

Autonomous Machine Learning in Oracle Analytics

Enterprise Predictive Analytics for Everyone

Tim Vlamis

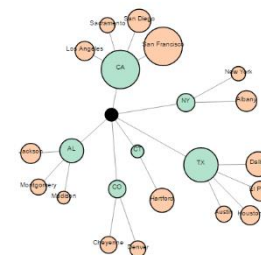
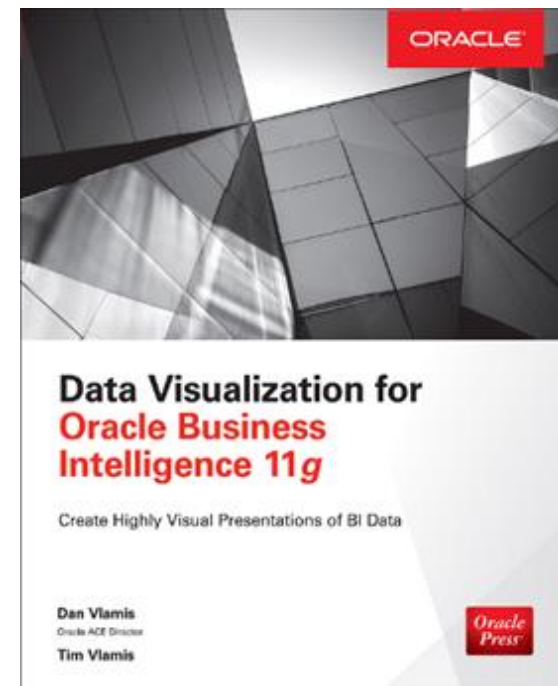
Dan Vlamis

Vlamis Software Solutions

August 19, 2020

VlamiS Software Solutions

- Founded in 1992 in Kansas City, Missouri
- 400+ Enterprise Clients
- Consults in :
 - Enterprise Business Intelligence & Analytics
 - Analytic Warehousing
 - Machine Learning and Predictive Analytics
 - Data Visualization
 - ETL and data integration
- Multiple Oracle ACEs, consultants average 15+ years
- Creators of the [Force Directed Graph Plugin](#) on [Oracle Analytics Library](#)
- www.vlamiS.com (blog, papers, newsletters, services)
- Co-authors of book “Data Visualization for OBI 11g”





Oracle Analytics New Features Webcasts

- 11.1.1.3 (August 2010)
 - Huge release with major new functionality
- 11.1.1.5 (May 2011)
 - BIWA TechCast May 25: [Oracle BI 11.1.1.5 - New Features](#)
- 11.1.1.6 (February 2012)
 - BIWA TechCast Feb 29: [Oracle BI 11.1.1.6 New Features](#)
- 11.1.1.6.2 and 11.1.1.6.2 BP1 (May 2012 and June 2012)
 - BIWA TechCast Aug 1: [Oracle BI 11.1.1.6.2 BP1 New Features](#)
- 11.1.1.7 (April 2013)
 - BIWA TechCast Apr 5 [Oracle BI 11.1.1.7 New Features](#)
- 11.1.1.9 (May 2015)
 - BIWA TechCast Jun 4, 2015 [Oracle BI 11.1.1.9 New Features](#)
- 12.2.1 (October 2015)
 - BIWA TechCast Nov 19, 2015 [Oracle BI 12.2.1 New Features](#)
- BICS March 2016 (March 2016)
 - Webcast Apr 5, 2016 BICS [BICS March 2016 New Features](#)
- 12.2.1.1 (June 2016)
 - BIWA TechCast Jul 19, 2016 [Oracle BI 12.2.1.1 New Features](#)
- 12.2.1.2 and DVD 2.0 (October 2016)
 - Webcast Nov 22, 2016 [DVD 12.2.2 and OBIEE 12.2.1.2. New Features](#)
- DVD 3.0 (June 2017)
 - Webcast Aug 29, 2017 [DVD 3.0 New Features](#)
- DVD 4 (October 2017)
 - Webcast Nov 9, 2017 [DVD 4 New Features](#)
- OAC 4 (December 2017)
 - Webcast Dec 19, 2017 [OAC 4 New Features](#)
 - Webcast Apr 4, 2018 [Migration from OBIEE to OAC Webinar](#)
- 12.2.1.4 (June 2018)
 - Webcast Jun 19, 2018 [Oracle BI 12.2.1.4 New Features](#)
- OAC 18.3.3 (September 2018)

Autonomous Machine Learning in Oracle Analytics: Enterprise Predictive Analytics for Everyone

Webcast Aug 19, 2020

- OAC 105.3 (June 2019)
 - Webcast Jul 25, 2019 [OAC 105.3 New Features](#)
- OAC 105.4 (October 2019)
 - Webcast Nov 5, 2019 [OAC 105.4 New Features](#)
- OAC 5.5 (January 2020)
 - Webcast Mar 5, 2020 [OAC 5.5 New Features](#)
- OAS 5.5 (January 2020)
 - **Webcast Mar 11, 2020 [Migrating OBIEE to OAS 5.5](#)**
- **OAC 5.6 (May 2020) OAC 5.7 (July 2020)**
 - **Webcast Aug 5, 2020 [OA 5.6 & 5.7 New Features](#)**

All listed on website at <http://www.vlamiS.com/oaweinars/>



OAC 5.6 and 5.7 New Features Overview

Visualizations

1. Pivot table column & width controls (5.6)
2. Background images in visuals (5.6)
3. Waterfall Visualization Controls (5.6)
4. Waterfall Bridge Report (5.7)
5. Spark Chart on Performance Tiles (5.7)

Map Enhancements

6. Digitized Custom Map Layers (5.6)
7. Map controls (5.6)
8. Map Warnings and performance with large data sets (5.6)

Filtering

9. List Box Filter Viz (5.7)
10. Filter Options in Properties (5.7)
11. Canvas Property: Show Use as Filter (5.7)
12. Top/Bottom N for attributes (5.7)

Front End

13. Duplication of Custom Calculations (5.6)
14. Canvas tab enhancements (5.6)
15. Layout Container & Viz spacing Enhancements (5.7)

