



COLLABORATE 18

TECHNOLOGY AND APPLICATIONS FORUM
FOR THE ORACLE COMMUNITY

Do the Mashup

How to Get BI Self-Service Data Connections Right

Remember to complete your evaluation for this session within the app!

Session ID:

10467

Prepared by:

Cathye Pendley

Vice President Consulting
Services

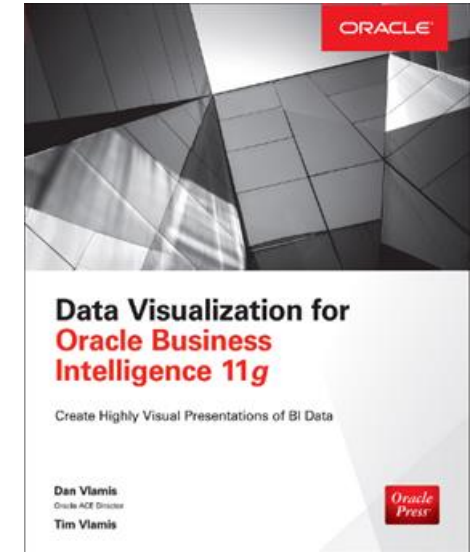
Vlami Software Solutions

April 22, 2018

#C18LV

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)
- Creators of the [Force Directed Graph Plugin](#) on the [Oracle Analytics Library](#)
- 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 APPROVED
EDUCATION CENTER

ORACLE Gold
Partner

Specialized
Oracle Business Intelligence
Foundation Suite 11g

Cathye Pendley background

#C18LV

Cathye Pendley – VP Consulting Services

- 25+ years in business intelligence and dimensional modeling
- Successfully lead and implemented 50+ business intelligence projects
- Serves Analytical and Data Summit committee
- Presents at numerous conference
- Mentor women interested in working in the technology industry

Vlamis Presentations

#C18LV

Presenter	Location	Time	Title
Cathye Pendley	Banyan D	Sunday 12:30pm	Do the Mashup: How to Get BI Self Service Data Connections Right
Tim Vlamis	South Pacific J	Sunday 3:00pm	Future Proof Your Career: What Every Executive Needs to Know About Adaptive Intelligence
Tim Vlamis Arthur Dayton	Banyan A	Monday 11:00am	Introduction to Machine Learning in Oracle Analytics Cloud
Dan Vlamis	Banyan B	Monday 4:15pm	Getting from Answers/Dashboards to Data Visualization
Arthur Dayton Dan Vlamis	Banyan E	Wednesday 11:00am	Using Node.js to Make OBIEE the Application You Always Wanted It to Be

