



Introduction to Machine Learning in Oracle Analytics Cloud

MAY 16 & 17, 2018

CLEVELAND PUBLIC AUDITORIUM, CLEVELAND, OHIO

WWW.NEOOUG.ORG/GLOC

Introduction to Machine Learning in Oracle Analytics Cloud

Great Lakes Oracle Conference 2018

Tim Vlamis

tvlamis@vlamis.com

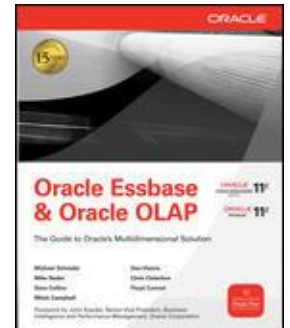
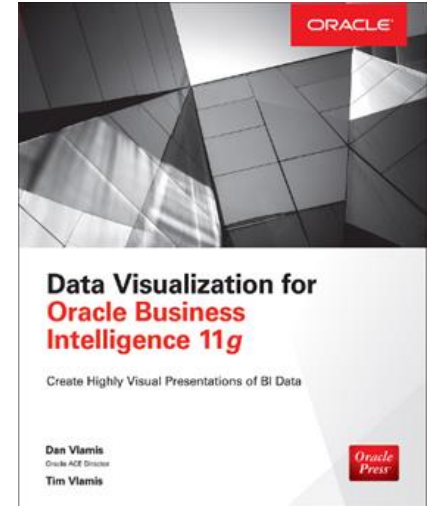
VlamiS Software Solutions

- VlamiS Software founded in 1992 in Kansas City, Missouri
- Developed 200+ Oracle BI and analytics systems
- Specializes in Oracle-based:
 - Enterprise Business Intelligence & Analytics
 - Analytic Warehousing
 - Data Mining and Predictive Analytics
 - Data Visualization
- Multiple Oracle ACEs, consultants average 15+ years
- www.vlamiS.com (blog, papers, newsletters, services)
- Co-authors of book “Data Visualization for OBI 11g”
- Co-author of book “Oracle Essbase & Oracle OLAP”
- Oracle University Partner
- Oracle Gold Partner

 EDUCATION RESELLER

 Gold
Partner


Specialized
Oracle Business Intelligence
Foundation Suite 11g





Presenter Background

Tim Vlami – Vice President & Analytics Strategist

- 30+ years in business modeling and valuation, forecasting, and scenario analyses
- Oracle ACE 
- Instructor for Oracle University's Data Mining Techniques and Oracle R Enterprise Essentials Courses
- Professional Certified Marketer (PCM) from AMA
- MBA Kellogg School of Management (Northwestern University)
- BA Economics Yale University
- @TimVlami