Data Sets and Connectivity

16. Simplified Data Set Editor (5.6)
17. Regular Expressions in Data Prep (5.6)
18. Remote Data Gateway support for MySQL (5.6)
19. Microsoft Azure SQL Database Certification (5.6)
20. Direct Connectivity to EPM Cloud (5.6)
21. Improved LIVE Data sets (5.7)

Machine Learning

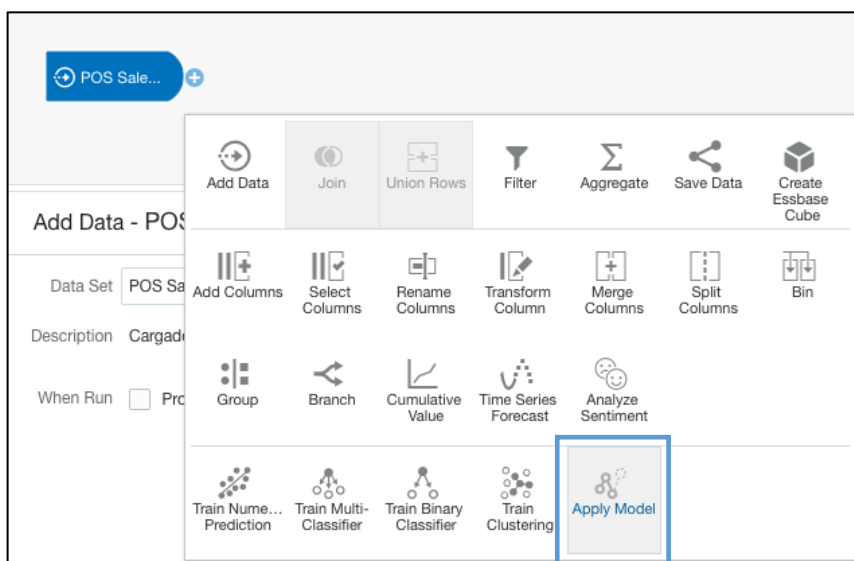
22. Additional outputs for ML Models (5.6)
23. Register & Apply OML Models (5.7)
24. Database Analytics Operation (5.7)
25. Un-pivot, Sampling (5.7)
26. Dynamic Clustering, Dynamic Prediction (5.7)

Administration

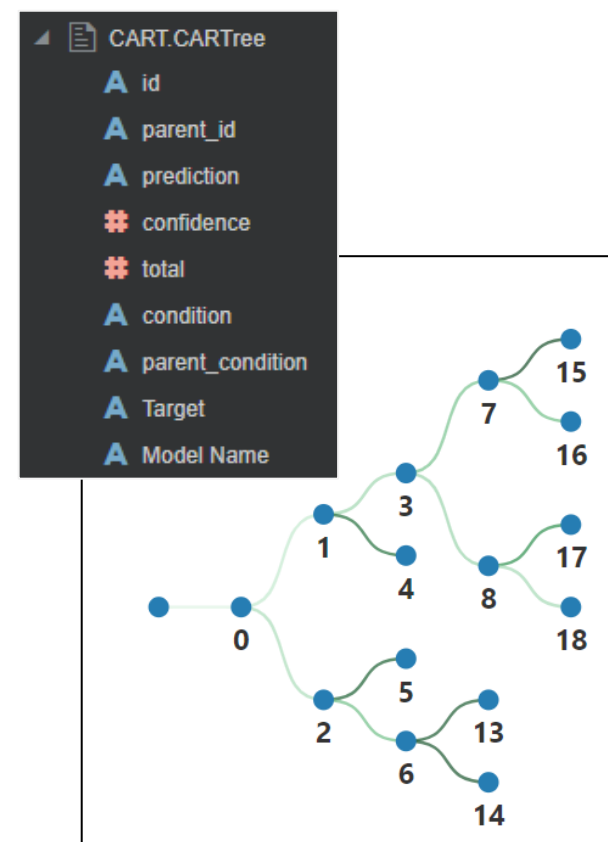
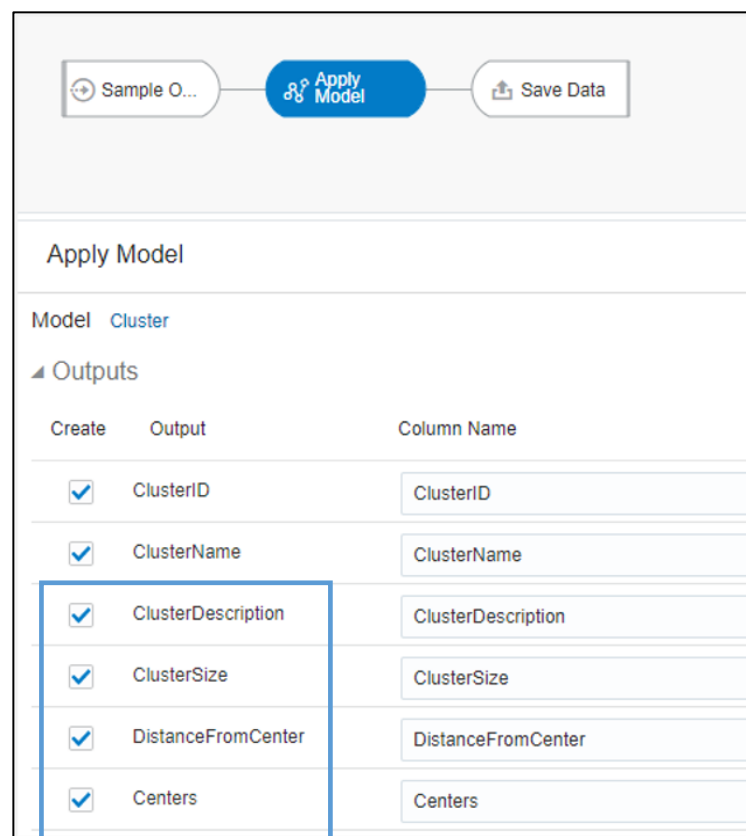
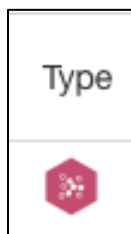
27. Data Actions to Publisher Report (5.6)
28. Access Controls – permissions (5.6)
29. Mobile Day by Day Admin Mode (5.6)
30. System Settings UI Improvements (5.6)
31. Configure Access Control Rule when create OAC instance (5.6)

ML Output enhancements

- When using Data Flow and applying a Clustering model you can leverage additional output columns that surface values in the model that can be visualized with the network chart for deeper insights.



