

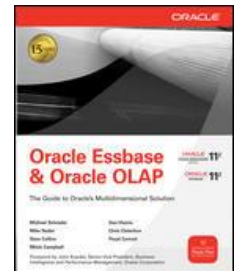
Blazing BI: the Analytic Options to the Oracle Database IAOUG September 5, 2013

**Dan VlamiS
VlamiS Software Solutions
816-781-2880
<http://www.vlamiS.com>**



Dan VlamiS and VlamiS Software Solutions

- VlamiS Software founded in 1992 in Kansas City, Missouri
- Developed more than 200 Oracle BI systems
- Specializes in ORACLE-based:
 - Data Warehousing
 - Business Intelligence
 - Design and integrated BI and DW solutions
 - Training and mentoring
- Expert presenter at major Oracle conferences
- www.vlamiS.com (blog, papers, newsletters, services)
- Developer for IRI (former owners of Oracle OLAP)
- Co-author of book “Oracle Essbase & Oracle OLAP”
- Beta tester for OBIEE 11g
- Reseller for Simba and Nokia map data for OBIEE
- HOL Coordinator for BIWA Summit 2013

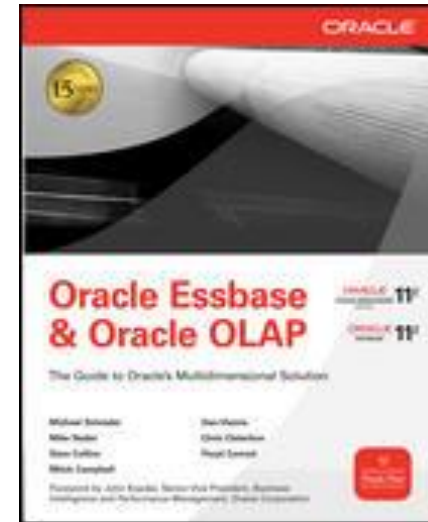




Oracle Essbase & Oracle OLAP: The Guide to Oracle's Multidimensional Solution

- Published by Oracle Press

- Dan VlamiS
- Chris Claterbos
- Michael Nader
- David Collins
- Floyd Conrad
- Mitchell Campbell
- Michael Schrader



- Covers both Oracle Essbase and Oracle OLAP
- 500 Pages



Analytical Options to Oracle Database

- Oracle OLAP
 - Defines a multi-dimensional data structure that allows information for highly complex calculations to be done quickly.
 - Fast query performance and incremental update
 - Simplified access to analytic calculations
- Oracle Advanced Analytics (Data Mining & R)
 - Refers to the process of automatically sifting through data to find hidden patterns and make predictions.
 - Series of highly advanced algorithms and procedures.
 - Extends the “R” language to the Oracle Database
- Oracle Spatial
 - Provides the capability of relating data to geo positional coordinates, objects, and constructs.
 - Allows the construction and analysis of network topologies.



Spectrum of Oracle DB BI & Analytics

OLAP

Summaries,
hierarchies and
dimensional data

“Analysis”

What is the
average
income of
mutual fund
buyers,
by region,
by year?

Data Mining & R

Knowledge discovery
of hidden patterns

“Insight & Prediction”

Who is likely to
purchase a mutual
fund in the next 6
months and why?

Spatial

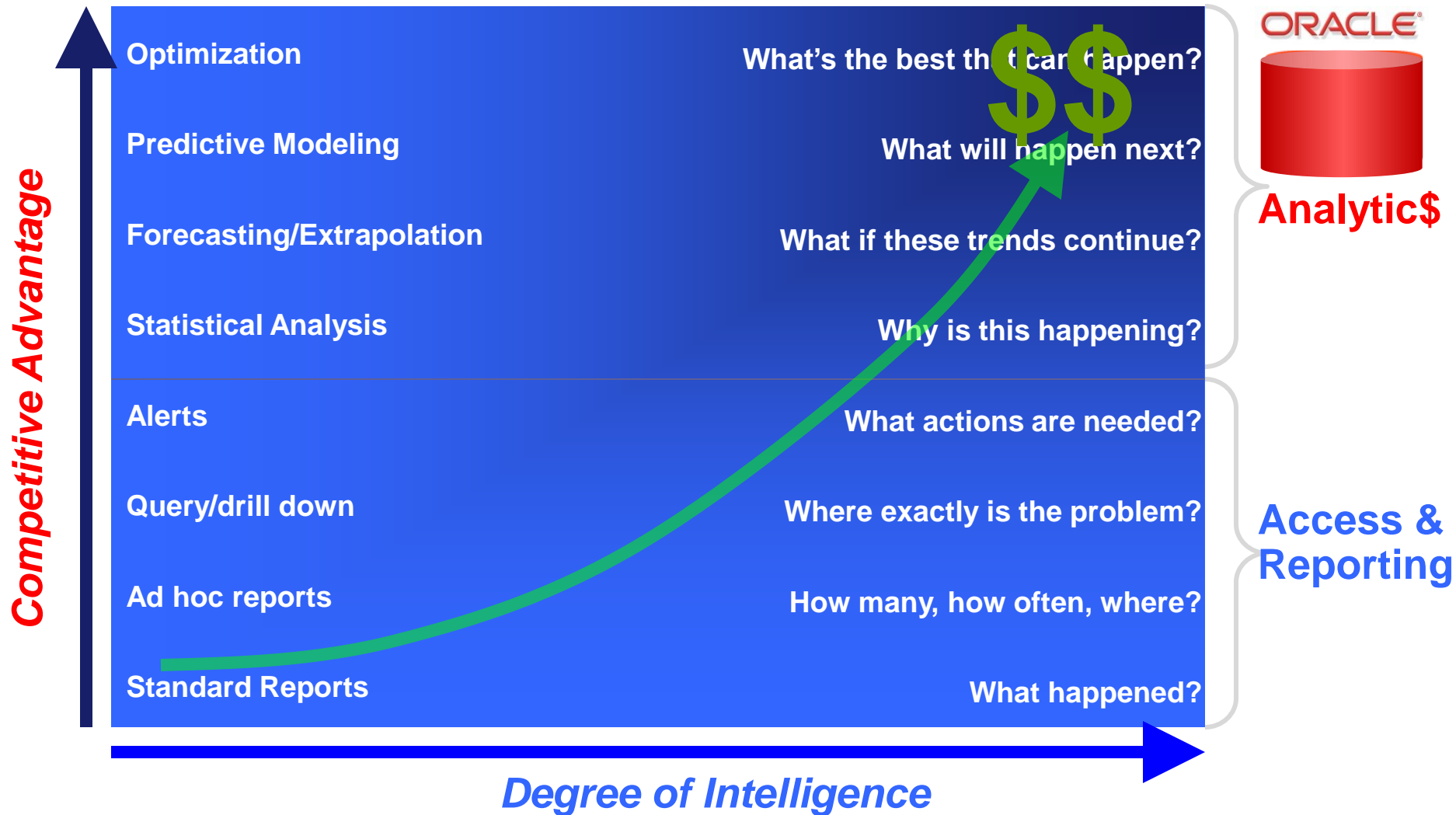
Spatial relationships
between data

“Location”

Where were
mutual funds
purchased
in the last 3
years?



Competitive Advantage of BI & Analytics



Source: Competing on Analytics, by T. Davenport & J. Harris



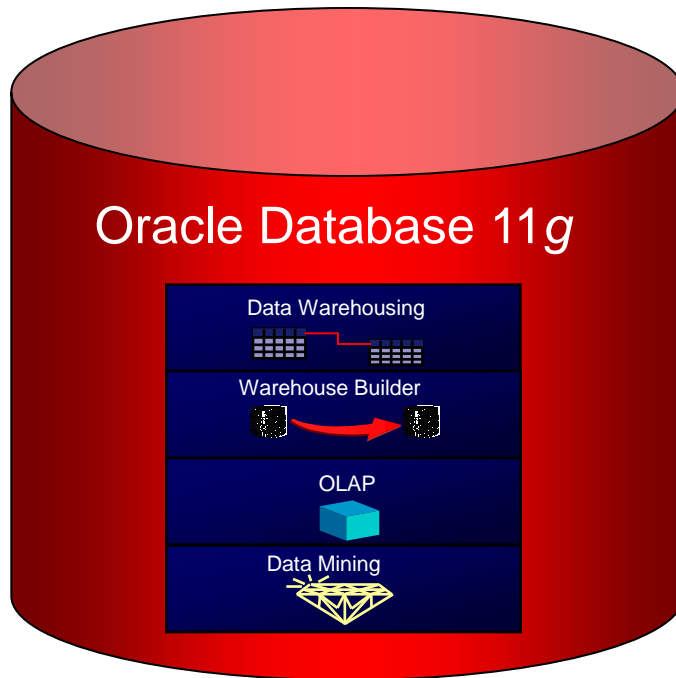
Why OLAP for BI?

