Oracle Data Warehousing With Erwin

Donald Soulsby

SANDHILL

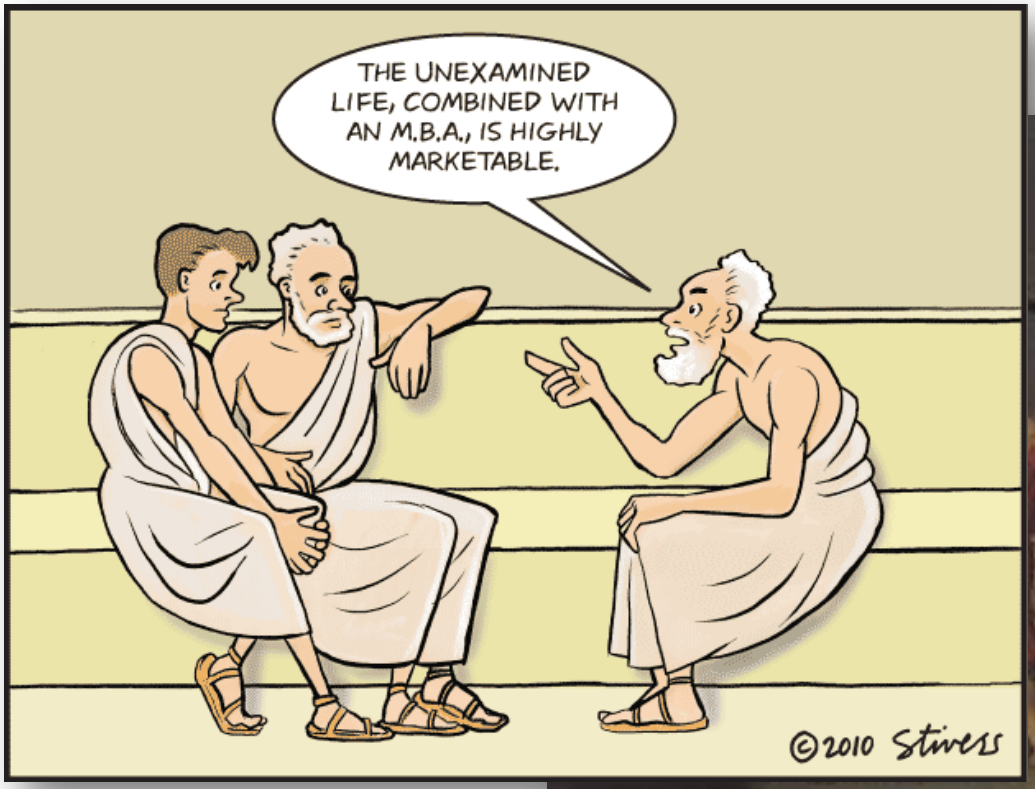
Agenda

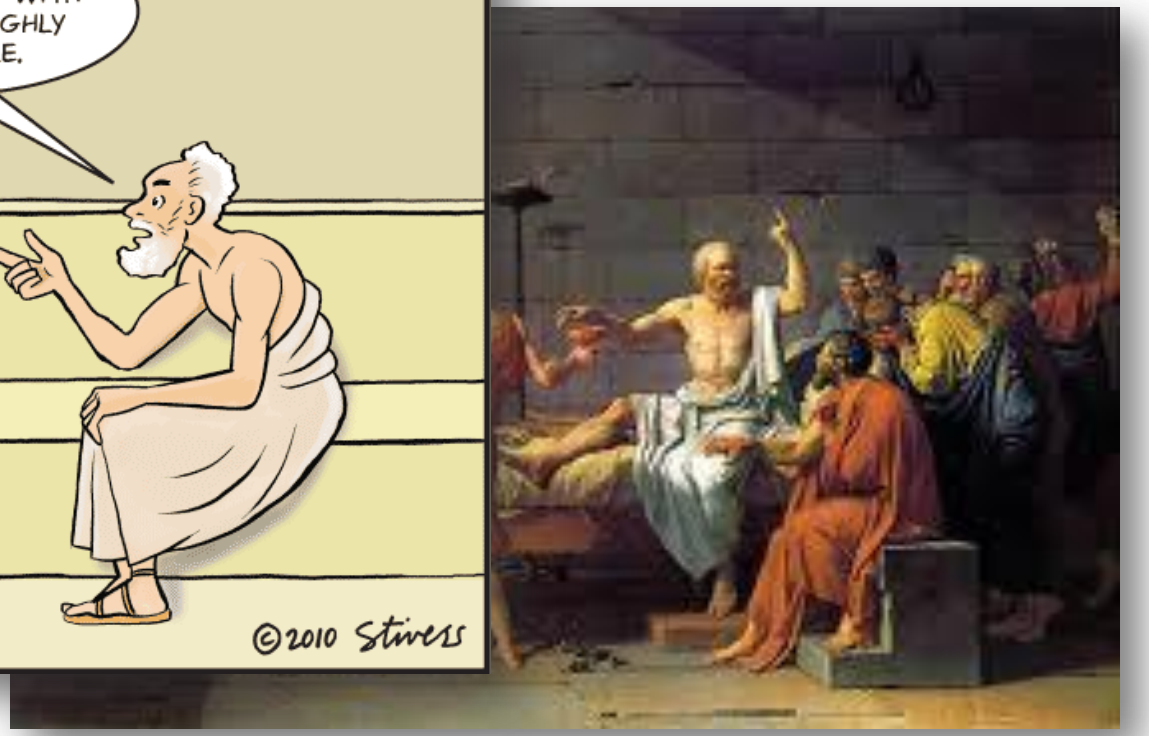
- Dimensional Modeling:
  - Fundamentals
  - Characteristics
  - Modeling functions within CA ERwin Data Modeler
- Erwin model metadata extensions
  - Oracle Warehouse Builder

SANDHILL

An unexamined life is not worth living." -

-- **Socrates**





SANDHILL

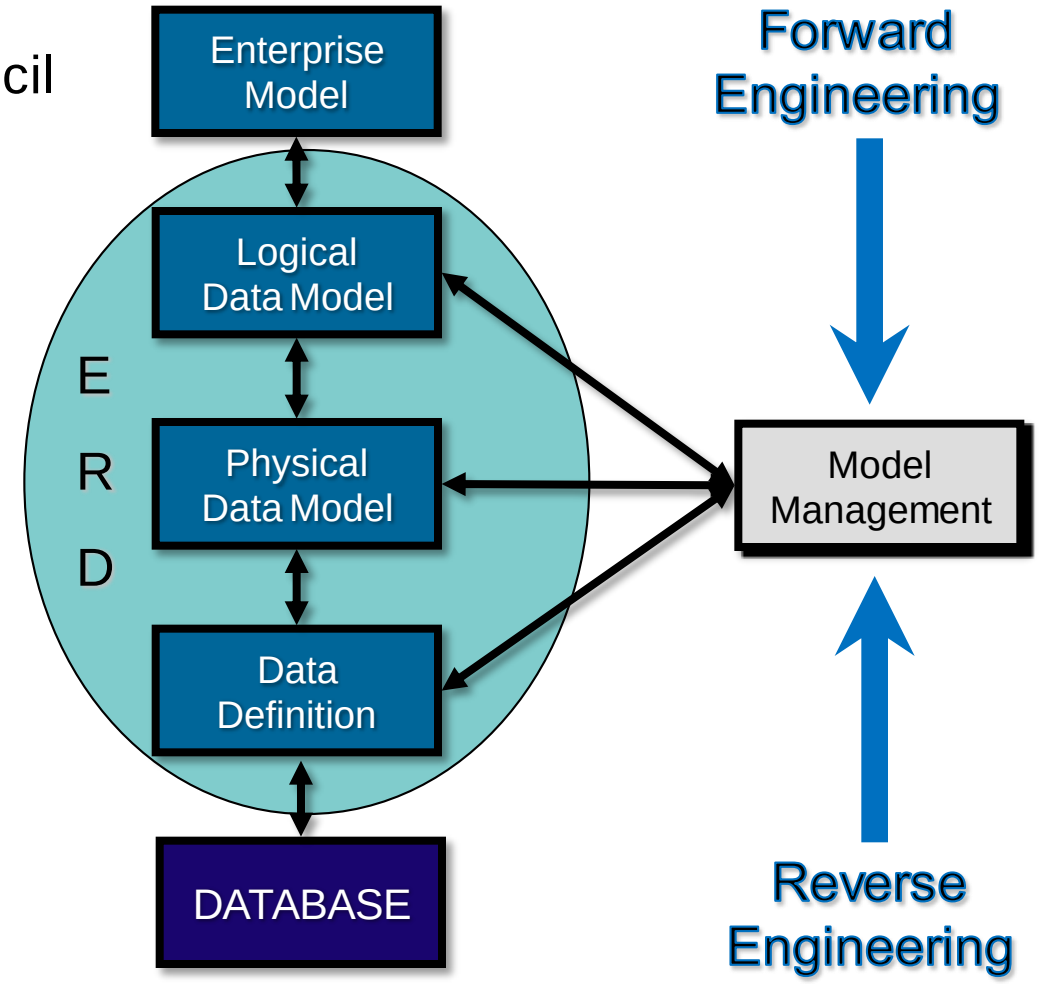
Design Layers

Governance Council

Data Architect

Data Modeler

DBA/Tester



SANDHILL

Why data model?

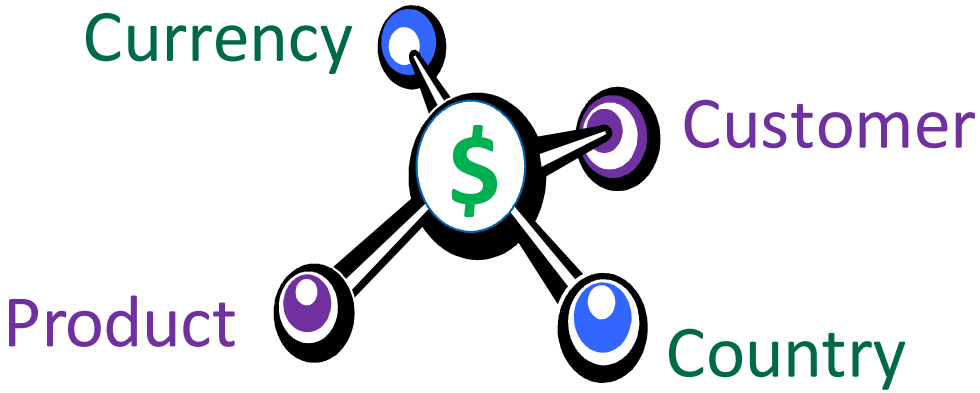
Data modeling enables you to:

- provide graphical/lexical documentation of the business area under examination
- follow principles of object reuse (inclusion) by making structures as applicable across enterprise
- permits analysis of data structures in isolation of business processes
- identify areas of agreement and contention between business and technical parties
- evolve an enterprise data foundation/architecture and provide support for model-driven development

SANDHILL

Transaction System

Point In Time



Transaction

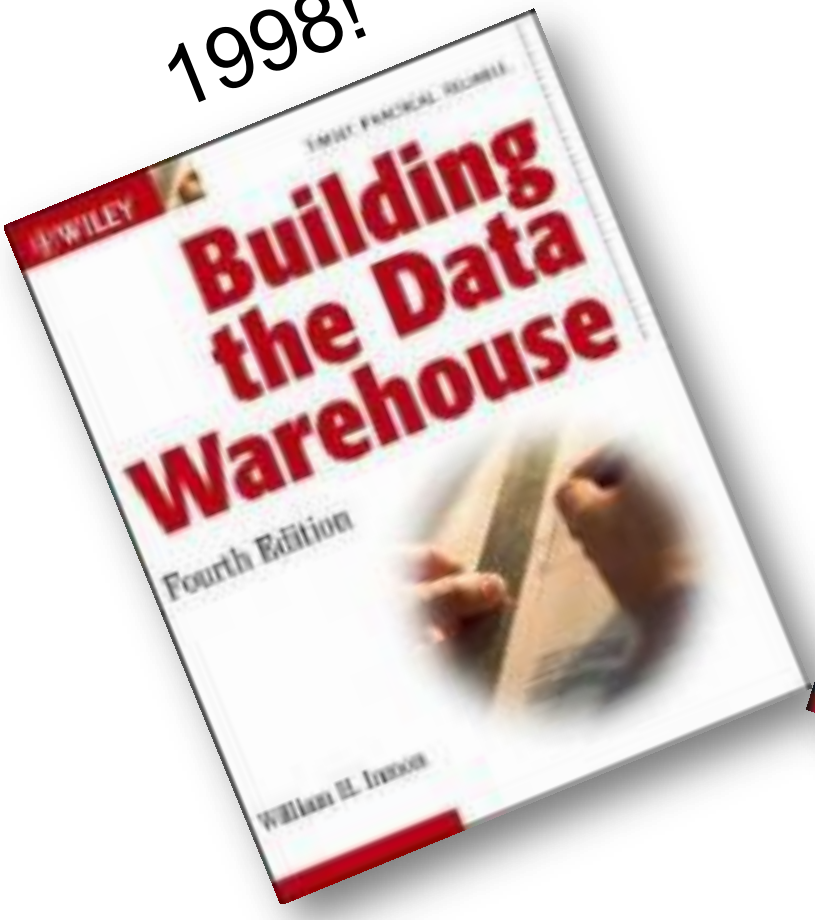
A Transaction is relationship of

Master Data + Reference Files

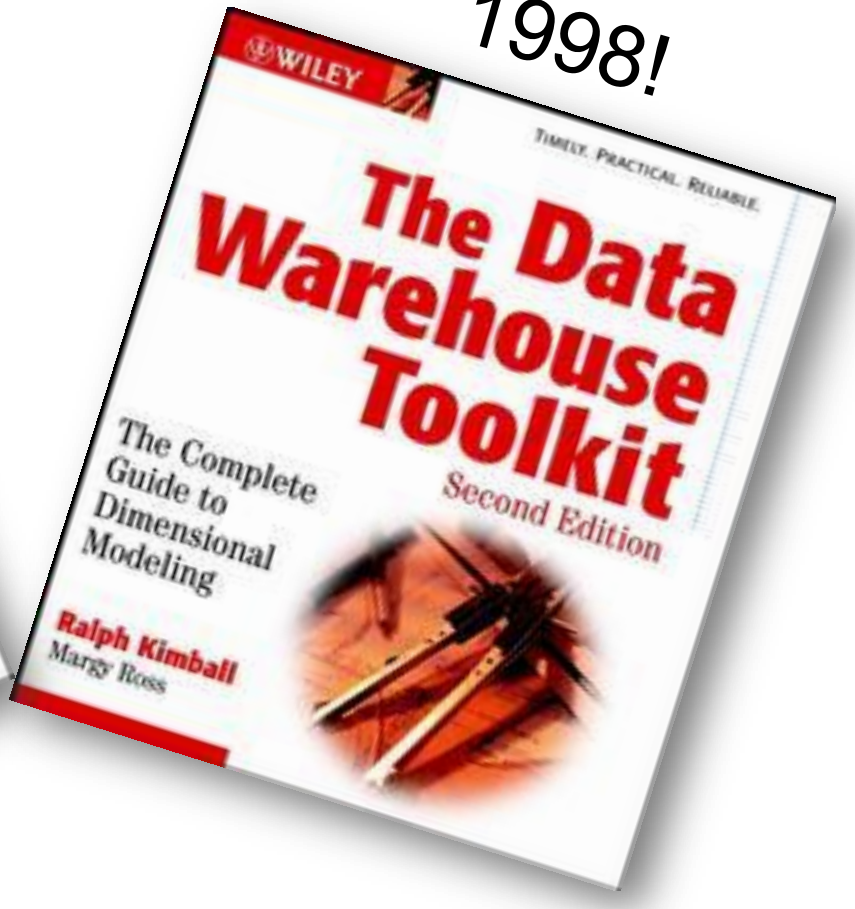
+ time stamp + volumes + secondary descriptors

SANDHILL

1998!