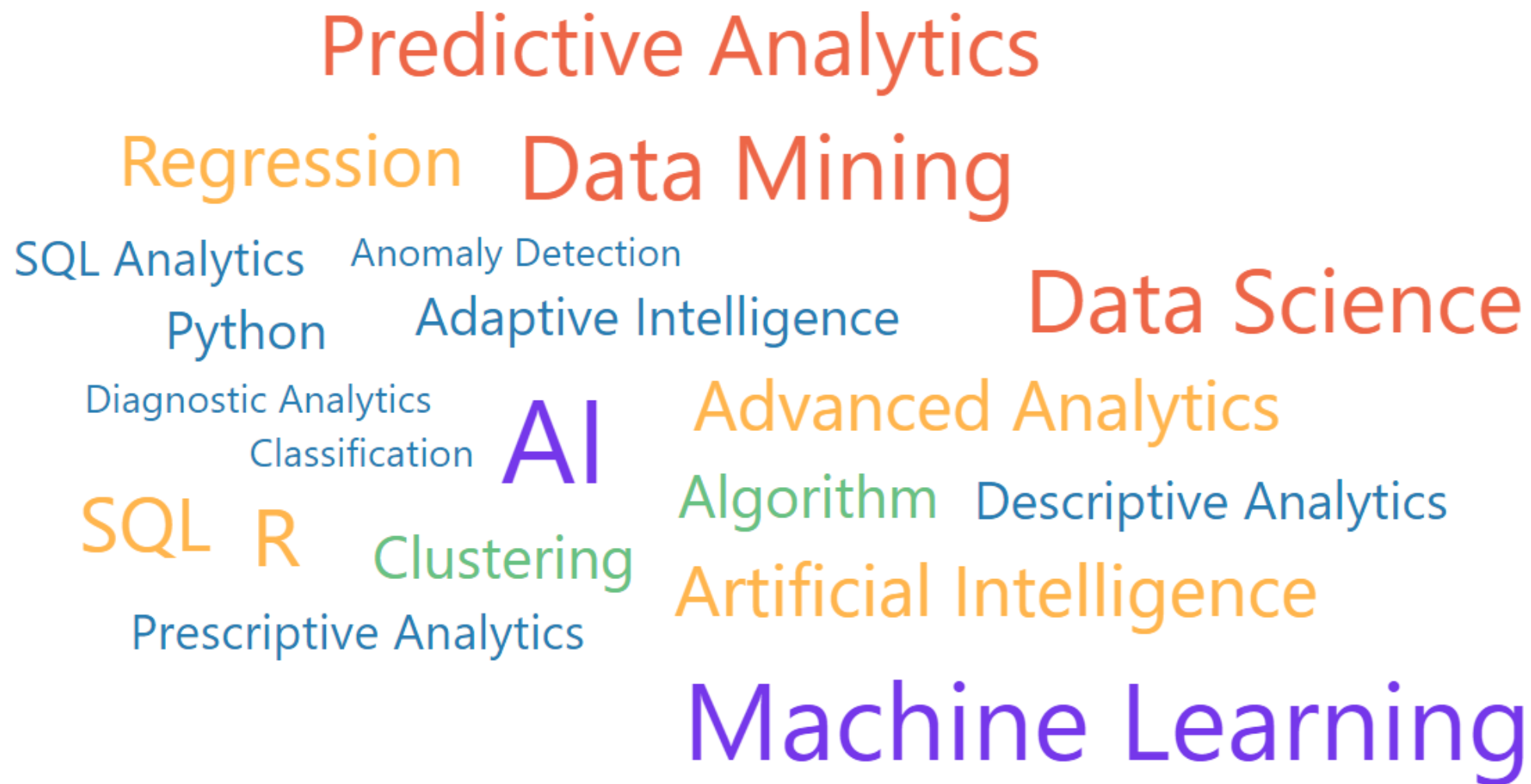


Presentation Agenda

- Machine Learning in OAC
 - Automated One-click algorithms
 - R-based column calculations in OBIEE 12c and OAC
 - Data Flow models
 - Languages and environments
- Lessons learned at Darden (Olive Garden)
- Lessons learned with other restaurants/service industries
- Lessons learned with manufacturing
- Recommendations for designing and implementing a machine learning strategy



Many Words Used for Similar Concepts





Many Options for ML in OAC

- Work in Data Visualization Canvas
 - One-click options
 - Use built-in scripts via My Calculations
 - Upload and call custom scripts
- Build data flows in DV interface with Machine Learning nodes
- Work in “edit formula” in Criteria Tab in Answers Analysis
- Develop custom scripts in Repository
- Connect to Oracle Advanced Analytics in Oracle Database Cloud Service (High Performance or Extreme Performance)
 - Oracle Data Mining
 - Oracle R Enterprise
- Connect to Oracle Machine Learning in Autonomous Data Warehouse Cloud Service



OAC and Data Visualization Desktop

- Oracle Advanced Analytics – Mostly R based
- Oracle Machine Learning – Mostly Python based
- You don't need to know either R or Python to use!