- BI often presents data dimensionally
- Dimensions are natural way to look at data
 - By, across, over, time, geography, product
 - Comparison of multiple dimension values
- Multi-dimensional storage of data speeds analysis
- Natural to express dimensional comparisons
 - Share of parent
 - Compared to last year
- Allows for hierarchical dimensions with multiple levels
 - E.g. by country, drill to state, drill to city



Oracle OLAP

Leveraging Core Database Infrastructure

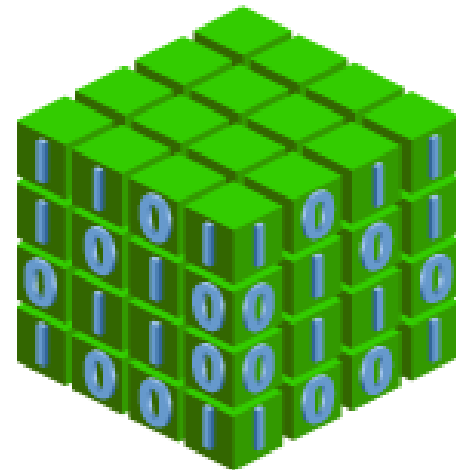


- Single RDBMS-MDBMS process
- Single data storage
- Single security model
- Single administration facility
- Grid-enabled
- Accessible by any SQL-based tool
- Embedded BI metadata
- Connects to all related Oracle data



Oracle OLAP

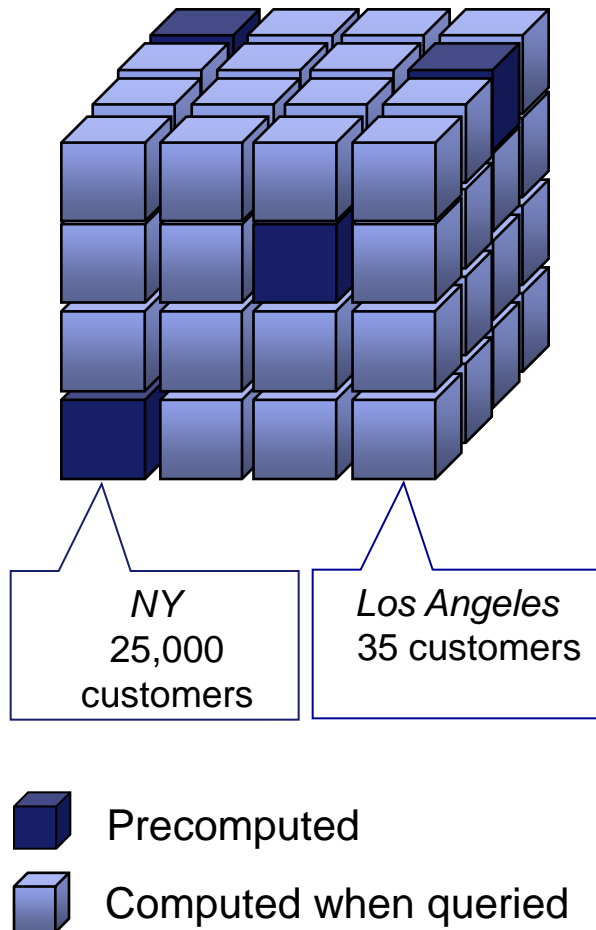
- A summary management solution for SQL based business intelligence applications
 - An alternative to table-based materialized views, offering improved query performance and fast, incremental update
- A full featured multidimensional OLAP server
 - Excellent query performance for ad-hoc / unpredictable query
 - Enhances the analytic content of Business intelligence application
 - Fast, incremental updates of data sets





Cost Based Aggregation

Pinpoint Summary Management



- Improves aggregation speed and storage consumption by pre-computing cells that are most expensive to calculate
- Easy to administer
- Simplifies SQL queries by presenting data as fully calculated



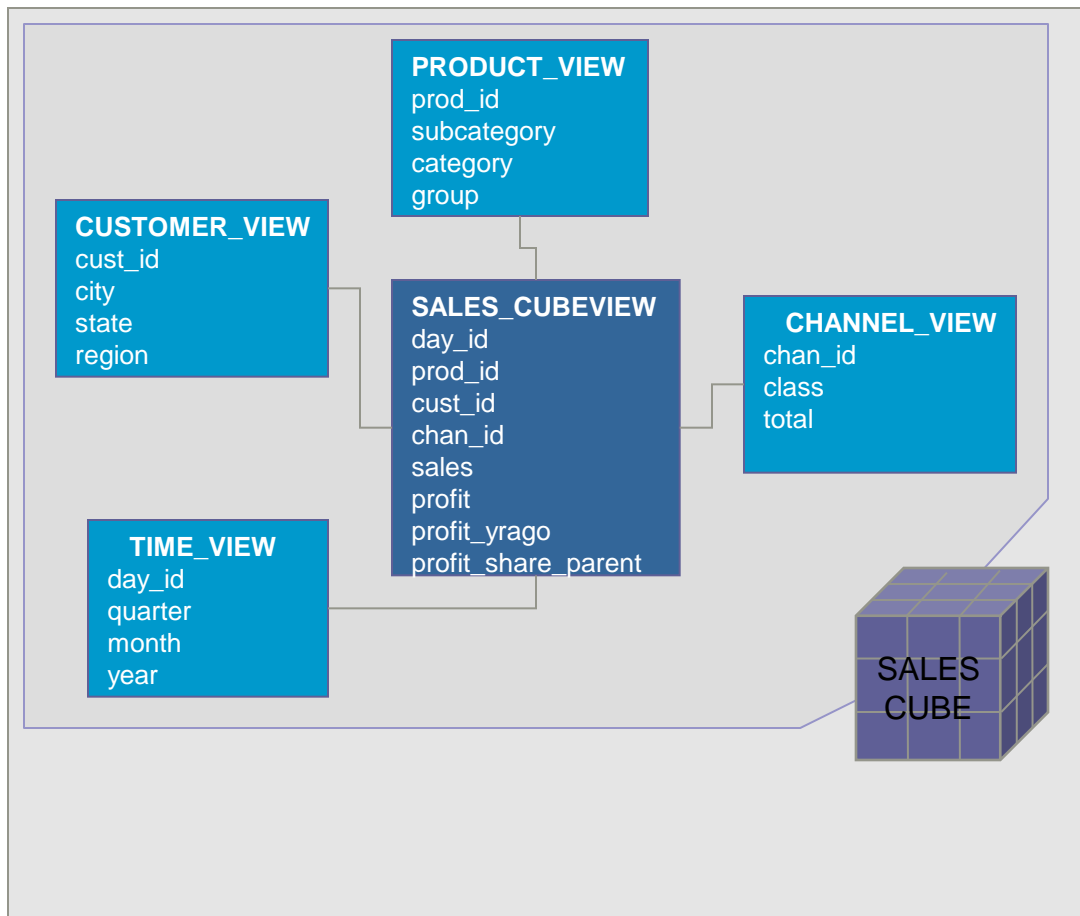
One Cube Accessed Many Ways...

- One cube can be used as
 - A summary management solution to SQL-based business intelligence applications as cube-organized materialized views
 - A analytically rich data source to SQL-based business intelligence applications as SQL cube-views
 - A full-featured multidimensional cube, servicing dimensionally oriented business intelligence applications



Cube Represented as Star Model

Simplifies Access to Analytic Calculations



- Cube represented as a star schema
- Single cube view presents data as completely calculated
 - Analytic calculations presented as columns
 - Includes all summaries
- Automatically managed by OLAP



Essbase vs. Oracle OLAP

Essbase

- Separate server
- List price* \$184K/CPU
- Separate admin
- Administer by LoB
- Must build cubes
- Part of middle tier
- Excellent writeback
- Query via MDX, XML/A

Oracle OLAP

- Built into Oracle DB
- List price* DB + \$23K/CPU
- Admin same as Oracle DB
- Administer by IT
- Must build cubes
- Part of server tier
- Limited writeback
- Query via SQL (now MDX)

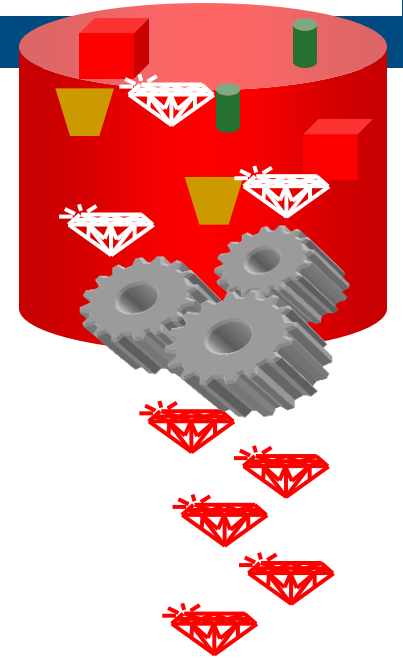
* <http://www.oracle.com/us/corporate/pricing/index.html>



What is Data Mining?