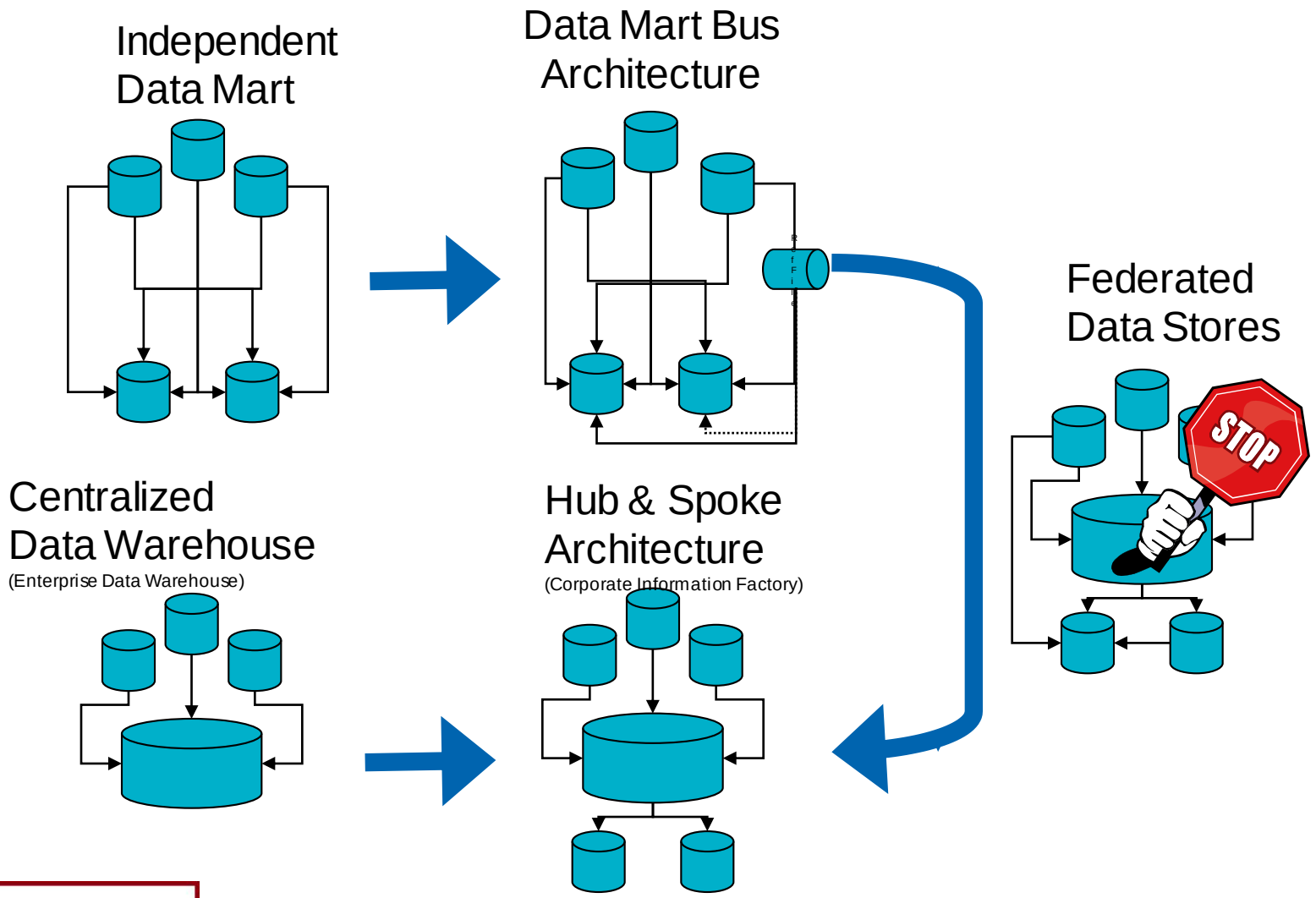


1998!



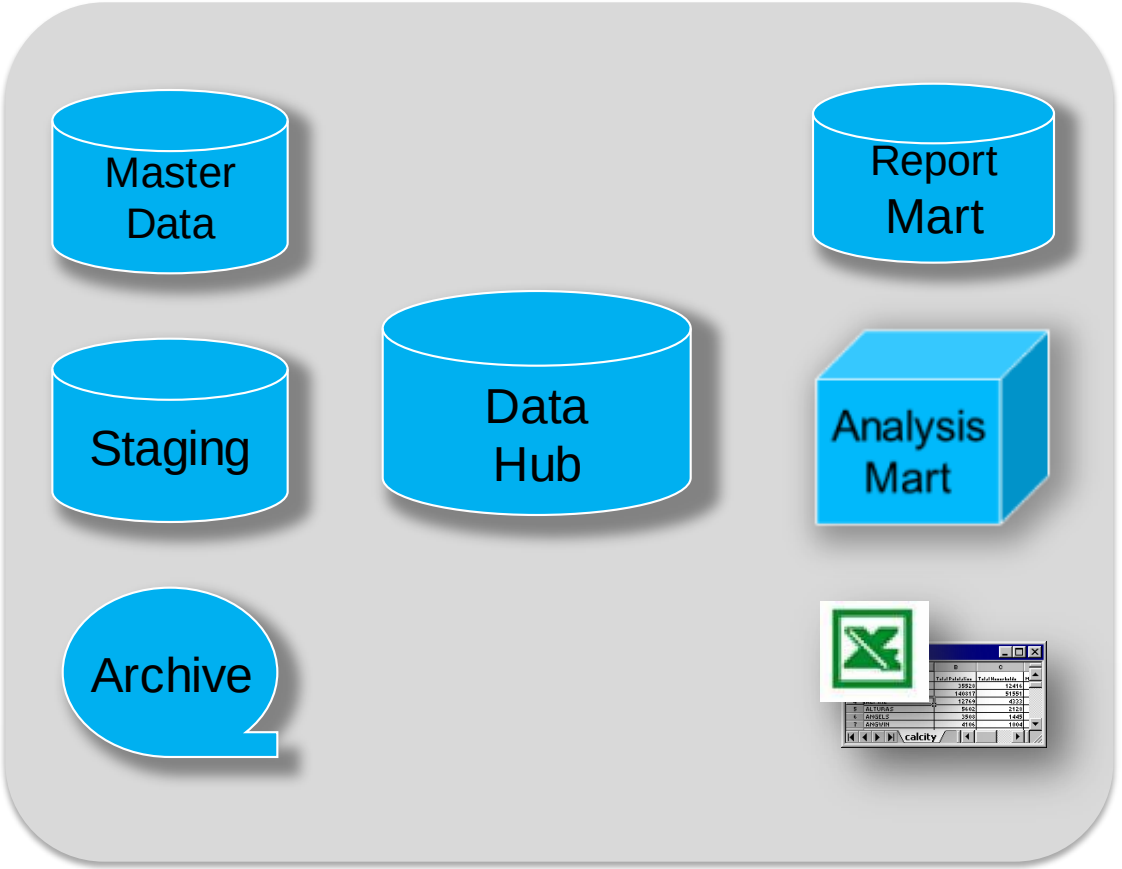
SANDHILL

Architecture choices



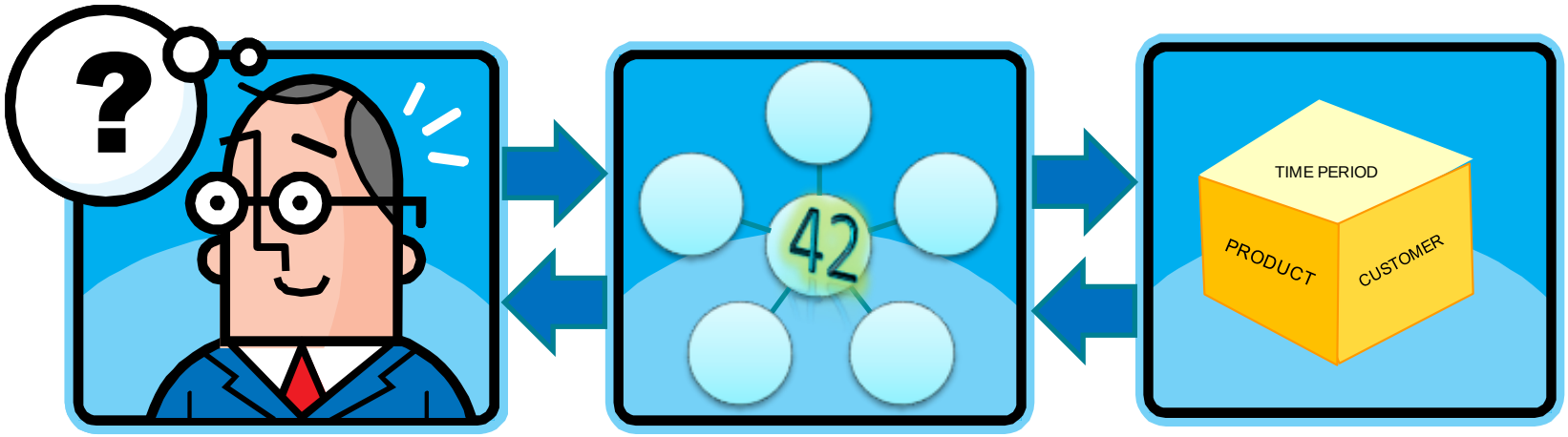
SANDHILL

Data Warehousing Models



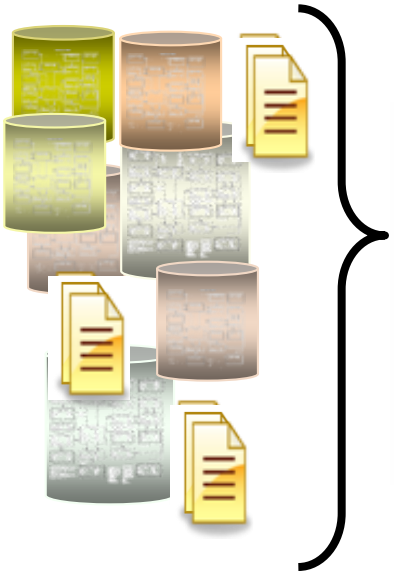
SANDHILL

Continuous Insight – Business Intelligence

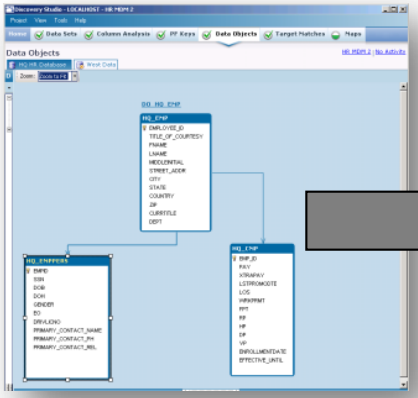


SANDHILL

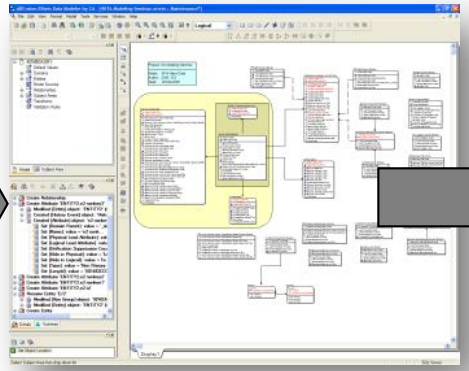
Data Profiling

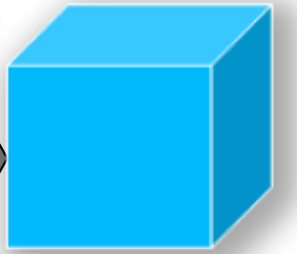


ERwin Data Profiler



ERwin Data Modeler

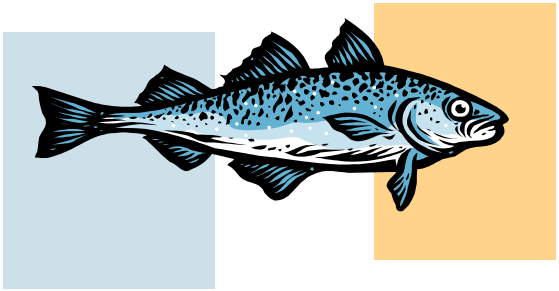




- Profile data
- Discover primary-foreign keys
- Identify orphan rows
- Find overlapping columns

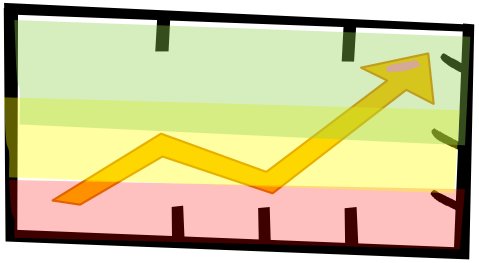
SANDHILL

Data Quality



Snapshot

DQ %



Time

Data Quality Scorecard

	C	A	T	S
ERP	%	%	%	%
CRM	%	%	%	%

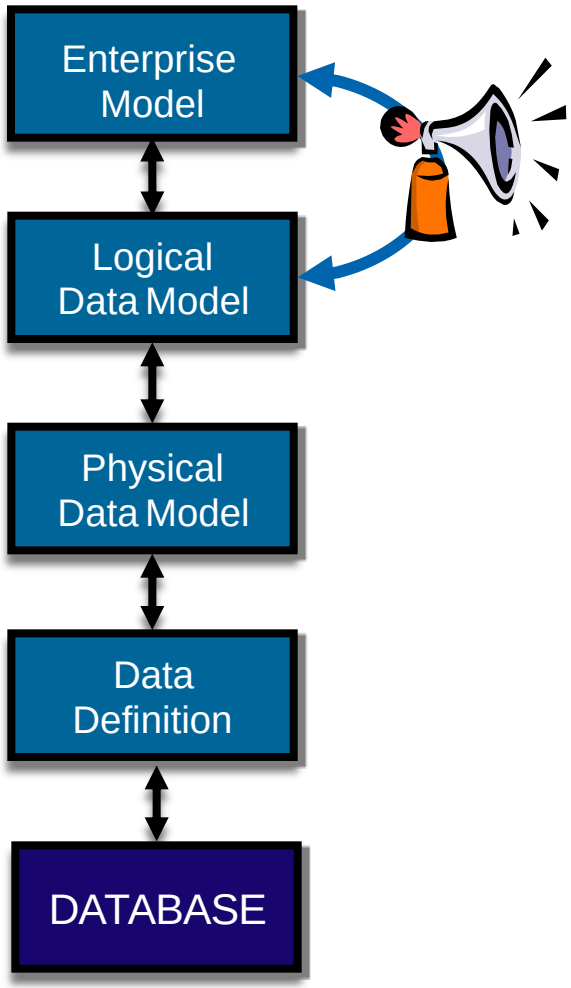
SANDHILL

Data warehouse modeling enablement

- Conduct a disciplined review of the existing system by the warehouse business and technical experts
- Partition the analysis by subject area to validate scope and objectives (agreement and contention)
- Deliver standards-based documentation of the business and technical warehouse data
- Expose and document the business and technical metadata
- Evolve the enterprise data foundation

