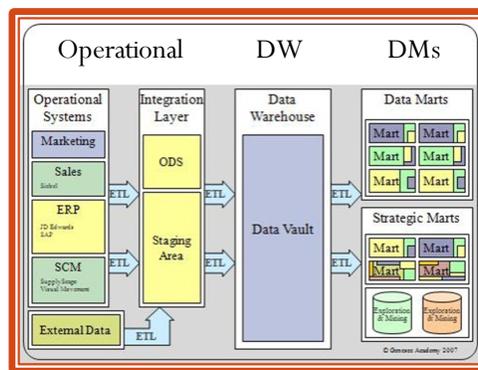


## Course Description

# DATA MODELING IMMERSION

## Modeling: Operational, Data Warehousing & Data Marts



**GENESEE ACADEMY, LLC**

2013

Course Developed by: Hans Hultgren

# DATA MODELING IMMERSION

---

## Modeling: Operational, Data Warehousing & Data Marts

### Overview

Data Modeling is the process of database design. As with most design processes, data modeling is both an art and a science. Art: in that it is a creative process whereby we analyze, design and model specific solutions based on unique requirements. Science: in that we apply modeling approaches, methodologies and specific modeling patterns based on the constraints, variables and performance characteristics of the database we are creating.

Our *operational systems* (core business applications), our *data warehouse*, and our *data marts* each have very different constraints, variables and performance characteristics.

This course teaches data modeling and covers the main data modeling techniques related to each of these three major architectural layers. In this class you will learn data modeling using the normalized (3NF) modeling approach for operational systems, the ensemble data vault modeling approach for the data warehouse, and the dimensional (Star Schema) modeling approach for data marts.

This is a two (2) day course delivered in the classroom.

## Course Description

This course covers the core principles of data modeling through lectures and hands-on labs and exercises. Providing a solid overview of current techniques for modeling operational systems, data warehouses, and data marts.

In covering these areas, the course considers data modeling and design using normalized data modeling **3<sup>rd</sup> Normal Form** for Operational Systems, Ensemble **Data Vault** modeling for the Data Warehouse, and **Dimensional** modeling (Star Schema) for Data Marts.

**3NF Modeling.** The course will start with the core fundamentals of database design including identifying the core entities and relationships, creating logical model designs, defining attributes and key structures, and developing entity relationship diagrams (ERDs).

Within the scope of this modeling approach, the lessons will cover business rules, normalization, validation rules, reference tables, key constraints, identifying and non-identifying relationships, recursive relationships, redundancy, subtypes and supertypes, and relationship cardinality. Students will become comfortable with the core rules and best practices approach to 3NF modeling.

**DataVault Modeling.** The training will continue with the fundamentals of data warehouse modeling for the enterprise data warehouse (EDW). Within this section the core concepts of data warehousing are presented including the focus on integrated, non-volatile, time-variant and subject oriented data. The Data Vault modeling approach is presented including business keys (Hubs), relationships (Links) and context/history (Satellites).

The course continues with core business concepts, ensembles, unified decomposition, concept constellations, and natural business relationships. The course considers how to model your enterprise data warehouse, modeling techniques for agility, operational support, auditability, and enterprise data integration. Lessons, exercises and labs are focused on best practices for architecting and modeling your data warehouse for long term success.

**Dimensional Modeling.** Next the training will cover dimensional data modeling based on current best practice interpretations of the Kimball Star Schema dimensional modeling approach. First lessons will begin with the fundamentals of dimensional modeling including the purpose and structure of Facts and Dimensions, denormalization, the concept of Slowly Changing Dimensions (SCDs) and the main Dimension Types (Type 2 and also covering Type 0, 1, 3, 4 and 6), and then the modeling and design of solid dimensions and encouraged forms of Star Schemas.

The course continues with defining and designing Snow Flake models, the encouraged and acceptable practices for deploying these concepts. Lastly, the course will cover physical data model considerations and DW/BI deployment topics.

## Classroom Course Outline

The classroom days are scheduled in a location that includes table space for each student to work, white boards, flip charts, and a presentation projector used during the lectures. The classroom days are engaging and highly interactive with the students. These days include lectures, labs and exercises.