ORACLE®

- 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*)

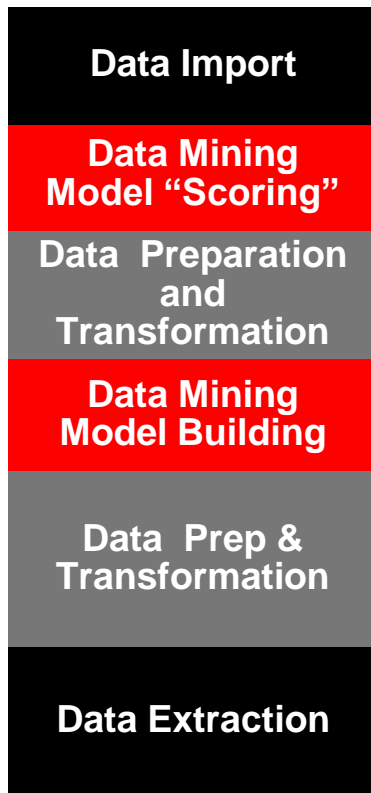




In Database Data Mining



Traditional Analytics



Hours, Days or Weeks



Oracle Data Mining



Secs. Mins or Hours



\$avings

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



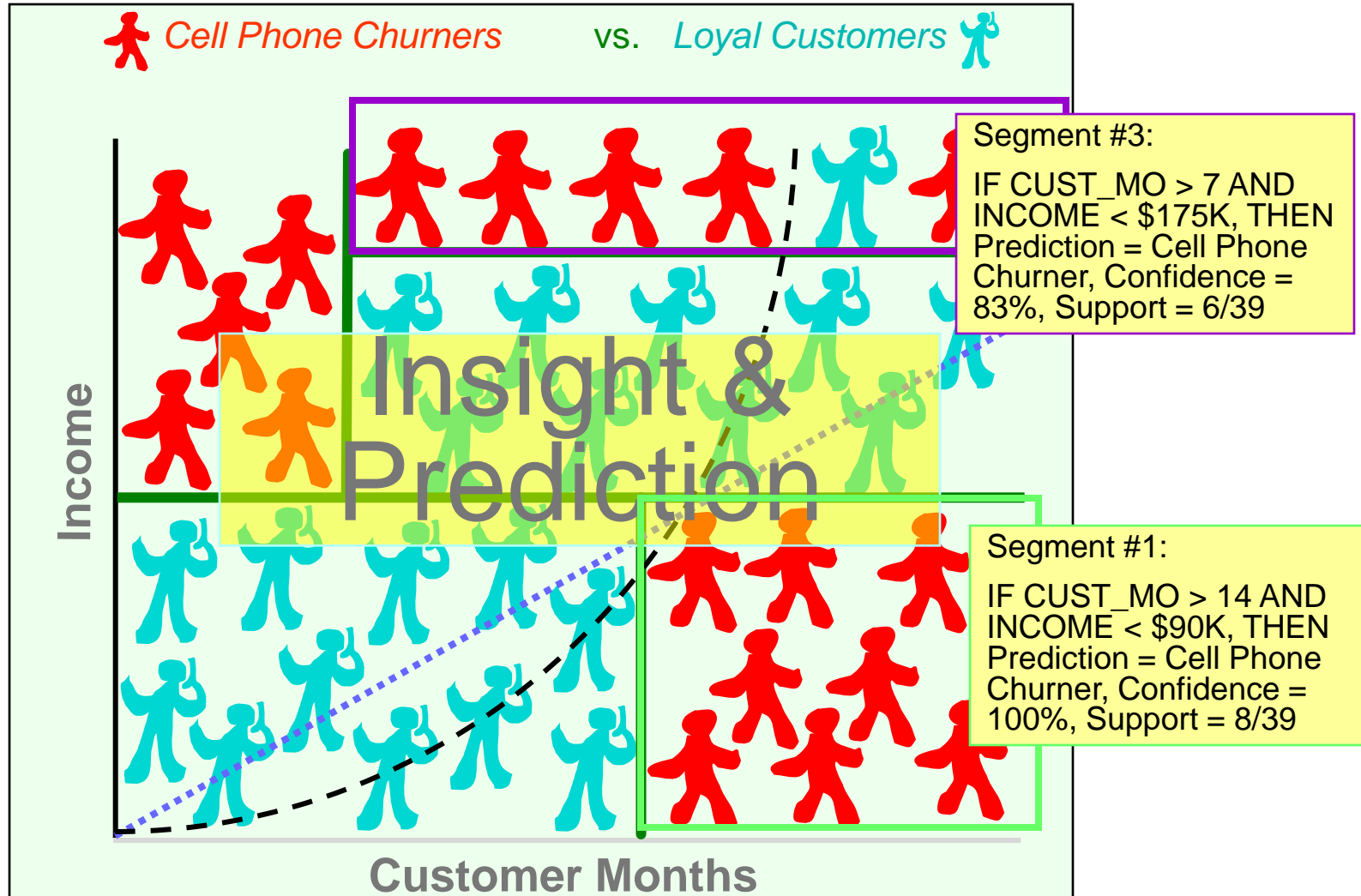
Oracle Data Mining

- Oracle Data Mining is an option for the Enterprise Edition of the Oracle Database.
- A collection of APIs and specialized SQL functions.
- Includes a large number of specialized algorithms and built-in procedures.
- Makes use of many built-in capabilities of the Oracle Database
- ODM typically refers to “Oracle Data Mining”



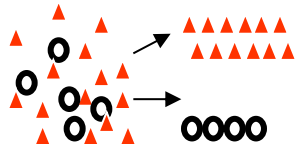
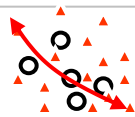
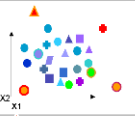
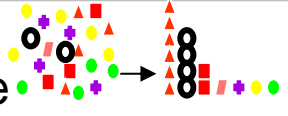
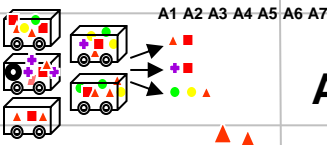
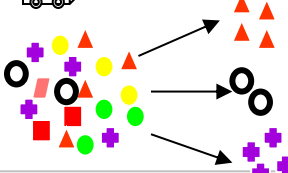
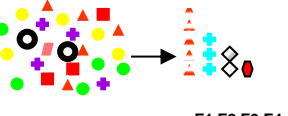
Data Mining Provides