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

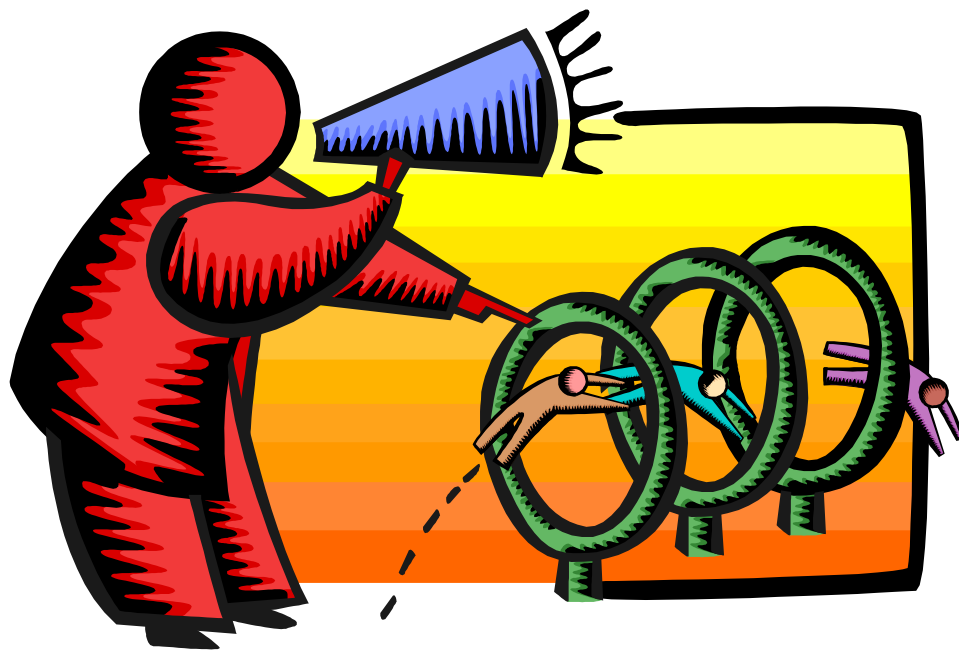


Oracle Database Cloud Service

- Oracle Advanced Analytics option to Oracle Database
 - Included with DBCS High Performance and Extreme Performance
 - Comprised of Oracle Data Mining and Oracle R Enterprise
- Oracle Data Miner SQL Developer Extension is fastest way to learn and develop predictive analytics/machine learning at enterprise scale