The classroom course runs 09:00 to 17:00 each day. Students arrive for coffee and introductions beginning at 08:30. Class starts promptly at 09:00. The classroom schedule:

### DAY 1

---

08:30	Arrive, Coffee and Introductions
09:00	Introduction to Data Management <ul style="list-style-type: none"><li>• History and Current Environment</li></ul>
09:30	Data Modeling Primer <ul style="list-style-type: none"><li>• Database design principles</li><li>• Data modeling approaches &amp; techniques</li></ul>
10:15	Break
10:30	Beginning with the Operational <ul style="list-style-type: none"><li>• 3<sup>rd</sup> Normal form</li><li>• The Operational ERD</li><li>• <i>Exercises</i></li></ul>
12:00	Lunch
13:00	Modeling Lab: 3NF <ul style="list-style-type: none"><li>• Lab work, presentations and discussions</li></ul>
14:30	Break
14:45	Moving to the Dimensional Model <ul style="list-style-type: none"><li>• The Presentation Layer, DWBI</li><li>• Understanding the Star Schema</li><li>• <i>Exercises</i></li></ul>
16:00	Modeling Lab: Star Schema <ul style="list-style-type: none"><li>• Lab work completed in groups</li></ul>
17:00	End of day 1

## Day 2

---

- 08:30 Arrive, Coffee and Introductions
- 09:00 Star Schema Modeling Lab Presentations
- Lab presentations and discussions
- 09:30 Recap Day One Material
- Review exercises and labs, Q& A
- 09:45 Dimensional Modeling and the DW
- The EDW & the Federated DW
  - Type 2 Dimensions & the Snow Flake
  - Moving to the EDW
- 10:15 Break
- 10:30 Data Modeling for the EDW
- EDW requirements and constraints
  - Optimized techniques for specific criteria
  - Ensemble & Data Vault Modeling
  - Hubs, Links and Satellites
  - *Exercises*
- 12:00 Lunch
- 13:00 Modeling Lab: Ensemble Data Vault
- Lab work
- 14:30 Break
- 14:45 Data Vault Lab Presentations
- Presentations and Discussions
- 15:30 Data Modeling – putting it all together
- Applying approaches in practice
  - Pros and Cons
  - *Workbook exercises*
- 16:30 Wrap up the course together
- Q&A / In Class Assessment
- 17:00 Class Dismissed

## Course Materials

This course includes detailed course materials to help support the learning process. Beyond the classroom the course materials, exercises, cases and lab books are designed to be used as references and guides to support your ongoing agility efforts. Materials for this course include:

Class Presentation Decks	Printed and bound 90 pages.
Workbook	Printed 3 x 10 pages.
Labs	Printed 2 x 8 pages.

## Target Audience

This course is intended for persons working with any form of modeling. This includes both business and technical resources. Those working with data modeling, data architecture, information modeling, master data management, and all business intelligence and data warehousing professionals. The class is for persons who are involved with the entire data architecture from operational to data warehouse to data marts. For this reason this is a great class for people involved in the data warehousing business intelligence (DWBI) program. Stakeholders who work for the business itself are good candidates such as those who understand the business models, those who gather requirements, design solutions, perform information modeling, data modeling and also mart design.

The class is perfect for DWBI management, project managers, all forms of modelers and architects including Information Modelers, Data Modelers, Data Architects, Information Architects, and Business Analysts. Also for Business Intelligence and Data Warehousing managers, Master Data Management (MDM) professionals, designers, project leads and project owners, DW DBAs, Data Mart designers, Integration Experts, ETL Developers, and Functional Area Representatives.

## Logistics and Fees

All students should be enrolled at least one (1) week before the classroom course day. Please plan for this lead time in your schedule when you register for the course.

Plan to arrive to the training location by 08:30 for the classroom course day. For planning purposes the end of class is 17:00 however it is typical that students have questions after the end of class. The instructor is available until 18:00 to discuss questions.

Location specific logistics and course fees for your class can be found here:

[www.GeneseeAcademy.com/course-schedule](http://www.GeneseeAcademy.com/course-schedule)