Day 1

Dimensional Modeling Concepts

- Dimensional Modeling in Context
  - Business Intelligence Defined
  - Data Warehousing Defined
  - Data Mart Defined
- Dimensional Modeling Basics
  - Dimensional Model Defined
  - Dimensional Modeling Defined
  - Business Metrics and Measures Defined
  - Business Metrics Examples
  - Dimensional Data Models
- Comparing E-R and Dimensional Models
  - A Quick Review of E-R Modeling
  - Introduction to Dimensional Models
  - Relational with Additional Constraints
  - A Basis for Comparison
  - Relational for Transaction Processing
  - Dimensional Data for Business Analysis
  - Conformed Dimensions
- Concepts Summary
  - Review of Some Key Points

Module Two

Requirements Gathering for Dimensional Models

- Business Context for Data Modeling
  - Business Value
  - Business Alignment
  - Business Process Alignment
- Business Questions as Requirements Models
  - A Framework for Business Questions
  - Examples
  - Refining Business Questions
- Fact/Qualifier Analysis
  - From Business Questions to Data Requirements
  - Mapping Business Questions
- Requirements Gathering Summary
  - Process Review

Module Three

Logical Dimensional Data Modeling
• Modeling Meters and Measures
  o A Group of Related Business Measures
• Modeling Dimensions
  o Adding Dimensions from Qualifiers
  o Dimension Hierarchy
  o Refining the Dimensions
  o Completing the Dimensions
• More about Meters and Measures
  o Granularity and the Meter
  o Granularity and the Measures
  o Completing the Meter
• Model Verification
  o Testing the Model
• Logical Modeling Summary
  o Process Review

Module Four
From Logical Model to Star Schema

• Star Schema Dimensions
  o Naming the Dimensions
  o Modeling Dimension Tables
  o Defining Dimension Table Keys
• Star Schema Fact Tables
  o Modeling the Fact Table
  o Defining the Fact Table Key
  o Supporting Calculated Measures
  o Semi-Additive and Non-Additive Facts
• Star Schema Design Challenges
  o Slowly Changing Dimensions
  o Degenerate Dimensions
  o Junk Dimensions
  o Difficult Situations
• Modeling Process Summary
  o From Business Requirements to Star Schema

Module Five
Dimensional Data and Business Analysis

• Delivering Business Value
  o Data Enabled Business Analysis
  o Collecting, Analyzing, and Using Business Metrics
• Effective Dimensional Modeling
  o Critical Success Factors
  o Mistakes to Avoid
  o References and Learning Resources
Day 2

Advanced Dimensional Modeling: Techniques for Practitioners

Offered by Chris Adamson through TDWI Onsite Education

Module One
Fundamentals

- Dimensional Modeling
  - Process measurement
  - Uses beyond database design
- Information Architectures and Dimensional Data
  - Inmon’s CIF Architecture
  - Kimball’s Dimensional Bus
  - Stand-alone Data Marts
- Fundamental Terms and Best Practices
  - Facts and dimensions
  - Surrogate keys, natural keys and slow change processing
  - Fundamental concepts including grain, sparsity and additivity
  - Best practices for fact tables and dimension tables
- Implementations
  - Relational (star and snowflake)
  - Multidimensional (cube)
  - Cubes as primary store
  - Cubes as derived data stores

Module Two
Multiple Stars

- Multiple star solutions
- Designing multiple fact tables
  - Identifying multiple processes
  - Differences in dimensionality
  - The pitfalls of single fact table design
- Using multiple stars
  - How not to query multiple fact tables
  - The concept of drilling across
  - What you need to know about your query and reporting tools
- Conformance and business value
  - High impact business questions span processes
  - The concept of conformance
  - Ensuring subject areas work together
  - Enabling incremental implementation
Module Three
Advanced Fact Table Design

- Transaction schemas
  - Transaction grain
  - Shortcomings of transaction designs
- Periodic Snapshots
  - Snapshot grain and period
  - Semi-additivity, density, and impact on BI
  - Building both transaction and snapshot schemas
  - Snapshots and averages
- Accumulating Snapshots
  - Studying process efficiency
  - Accumulating metrics in a single row
  - Lag analysis
  - Impacts on slow change processing and data integration
  - Building both transaction and accumulating snapshots
- Factless Fact Tables
  - Processes that seem to lack metrics
  - Factless fact tables that track events
  - Pros and cons of adding constant-value fact
  - Factless fact tables that track conditions
  - Comparing conditions to actual events
- Heterogeneous Attributes
  - Attributes that vary based on category
  - The impact of modeling a single set of attributes
  - Core and custom dimensions
  - Core and custom fact tables
  - Alternatives to core and custom solutions