In Data Flow, choose your Data Set and add the Apply Model step. Choose a **Clustering** Model type and you will automatically see the additional clustering options.





Register ML Model from Oracle DB/ADW

Register external ML Model

Create

Import Project/Flow...

Data Set Management

Open Data Modeler

Register ML Model

Open Classic Home

Customize Home Page...

Select ADW/Oracle DB Connection

Register an ML Model

From a Connection

Type	Name	Description
	BugDB_Conn	BugDB_Conn
	BugDB_Connection	
	bur10347	
	DemoADW	
	DemoOML	

Select a Model to Register

Type	Name	Last Modified
	CLUS_OC_1_5	Jun 15, 2016
	CLUS_SAMPLE	Oct 21, 2019
	CLUS_SAMPLE	Oct 24, 2019
	H_CLAS_SAMPLE	Oct 22, 2019
	H_REGR_SAMPLE	Nov 12, 2019
	CLUS_SAMPLE	Oct 17, 2019
	CLAS_SAMPLE	Oct 16, 2019
		Feb 18, 2020
	SAMPLE	Oct 23, 2019

Name

Description

Model Info

Model Name

Model Class

Algorithm

Description

Owner

Created On

Target

Input Columns

Output Columns

Parameters

Cancel Register

Register ML model to score OAC datasets

Several model details available



OML Integration – Apply OML model

Score OAC dataset using ML model

MINING_... Filter Apply Model Save Data

Apply Model

Model NB_SH_CLAS_SAMPLE

Outputs

Create	Output	Column Name
<input checked="" type="checkbox"/>	Prediction	Prediction
<input checked="" type="checkbox"/>	PredictionProbability	PredictionProbability
<input checked="" type="checkbox"/>	PredictionCost	PredictionCost
<input checked="" type="checkbox"/>	PredictionDetails	PredictionDetails

Model	Input
AGE	AGE
HOME_THEATER_PACKAGE	HOME_THEATER_PACKAGE
CUST_GENDER	CUST_GENDER

Apply Model lists OAC built and External Models

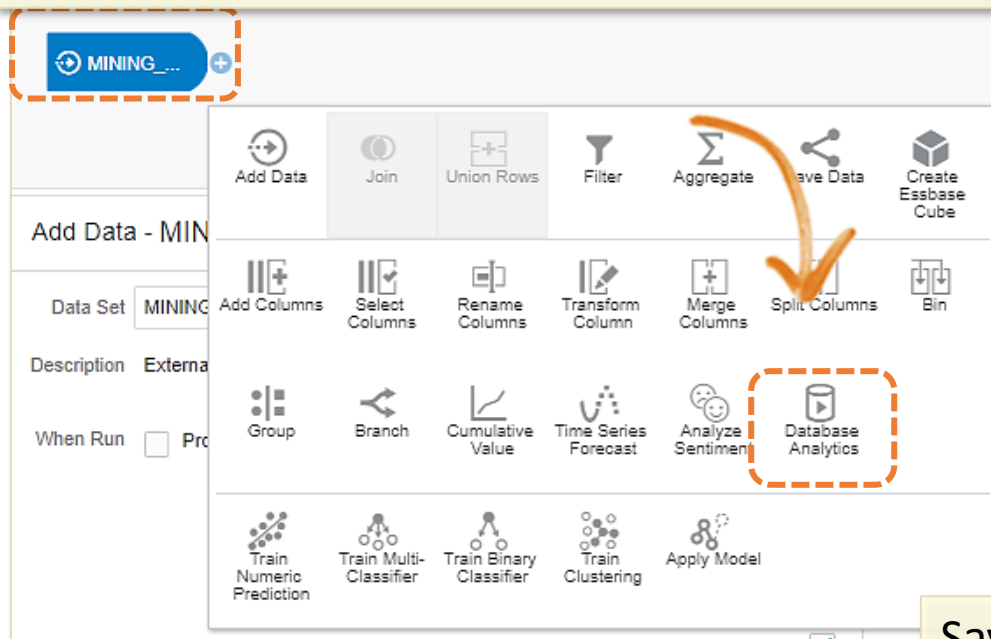
Type and Source differentiate OAC built vs External Model

Select Model				
Type	Name	Target	Author	Source
	Predict Age using CART	AGE	admin	Oracle Analytics Model
	NB_SH_CLAS_SAMPLE	AFFINITY_CARD	admin	Oracle Database
	ODM_LTV_BIN	M_LTV_BIN	admin	Oracle Database



Advanced Analytics : Database Analytics

Database Analytics Node available in Data Flows only for Oracle ADW/Oracle DB Datasets



Saves results and executes directly in underlying database

Select Database Analytics

Four Database Analytics Operations available

Analytics Operation	Description
Dynamic Anomaly Detection	This script performs dynamic anomaly detection on the input data without a pre-d
Dynamic Clustering	This script performs dynamic clustering on the input data without a pre-defined m
Sampling Data	The Oracle database ora_hash clause instructs the database to select a random
Un-pivoting Data	The DB Unpivot clause allows to transpose data that is stored in columns, into row