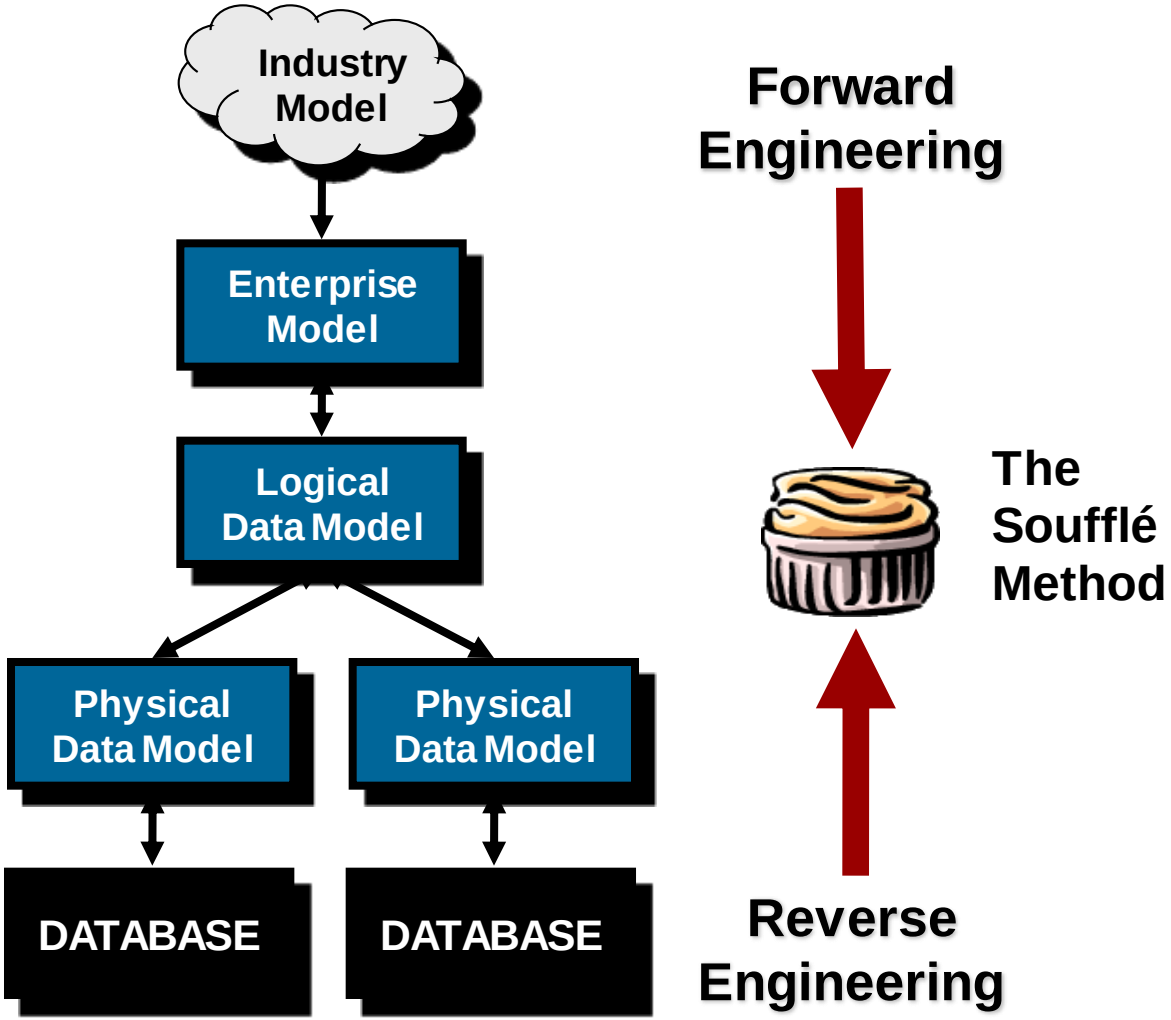
Transformation NOT Translation

Unambiguous Business Terminology



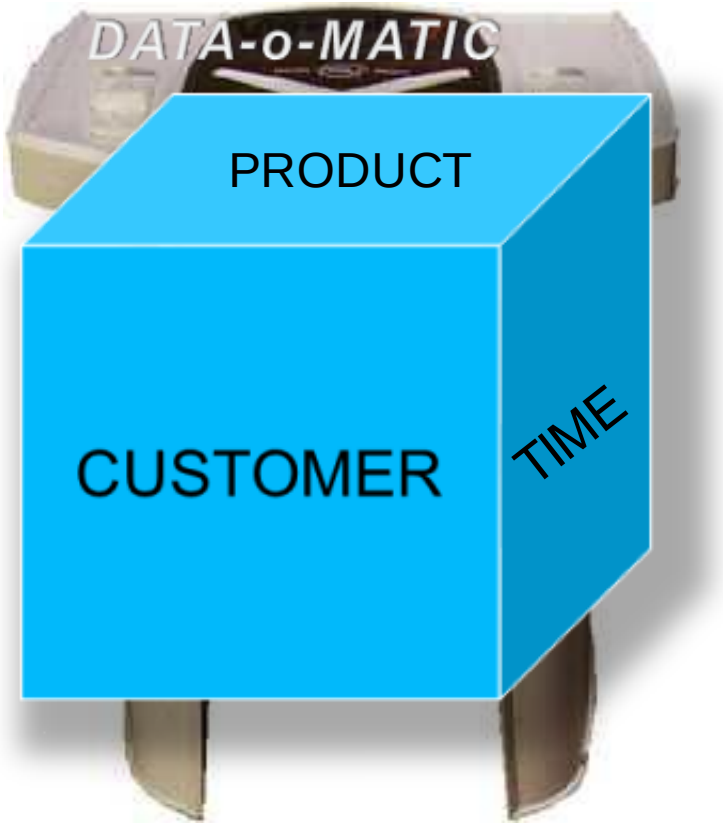
SANDHILL

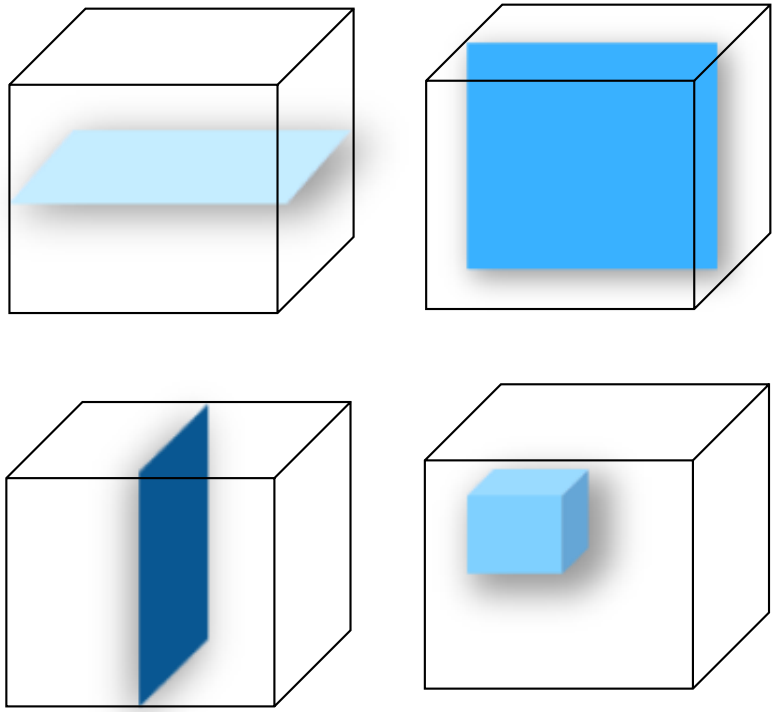
Reference / Master Data Harmonization



SANDHILL

The Cube..... It slices, it dices

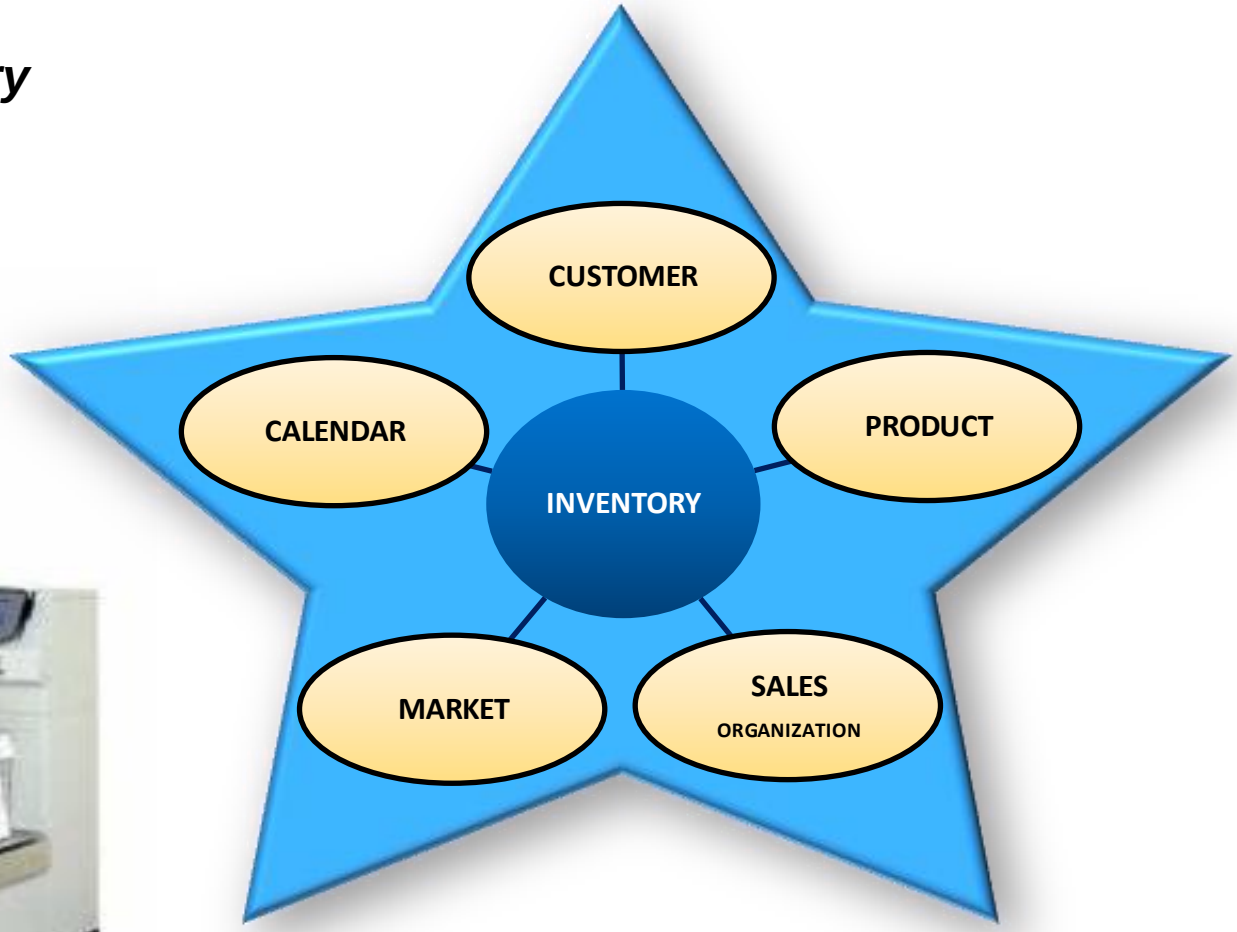




SANDHILL

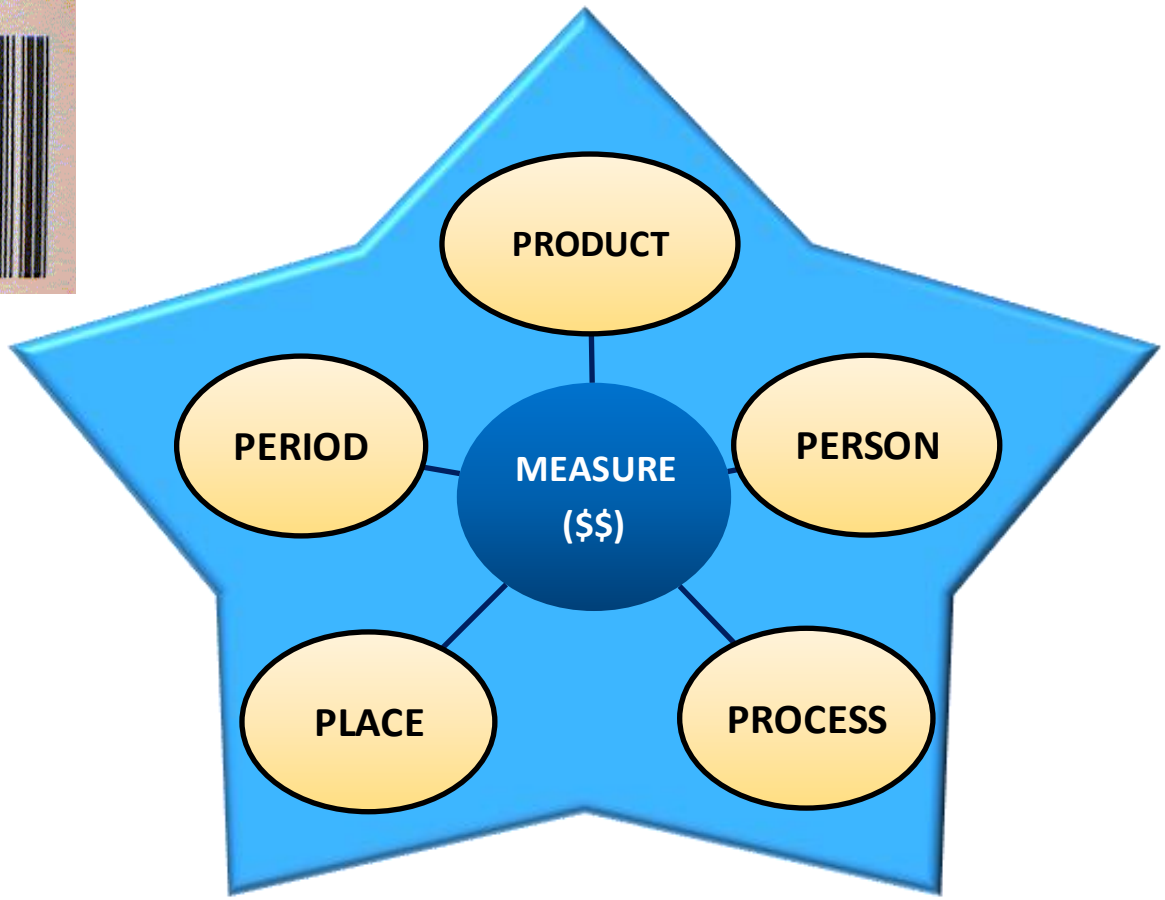
Self-Serve Business Information

I want to see Inventory
by Customer
by Product
by Day



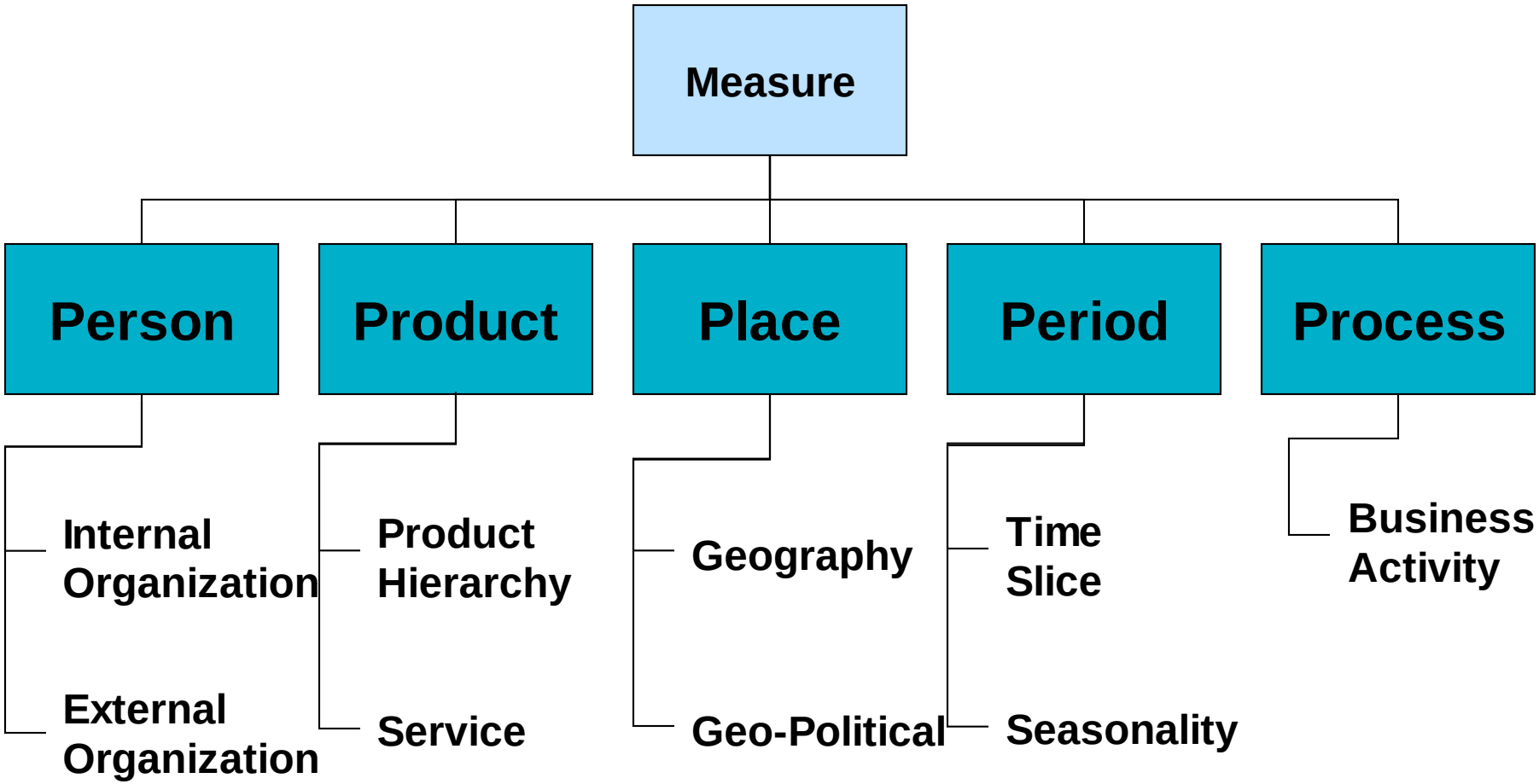
SANDHILL

Universal “P” Codes



SANDHILL

Generalized Model



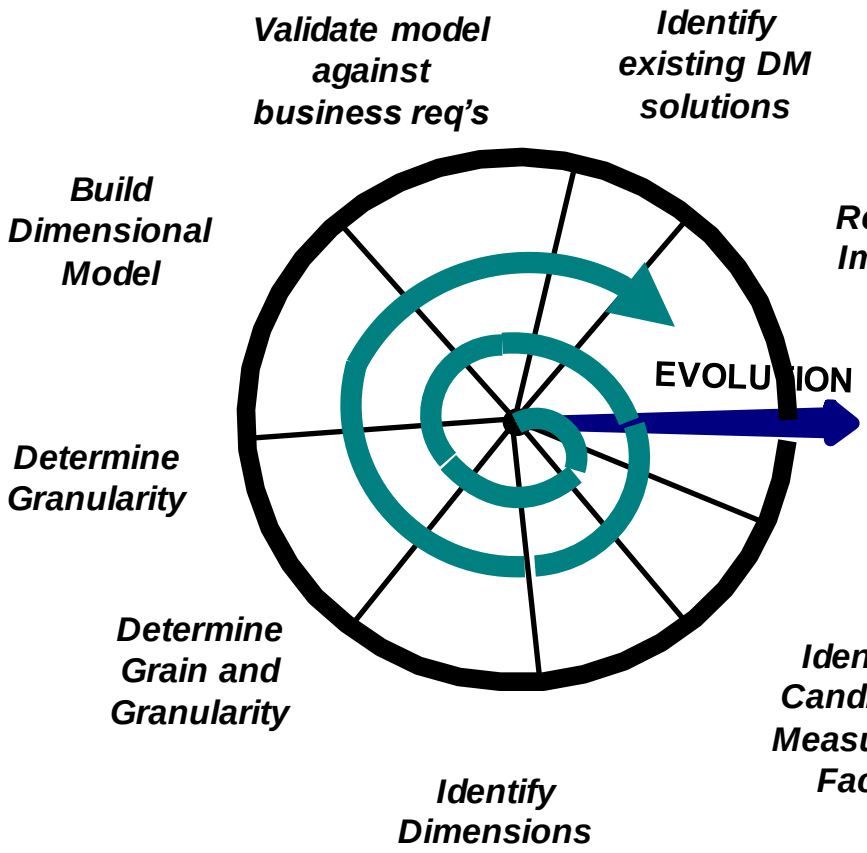
SANDHILL

Dimensional model application

- Usage of data within the dimensional model and supporting data mart (or data warehouse)
  - Confirm assumptions and by simple querying
  - Analyze behavior of different players in the business process
  - Answer to non-trivial questions
  - Perform statistical analysis of cross-departmental data
  - Analyze business results in a historical context