Module Four
Design and Data Integration

- The Data Integration Process
  - Loading a dimension table
  - Loading fact tables
  - Key lookup processing
- Designing to aid key management
  - The impact of slow changes on key lookups
  - Adding attributes to aid lookup process
- Designing to reduce slow change bottlenecks
  - The bottlenecks of slow change processing
  - How dimension checksums aid slow change processing
- Specifying data transformation rules
  - How to capture standard translations and data quality rules
  - Why this must be done at design-time
- Invalid or late reference data
• Receiving valid facts with invalid or missing reference data
  • Adding rows to dimension tables so that facts can be loaded
  • Adjusting facts when reference data arrives

• Adding columns to support QA
  • Housekeeping columns in dimension tables
  • Housekeeping dimensions for fact tables

**Day 3**

**Module Five**
**Advanced Dimension Design**

• Understanding hierarchies
  • Drilling with and without hierarchies
  • Multiple hierarchies in one dimension
  • Impact of hierarchies on BI, ETL and DBA perspectives
  • Why to model hierarchies
  • Snowflake schemas

• Dimension Reuse
  • Modeling roles
  • Querying with roles

• Nulls
  • Problems introduced by NULL values
  • Avoiding NULL dimension attributes
  • Avoiding NULLS with optional relationships
  • NULL facts

**Module Six**
**Advanced Slow Change Processing**

• Time-stamped Dimensions
  • Why type 2 is not good enough
  • Time-stamped dimensions (transaction dimensions)
  • Using with a fact table
  • Hybrid attributes that behave like facts and dimensions

• Mini-dimensions (Types 4 & 5)
  • Large and expanding dimension tables
  • The use of a mini-dimension to stem growth
  • Impact on schema capability
  • Type 4 and Type 5 slow changes
  • Loading the mini-dimension

• Current and Previous (Type 3)
  • Limited access to unchanged value

• Hybrid Responses
  • Tracking both Type 1 and Type 2 responses
  • For dimension Attributes (Type 6)
  • For entire dimension tables (Type 7)
Module Seven
Bridge Table Design

- The Dimension Bridge
  - When a single fact needs to reference more than one dimension row
  - The “flattening” option
  - The bridge table
    - Use of the bridge
    - Avoiding double counting with an allocation factor
    - Avoiding double counting by hiding the bridge (sandbox)

- The attribute bridge
  - A dimension attribute repeats for a single dimension row
  - The attribute bridge
  - Use of the bridge and avoidance of double-counting
  - Bridge table vs. factless fact table

- The hierarchy bridge
  - Recursive relationships (ragged hierarchies, unbalanced hierarchies, variable depth hierarchies and instance hierarchies)
  - Flattening and backfilling
  - The hierarchy bridge and its structure
  - Using the bridge to “roll up” a recursive hierarchy
  - Using the bridge to “roll down” the hierarchy
  - ETL implications of hierarchy change and slow changes

Module Eight
Scaling Dimensional Designs

- Scaling and scope
  - Conformance across subject areas
  - Conformed dimensions that are not identical
  - Conformed rollups and overlapping dimensions
  - An advanced conformance matrix

- Derived schemas and performance
  - Merged fact tables
  - Pivoted fact tables
  - Set operations on fact tables
  - Sliced or partitioned fact tables

- Aggregate schemas and performance
  - N-way aggregate design
  - Aggregate portfolio and impact on throughput
  - Aggregates and aggregate navigation

Module Nine
Design Tasks and Deliverables

- Design and data architecture
  - Dimensional design as architecture task
• Design activities
  o Key design activities
  o Resources required for design success
  o Estimating duration of design activities
  o Requirements, top level design, detailed design and roadmap

• Templates for Capturing Business Requirements
  o Subject area definition
  o Metric groups and metric definition
  o Business hierarchies
  o Conformance

• Templates for Capturing Design Requirements
  o Key elements of dimension table design
  o Key elements of fact table design
  o Documenting what is not visible
  o The conformance matrix
  o The design review

• Detailed design
  o Attribute level design of fact and dimension tables
  o Technical design elements (DBA perspective)
  o Business definitions (business perspective)
  o Source to target mappings (integration perspective)

• Roadmap
  o Project definitions
  o Managing scope of implementations

• Implementation approaches
  o Time-boxed projects with iteration
  o Agile projects (with and without up-front discovery)
  o Typical team member roles for implementation projects

Appendix A
Exercises

• Ten modeling exercises designed to reinforce concepts of each module
  o Two are completed during the class (one each day)
  o Students can work through remaining exercises on their own

• Exercise sessions
  o Students are divided into groups of 4-5 to complete the exercises
  o Flip chart or white board must be provided for each group
  o Exercise is discussed by class as a whole upon completion

Appendix B
Exercise Solutions

• Solution pages for each exercise
• Example models and discussion points