DB Analytics – Un-pivoting Data

ab QUANTILE	ab COL_2003	ab COL_2004	ab COL_2005	ab COL_2006	99 COL_2007	99 COL_2008
MED	8.	7.	6.	5.	6	5
MED	3.	2.	4.	0	0	1
MED				1.	3	0
MED			0	0	0	0
MED			0	0.	0	0
MED			8	8.	8	6
MED	24	25	24	25	25	25
MED		0.	0.	0.	0	0
MED					2	2
MED				0	0	0
MED		11	11	11	11	11
MED					7	12
MED					0	0
MED			0	0	0	0
MED	0	0	0	0	0	0
MED					7	0

ab Quantile	ab Geography	ab Income_Group	ab Year	ab Value
MED	AT	QU1	COL_2005	6.
MED	AT	QU1	COL_2006	5.
MED	AT	QU1	COL_2007	6
MED	AT	QU1	COL_2008	5
MED	AT	QU1	COL_2009	5
MED	AT	QU1	COL_2010	5.099999999999999
MED	AT	QU1	COL_2011	5.5
MED	AT	QU1	COL_2012	5.4
MED	AT	QU1	COL_2013	5.7
MED	AT	QU1	COL_2014	6.1
MED	AT	QU1	COL_2015	6.5
MED	BE	QU1	COL_2003	3.
MED	BE	QU1	COL_2004	2.
MED	BE	QU1	COL_2005	4.
MED	BE	QU1	COL_2006	0
MED	BE	QU1	COL_2007	0
MED	BE	QU1	COL_2008	1

TAX_CON...

Database Analytics

Database Analytics

Analytics Operation: Un-pivoting Data

Outputs

Create	Output	Column Name
<input checked="" type="checkbox"/>	MappedAttribute1	Quantile
<input checked="" type="checkbox"/>	MappedAttribute2	Geography
<input checked="" type="checkbox"/>	MappedAttribute3	Income_Group
<input checked="" type="checkbox"/>	GeneratedAttribute	Year
<input checked="" type="checkbox"/>	AttributeValue	Value

Parameters

* MappedAttribute1 QUANTILE

The column to be included in the output dataset

MappedAttribute2 GEO

The column to be included in the output dataset

MappedAttribute3 INCOME_GROUP

The column to be included in the o/p dataset

* Start Column COL_2003

Starting column for unpivoting

* End Column COL_2015



What is Machine Learning?

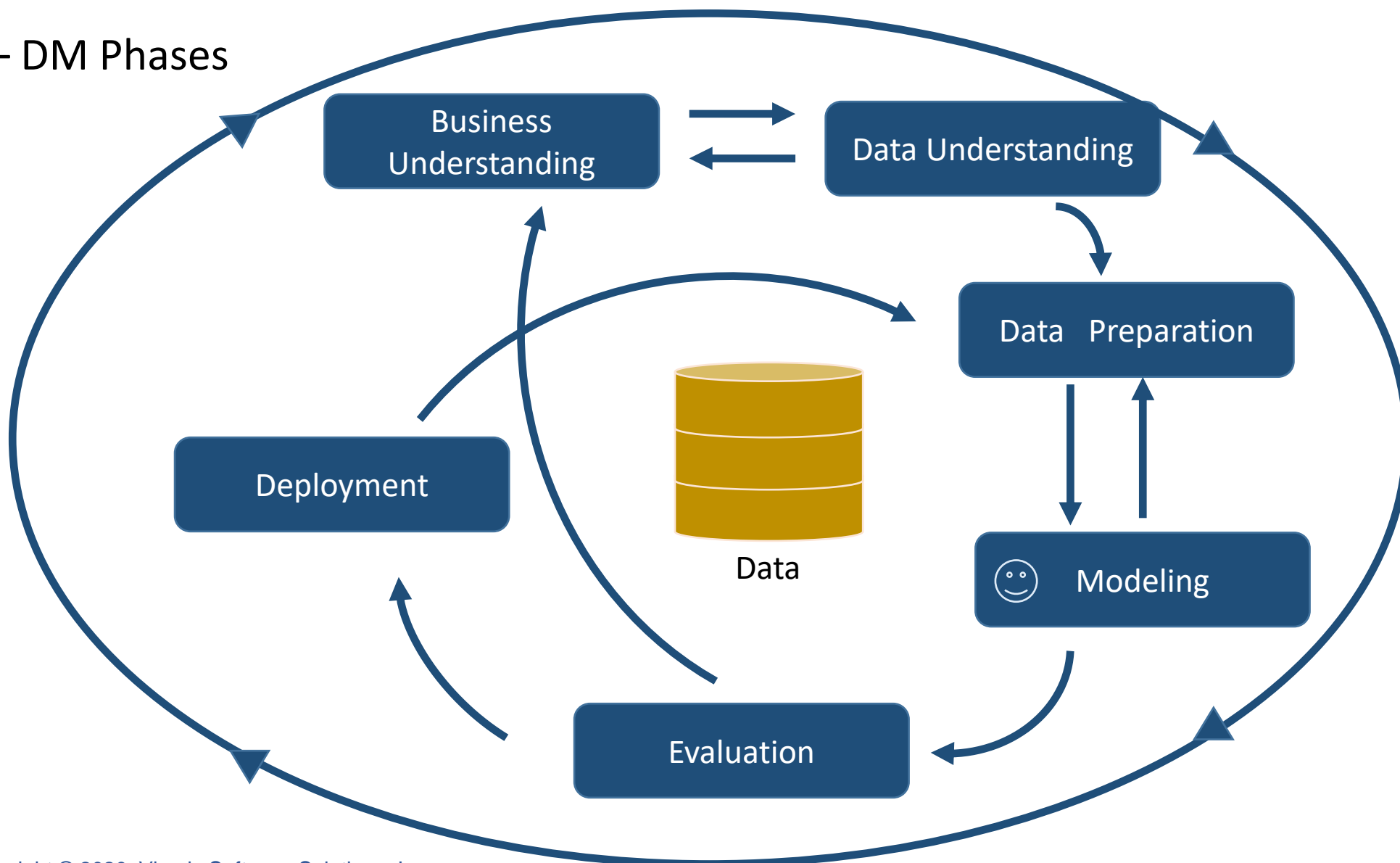
- The application of advanced analytic algorithms which automatically update their predictions over time.





Implications of Systems That “Learn”

CRISP – DM Phases





Implications of Systems that “Learn”

- Machine Learning implies a far higher level of involvement and connectedness.
- Being in production implies being at the center of the system.
- Deployment becomes even more important.



