Introduction to Data Warehousing

Course Summary

Description

This course provides students with the skills necessary to begin to design a successful data warehouse using multi-dimensional data modeling techniques. It is based on the following Ralph Kimball book: The Data Warehouse Lifecycle Toolkit, Second Edition, Wiley, ISBN: 0470149779, published on January 10, 2008. The students each get this book plus a set of printed PowerPoint slides. The course discusses industry-wide best practices concerning the design and construction of an Enterprise Data Warehouse featuring Dimensional Modeling techniques and Star Schemas. The class uses the free Computer Associates ERwin Community Edition data modeling software. If the customer has the licenses or would like to use the Try It versions--the class could use Embarcadero ER Studio, Power Designer, MS Access, MS SQL Server, Oracle, DB2 or InfoSphere Data Architect data modeling/table design software. The customer should talk with the instructor before the course concerning which data modeling software to use.

Objectives
At the end of this course, students will be able to:

- Understand and apply the concepts, processes and principles of data warehousing
- Identify the components of a data warehouse architecture
- Be familiar with data warehouse terminology
- Identify success and risk factors of data warehousing
- Place deliverables within the context of a comprehensive data warehousing process
- Take practical steps to begin a success data warehousing initiative

Topics

- Dimensional Modeling Primer
- Retail Sales
- Inventory
- Building The Data Warehouse

Audience

This course is targeted at technical staff, team leaders, project managers, and users who need to understand the basics of how to design a data warehouse using multi-dimensional data modeling techniques

Prerequisites

Students should have at least some experience with any relational database management system.

Duration

Two days

Introduction to Data Warehousing

Course Outline

I. Data Warehousing Primer
- A. Operational Systems (OLTP)
- B. Analytical Processing (OLAP)
- C. Data Warehousing Requirements
- D. Data Warehousing Team Responsibilities
- E. Data Warehousing Components
- F. Operational Source Systems
- G. Data Staging Area
- H. Extraction
- I. Transformation
- J. Cleaning
- K. Conforming
- L. Loading
- M. Data Presentation Area
- N. Data Access Tooling
- O. Data Warehouse Terminology
- P. Star Schema
- Q. On-Line Analytical Processing (OLAP)
- R. Cubes
- S. What Is Metadata?
- T. Staging Meta Data
- U. DBMS Meta Data
- V. Data Access Tooling Meta Data
- W. What Is a Fact?
- X. What Is a Dimension?
- Y. Dimensional Modeling Myths
- Z. Avoiding Common Pitfalls

II. Retail Sales
- A. Retail Sales Case Study
- B. The Kimball 4 Step Design Process
- C. Select the Business Process
- D. Declare the Granularity of the Fact Table
- E. Choose the Dimensions
- F. Identify the Facts
- G. Non-Additive Facts
- H. Date Dimension
- I. Product Dimension
- J. Sales Amount and Quantity By Department Report
- K. What Is a Drill Down?
- L. What Is a Roll Up?
- M. Location (Geographical or Store) Dimension
- N. Promotion Dimension
- O. What Is a Factless Fact Table?
- P. What Is a Degenerate Dimension?
- Q. Star Schema Extensibility
- R. What Is Snow Flaking?
- S. Too Few or Too Many Dimensions
- T. Surrogate Keys Versus Natural Keys

III. Inventory
- A. Inventory Periodic Snapshot Fact Table
- B. Semi-Additive Facts
- C. Enhanced Inventory Facts
- D. Inventory Transaction Fact Table
- E. Inventory Accumulating Snapshot Fact Table
- F. Shared Common Dimensions
- G. The Data Warehouse Bus Matrix
- H. What Are Conformed Dimensions?
- I. What Are Conforming Roll-Up Dimensions?
- J. What Are Conforming Dimension Subsets?
- K. What Are Conformed Facts?

IV. Building The Data Warehouse
- A. Data Warehouse Life Cycle Road Map
- B. DW Project Planning and Management
- C. DW Project Scoping
- D. DW Project Justification
- E. DW Project Business Staffing
- F. DW Project Business or IT Staffing
- G. DW Project IT Staffing
- H. DW Requirements Planning
- I. Collecting DW Requirements
- J. DW Effective Interviewing Techniques
- K. DW Interview Wrap-up
- L. DW Post Interview Documentation
- M. DW Requirements Prioritization and Consensus
- N. DW Business Impact and Feasibility Prioritization of Requirements in Quadrants
- O. DW Technical Architecture in 8 Steps
- P. DW Tool Suite Selection and Installation
- Q. DW Star Schema Physical Design Considerations
- R. DW Star Schema Aggregation Strategies
- S. DW Star Schema Indexing Strategies
- T. DW Data Staging Physical Design Considerations
- U. DW Dimension Table Staging Considerations
- V. DW Master Dimension Cross-Referencing Strategies
- W. DW Fact Table Staging
- X. DW Analytics Specification
- Y. DW Deployment
- Z. DW Maintenance and Growth
- AA. Ten Common DW Design Mistakes to Avoid
- BB. Data Warehousing Labs