Better Information, Valuable Insights and Predictions





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



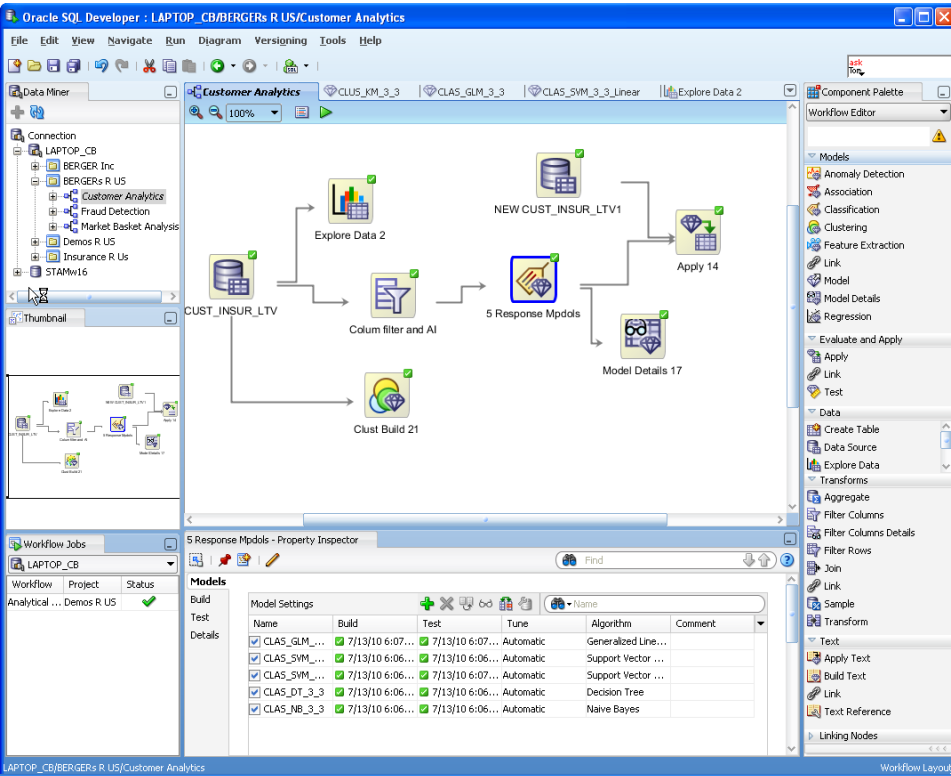
11g Statistics & SQL Analytics (Free)

- **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, median, stats_mode, variance, standard deviation, 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

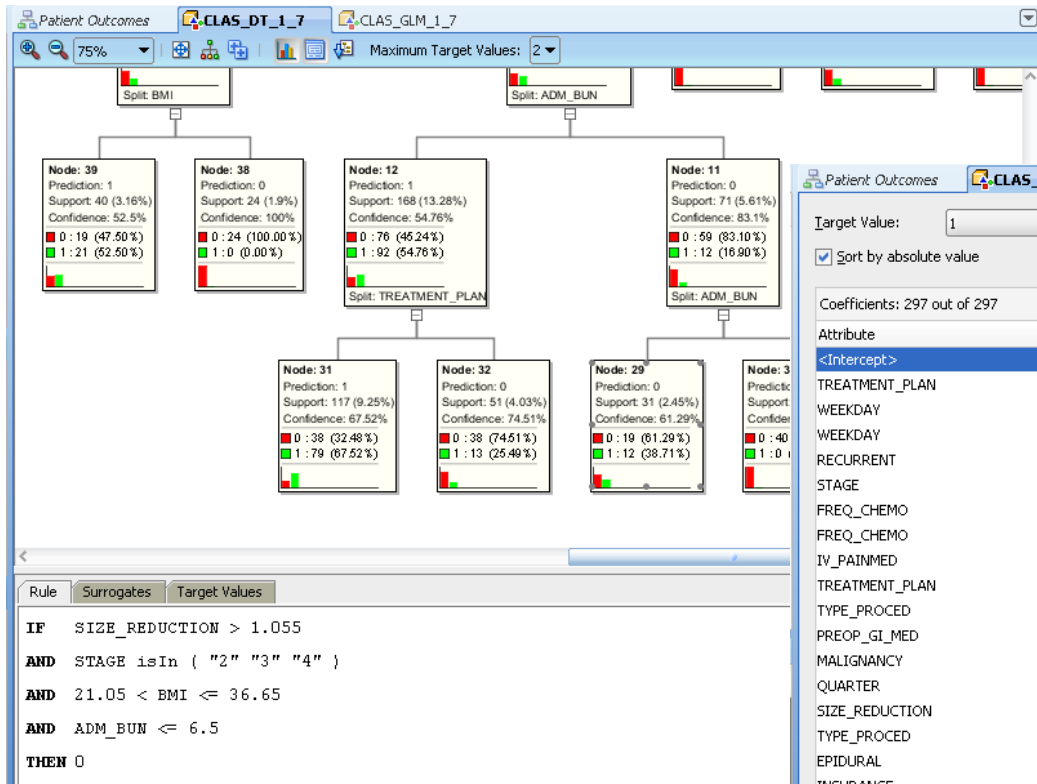
Oracle Data Miner 11g



- Oracle Data Miner is a front end GUI for Oracle Data Mining.
- Extension for Oracle SQL Developer 3.x, a free utility program from Oracle that facilitates interaction with databases.
- Functions as an object oriented programming interface for designing data mining processes and procedures.

Understand Model Details

- Interactive model viewers



Target Value: 1

Sort by absolute value

Fetch Size: 10,000

Coefficients: 297 out of 297

Attribute	Value	Coefficient	Standardized Coeffi...	Exp(Coefficient)
<Intercept>	NULL	-1.83481346	0	6.26396556
TREATMENT_PLAN	Chemo_only	-0.46513283	0.11735002	1.59222567
WEEKDAY	W	-0.40697858	0.0869471	1.50227193
WEEKDAY	Th	-0.34941526	0.05883753	1.418238
RECURRENT	1	-0.33993936	0.07348783	1.4048624
STAGE	3	0.29916993	-0.06150948	0.74143341
FREQ_CHEMO	1	0.29378459	-0.06262496	0.74543705
FREQ_CHEMO	0	-0.26376819	0.05597178	1.30182638
IV_PAINMED	DEM	-0.26085980	0.036163	1.29804567
TREATMENT_PLAN	Chemo&Radiation	-0.25534174	0.03324906	1.2909027
TYPE_PROCD	closed	0.25466832	-0.01992872	0.77517356
PREOP_GI_MED	1	0.25194913	-0.06873117	0.77728428
MALIGNANCY	1	0.24061736	-0.05486614	0.78614238
QUARTER	A	0.23306129	-0.05746447	0.79210502
SIZE_REDUCTION	NULL	0.22915110	-0.15356344	0.79520837
TYPE_PROCD	1	-0.22759025	0.03846051	1.25557075
EPIDURAL	1	-0.22715954	0.05119796	1.25503009
INSURANCE	B	0.21168257	-0.05517357	0.80922152
OR_TRANSFUSIONS	1	0.20613024	-0.0550411	0.81372709
TYPE_ABX	Cipro	0.20248206	-0.02044382	0.81670114
EKG	SB	0.19228831	-0.02216336	0.82506896
IV_PAINMED	TORD	-0.19105185	0.01912802	1.21052222
INCISION	KNEE	-0.18882816	0.01878139	1.20783338
INSURANCE	C	0.18859100	-0.02710814	0.82812514
WT_LOSS_TIME	NULL	-0.17535293	0.11368976	1.19166672
WEEKDAY	Sa	0.17096336	-0.02674837	0.84285246

Details Coefficients Compare Settings



Oracle Data Mining & OBI 11g

ODM's predictions & probabilities are available in the Database for reporting using Oracle BI EE and other tools

8.4 Oracle Datamining						Home Catalog Dashboards New Open Signed In As				
LTV Prediction LTV Details Classification Tree LTV Probabilities What If Scoring Geo LTV P						Map				
Classification Tree										
Page Information (click to collapse or expand)										
Classification Tree						ails				
Time run: 12/9/2011 1:03:03 PM						3:03 PM				
20 Actual Unit Price										
		2008	2009	2010	Grand Total					
0 - All Individuals	MEDIUM	9,302	9,302	9,382	9,331					
1 - M_MARITAL_ST in 'DIVORCED', 'SINGLE'	MEDIUM	9,207	9,329	9,421	9,322					
2 - M_CRDT_RATE <= 657.5	LOW	9,225	9,164	9,377	9,261					
12 - M_INCOME_LVL in 'LEVEL 5', 'LEVEL 6', 'LEVEL 7', 'LEVEL 8', 'LEVEL 9'	MEDIUM	8,904	9,131	9,670	9,261					
13 - M_INCOME_LVL in 'LEVEL 1', 'LEVEL 2', 'LEVEL 3', 'LEVEL 4'	LOW	9,345	9,176	9,259	9,261					
3 - M_CRDT_RATE > 657.5	MEDIUM	9,193	9,462	9,454	9,370					
14 - M_MONTHS_CONTACT <= 12.5	VERY HIGH	8,815	9,418	8,690	8,951					
4 - M_MONTHS_CONTACT > 12.5	MEDIUM	9,242	9,468	9,543	9,421					
7 - M_MARITAL_ST in 'MARRIED', 'WIDOW'	HIGH	9,397	9,276	9,343	9,341					
1- Revenue										
		2008	2009	2010	Grand Total					
0 - All Individuals	MEDIUM	16,500,000	15,000,000	18,500,000	50,000,000					
1 - M_MARITAL_ST in 'DIVORCED', 'SINGLE'	MEDIUM	8,155,247	7,589,505	9,289,014	25,033,766					
2 - M_CRDT_RATE <= 657.5	LOW	3,560,875	3,340,550	4,015,646	10,917,071					
12 - M_INCOME_LVL in 'LEVEL 5', 'LEVEL 6', 'LEVEL 7', 'LEVEL 8', 'LEVEL 9'	MEDIUM	938,983	889,059	1,189,016	3,017,058					
13 - M_INCOME_LVL in 'LEVEL 1', 'LEVEL 2', 'LEVEL 3', 'LEVEL 4'	LOW	2,621,892	2,451,491	2,826,630	7,900,013					