  - Use data to discover trends and relationships
  - Use data to help with new decisions in strategic or commercial areas
  - Help to identify functional relationships, in addition to analyzing the broader business processes, such as customer purchase behaviors.

SANDHILL

Dimensional Modeling Process



Starting Point

- Groups of end-users performing specific business data analysis tasks

SANDHILL

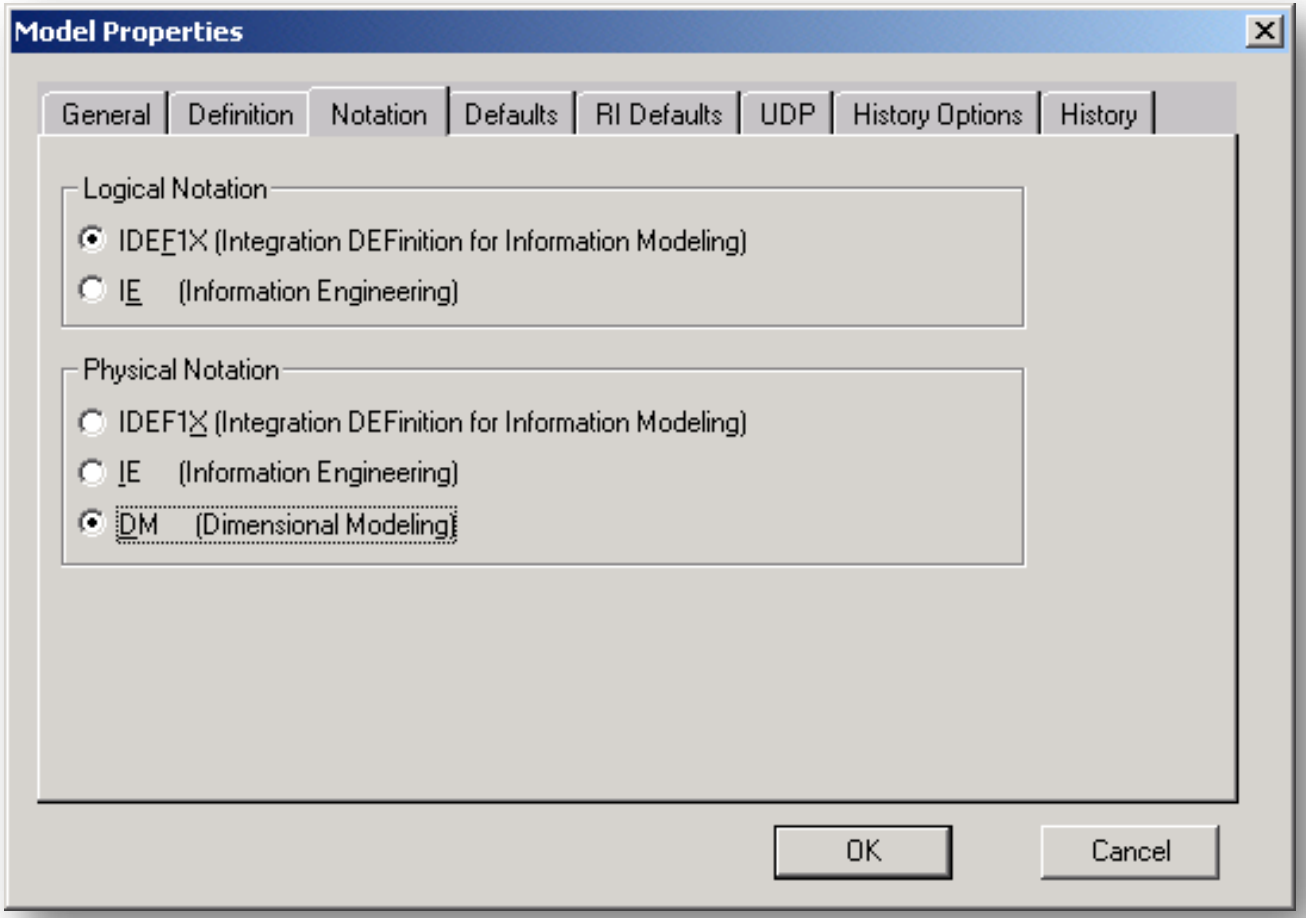
Fundamental activities

- Identify and interpret the business question
- Identify candidate measures and dimensions
- Identify grains and granularities of the Dimensional Model model
- Identify dimension hierarchies and aggregation levels
- Build the initial normalized Dimensional Model model

SANDHILL

ERwin - activating dimensional features

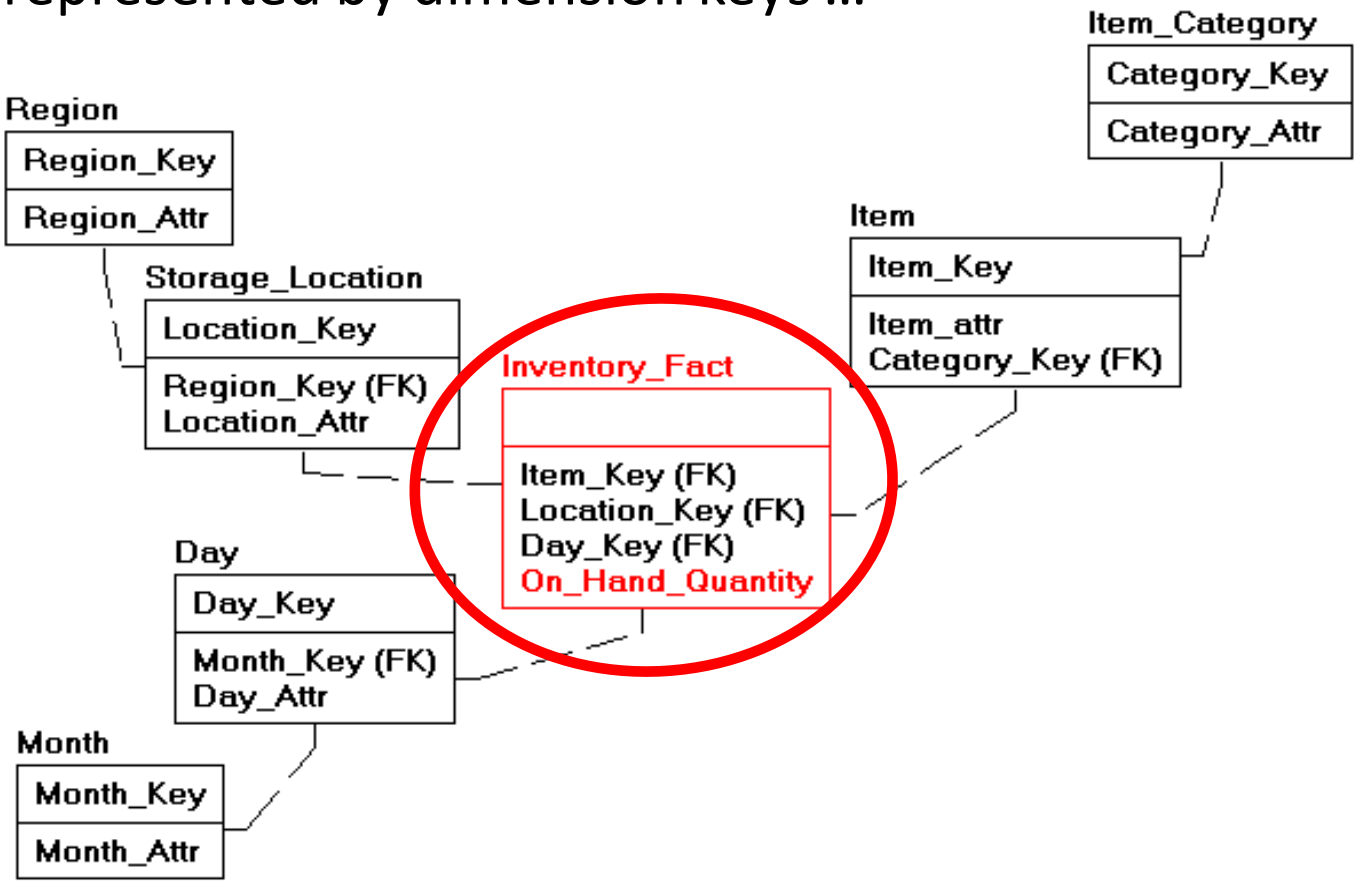
Classic model - note physical options



SANDHILL

Fact

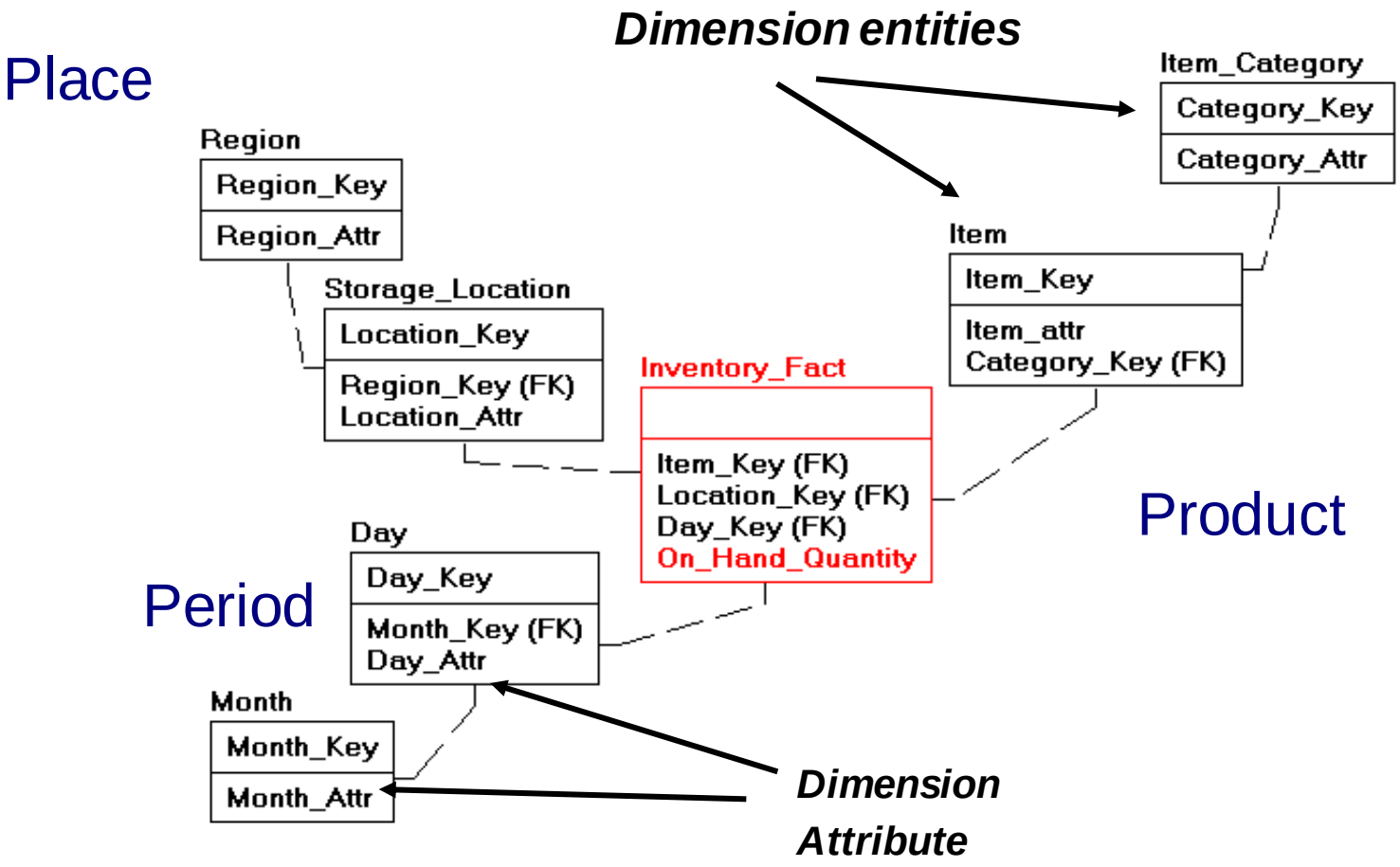
A Fact is a collection of related measures plus their associated dimensions, represented by dimension keys …