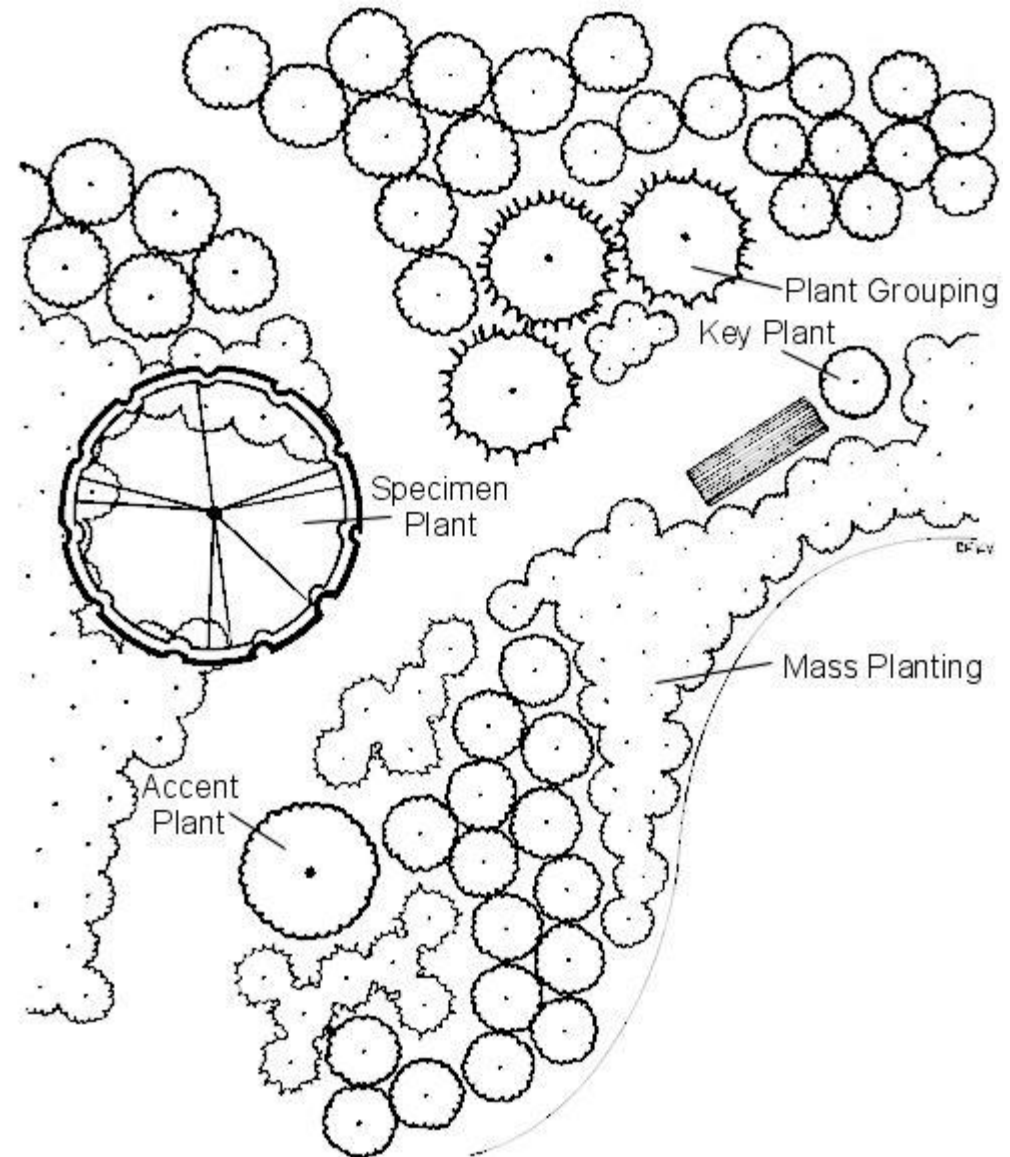
Demo





Grow Your Analytics Naturally

- Start with a general plan





Grow Your Analytics Naturally

- Start with a general plan
- Enhance existing features





Grow Your Analytics Naturally

- Start with a general plan
- Enhance existing features
- Prepare one bed at a time





