Course Agenda, Day 1 of 2

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

Module Four
Design and Data Integration

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

Course Agenda, Day 2 of 2

Module Five
Advanced Dimension Design

• Understanding hierarchies
  o Drilling with and without hierarchies
  o Multiple hierarchies in one dimension
  o Impact of hierarchies on BI, ETL and DBA perspectives
  o Why to model hierarchies
  o Snowflake schemas
• Dimension Reuse
  o Modeling roles
  o Querying with roles
• Nulls
  o Problems introduced by NULL values
  o Avoiding NULL dimension attributes
  o Avoiding NULLs with optional relationships
  o NULL facts

Module Six
Advanced Slow Change Processing

• Time-stamped Dimensions
  o Why type 2 is not good enough
  o Time-stamped dimensions (transaction dimensions)
  o Using with a fact table
  o Hybrid attributes that behave like facts and dimensions
• Mini-dimensions (Types 4 & 5)
  o Large and expanding dimension tables
  o The use of a mini-dimension to stem growth
  o Impact on schema capability
  o Type 4 and Type 5 slow changes
  o Loading the mini-dimension
• Current and Previous (Type 3)
  o Limited access to unchanged value
• Hybrid Responses
  o Tracking both Type 1 and Type 2 responses
  o For dimension Attributes (Type 6)
  o For entire dimension tables (Type 7)

Module Seven
Bridge Table Design

• The dimension bridge
  o When a single fact needs to reference more than one dimension row
o The “flattening” option
  o The bridge table
  o Use of the bridge
  o Avoiding double counting with an allocation factor
  o Avoiding double counting by hiding the bridge (sandbox)
• The attribute bridge
  o A dimension attribute repeats for a single dimension row
  o The attribute bridge
  o Use of the bridge and avoidance of double-counting
  o Bridge table vs. factless fact table
• The hierarchy bridge
  o Recursive relationships (ragged hierarchies, unbalanced hierarchies, variable depth hierarchies and instance hierarchies)
  o Flattening and backfilling
  o The hierarchy bridge and its structure
  o Using the bridge to “roll up” a recursive hierarchy
  o Using the bridge to “roll down” the hierarchy
  o ETL implications of hierarchy change and slow changes

Module Eight
Scaling Dimensional Designs

• Scaling and scope
  o Conformance across subject areas
  o Conformed dimensions that are not identical
  o Conformed rollups and overlapping dimensions
  o An advanced conformance matrix
• Derived schemas and performance
  o Merged fact tables
  o Pivoted fact tables
  o Set operations on fact tables
  o Sliced or partitioned fact tables
• Aggregate schemas and performance
  o N-way aggregate design
  o Aggregate portfolio and impact on throughput
  o Aggregates and aggregate navigation

Module Nine
Design Tasks and Deliverables

• Design and data architecture
  o Dimensional design as architecture task
  o Architecture and incremental implementation
  o Implementation without up-front architecture activity
• Design activities
  o Key design activities
  o Resources required for design success
• Estimating duration of design activities
  o Requirements, top level design, detailed design and roadmap
• Documenting requirements (Conceptual design)
  o Subject area definition
  o Metric groups and metric definition
  o Business hierarchies
  o Conformance
• Top level design
  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

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