Four Realms of Analytics

Probability Based

**Diagnostic
Analytics**

**Predictive
Analytics**

Rules Based

**Descriptive
Analytics**

**Prescriptive
Analytics**

Past

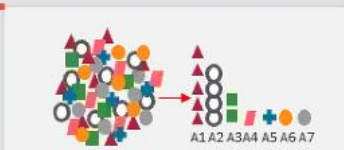
Future



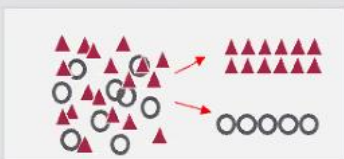
Machine Learning Use Cases

Algorithms **automatically** sift through large amounts of data to discover hidden patterns, new insights and make predictions

Supervised Learning

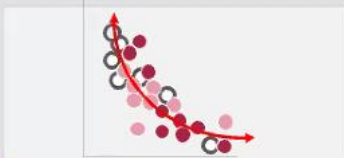


Identify most important factor (Attribute Importance)



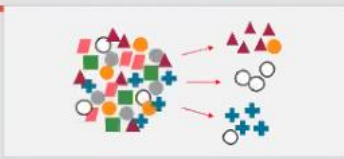
Predict customer behavior (Classification)

Find profiles of targeted people or items (Classification)



Predict or estimate a value (Regression)

Unsupervised Learning



Segment a population (Clustering)



Find fraudulent or "rare events" (Anomaly Detection)



Determine co-occurring items in a "basket" (Associations)



Many Options for ML in OA

- Oracle Analytics
 - Work in OAC Data Visualization Project
 - One-click options
 - Use built-in scripts via My Calculations
 - Automated data enrichment in Prepare tab
 - Use Explain for data profiling and unsupervised learning
 - Train and apply models in Data Flows
- Oracle Database (ADW, Exadata, DBCS, etc.)
 - Connect to Oracle Machine Learning in Autonomous Data Warehouse Cloud Service
 - Zeppelin Notebooks
 - Connect to Oracle Machine Learning in Oracle Database
 - Data Miner or RStudio



Many Options for ML in OAC

- 4 Different interfaces with OAC
 - Data Visualization Project
 - Data Flows train and apply models
 - Machine Learning Inspect Models
 - Classic interface Answers & Dashboards
- 4 Different Places in DV Projects
 - Visualization on Canvas right click
 - “My Calculations” custom script
 - “Explain”
 - Prepare machine learning enrichments



Demo of Single-click ML with use cases and limitations

- Clustering
- Outliers
- Trend
- Forecast



Single-click OAC ML Don'ts

- Do not overinterpret
- Do not be afraid to use
- Do not spend a lot of time when you have more powerful tools



Explain feature has power

- Demo of “Explain” with use cases and limitation

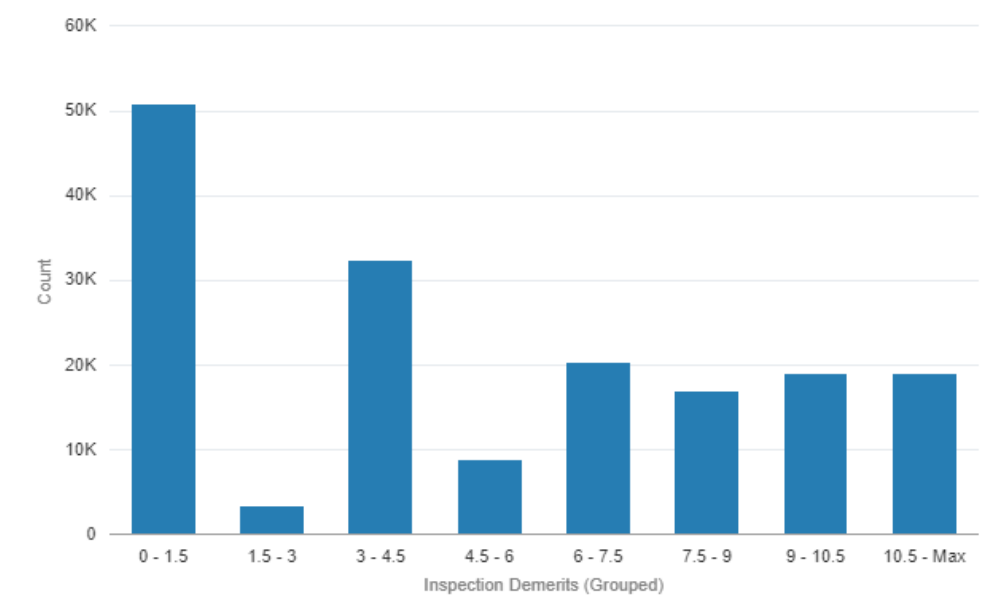
Basic Facts about Inspection Demerits

What are the values of Inspection Demerits and how do they relate to each other?

Anomalies of Inspection Demerits

What groups in the data exhibit unexpected results for Inspection Demerits?

Basic facts about Inspection Demerits



Inspection Demerits is a Numeric Measure, whose sum across 170,463 rows is 979,790.00. The values of Inspection Demerits on each row range from 0.00 to 250.00 and is 5.00 on average.

The charts below summarize the values of Inspection Demerits by the measures in this data set. Click the checkmarks above any of the visuals to add them to your project when done.