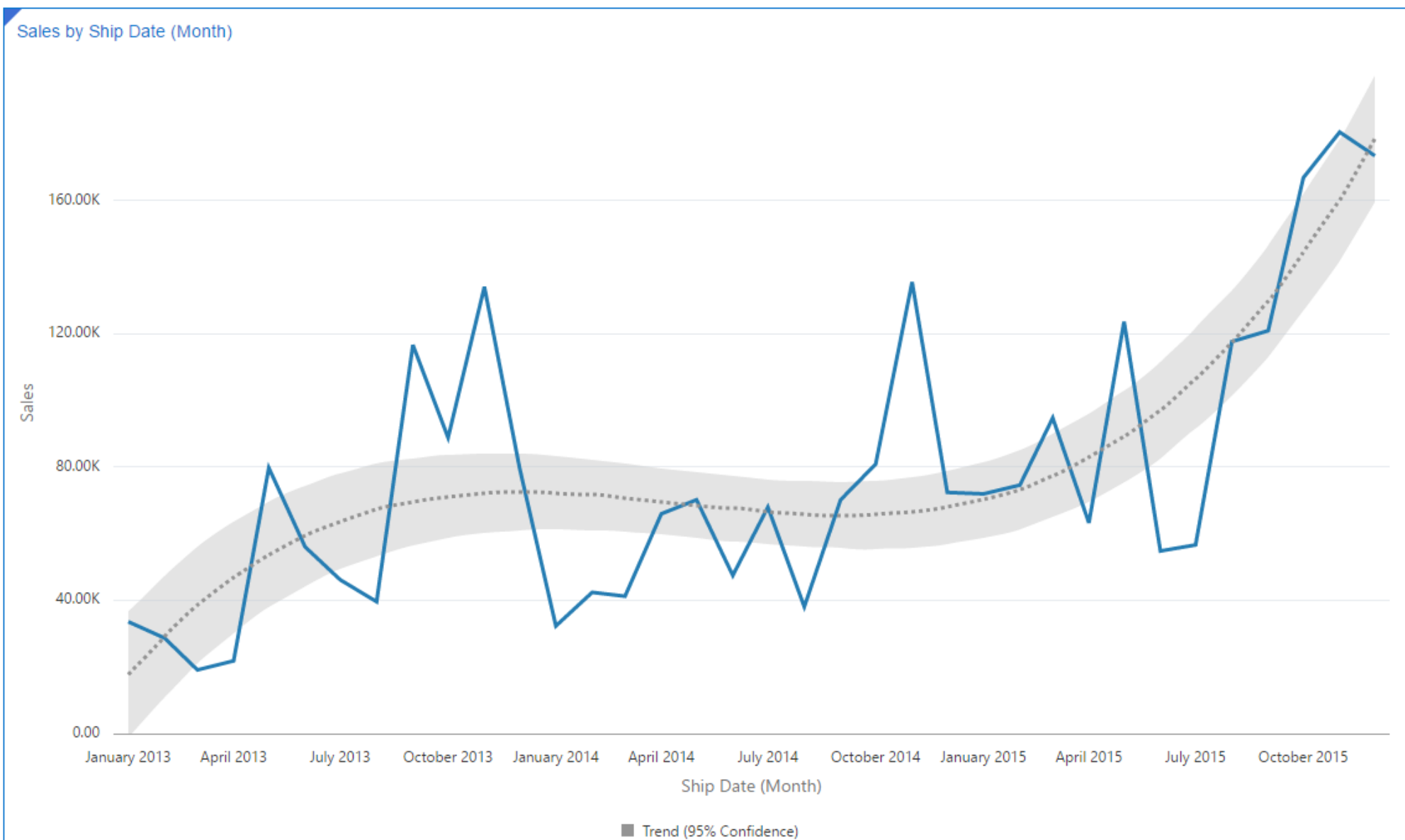
Grow Your Analytics Naturally

- Start with a general plan
- Enhance existing features
- Prepare one bed at a time
- **Mulch and weed early on**





Trend with Confidence Interval





ETS Exponential Smoothing

- Default for forecast()
- Can choose from Additive or Multiplicative Methods
 - For example series of (5, 10, 15):
 - Additive methods use discrete differences
 - growth is constant +5 units
 - Multiplicative method uses percentage or marginal differences
 - growth is a decreasing percentage
- Use Box-Cox transform? ($\lambda=0$ is a log transform)
 - Transforms data to achieve a more normal distribution
- Use Trend Dampening? (past results are as important as recent)
- Output forecast or upper/lower bounds?
- Set Prediction Interval (0-100)



ARIMA Model

- **AutoRegressive Integrated Moving Average**
- Powerful algorithm for series analysis and prediction
- Three parameters (p, d, q)
 - Auto regression (how reliant series values are on previous series values). AR(0) is white noise.
 - Integrated (degree of AR differencing, Random Walk)
 - Moving average (smoothing function)
- Set Prediction Interval (0-100)
- Choose AIC method
- Output forecast or upper/lower bounds
- Know the name Rob Hyndman for ARIMA in R

<https://www.otexts.org/fpp/>



ARIMA Model

- Set Prediction Interval (0-100)
- Choose Information Criterion method AIC or BIC (use auto)
 - Helps you determine which models have the lowest error
- Output forecast or upper/lower bounds



Questions to ask yourself

- Do we have people currently on staff who want to execute R or Python models inside the BI system?
- 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 with have to be grown over time?
- Do we want to start with training or a defined use case?

