enterprise



Applications Powered by Oracle Data Mining—Integration Opportunities

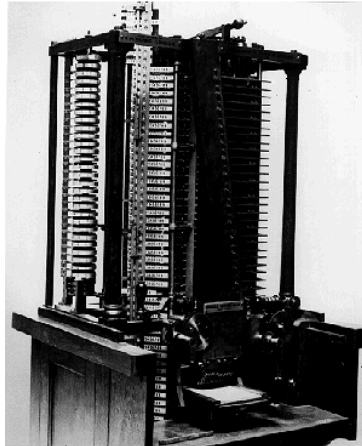
- Financial applications
- Expense reporting
- Network monitoring
- Healthcare applications
- “Green” applications
- Higher Education
- Insurance vertical
- Retail
- ISV Partners
- More...

Analytical Database



- Oracle Exadata + Oracle Data Mining
- Higher users expectations from information managed in Oracle
- — *“You (Oracle) should be able to know this!”*

From Computer Desktop Encyclopedia
Reproduced with permission.
© 2000 Charles Babbage Institute, University of Minnesota



Additional Information

- ODM preso and demo(s) posted www.oraclebiwa.org
 - Webcast: July 22, 2009, [Oracle Data Mining Overview and Demos](#) by Charlie Berger ([slides](#), [recording 37MB](#))
- [OTN ODM web site](#):
 - [Oracle Data Mining 11gR1 presentation](#)
 - [Oracle Data Mining 11gR1 data sheet](#)
 - [Oracle Data Mining 11gR1 white paper](#)
 - [Anomaly Detection and Fraud using ODM 11gR1 presentation](#)
 - [OTN Discussion Forum](#)



Oracle Data Mining



Oracle BIWA SIG—Like Minded Users

- **BIWA TechCasts** (45-min webcasts + Q&A)
 - Any Oracle professional may submit abstracts for
 - **Audience is technical**
 - Live demos are strongly encouraged
 - Visit: www.oraclebiwa.org to submit
 - Apple iPod awarded to “best new presenter”
(see www.oraclebiwa.org for details)



• BIWA Training Days @ Collaborate 2010

- **“Get Analytical with BIWA Training Days”**
- April 18-22, 2010
- Las Vegas, Nevada



- Call for Presentations Open Now!
- REGISTER with **“BIWA2010”** for IOUG Special Member Rate



Wednesday TechCast Series

Example topics of particular interest to BIWA summit attendees include, but are not limited to the following:

Data Access and Data Integration

- Data quality
- Extract, transform, load (ETL)
- Accessing distributed data
- SOA integration

Data Warehouses

- Data Governance
- Master Data Management
- Partitioning
- Tuning warehouse
- Faster cubes for faster information
- Managing images

Reporting and BI Dashboards

- Better reports & better information
- Custom BI environments
- Real-time analytics
- Interactive dashboards & EPM
- OBI EE, Essbase & Oracle Database

Advanced Analytics

- Predictive analytics and modeling
- Data mining and text mining
- SQL Statistical functions
- Fraud detection
- Market basket analysis
- Churn and retention strategies
- Building & using OLAP “cubes”
- What if? Analysis
- Leveraging spatial data
- Time series and forecasting
- Harvesting more insight from data “Best practices”

Case Studies

Tips & Tricks



ORACLE®