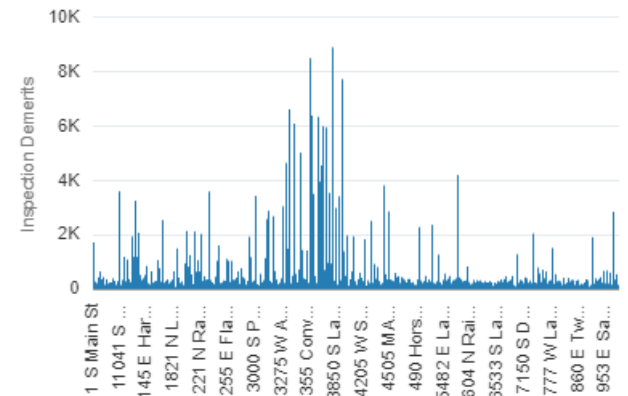
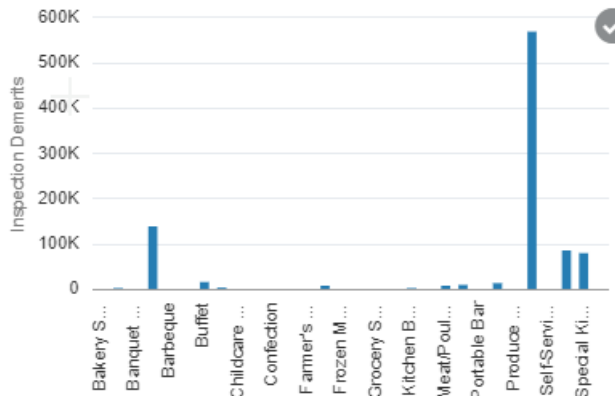
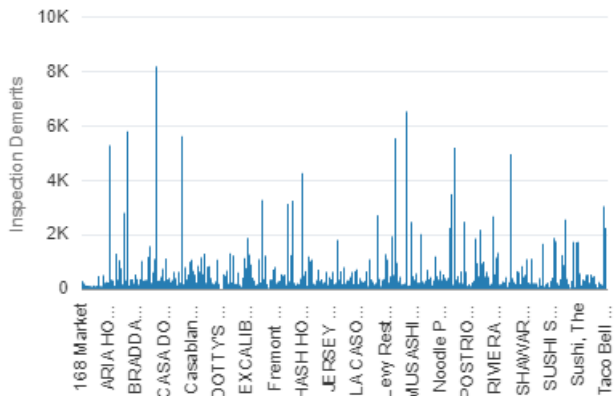
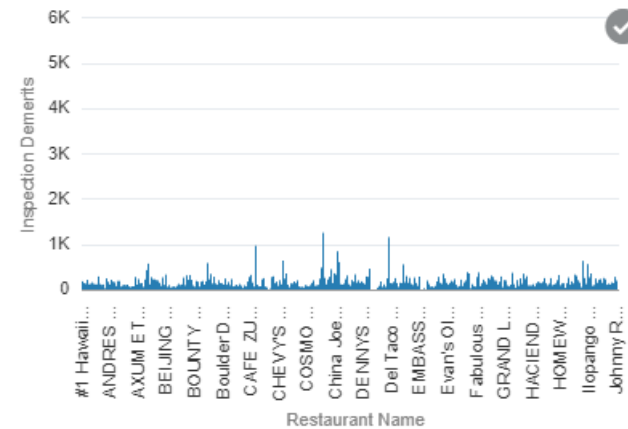
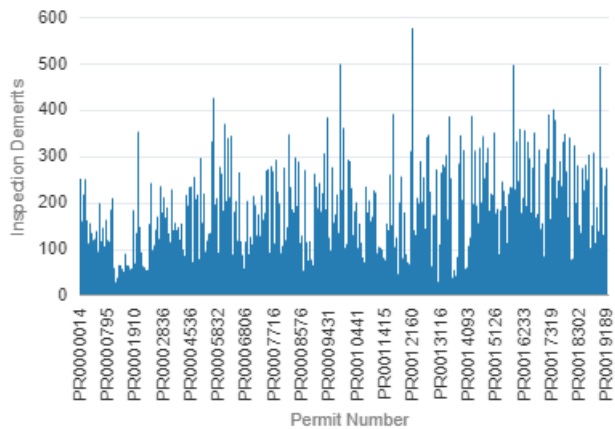
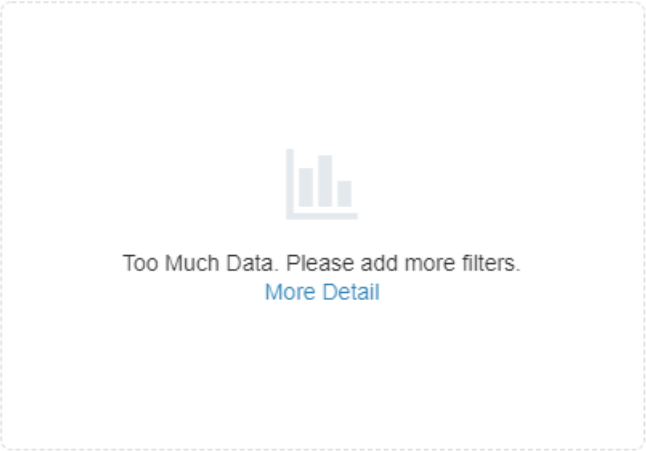
Basic Facts about Inspection Demerits

What are the values of Inspection Demerits and how do they relate to each other?

Anomalies of Inspection Demerits

What groups in the data exhibit unexpected results for Inspection Demerits?

The charts below summarize the values of Inspection Demerits by the measures in this data set. Click the checkmarks above any of the visuals to add them to your project when done.



Basic Facts about Inspection Demerits

What are the values of Inspection Demerits and how do they relate to each other?

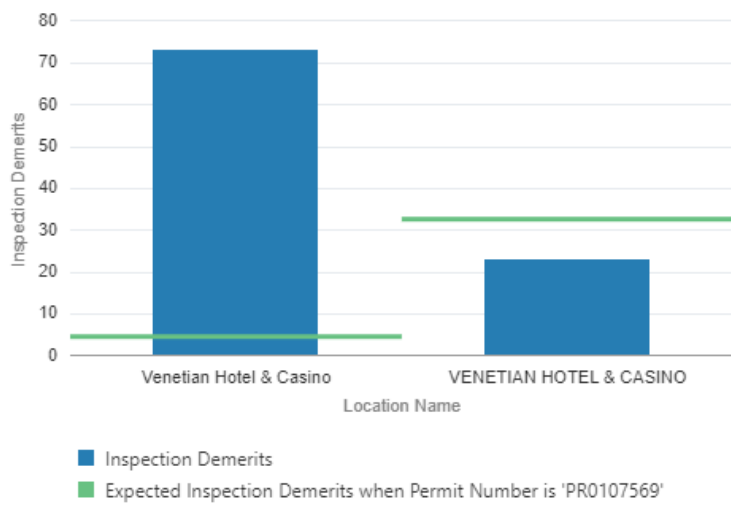
Anomalies of Inspection Demerits

What groups in the data exhibit unexpected results for Inspection Demerits?