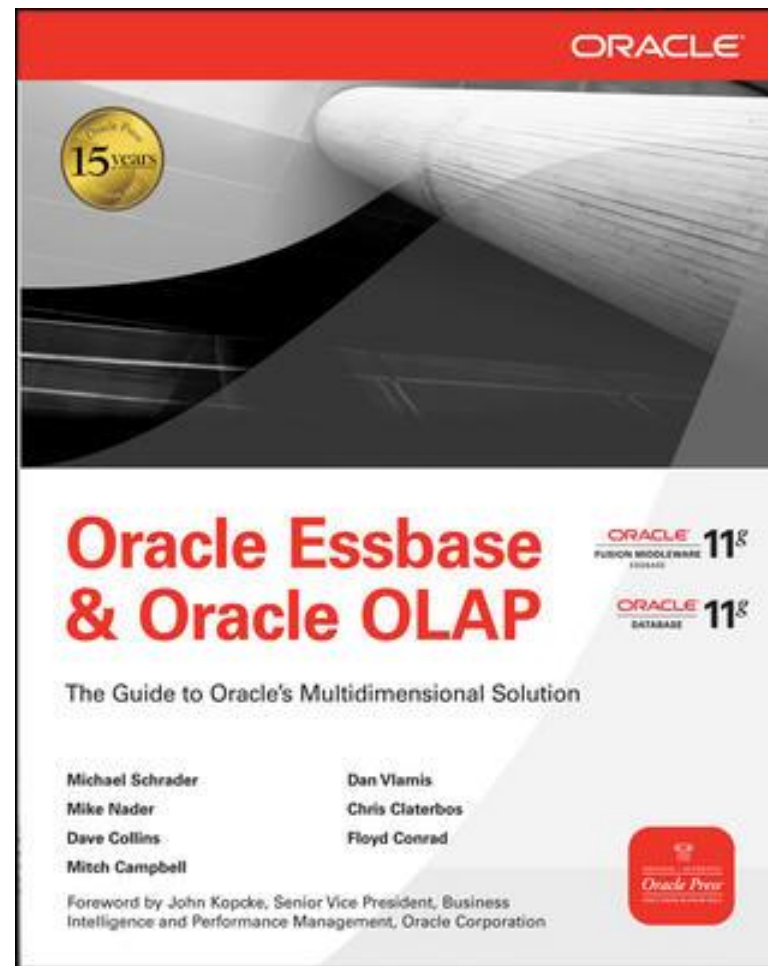
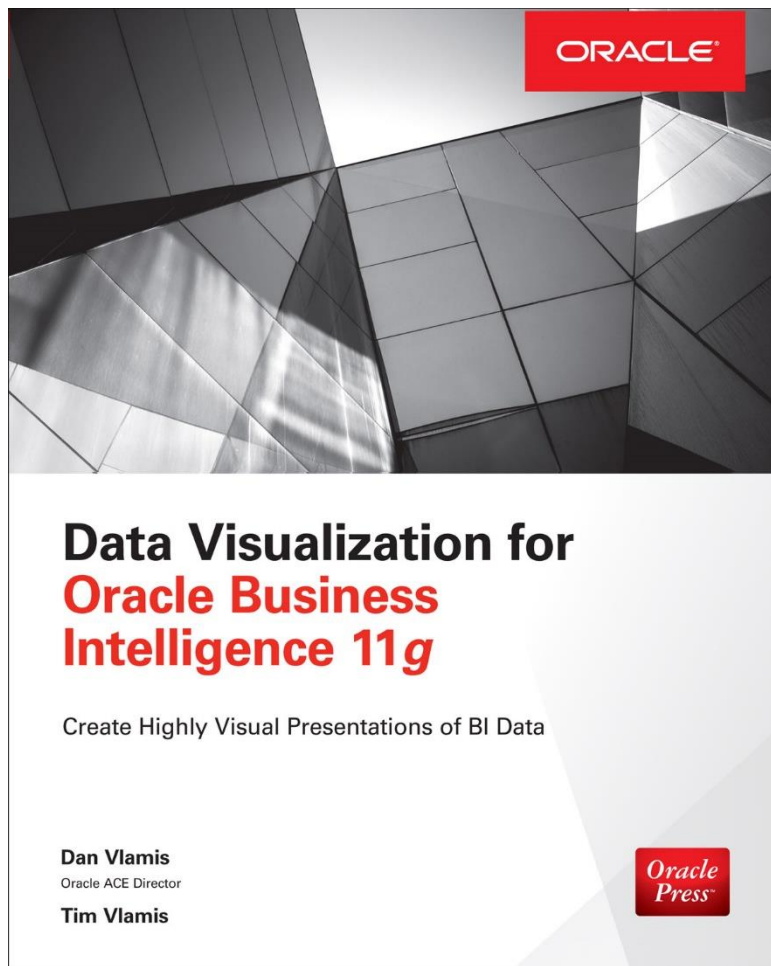


Advice

- Start simply and build
- Make sure that you have date type fields
- Make sure data is clean and consistent
- Be careful of months (days and weeks often work better)
- Negative values can throw off some models
- Outliers can have very large effects
- Don't set Prediction Interval too high

