



SOFTWARE SOLUTIONS

Three Steps Ahead: Making and Visualizing Predictions and Forecasts

ODTUG Kscope 18

Tim VlamiS

June 12, 2018

@VlamiSSoftware



ODTUG
Kscope18



ORLANDO, FLORIDA • JUNE 10-14

PLEASE
FILL OUT YOUR
EVALUATIONS

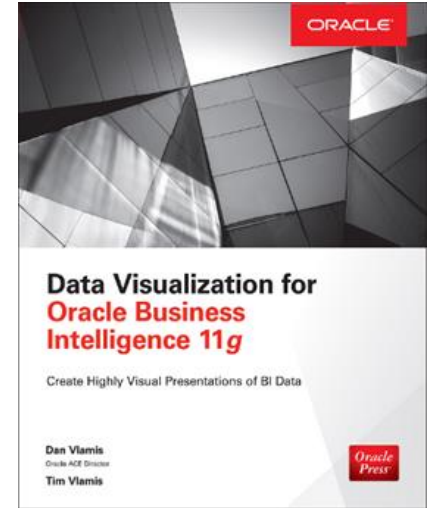
Vlamis Software Solutions

- 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

ORACLE EDUCATION RESELLER

ORACLE Gold Partner

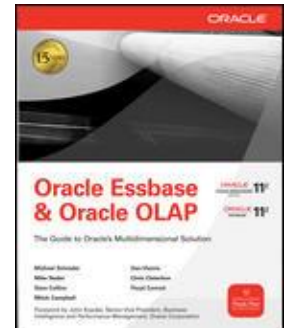
Specialized
Oracle Business Intelligence
Foundation Suite 11g



ORACLE
ACE Director


ORACLE
ACE

ORACLE
ACE Associate





Vice President & Analytics Strategist

- 30+ years in business modeling and valuation, forecasting, and scenario analyses
- Oracle ACE 
- Instructor for Oracle University's Predictive Analytics, 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



Vlami Presentations

Presenter	Location	Time	Title
Tim Vlami	Northern Hemisphere A1 Fifth Level	Monday 1:15pm	Future Proof Your Career: What Every Executive Needs to Know About Adaptive Intelligence
Arthur Dayton	Northern Hemisphere A1 Fifth Level	Monday 2:30pm	Advanced Visualizations for Data Discovery and Exploration in Oracle DV
Tim Vlami	Northern Hemisphere A2 Fifth Level	Tuesday 9:00am	Three Steps Ahead: Making and Visualizing Predictions and Forecasts
Tim Vlami Mike Durran	Northern Hemisphere A3 Fifth Level	Wednesday 10:15am	The Many Ways to Get Your Data into Oracle Analytics Cloud
Arthur Dayton	Asia 4, Lobby/ Third Level	Wednesday 2:15pm	Do the Mashup: How to Get BI Self-Service Data Connections Right



Forecasting Today...

- Predictions are the holy grail of BI systems and initiatives.
- Most all corporations have need for forecasting.
- Typical forecasting systems
 - Are stand alone or from ERP (not integrated to BI system)
 - Tend to use straight line or heuristic calculations
 - Are not always integrated into the business
 - Are often tied directly to the budgeting process
- High level of angst surrounding forecasts



Forecasting Should...

- Be integrated with rest of BI system.
- Be another series of measures that are revealed in the context of historic information.
- Employ a consistent and known methodology.
- Be a part of the Common Enterprise Model when possible.
- Have visibility across functional areas and roles in corporations.
- Be regularly tested against actual results.



Forecasts Should be Based on Evidence





Forecasting Methodologies

- Rule-based heuristic (last period, last period +5%, etc.)
- Cross-sectional methodologies (point in time)
- Time series (time sequenced data series)
- Averages (moving, weighted, etc.)
- Linear and Non-linear regressions (line fitting)
- Transforms, projections, min/max
- Mixed models



Methodologies for Today

- Forecast Function in OAC, DVD, and OBI 12c
 - ETS (Triple Exponential Smoothing – Error Trend Seasonal)
 - ARIMA (Auto Regressive Integrated Moving Average)
- Oracle R Distribution OBIEEAdvancedAnalytics Package
- OAC Data Flow Time Series Forecast node
- Other OBI Time Series Functions
 - AGO
 - TODATE
 - PERIODROLLING