#	M23 Full Rule	Predicted LTV	# of Cust	1- Revenue	Trend
12	M_MARITAL_ST in 'DIVORCED', 'SINGLE' ; AND M_CRDT_RATE <= 657.5; AND M_INCOME_LVL in 'LEVEL 5', 'LEVEL 6',	MEDIUM	0		
13	M_MARITAL_ST in 'DIVORCED', 'SINGLE' ; AND M_CRDT_RATE <= 657.5; AND M_INCOME_LVL in 'LEVEL 1', 'LEVEL 2',	LOW	0		
14	M_MARITAL_ST in 'DIVORCED', 'SINGLE' ; AND M_CRDT_RATE > 657.5; AND M_MONTHS_CONTACT <= 12.5	VERY HIGH	0		
15	M_MARITAL_ST in 'DIVORCED', 'SINGLE' ; AND M_CRDT_RATE > 657.5; AND M_MONTHS_CONTACT > 12.5;	MEDIUM	0		
16	M_MARITAL_ST in 'DIVORCED', 'SINGLE' ; AND M_CRDT_RATE > 657.5; AND M_MONTHS_CONTACT > 12.5;	LOW	0		
17	M_MARITAL_ST in 'DIVORCED', 'SINGLE' ; AND M_CRDT_RATE > 657.5; AND M_MONTHS_CONTACT > 12.5;	MEDIUM	0		
18	M_MARITAL_ST in 'DIVORCED', 'SINGLE' ; AND M_MONTHS_CONTACT > 12.5; AND M_INCOME_LVL in 'LEVEL 1', 'LEVEL 2',	HIGH	18	48,866	
19	M_MARITAL_ST in 'MARRIED', 'WIDOW' ; AND M_INCOME_LVL in 'LEVEL 1', 'LEVEL 2', 'LEVEL 3', 'LEVEL 4' ;	MEDIUM	0		
20	M_MARITAL_ST in 'MARRIED', 'WIDOW' ; AND M_INCOME_LVL in 'LEVEL 1', 'LEVEL 2',	HIGH	0		



Oracle Data Mining & Spatial

ORACLE Business Intelligence

8.4 Oracle Datamining

Search All Advanced Administration Help

Home Catalog Dashboards New Open Signed In As

LTV Prediction LTV Details Classification Tree LTV Probabilities What If Scoring **Geo LTV Prediction** Regression Regression Model Results Map

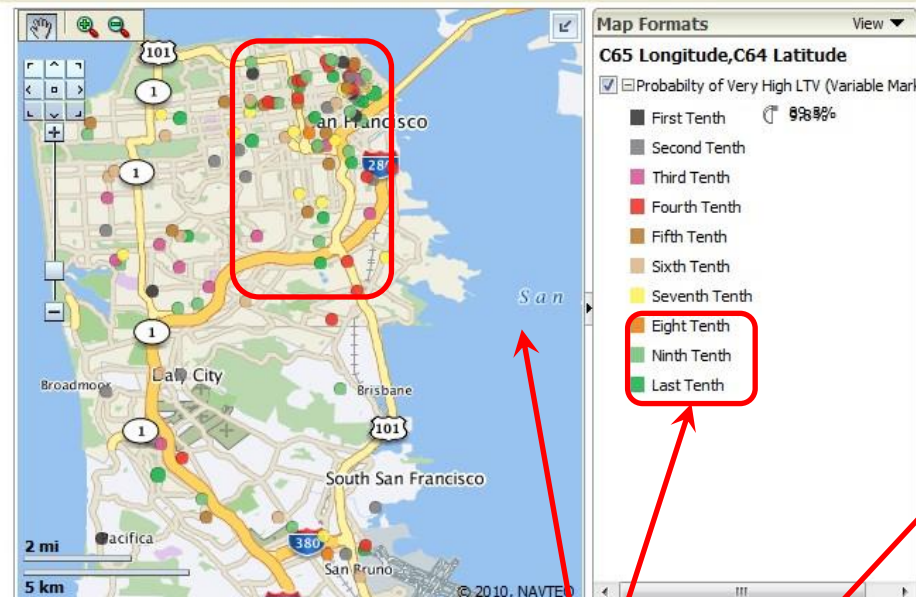
Geo LTV Prediction

[Return to Main Index page](#)

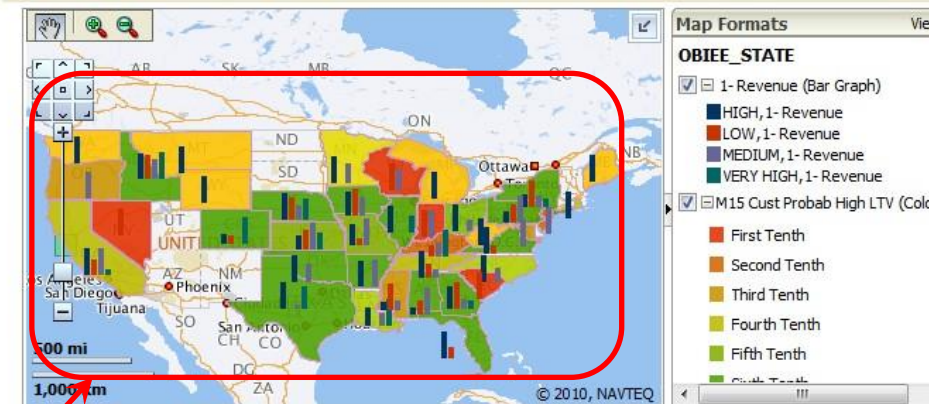
Page Information (click to collapse or expand)

Description : This dashboard combines the results of an Oracle Data Mining Classification model with a geographical visual representation of the probabilities in OBIEE maps.

Location LTV Prediction
Time run: 12/9/2011 12:55:35 PM



Geo LTV Prediction
Time run: 12/9/2011 12:55:35 PM



1- Revenue

	HIGH	LOW	MEDIUM	VERY HIGH	Grand Total
Alabama	32,041	62,229	124,009		218,279
Alaska	35,974			127,974	163,948
Arkansas	52,224		56,405		108,629
California	3,215,891	1,820,802	3,142,887	509,122	8,688,700

Customer "most likely" be
be HIGH and VERY HIGH
value customer in the future

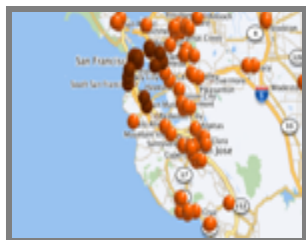


What is Spatial Data?

- Business data that contains or describes location
 - Street and postal address (customers, stores, factory, etc.)
 - Sales data (sales territory, customer registration, etc.)
 - Assets (cell towers, pipe lines, electrical transformers, etc.)
 - Geographic features (roads, rivers, parks, etc.)
- Anything connected to a physical location
- Any data sets that contain “link and node” relationships between data objects. Can be directional or non-directional.