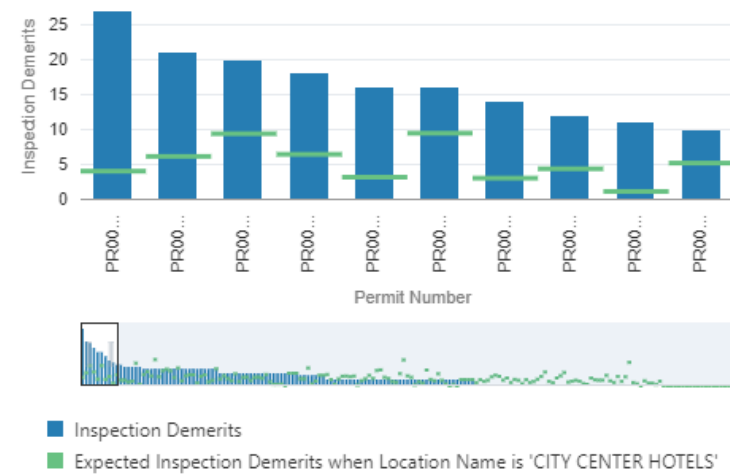
Anomalies of Inspection Demerits

231 combinations of 22 dimensions are being analyzed. Here are the top outliers for Inspection Demerits

When Permit Number is PR0107569, we expected Inspection Demerits for Location Name: Venetian Hotel & Casino to be 4.65, however, it is 73.00, representing a difference of 68.35.



When Location Name is CITY CENTER HOTELS, we expected Inspection Demerits for Permit Number: PR0023967 to be 4.08, however, it is 27.00, representing a difference of 22.92.





OAC Data Flows

- Requires some knowledge of data mining/machine learning
- Build/train models in data flows
- Use classification for single/multi-class predictions
 - Churn models of customer loyalty
 - Predict buy/not buy specific products
- Use regression models for prediction of continuous values
 - Lifetime customer value prediction
 - Next year customer purchases
- Use clustering to segment members into groups
 - Customer segmentation based on history/buying behavior
- Use sentiment analysis to understand comments



Building Models in Data Flows (demo)

- Can use OAC models or OML models from ADW/DBCS
- Classification methods and algorithms in OAC
- Interpretation of results
 - Confusion matrix
 - Accuracy, precision, and recall
 - ROC curve
- Visualizing classification results
- Applying classification models
- Use cases



Other ML Models in OAC Data Flows

- Sentiment analysis
- Regression
- Clustering



Oracle ML in Autonomous Data Warehouse

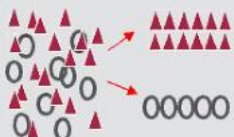
- In-database machine learning
 - Don't move the data
 - Extreme power and scalability
 - Extreme flexibility and extendibility
- Zeppelin Notebooks
 - Great for collaboration
 - Powerful tool in the hands of knowledgeable
 - Built-in visualization capability
 - Script development and management
- Oracle Data Mining Algorithms



Oracle DB Machine Learning Algorithms

CLASSIFICATION

Naïve Bayes
Logistic Regression (GLM)
Decision Tree
Random Forest
Neural Network
Support Vector Machine
Explicit Semantic Analysis
XGBoost*



CLUSTERING

Hierarchical K-Means
Hierarchical O-Cluster
Expectation Maximization (EM)



ANOMALY DETECTION

One-Class SVM
MSET-SPRT*



TIME SERIES

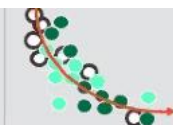
Forecasting - Exponential Smoothing
Includes popular models
e.g. Holt-Winters with trends,
seasonality, irregularity, missing data



Includes support for Partitioned Models, Transactional data and aggregations, Unstructured data, Geo-spatial data, Graph data. etc,

REGRESSION

Linear Model
Generalized Linear Model
Support Vector Machine (SVM)
Stepwise Linear regression
Neural Network
XGBoost*



ATTRIBUTE IMPORTANCE

Minimum Description Length
Principal Comp Analysis (PCA)
Unsupervised Pair-wise KL Div
CUR decomposition for row & AI



ASSOCIATION RULES

A priori/ market basket

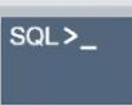


PREDICTIVE QUERIES

Predict, cluster, detect, features

SQL ANALYTICS

SQL Windows
SQL Patterns
SQL Aggregates



* New in 20c

FEATURE EXTRACTION

Principal Comp Analysis (PCA)
Non-negative Matrix Factorization
Singular Value Decomposition (SVD)
Explicit Semantic Analysis (ESA)

TEXT MINING SUPPORT

Algorithms support text
Tokenization and theme extraction
Explicit Semantic Analysis (ESA) for
document similarity



STATISTICAL FUNCTIONS

Basic statistics: min, max,
median, stdev, t-test, F-test, Pearson's,
Chi-Sq, ANOVA, etc.



R & PYTHON * Coming soon

Third-party R & Python Packages
through Embedded Execution
Spark MLlib algorithm integration



MODEL DEPLOYMENT & MONITORING * Coming soon

SQL—1st Class Objects
Oracle RESTful API (ORDS)
OML Web Services (for Apps)





Demo of Oracle Machine Learning in ADW



Oracle Data Miner in Database Cloud Service

- GUI for building predictive analytics workflows
- Build scripts for oracle database without coding
- Powerful built-in visualizations for interpretation