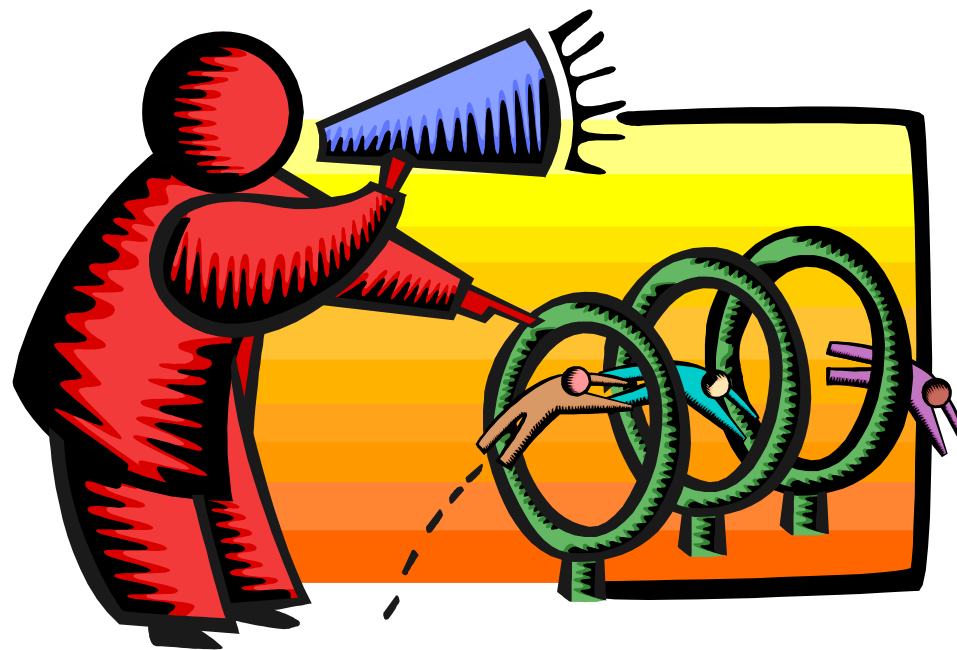
Presentation Agenda

- Overview of self-service data mashup capabilities
 - Data warehouse data
 - Big Data
 - External/departmental data (spreadsheets)
 - Cloud Applications
- Connecting to data sources
- Editing data sources in DV project initiation
- Using the prepare tab in DV projects
- Data Flows
- Data Modeler

#C18LV

Demo Data Connections

#C18LV



Different Strategies

- If it's worth doing, it's worth doing right.



- The perfect is the enemy of the good.



#C18LV

Basic Principles

#C18LV

- Analysts need data, do not deprive them
- Data modeling is an art, (but not everyone is an artist)
- Expediency versus Efficiency
- The lessons of **WordArt** and website development
- Work front to back
- Regularly review and choose the most important data for hardening

Mashup Tips

#C18LV

- Use Excel .xlsx files, not .csv formatted files
 - Data Visualization will read some Excel formats
 - Date and time
 - Numeric
 - CSV files have to have aggregation rules individually applied
- Fix formatting early
- Create columns for both sum and average when appropriate
- Set a strategy for suspect outlier records upfront
 - Use Excel filters and sorts to fix or scrub suspect records in advance

Creating Data Sources with Excel

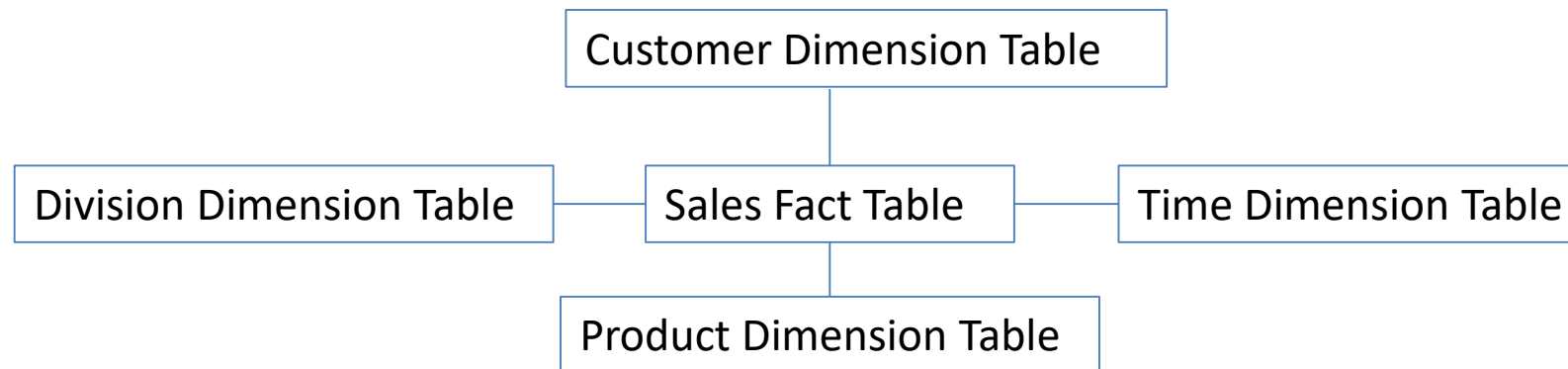
#C18LV

- First row must contain column names
- All column names must be only in the first row
- Column names must be unique
- All data must be on a single data sheet
- Avoid “reserved” column names like “Count”
- Column names should not have special characters
- Rows 2-x should have record data only
- Do not include any “inter-row” data such as totals, sub-totals, sections, etc.
- Do not use hierarchical columns
- Do not use any “merged” cells
- Repeat data in every row (not just first instance)