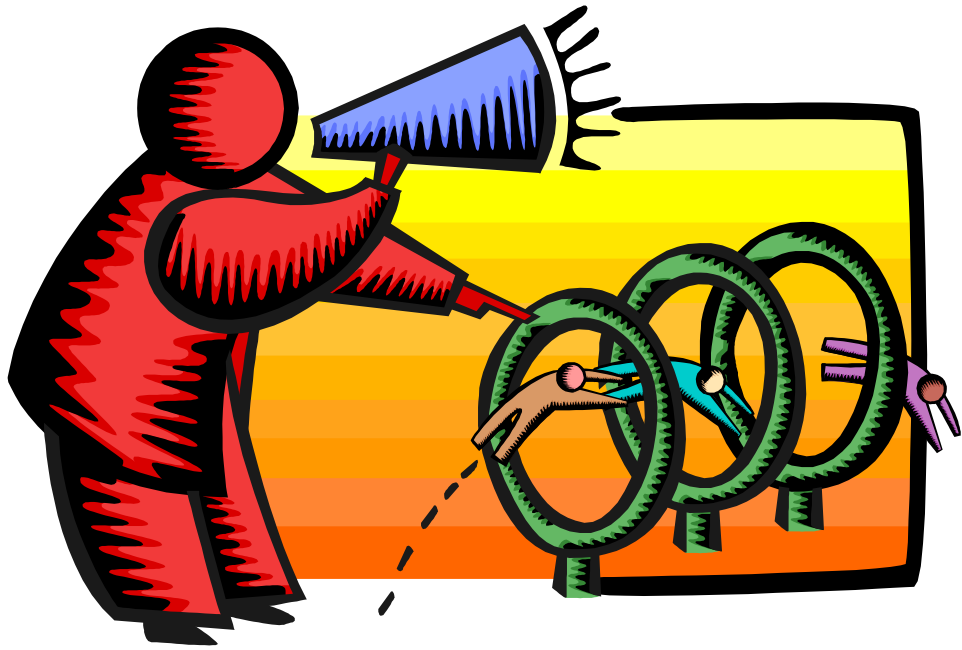
Exponential Smoothing

- Methodology for smoothing data and preferencing more recent periods when doing time series forecasts
- Similar conceptually to a weighted moving average
- Weights decline according to an exponential function.
 $\{1, (1-\alpha), (1-\alpha)^2, (1-\alpha)^3, \dots\}$
- Higher values give more weight to more recent periods

- Single (weighted average of most recent observation and the most recent smoothed statistic)
- Double (trend either up or down)
- Triple (period effect)



Demo





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



Confidence Interval and Prediction Interval

- Confidence Interval is associated with a parameter (.01-.99)
 - Parameter is unknown but assumed to be non-random
 - CI gives the probability that a value within the range is not associated with a chance difference
 - Used in OAC with trend lines
 - Higher values give larger intervals
- Prediction Interval is associated with a random variable yet to be observed (.01-.99)
 - Probability of the predicted value of the variable lying within the range
 - Used in OAC with forecasts
 - Higher values give larger intervals
- <https://robjhyndman.com/hyndsight/intervals/>



Data Flow Time Series Forecast Node

- Uses Python NumPy and Pandas Arima(1,0,0) aka AR(1)
- Only three input parameters
 - Target column
 - Date column
 - Number of periods
- Get output column with forecasted values
- Runs at the individual day grain level



Time Series Functions in OAC and OBI

- Very powerful, accessible capability
- Time dimension must be designated
- Query results must be exact to pull from cache
- Can be “expensive” in processing
- Make sure that unique keys are defined at each level (“Jan13” rather than “Jan”)
- Does not require R



AGO function

- Defines a time-based offset
- Can nest multiple AGO statements (same level)
- Ago(<<Measure>>, <<Level>>, <<Number of Periods>>)
- Measure is a fact such as sales.
- Level is an optional term, default is set by the grain of the query (BY clause) or is specified in repository for level based measures.
- Number of periods is an integer specifying the offset value.



TODATE

- Time-based aggregation function.
- Calculates based on starting value to current.
- Can nest with AGO (same level)
- ToDate(<<Measure>>, <<Level>>)
- Measure is a fact such as sales
- Level is the time grain such as year or month



PERIODROLLING

- Defines a period of time contextually
- Performs an operation across a specified set of query grain periods
- PeriodRolling(<<Measure>>, <<Starting Period Offset>>, <<Ending Period Offset>>, <<[Hierarchy]>>)
- Measure is a fact such as sales
- Starting Period Offset is an integer value, use a minus sign (“-2” means 2 periods ago)
- Ending Period Offset defines the end of the period, use a zero for current period
- Hierarchy is an optional setting to specify which time hierarchy to use such as “fiscal”
- Use “unbound” for starting period offset to calculate total from beginning
- PeriodRolling uses either the query level grain of “measure” or the measure level for “measure” if it has been set in the Admin tool.