Natively Manage All Geospatial Data



"Points"



"Lines"



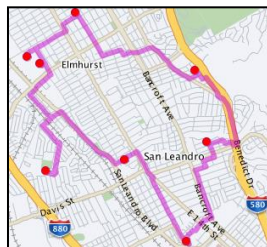
"Polygons"

Web Services
(OGC)

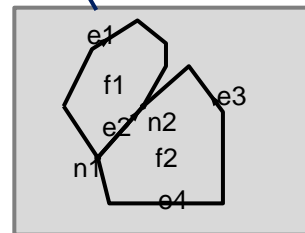
Geocoding
Routing



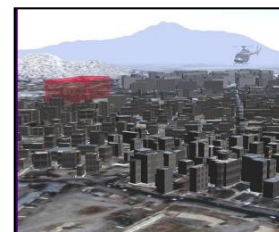
Rasters



Networks



Topologies



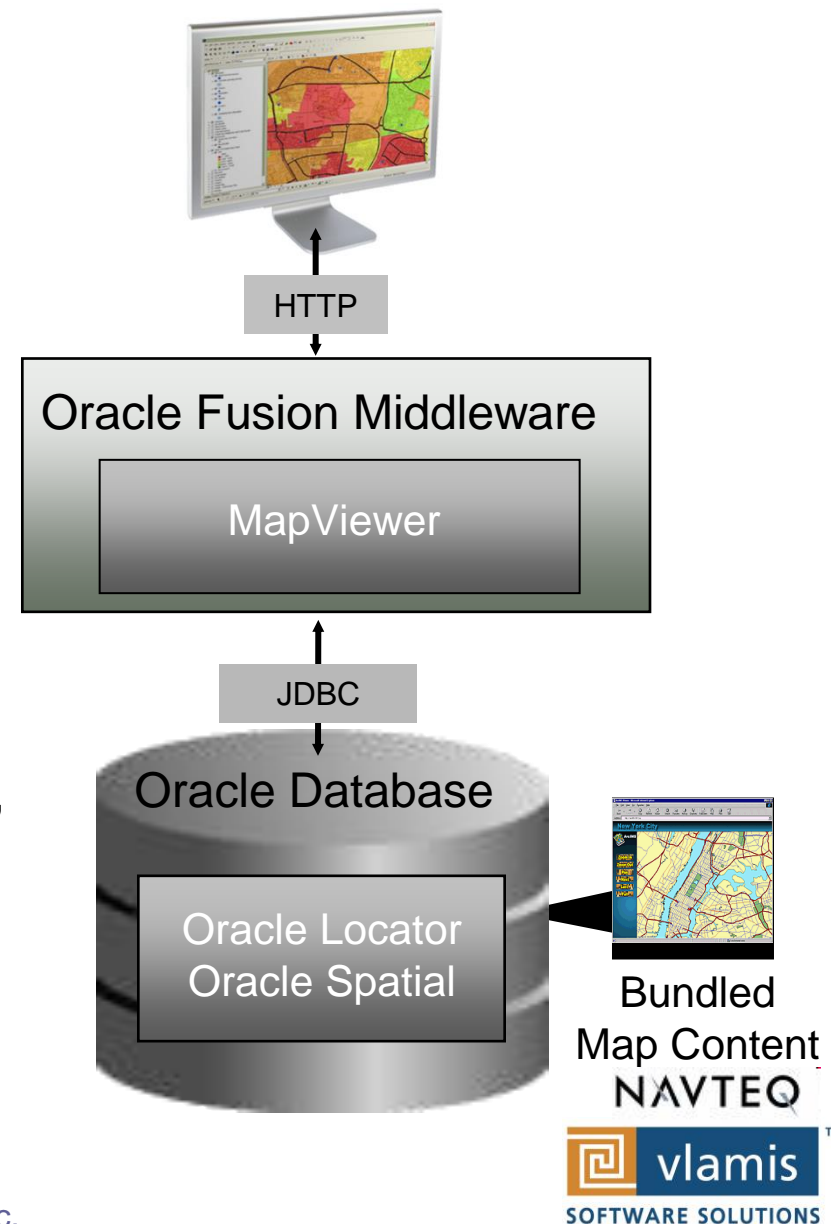
3D





Oracle Spatial Technologies

- **Oracle Locator:** Feature of Oracle Database XE, SE, EE
- **Oracle Spatial:** Priced option to Oracle Database EE
- **MapViewer:** Java application and map rendering feature of Oracle Fusion Middleware
- **Workspace Manager:** Long transactions feature of Oracle Database SE, EE
- **Bundled Map Content:** Major roads, administrative boundaries (city, county, state, country) - worldwide coverage from Navteq

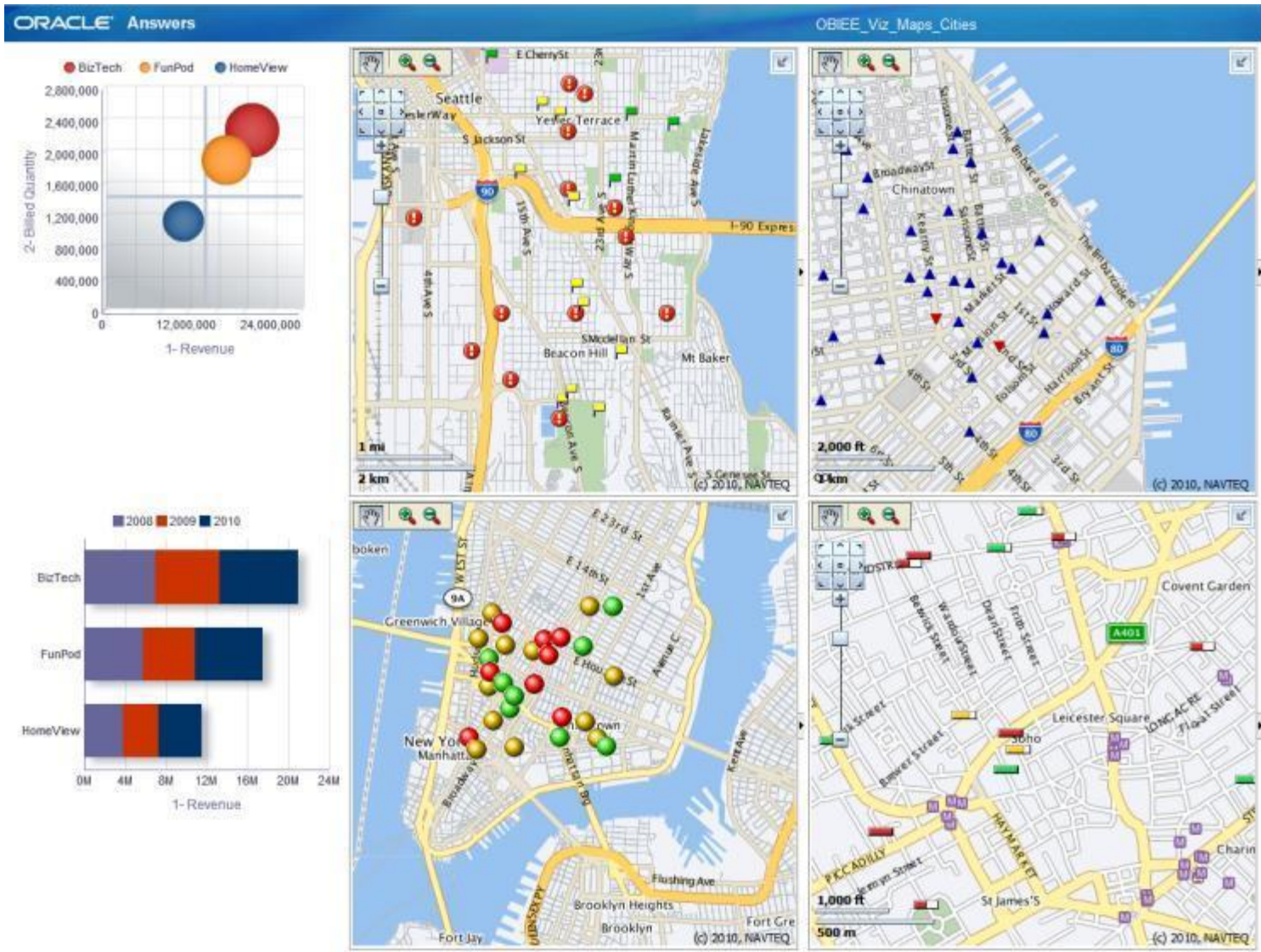




What is Spatial Data?

- Business data that contains or describes location
 - Street and postal address (customers, stores, factory, etc.)
 - Sales data (sales territory, customer registration, etc.)
 - Assets (cell towers, pipe lines, electrical transformers, etc.)
 - Geographic features (roads, rivers, parks, etc.)
- Anything connected to a physical location
- Any data sets that contain “link and node” relationships between data objects. Can be directional or non-directional.