OAC Loves Star Schema Data Sets

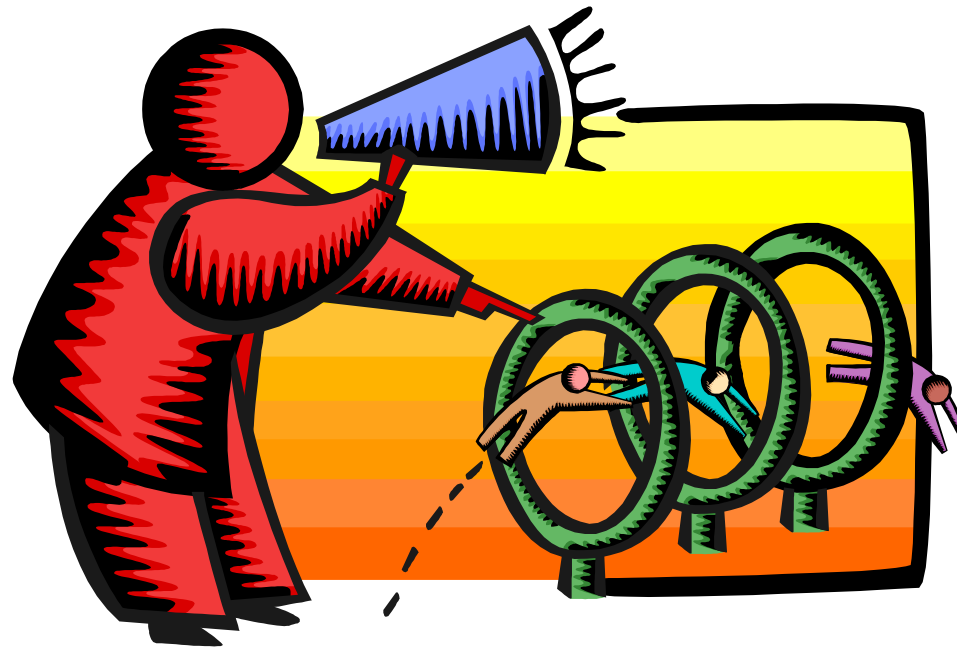
#C18LV

- Well-organized dimensions in attribute columns
- Facts in measure columns
- Denser data sets are better than highly sparse data sets
- Use automatically generated date columns if at all possible



Demo Data Upload and Mashup

#C18LV



Data Visualization Desktop Tips

#C18LV

- The size and processing power of your machine will influence:
 - The speed of processes
 - The size of data sets you can work import
 - The other programs you can run
 - The kinds of advanced processes you can run (machine learning, etc.)
- Suggested upper guidelines for “typical” PCs
 - About 500,000 to 1 million rows
 - About 50-70 columns