Advice

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

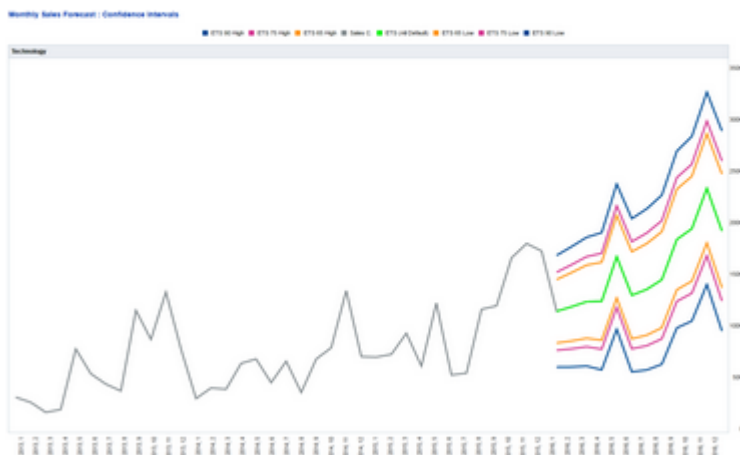


Oracle Data Visualization has Examples

- <http://www.oracle.com/technetwork/middleware/bi-foundation/data-visualization-2954126.html>

Forecast Syntax Examples

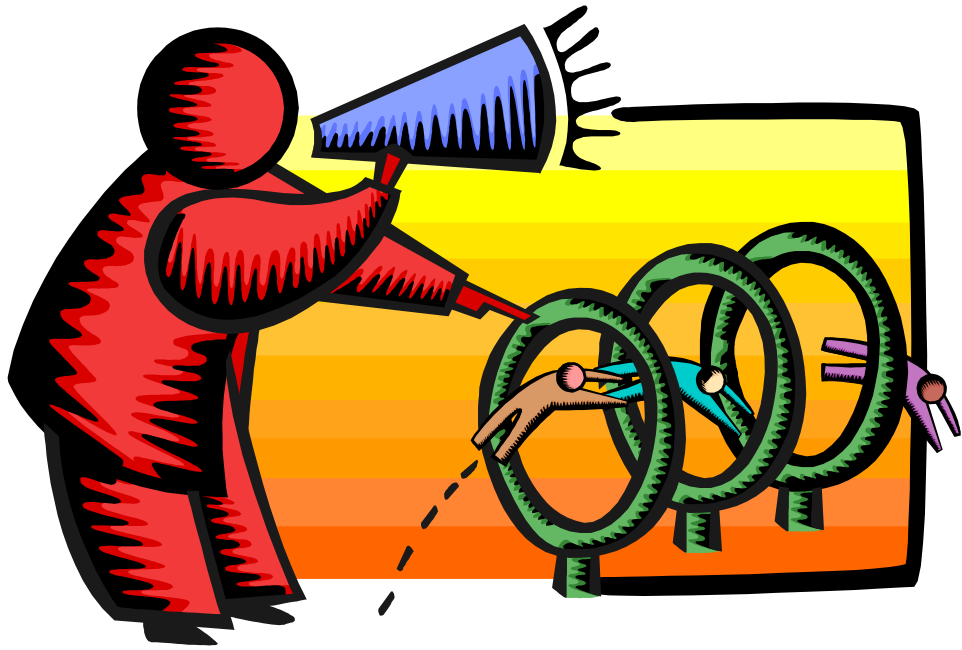
Leveraging various options in the Forecast function



 [Download Examples \[Video\]](#)