SANDHILL

Dimension

A dimension provides a certain business context to each measure

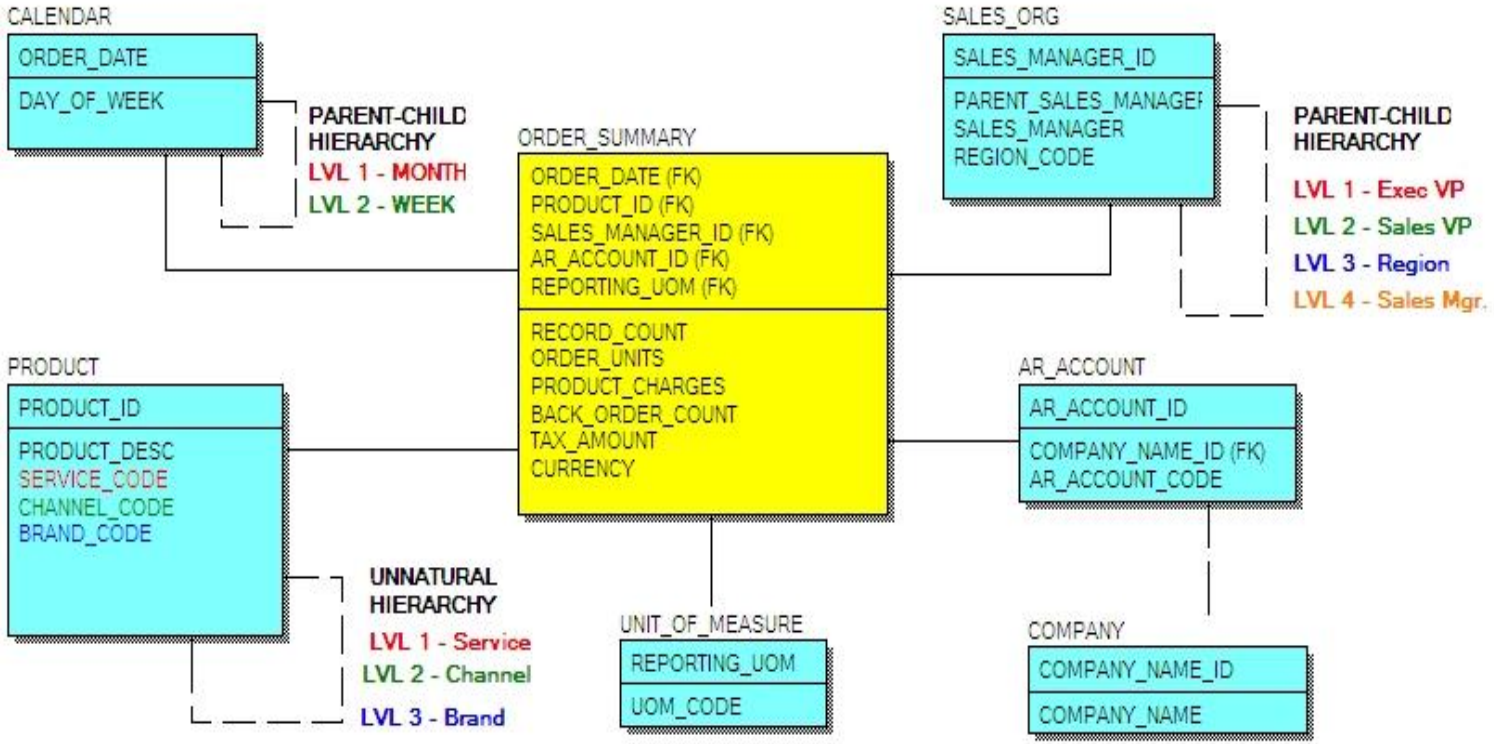


SANDHILL

Dimension Hierarchies

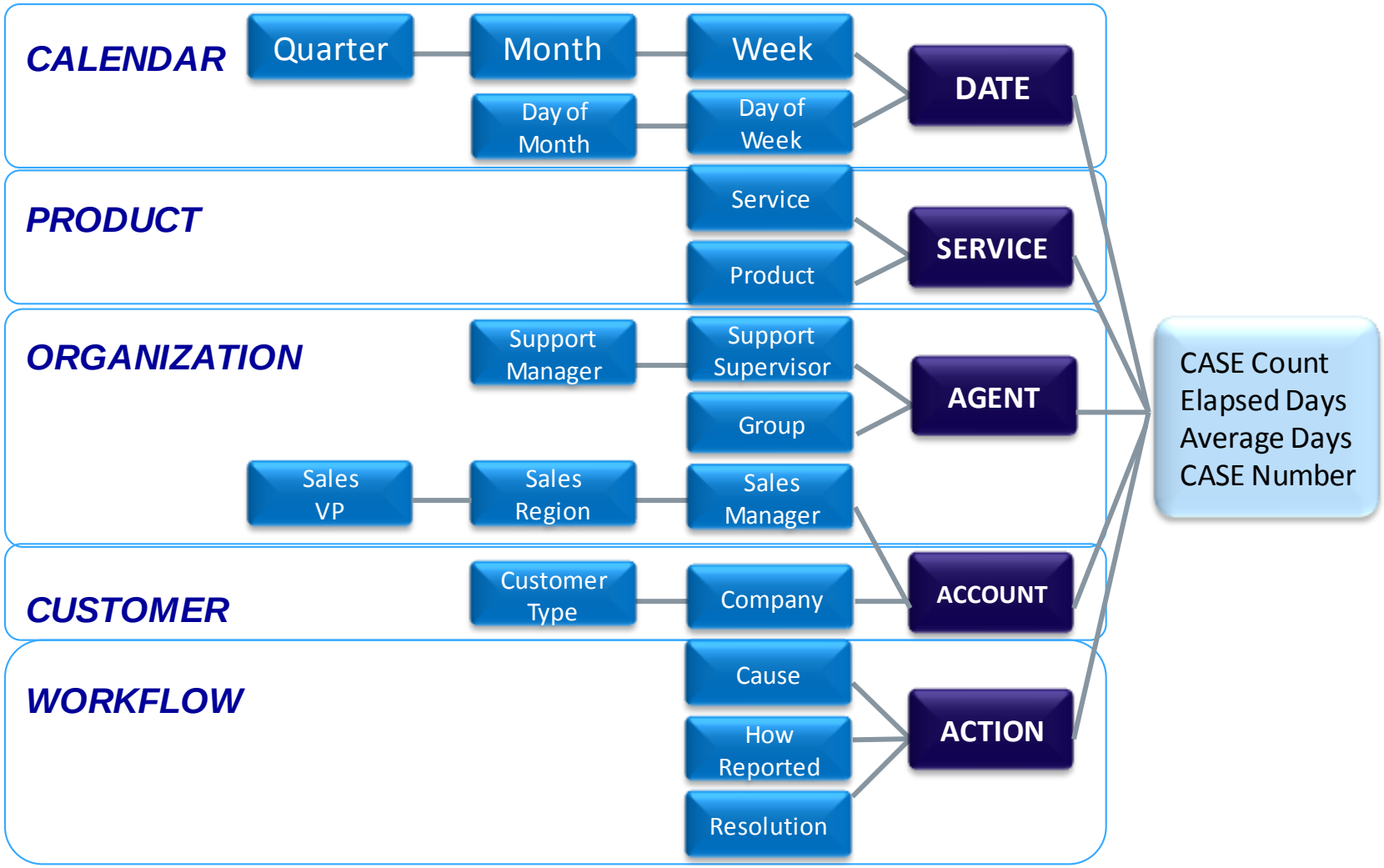
Dimension entities are related with each other through one or more dimension hierarchies

- Also known as *Aggregation paths* or *Drill hierarchies*



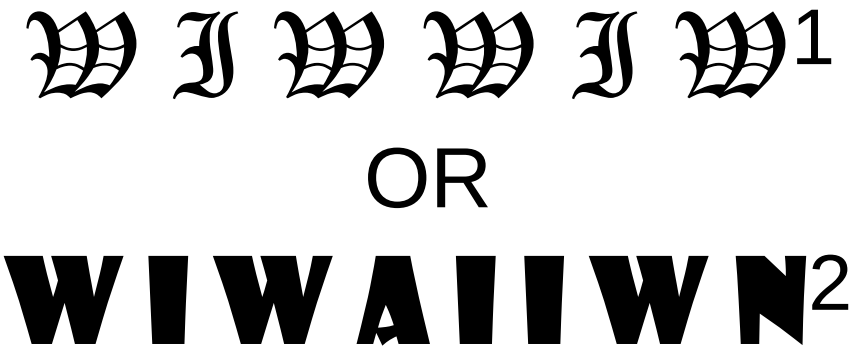
SANDHILL

Dimensional Model Drilldown



SANDHILL

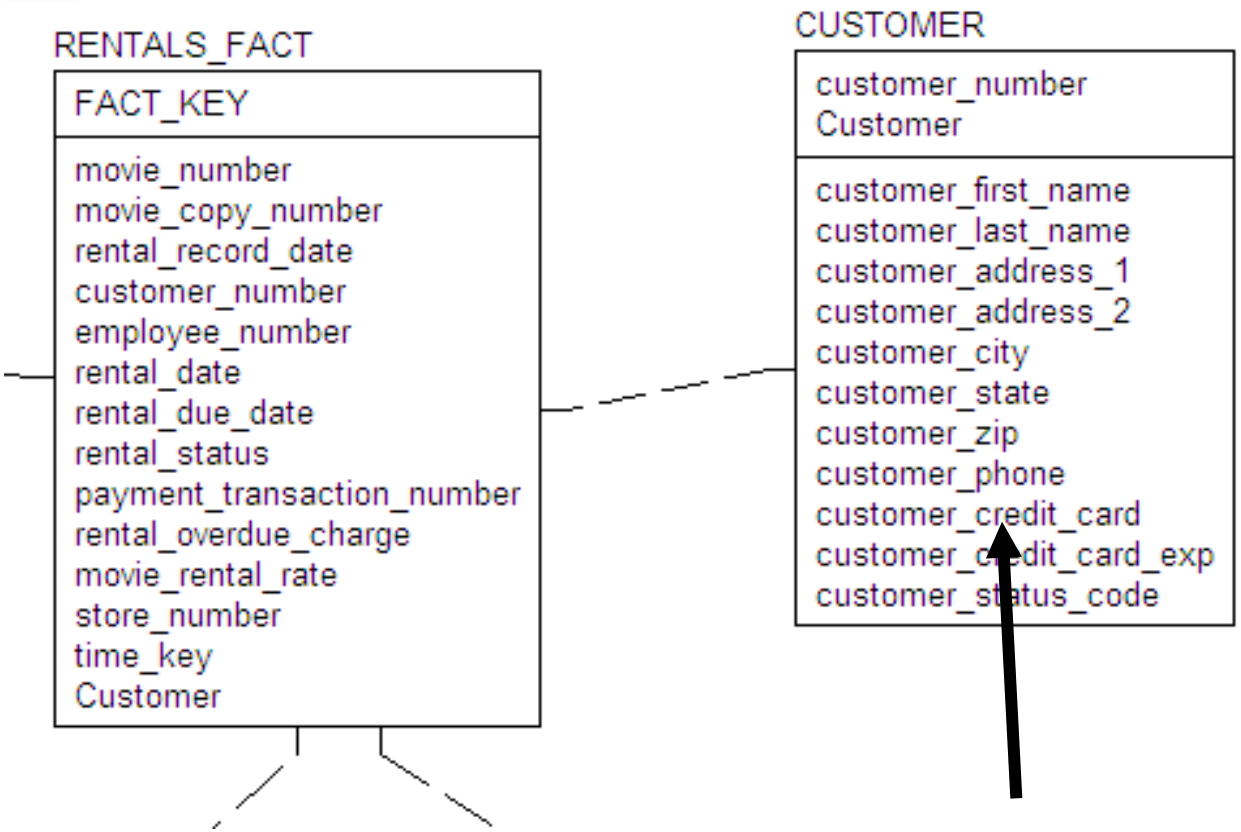
Slowly Changing Dimensions



1. What is was When it Was
2. What it Was as if it Was Now

SANDHILL

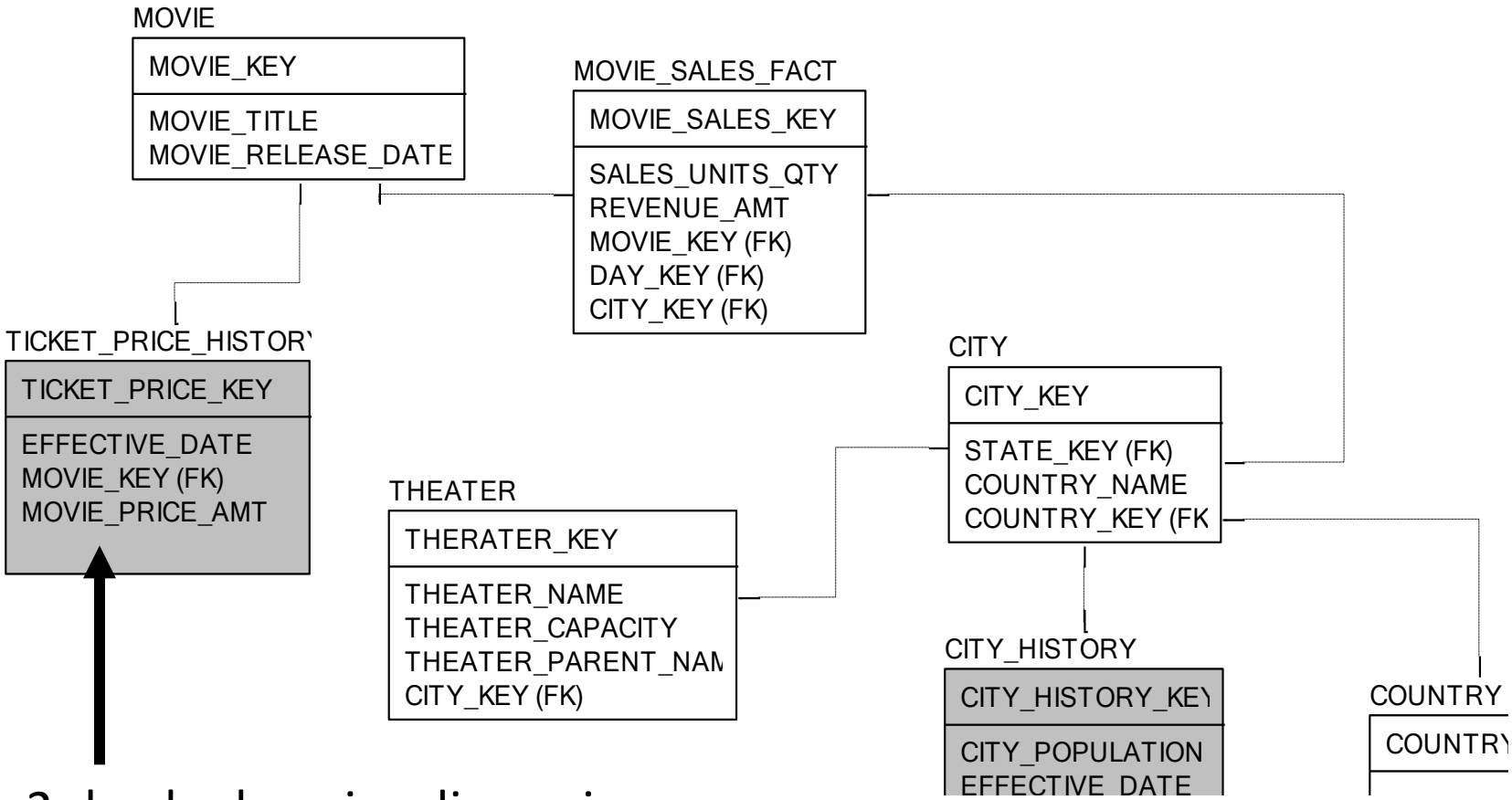
Modeling Slowly-Changing dimensions – Type 1



Type 1 slowly changing dimension, value does not change or is deemed to be constant