Depict and Detect Spatial Relationships



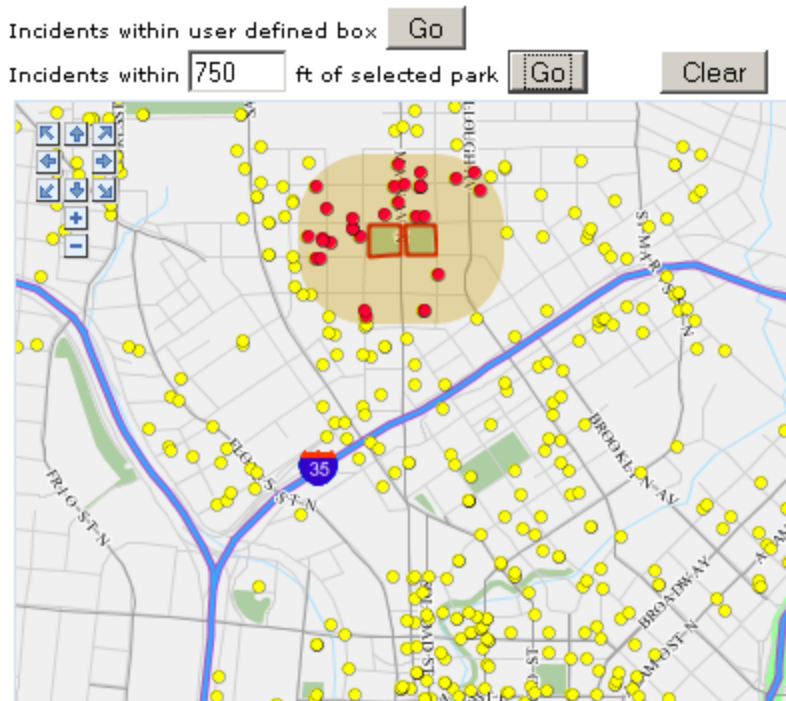


Why Maps are Powerful

Maps convey dense, multi-dimensional relationships in data faster and more intuitively than any other graphical display methodology.



Some Analysis Is Possible Only with Spatial Analytics



Show incidents within 750 ft
of selected park

Complaint Detail

Offense Desc	PD Desc	Date Key	Complaint Key	Service Area	Region
CRIMINAL MISCHIEF & RELATED OF	MISCHIEF, CRIMINAL 4, OF MOTOR	18-Feb-03	1026	28	Central
DANGEROUS DRUGS	CONTROLLED SUBSTANCE, POSSESSI	10-Nov-02	30099	28	Central
		10-Mar-03	40099	28	Central
HARRASSMENT 2	HARASSMENT,SUBD 1,CIVILIAN	02-Aug-03	1064	32	Central
	HARASSMENT,SUBD 3,4,5	04-Mar-03	1027	28	Central
		04-May-03	31027	28	Central
		04-Sep-03	41027	28	Central
		19-Sep-03	41028	28	Central
ROBBERY	ROBBERY,UNCLASSIFIED,OPEN AREA	09-Jan-04	41032	28	Central



Oracle Locator and Oracle Spatial

- Oracle Locator is a **feature** of both Oracle Standard and Enterprise Database Editions.
- Oracle Locator provides basic location functionality.
 - Point, line, and polygon spatial locations (SDO_GEOMETRY)
 - Spatial indexing
 - Spatial operators that use the spatial index for performing spatial inquiries.
- Oracle Spatial is an **option** for Oracle Database Enterprise Edition
 - Provides extensive support for advanced spatial processing and analytics including routing, vector and raster data, topology and network models, and more.



What is R?

CRAN Task Views

- R is an Open Source scripting language and environment for statistical computing and graphics
<http://www.R-project.org/>
- Popular alternative to SAS, SPSS & other proprietary statistical environments
- Around 2 million R users worldwide
- Thousands of R packages available



Bayesian	Bayesian Inference
ChemPhys	Chemometrics and Computational Physics
ClinicalTrials	Clinical Trial Design, Monitoring, and Analysis
Cluster	Cluster Analysis & Finite Mixture Models
DifferentialEquations	Differential Equations
Distributions	Probability Distributions
Econometrics	Computational Econometrics
Environmetrics	Analysis of Ecological and Environmental Data
ExperimentalDesign	Design of Experiments (DoE) & Analysis of Experimental Data
Finance	Empirical Finance
Genetics	Statistical Genetics
Graphics	Graphic Displays & Dynamic Graphics & Graphic Devices & Visualization
HighPerformanceComputing	High-Performance and Parallel Computing with R
MachineLearning	Machine Learning & Statistical Learning
MedicalImaging	Medical Image Analysis
MetaAnalysis	Meta-Analysis
Multivariate	Multivariate Statistics
NaturalLanguageProcessing	Natural Language Processing
OfficialStatistics	Official Statistics & Survey Methodology
Optimization	Optimization and Mathematical Programming
Pharmacokinetics	Analysis of Pharmacokinetic Data
Phylogenetics	Phylogenetics, Especially Comparative Methods
Psychometrics	Psychometric Models and Methods
ReproducibleResearch	Reproducible Research
Robust	Robust Statistical Methods
SocialSciences	Statistics for the Social Sciences
Spatial	Analysis of Spatial Data
SpatioTemporal	Handling and Analyzing Spatio-Temporal Data
Survival	Survival Analysis
TimeSeries	Time Series Analysis
gR	gRaphical Models in R



Oracle R Enterprise

- Part of the Advanced Analytics Option to the Oracle Database Enterprise Edition
- Provides transparent access to database-resident data from R
- Embedded R script execution through database managed R engines with SQL language integration
- Provides data and task parallelism and full power of Oracle database for R
- Enables advanced statistics for in-database execution
- Integrates R into the IT software stack
- Extends and enhances open source R



Oracle R Distribution



ability to dynamically load:

Intel Math Kernel Library (MKL)

AMD Core Math Library

Solaris Sun Performance Library



Oracle
Support

- Improved scalability at client and database for embedded R execution
- Enhanced linear algebra performance using Intel's MKL, AMD's ACML, and Solaris Sun Performance Library
- Enterprise support for customers of Oracle Advanced Analytics option, Big Data Appliance, and Oracle Linux
- Available as a free download from Oracle
- Oracle to contribute bug fixes and enhancements to open source R



Other R Offerings

- ROracle
 - Open source Oracle database interface driver for R based on OCI
 - Maintained by Oracle, optimizations and bug fixes released to open source community
- Oracle R Connector for Hadoop
 - R interface to Oracle Hadoop Cluster on Big Data Appliance
 - Access and manipulate data in HDFS, database, and file system
 - Write MapReduce functions using R and execute through R
- Rstudio
 - Popular open source user interface for R
 - Integrated Development Environment
- Rcommander
 - Extended GUI for R
 - Used for plugins (e.g. text mining)
 - Often used with Rstudio

R now integrated into OBIEE 11.1.1.7

8.13 Oracle R Enterprise

R Integration R End-User Interaction

R Workbench

R Sourcing from BI CEIM

R Sourcing from BI CEIM (end-user)

Custom Link

Home

Catalog

Favorites

Dashboards

New

Open

Signed In As Paulo Ro

Clear All Filters

Script Name

- ☐ Balloon Plot
- ☐ BusiestAirports-1
- ☐ Conditional Histogram
- ☐ Correlation Matrix Ellipses
- ☐ Correlation matrix circles
- ☐ Enhanced Bar Plot
- ☐ Heatmap
- ☐ Multipanel Geo Lattice Plot
- ☐ Multivariate Star Plot
- ☐ PredictDelays-build
- ☐ PredictDelays-score
- ☒ QCC-1
- ☐ RandomRedDots
- ☐ Times Series Forecasting
- ☐ Volcano Perspective Plot
- ☐ Volcano Perspective Plot2
- ☐ sincFunction

Param 1:

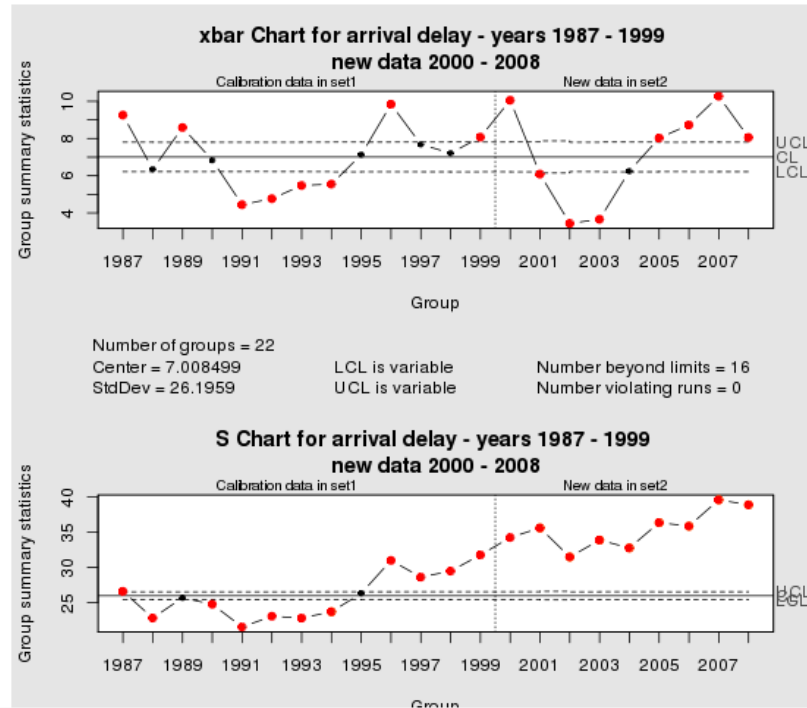
cursor(select q'ALL' UNIQUE_CARRIERS,q'both' GRA

Param 2:

cursor(select 550 "ore.png.height", 550 "ore.i

Apply

QCC-1 Graphics Result



QCC-1

Edit Script

```
function(dat) {  
  uniquecarrier <- dat$UNIQUE_CARRIERS  
  delayType <- dat$DELAY_TYPE  
  startYear <- dat$START_YEAR  
  endYear <- dat$END_YEAR  
  newDataStartYear <- dat$NEW_DATA_START_YEAR  
  newDataEndYear <- dat$NEW_DATA_END_YEAR  
  graphTypes <- dat$GRAPH_TYPES  
  ore.sync(table="ONTIME_S")  
  ore.attach()  
  library("qcc", lib.loc="/home/oracle/R/library")  
  if(graphTypes == "both") graphTypes <- c("xbar","S")  
  dat <- ONTIME_S  
  if(newDataStartYear != -1) {  
    dat <- dat[(dat$YEAR>=startYear & dat$YEAR<=endYear) |  
              (dat$YEAR>=newDataStartYear & dat$YEAR<=newDataEndYear),]  
  } else {  
    dat <- dat[dat$YEAR>=startYear & dat$YEAR<=endYear,]  
  }  
  if(uniquecarrier != "ALL") {  
    dat <- dat[dat$UNIQUECARRIER == uniquecarrier, ]  
  }  
  if (delayType=="arrival") {  
    dat <- dat[,c("ARRDELAY", "YEAR")]  
  } else {  
    dat <- dat[,c("DEPDELAY", "YEAR")]  
  }  
  names(dat) <- c("DELAY", "YEAR")  
  delayData <- ore.pull(dat)  
  library(qcc)  
  group <- qcc.groups(delayData$DELAY, delayData$YEAR)  
  numYears <- endYear - startYear + 1  
  par(mfrow=c(length(graphTypes),1))  
  for(gt in graphTypes) {  
    if (newDataStartYear == -1) { # not using new data comparison  
      title <- paste(gt, "Chart for", delayType, "delay -  
years", startYear, "-", endYear)  
      set1 <- group[1: numYears,]  
      obj <- qcc(set1, type=gt, title=title)  
    } else {  
      title <- paste(gt, "Chart for", delayType, "delay -  
years", startYear, "-", endYear)  
      set1 <- group[1: numYears,]  
      set2 <- group[numYears+1: numYears+numYears,]  
      obj <- qcc(set1, type=gt, title=title)  
    }  
  }  
}
```

R now integrated into OBIEE 11.1.1.7

R Workbench

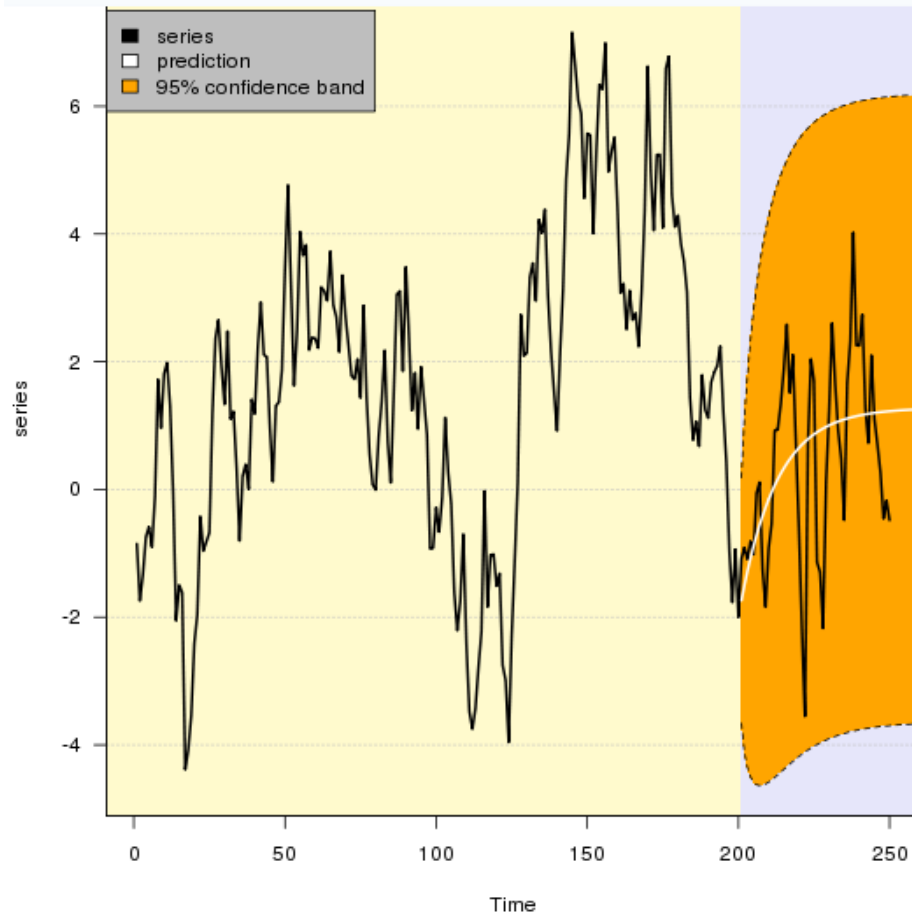
R Sourcing from BI CEIM

R Sourcing from BI CEIM (end-user)

R Results Object in RPD

BIP Sourcing from R

Quality Control Chart (BIP)



[Refresh](#)

```
require(gplots)

set.seed(120)
# simulate an AR(1) process
coefs <- 0.95
series <- arima.sim(list(ar=coefs),n=250)

# fit AR(1) with the 200 first data
model <- arima(series[1:200],c(1, # AR part
                                0, # I order
                                0)) # MA part

# make forecast from the model
forecast <- predict(model,80)

# compute the limits of the graph
ylim <- c( min(series[1:200],forecast$pred - 1.96 * forecast$se),
           max(series[1:200],forecast$pred + 1.96 * forecast$se))

# prepare the space where to plot
opar <- par(mar=c(4,4,2,2),las=1)
plot(series,ylim=ylim,type="n",xlim=c(1,250))
usr <- par("usr")

# split the figure in two parts
# - the part used to fit the model
rect(usr[1],usr[3],201,usr[4],border=NA,col="lemonchiffon")

# - the part used to make the forecast
rect(201,usr[3],usr[2],usr[4],border=NA,col="lavender")

abline(h= (-3:3)*2, col="gray", lty=3)

# draw a 95% confidence band
polygon( c(201:280,280:201),
         c(forecast$pred - 1.96*forecast$se,rev(forecast$pred + 1.96*f
         col="orange",
         lty=2,border=NA)

lines( 201:280, forecast$pred - 1.96*forecast$se, lty=2)
lines( 201:280, forecast$pred + 1.96*forecast$se, lty=2)
```



Oracle Test Drive

- Free to try out Oracle BI
- Go to www.vlamis.com/testdrive-registration/
- Runs off of Amazon AWS
- Hands-on Labs based on Collaborate 2012 HOLs
- Test Drives for:
 - Oracle BI
 - BI Publisher
 - Microsoft Excel against Oracle OLAP
 - Oracle Data Mining
 - Map Views in OBIEE
- Once sign up, you have private instance for 5 hours
- Available now



Thank You!

Thank You for Attending Session **Blazing BI with Oracle DB Analytical Options**

Presenter Information

Dan Vlami, President, dvlamis@vlamis.com

Vlami Software Solutions, Inc.

816-781-2880

For more information go to www.vlami.com