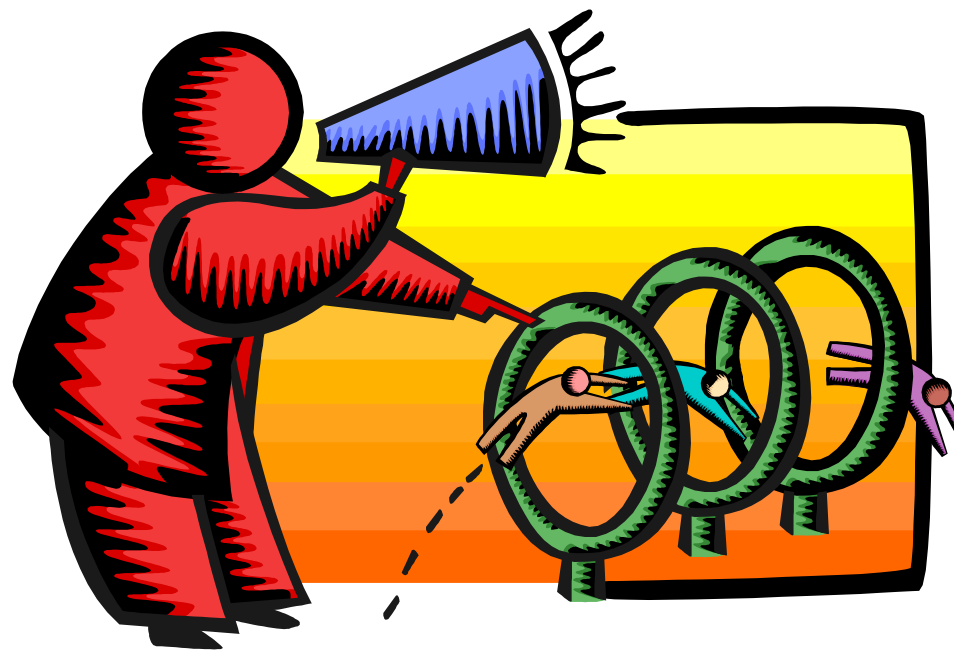
Using Data Flows for Mashups

- Think of data flows as modules
- Use data flows to avoid replicating work in every new project
- Use sequences to connect multiple data flows
- Data flow acts as visual representation document of data transformations
- Data flows are good for known issues (i.e. outlier removal)

#C18LV

Demo Data Flows

#C18LV



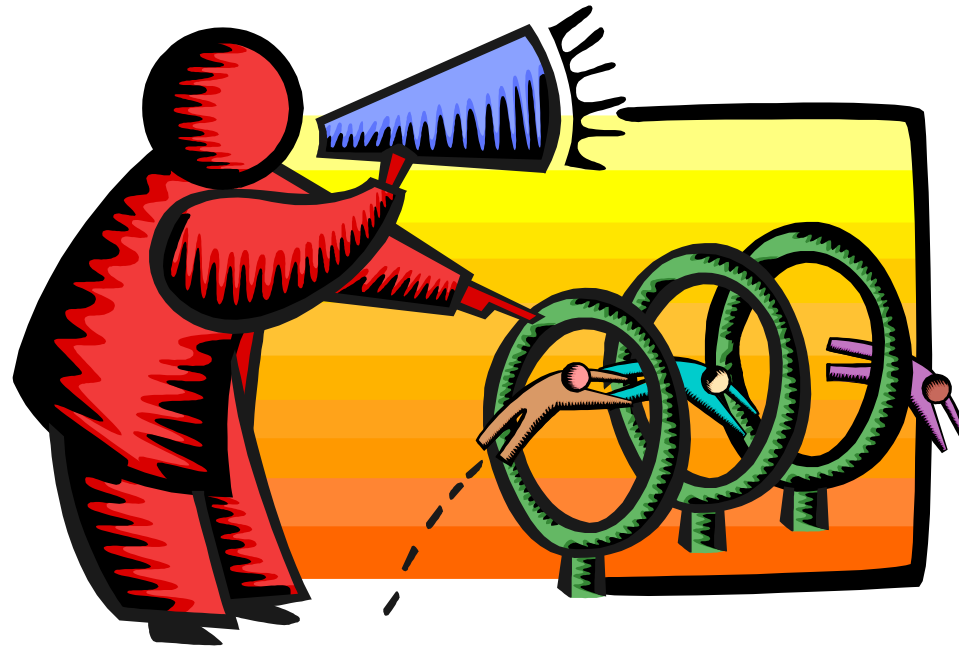
Data Source Blending using Data Modeler

#C18LV

- Best way to join and model data sets from multiple, dynamic sources
- Ability to define subject areas and bring data together
- Ability to “shape” and filter data
- Ability to create custom calculations using multiple sources

Demo Data Modeler

#C18LV



Progression of Data Prep Tools

#C18LV

- BI Admin Tool (.rpd file)
- OAC Thin Client Modeler
- Data Flows
- Data definition as part of a DV Project
- Data Preparation tab
- Add Custom Calculation

- ETL to warehouse
- Business layer in RPD file
- Custom Group and Measure (catalog object)
- Selection steps
- Custom column formula in Criteria tab



COLLABORATE 18

TECHNOLOGY AND APPLICATIONS FORUM
FOR THE ORACLE COMMUNITY

cpendley@vlamis.com

Session ID:

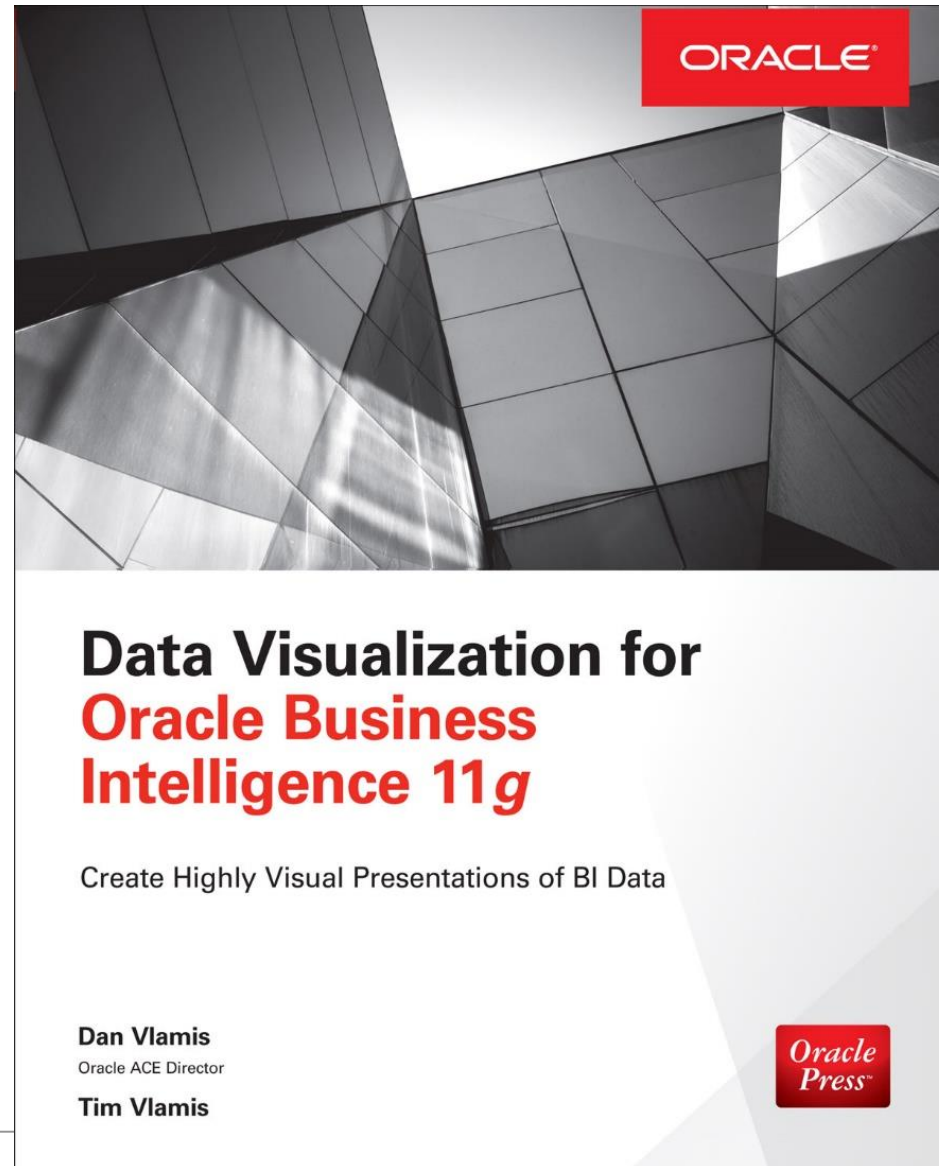
10467

Remember to complete your evaluation for this session within the app!

#C18LV

Drawing for Free Book

Add business card to basket
or fill out card



#C18LV



Analytics and Data Summit

All Analytics. All Data. No Nonsense.

SAVE THE DATE FOR
ANALYTICS AND DATA SUMMIT 2019
MARCH 12-14, 2019

www.AnalyticsAndDataSummit.org

Brought to you by:
BIWA User Community

Formerly called the BIWA Summit with the Spatial and Graph Summit.
Same great technical content – great new name!