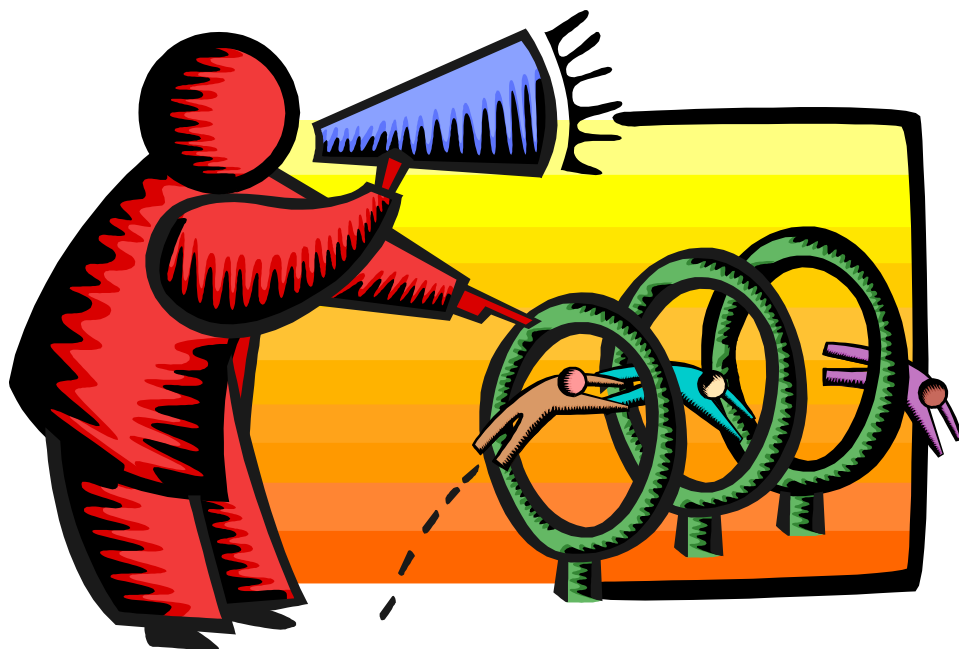


Oracle Press Books





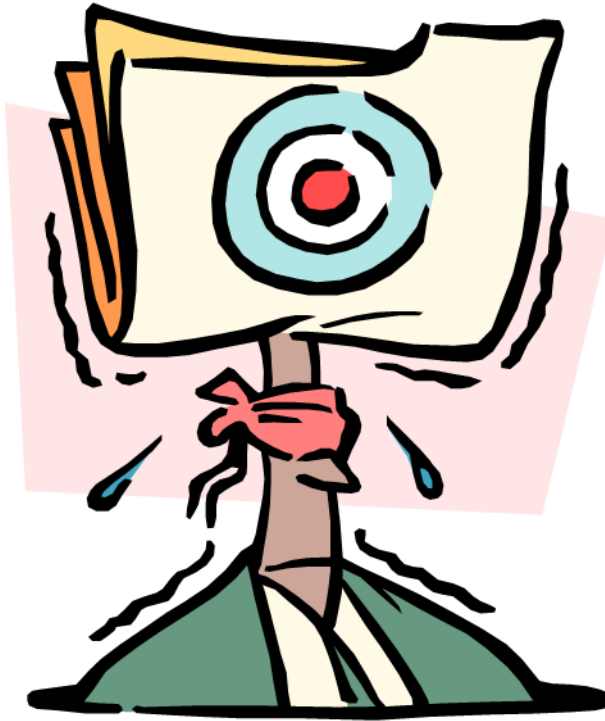
Demo Showed



- Create new project
- Ordering a graph
- Saving a visualization
- Adding and coloring by a dimension
- Formatting a visualization
- Navigating in a visualization
- Adding a line chart
- Marquee selection (keep selected)
- Adding a performance tile visualization
- Adding a map visualization and coloring
- Adding a treemap visualization
- Filtering visualizations
- Adding a calculation
- Data mashup



Questions?





Thank You!

Presenter Information

XXX, Consultant

VlamiS Software Solutions, Inc.

816-781-2880

xxx@vlamis.com

@twitterhandle

For more information go to www.vlamis.com