SANDHILL

Modeling Slowly-Changing Dimensions – Type 2

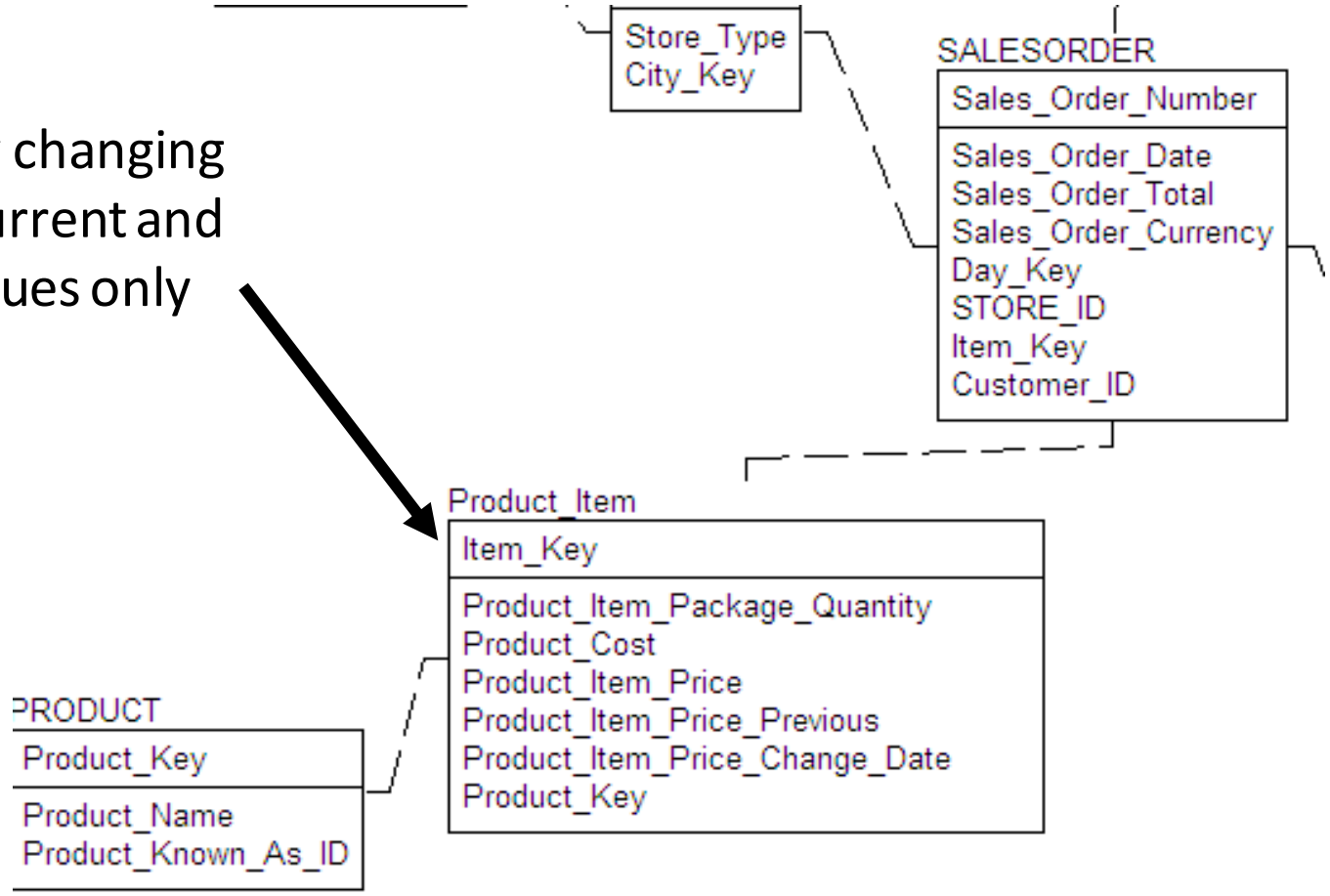


Type 2 slowly changing dimension,
add more rows

SANDHILL

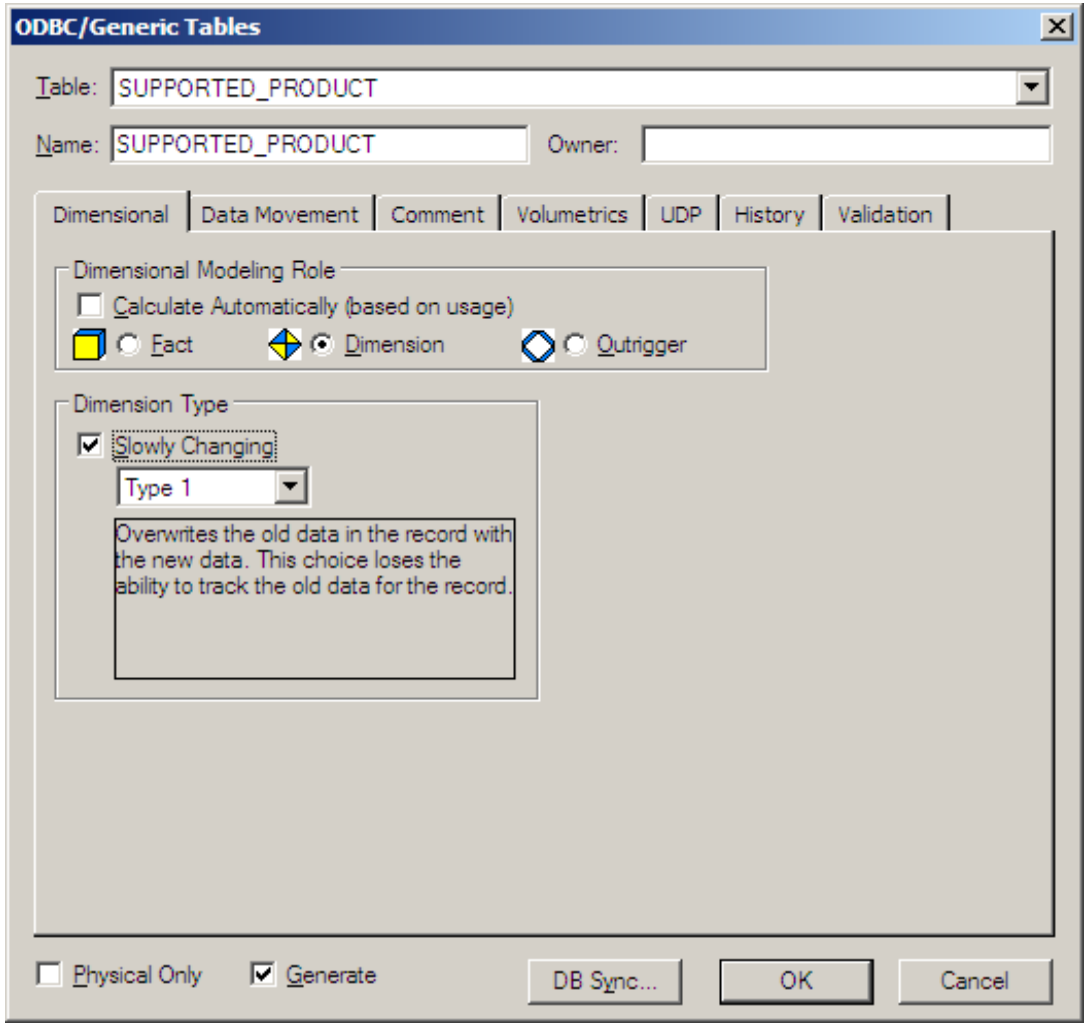
Modeling Slowly-Changing Dimensions – Type 3

Type 3 slowly changing dimension, current and previous values only

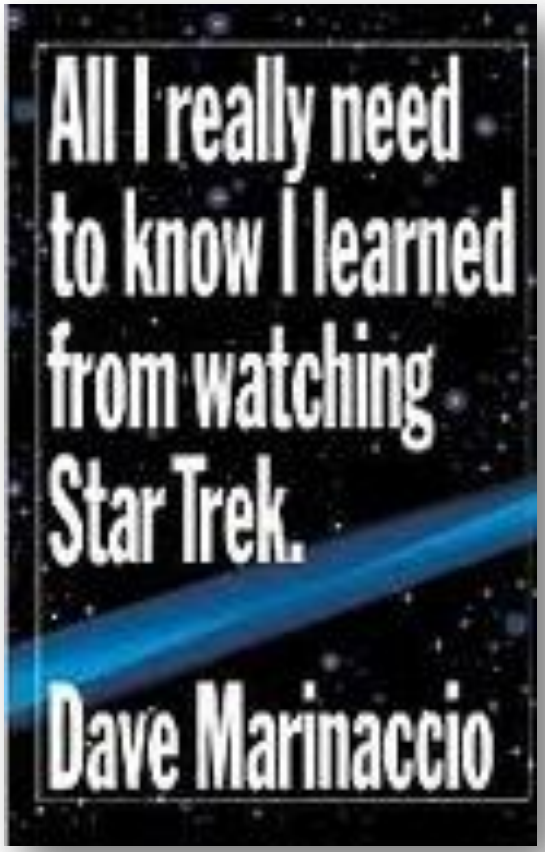


ERWin - Slowly Changing Dimensions

- In the Table properties editor, “Dimensional” tab
  - Check “Slowly Changing”
  - Select type
  - Note Automatic….