Demo of Oracle Data Miner



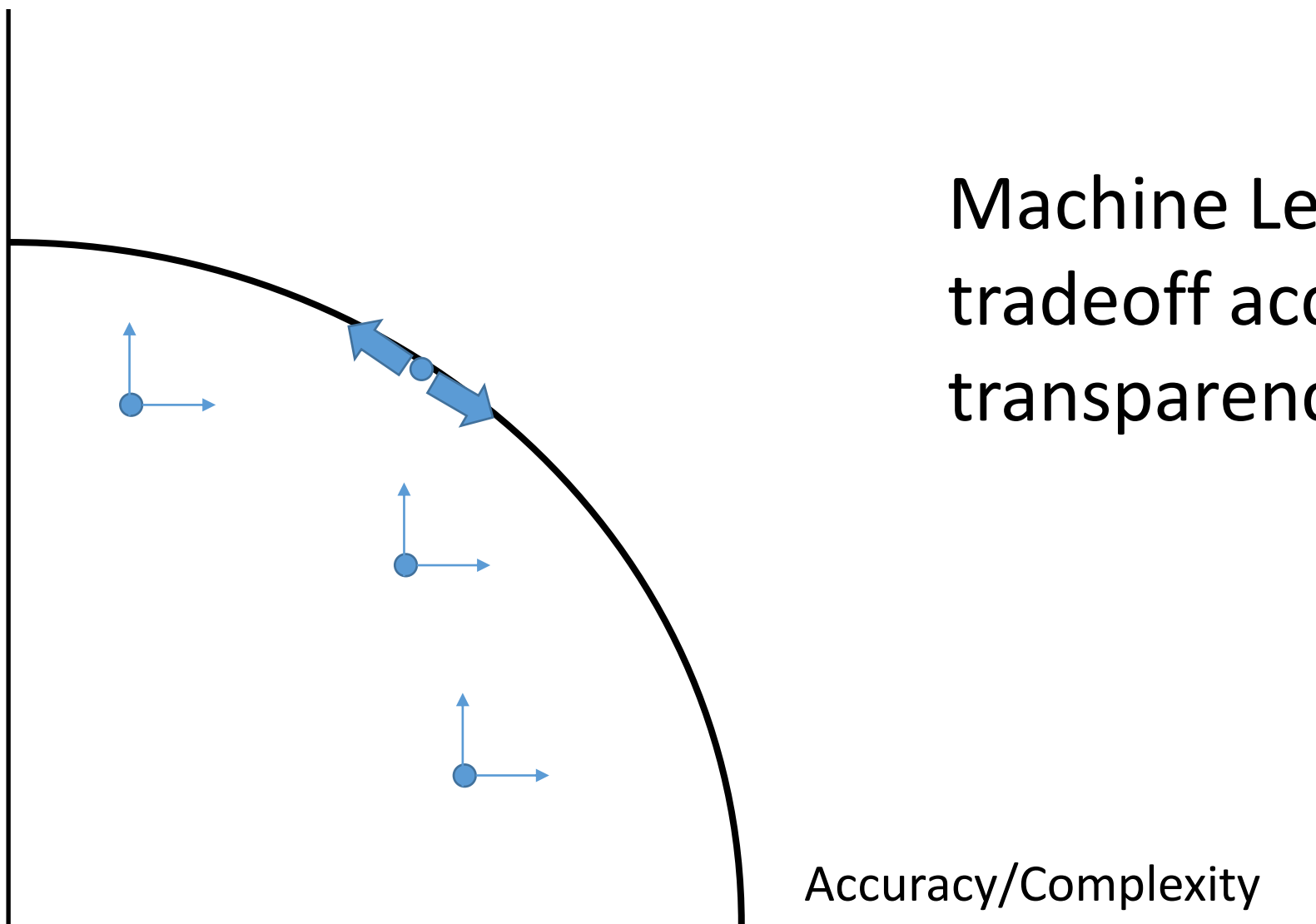
Tradeoffs Abound

- Explanatory power and transparency vs. Accuracy
- Automated data prep vs. conscious data shaping choices
- Clear visualizations vs. multi-dimensional transforms and relationships
- Ease of use vs. computational understanding



Horizon Function – Tradeoff Visualization

Transparency/Explanatory Power



Machine Learning models
tradeoff accuracy and
transparency.



Questions to ask yourself

- Where do you want to shape data sets for machine learning/predictive analytics?
- Do we currently have clean, consistent, accurate data?
- Do we have an executive champion who understands that systems have to be grown over time?
- Do we want to start with training, mentoring, or a defined use case?

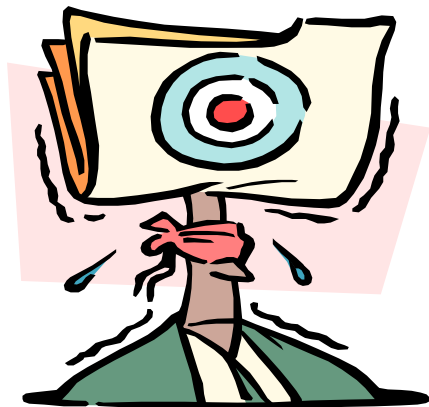


Advice

- Start simply and build
- Start (don't wait)
- Make sure data is clean and consistent
 - Outliers can have very large effects
 - Negative values can throw off some models
- Don't believe you need a lot of data
- Your best data is likely your internal data
- Let the data speak
- Promote solid interpretation and understanding of models and results



Resources and Questions?



Dan Vlami

Email: dvlamis@vlamis.com

Twitter: [@dvlamis](https://twitter.com/dvlamis)

Phone (816) 781-2880

Vlami Software Solutions

Twitter: [@vlamisoftware](https://twitter.com/vlamisoftware)

Let us know if you want to know about future webinars!

Tim Vlami

Email: tvlamis@vlamis.com

Twitter: [@timvlamis](https://twitter.com/timvlamis)

Vlami Home www.vlami.com Blog www.vlami.com/blog Prior OBIEE releases www.vlami.com/oaweinars

Blog article with links on Getting Started with Oracle Analytics <http://www.vlami.com/blog/2019/7/11/getting-started-with-oracle-analytics>

OAC What's New Documentation <https://docs.oracle.com/en/cloud/paas/analytics-cloud/acsw/index.html#ACSWN-GUID-CFF90F44-BCEB-49EE-B40B-8D040F02D476>

Oracle Analytics Library with plugins <http://bit.ly/OAClibrary> and demos <http://bit.ly/OACLIVE>

YouTube channel Oracle Analytics from SampleApp Team <https://www.youtube.com/user/EvolvingBI/videos>

YouTube video playlist on OAC 5.7 new features <http://bit.ly/OAC57Features>

Course on Oracle Analytics <https://www.udemy.com/course/augmented-analytics/>