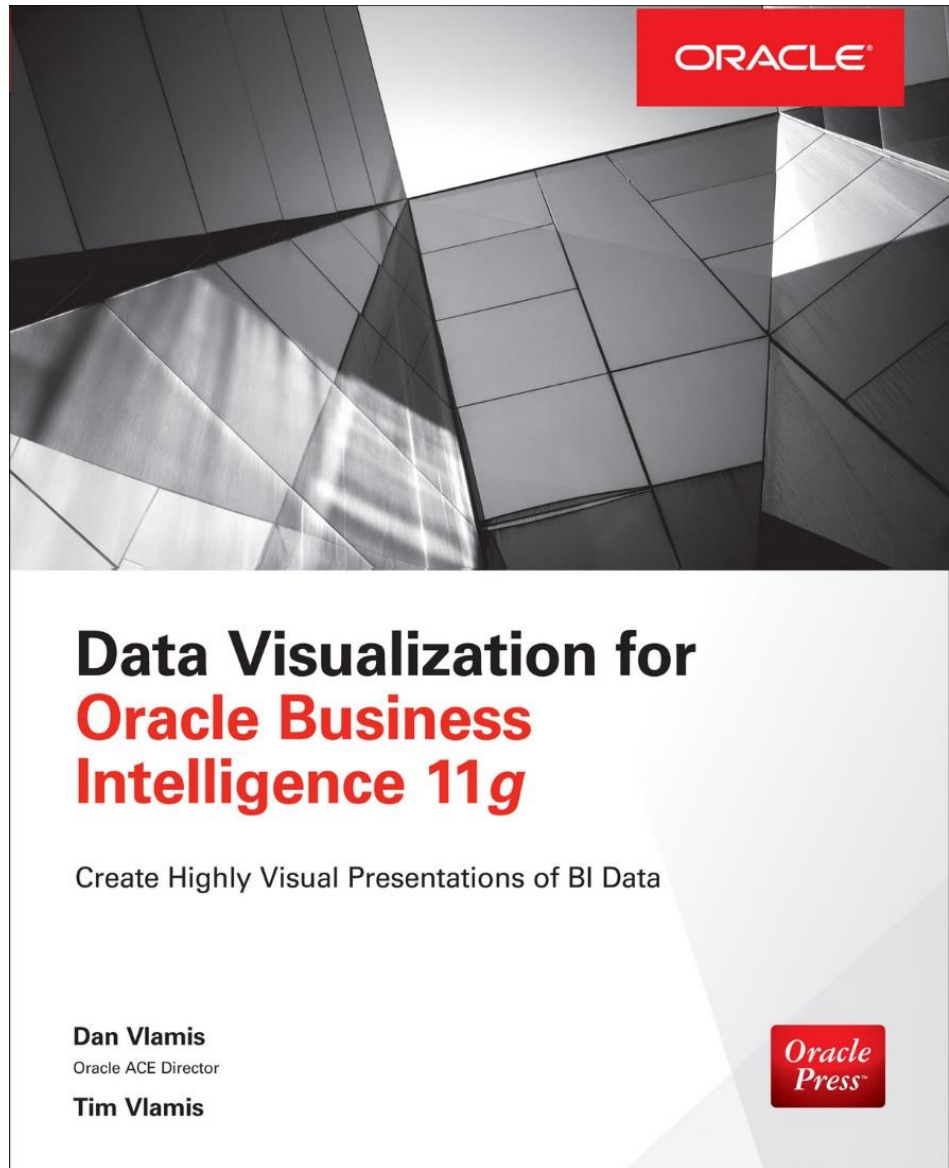
Demo





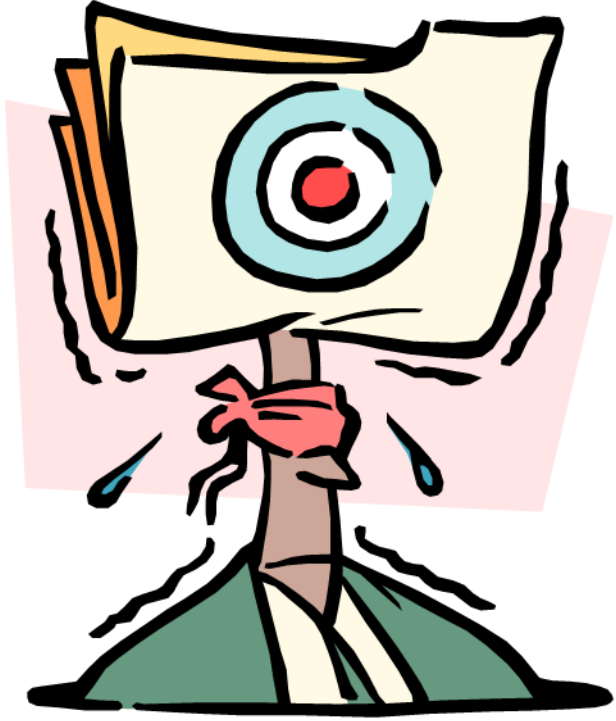
Drawing for Free Book

Add business card to basket
or fill out card





Questions?





Thank You!

Presenter Information

Tim VlamiS

VlamiS Software Solutions, Inc.

816-781-2880

tvlamis@vlamis.com

@timvlamis

For more information go to www.vlamis.com



ODTUG
Kscope18



ORLANDO, FLORIDA • JUNE 10-14

PLEASE
FILL OUT YOUR
EVALUATIONS