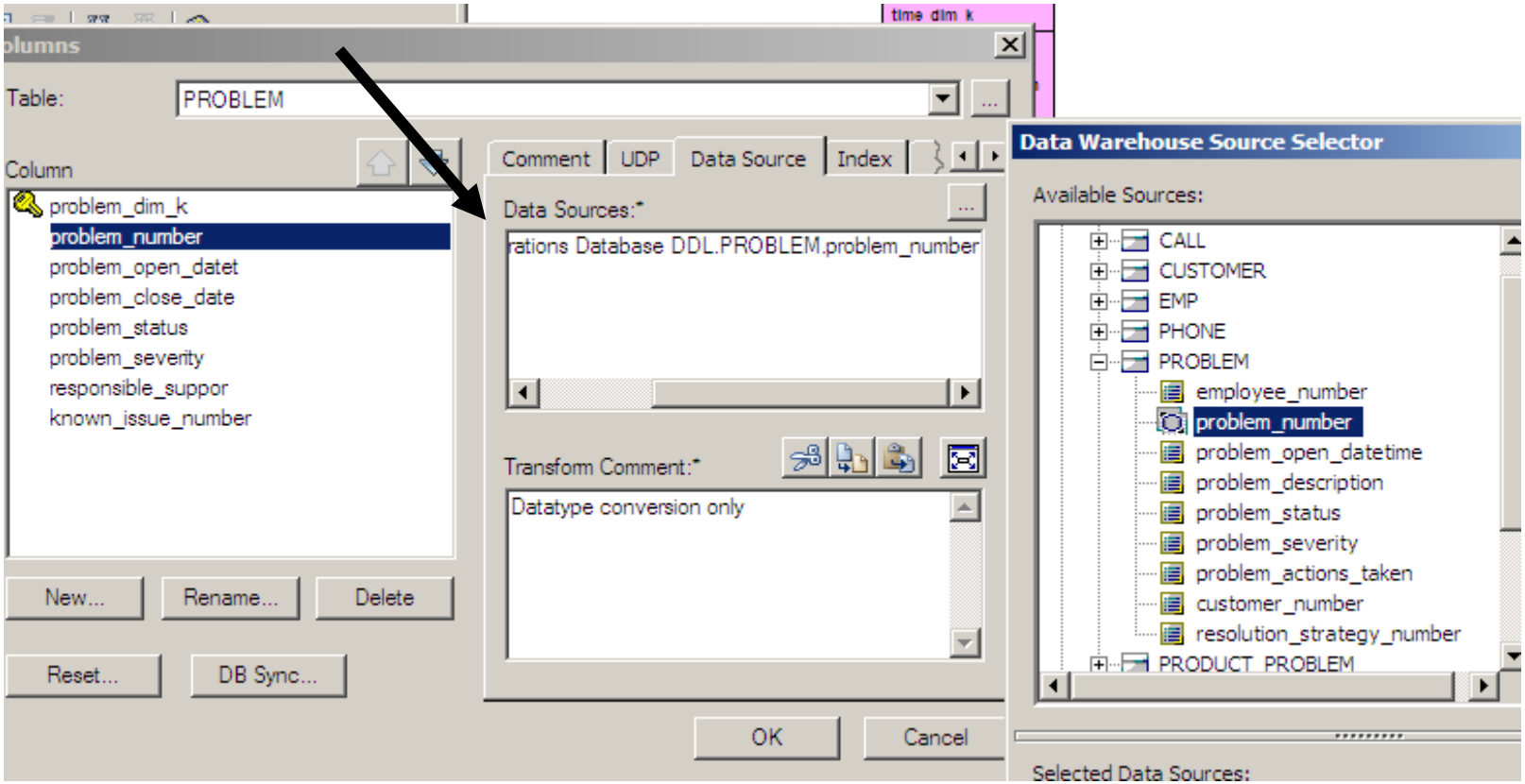
Slowly Changing Dimensions



SANDHILL

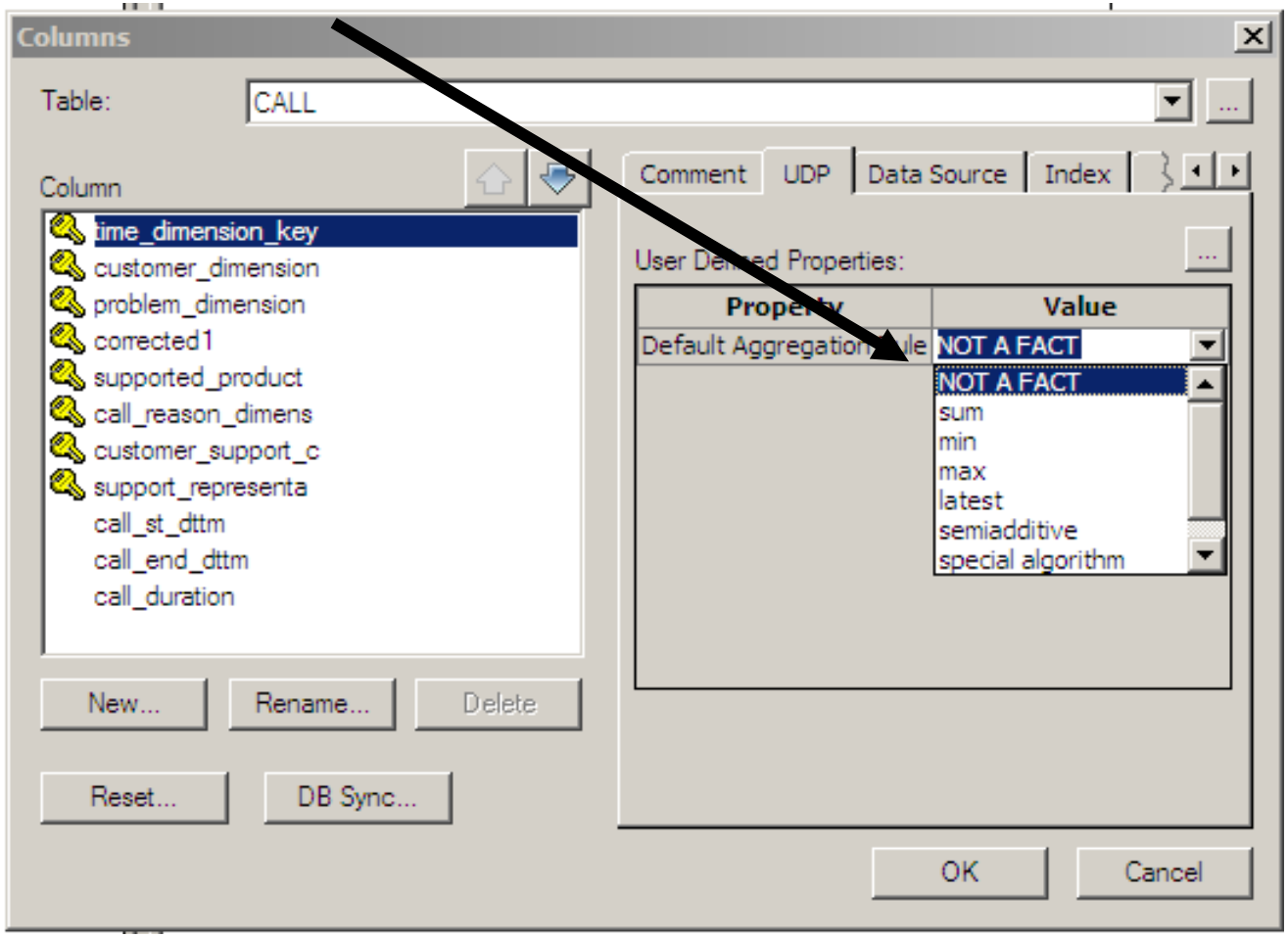
Creating Data Sources

Select the column(s) as source and a transform spec or comment



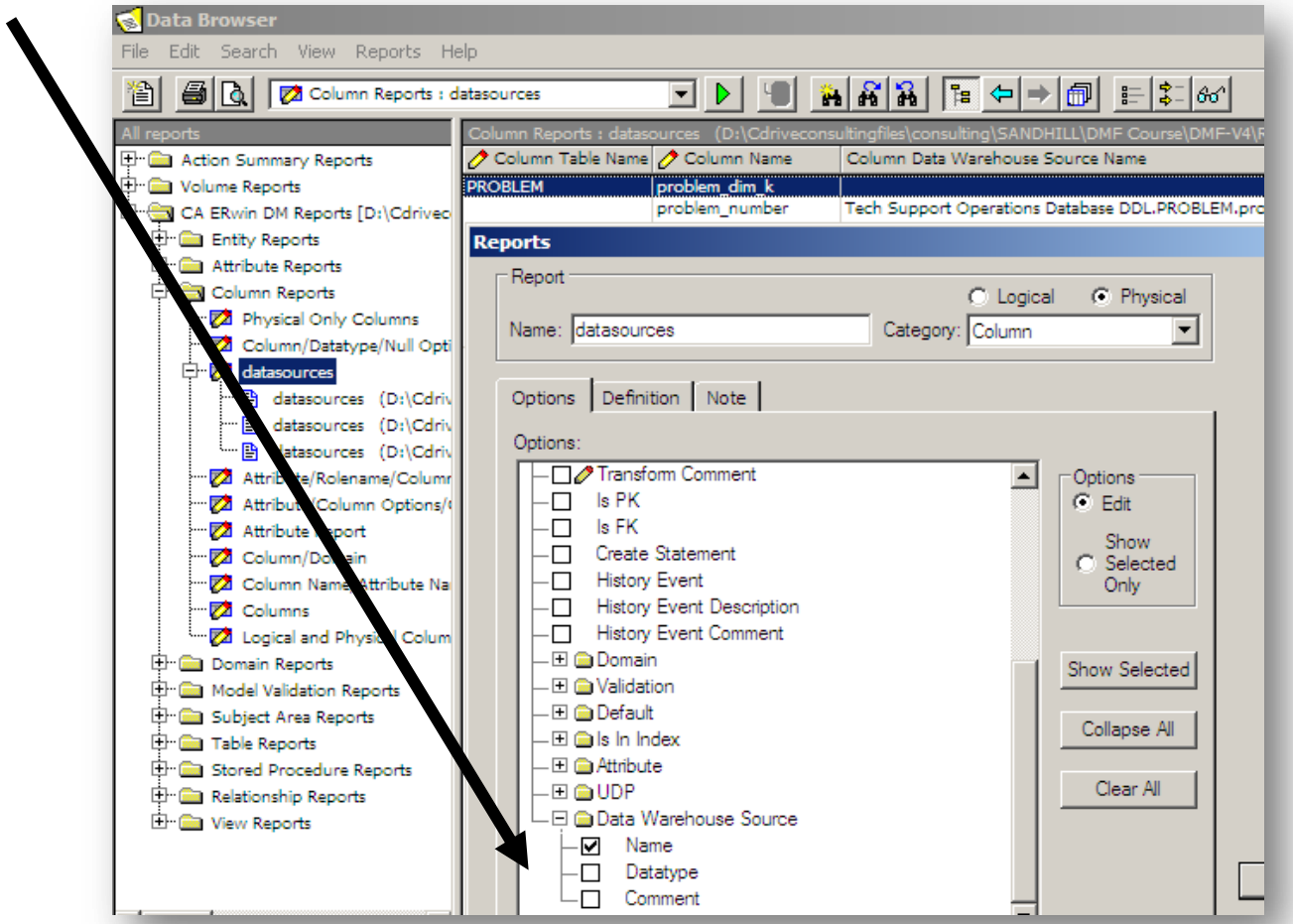
Creating Column Level UDP

Identify Aggregation Rules

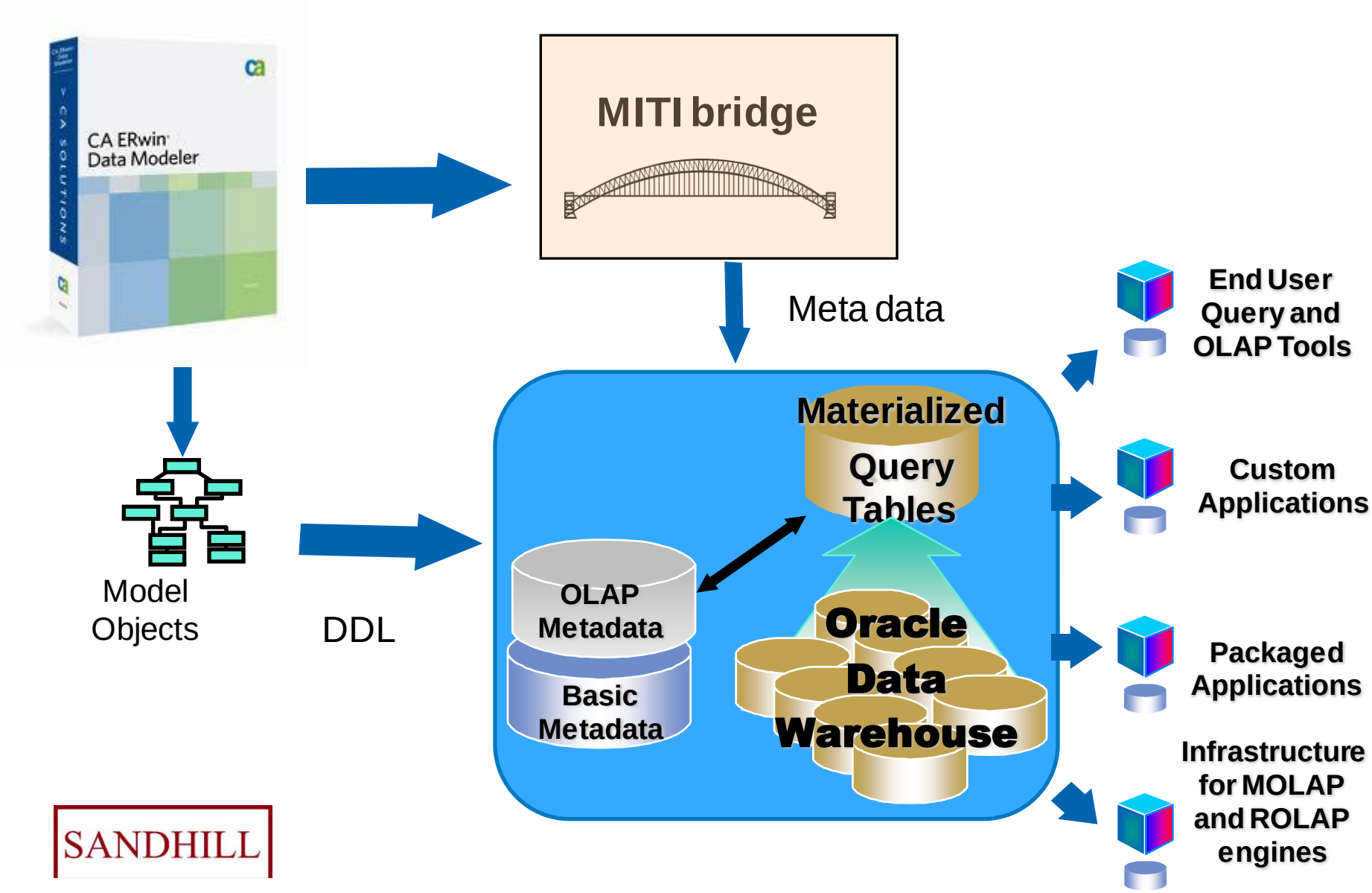


Publish a Data Browser Report

Report on Sales Orders data source



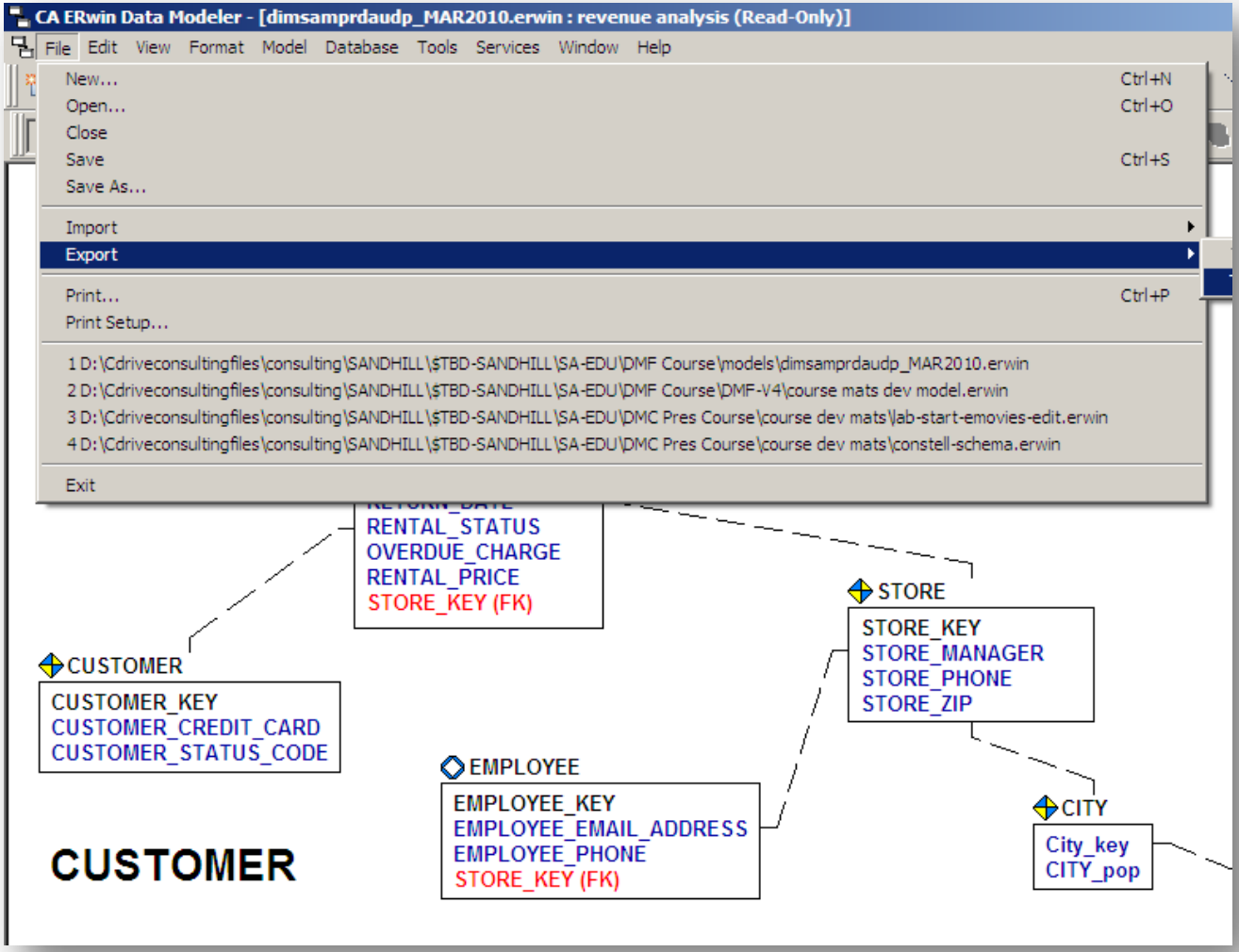
Meta Integration Bridge - MITI



SANDHILL

ERwin Metadata Export

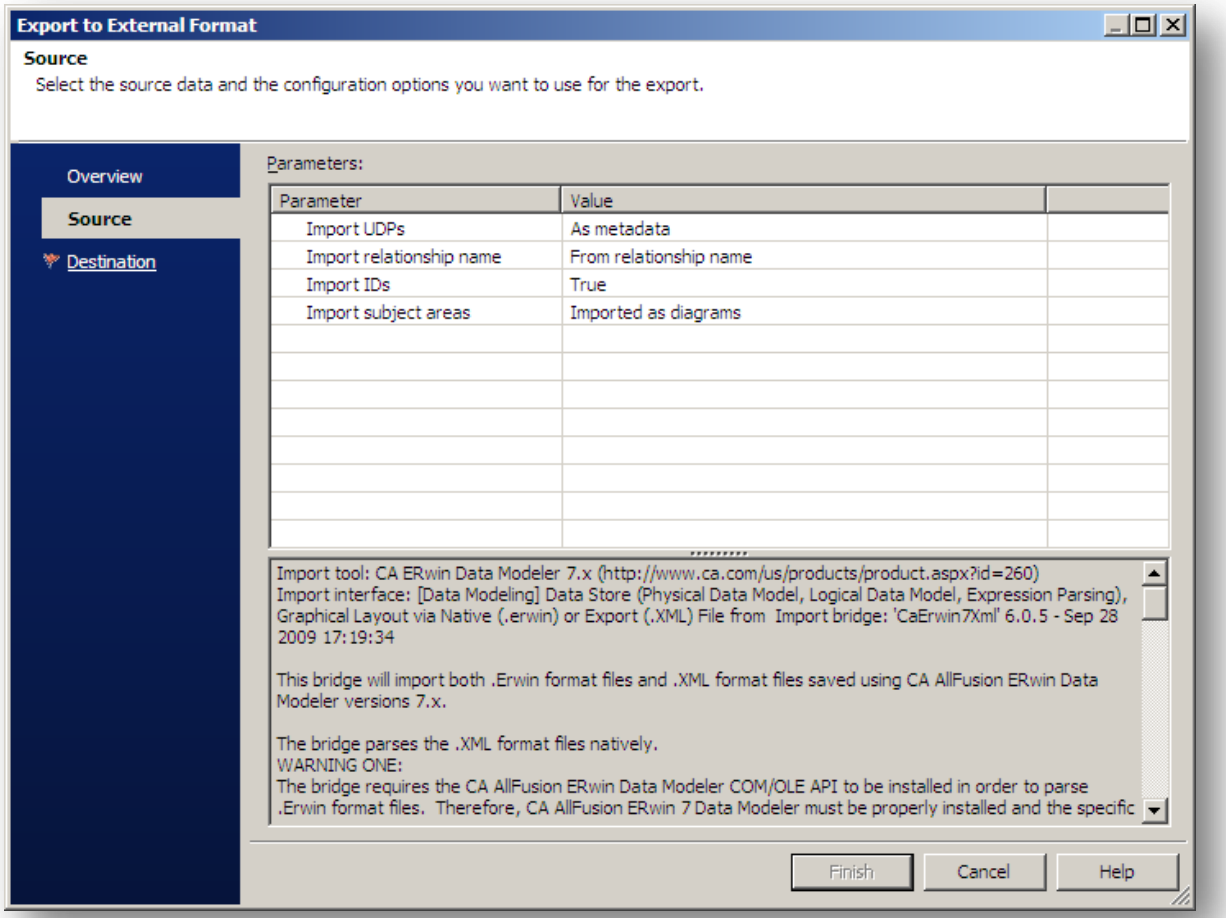
File…Export



SANDHILL

ERwin Metadata Export

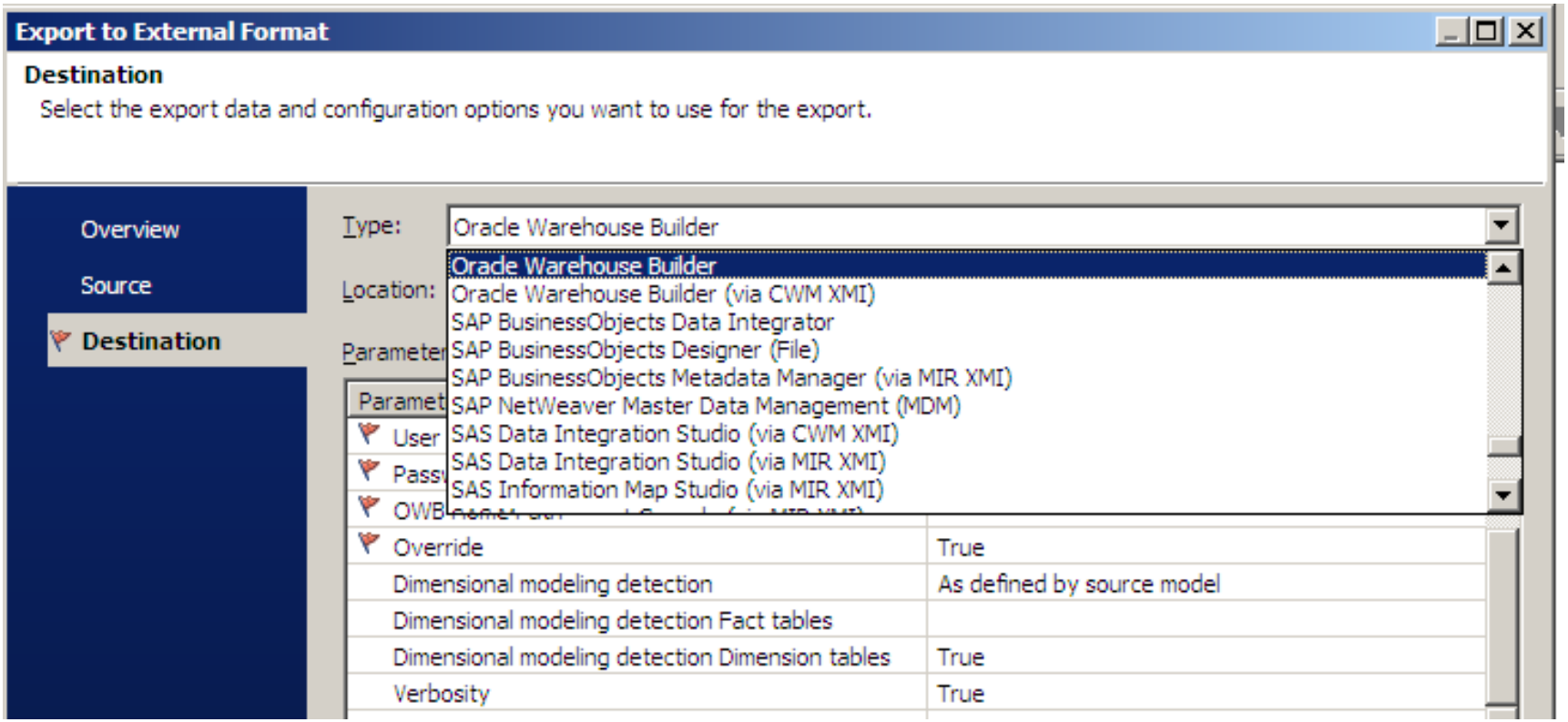
Source Model options…



SANDHILL

ERwin Metadata Export

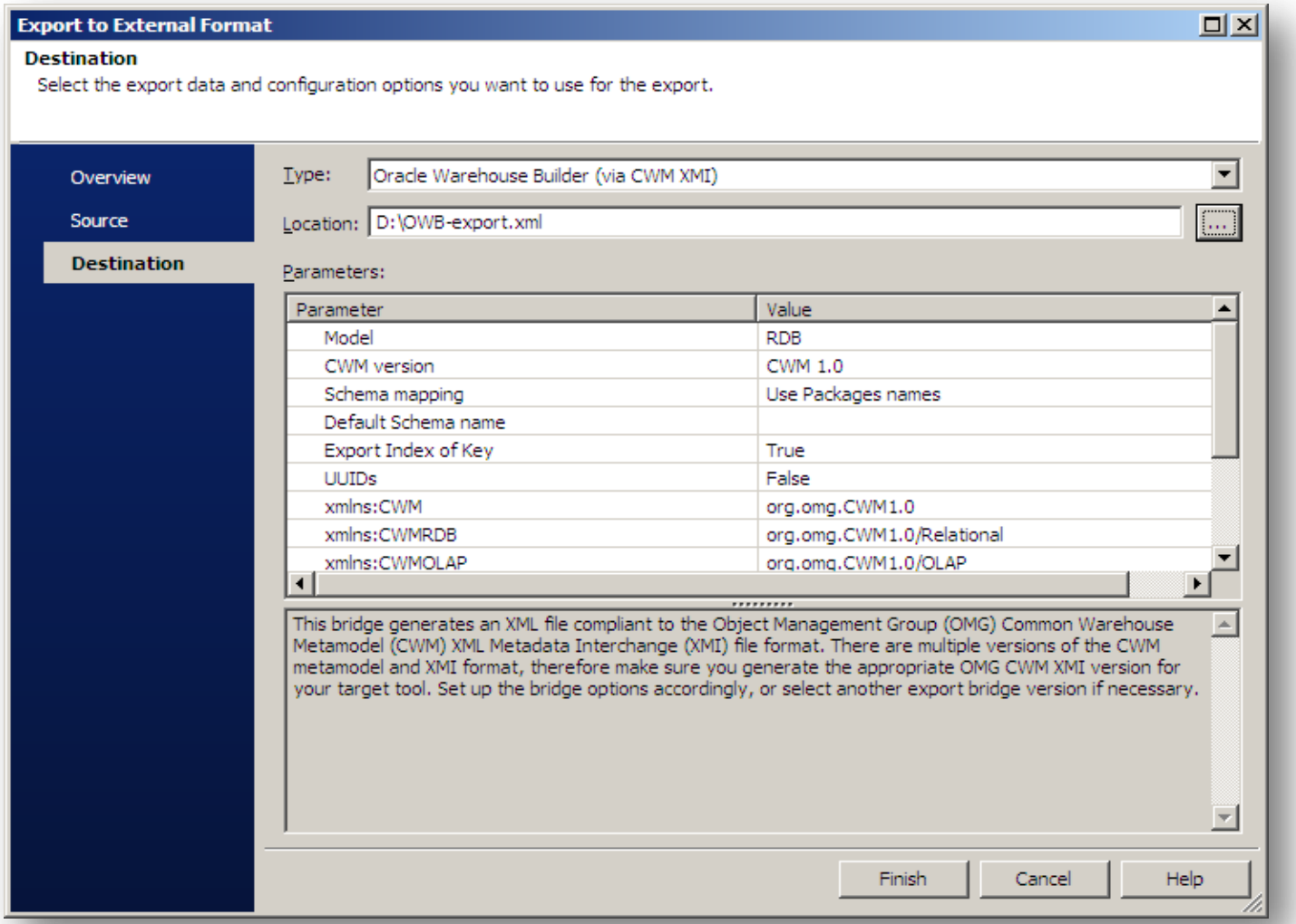
Target Project …note options



SANDHILL

ERwin Metadata Export

CWM Options…

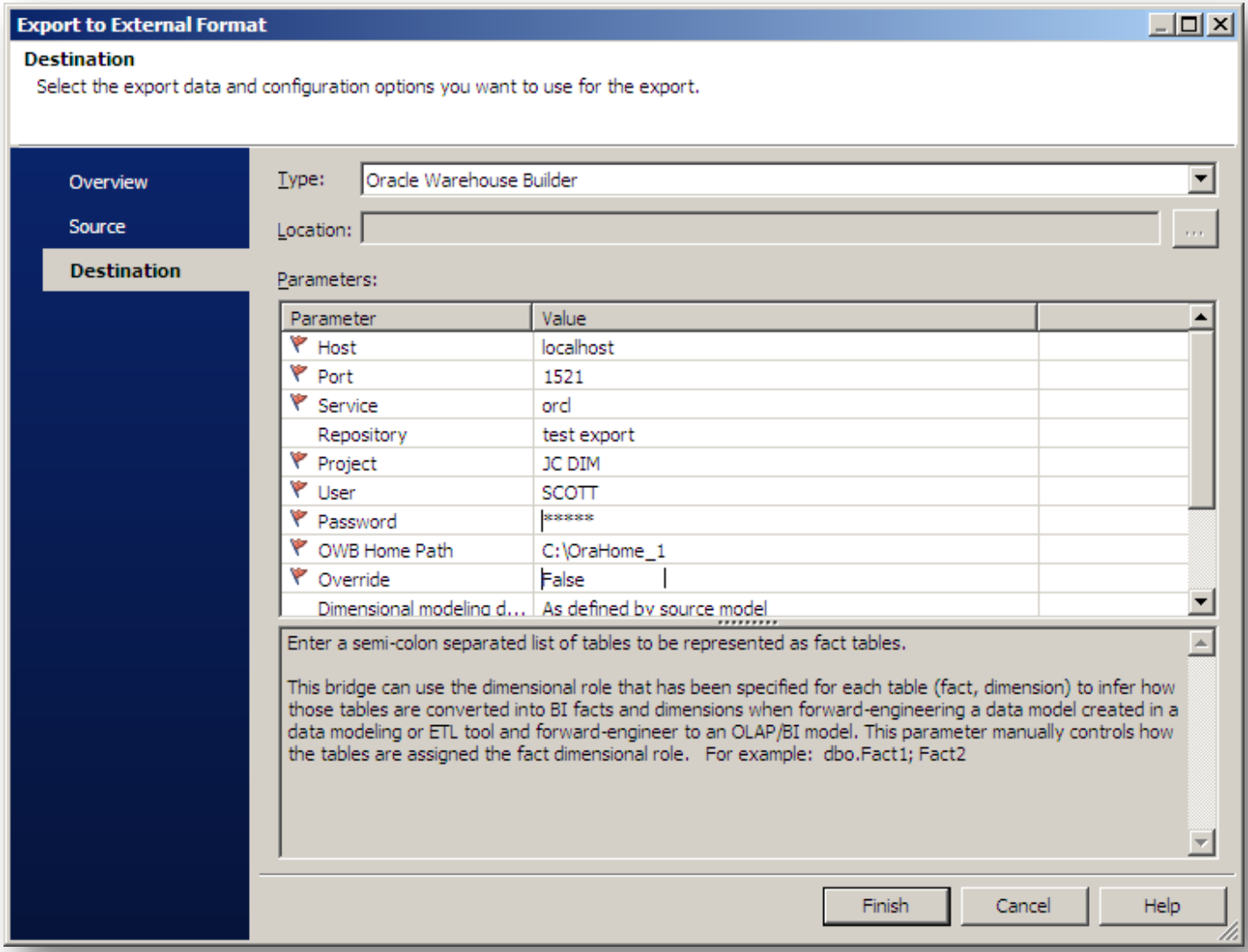


SANDHILL

ERwin Metadata Export

OWB

Options…



SANDHILL

ERwin – MITI – OWB Bridge Rules

- MITI bridge:
  - Uses the dimensional role that has been specified for each table (fact, dimension, outrigger) to infer how those tables are converted into BI facts and dimensions
  - Table's dimensional role (fact, dimension, outrigger) is to be determined from ERwin and the bridge may makes such a determination, even if the source model did not specify tables dimensional role.

ERwin – MITI – OWB Bridge

- 'As defined by source model' - Determine dimensional role only as defined in the source model.
- 'Autodetect from relational schema' - Determine dimensional role of fact and dimension tables based on foreign keys.
- The MITI algorithm determines a table to be a:
  - fact table - if the table has only incoming foreign keys
  - dimension table - if the table has with no foreign key relationships to any other tables
  - dimension table or outrigger table - if the table has outgoing foreign keys and depending on the minimum distance (number of foreign key relationships) to a fact table

Importing Metadata from an Object Management Group CWM

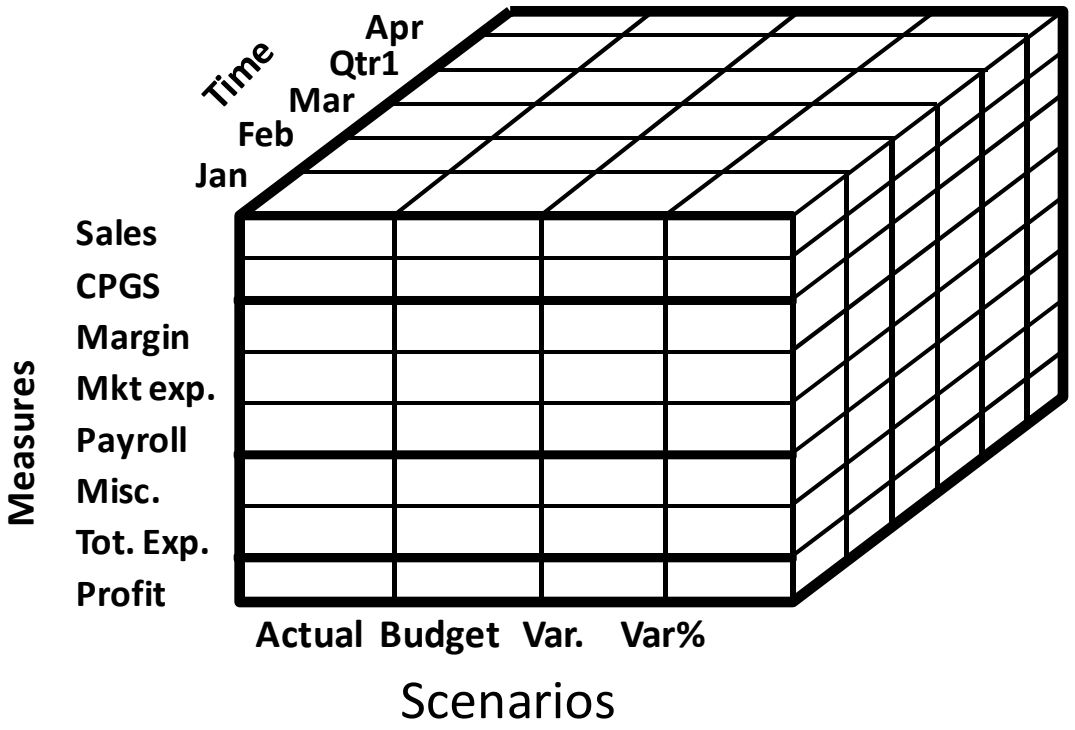
Oracle Warehouse Builder Transfer Wizard

OMG CWM Object	OWB Object
Package	Project
Schema	Module
Dimension	Dimension
Level	Level
Attribute	Level Attributes
Hierarchy	Hierarchy
Cube	Cube
Measure	Cube Measure
Table	Table
Column	Column
Foreign Key	Foreign Key
Unique Constraint/Primary Key	Unique Key
View	View
Column	Column

SANDHILL

Hyperion Storage method

- Non-relational - Stores data in Blocks
- Contents of a block determined by dense dimensions
- Block size affects
  - Data Load time
  - Calculation time
  - Retrieval time



SANDHILL

Oracle 11 BI Suite Enterprise Edition
Enterprise Business Model Administration

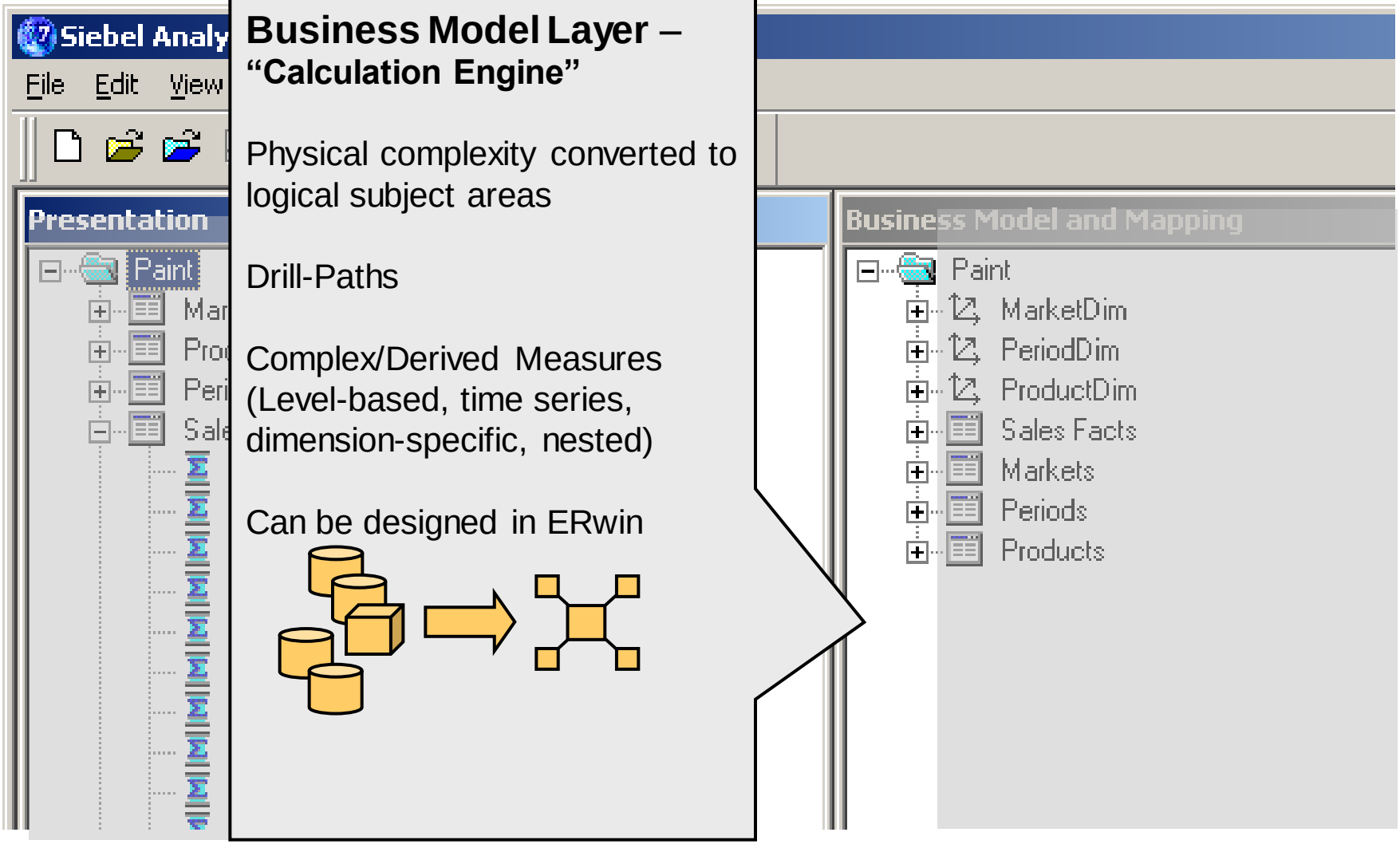
Business Model Layer –
"Calculation Engine"

Physical complexity converted to logical subject areas

Drill-Paths

Complex/Derived Measures (Level-based, time series, dimension-specific, nested)

Can be designed in ERwin



SANDHILL

Summary

- ERwin Model objects and metadata that Oracle can leverage:
  - Dimensional Modeling
    - Measures, Dimensions, Facts, and metadata (derivation and usage)
  - ETL mapping
  - Metadata publishing and extension
  - Dimension Hierarchy Characteristics

SANDHILL