

Course Description:

This course covers the business of the Data Vault modeling concepts along with the technical details of the Data Vault modeling architecture.

In the first part of the course, we cover: From why it exists, to the needs it serves, to the business impacts it has. We discuss (from a business level) the benefits, issues, and applicability to the business for utilizing the Data Vault Data Modeling techniques. From I.T. agility, to business needs to scalability and flexibility – we discuss the purpose of the Data Vault Modeling Architecture. It is imperative that business units (especially executives) discover, and participate in this section of the Data Vault course in order to discuss the applicability of the architecture in meeting their business needs.

In the next section of this course we cover topics including: Hubs, Links, and Satellites along with loading, and querying architectures. We dive in deep to provide information around the do's and don'ts (best practices) for Data Vaults. The primary purpose of this course is to prepare you for certification in Data Vault modeling. We cover the technical details around Data Vaults including business rules, delta detection, date stamping, staging, point in time tables, and querying (fetching data from the Vault in to Star Schemas or report collections).

Instructor:

Dan Linstedt

Audience Roles:

For the business section: ETL Designer, Data Modeler, Data Architect, Systems Architect, Business Sponsor, Executives, Project Managers, B.I. Architects, Report designers, Database Architects

For the Technical Section: ETL Designer, Data Modeler, Data Architect, Systems Architect, Project Managers, B.I. Architects, Report designers, Database Architects

Course Duration:

This course is three and ½ days in length with the certification exam taking place on the morning of the 4th day. Certification exam is ***not optional***.

Class Delivery:

On-Site, and at Genesee Academy only.

Lab Work:

We provide a lab-book with 6 example models: one of Northwind which we walk through during course lecture. The other 5 labs focus on different parts of modeling a Data Vault. We actively engage the students for 3 labs during class time; the remaining 2 labs can be done off-site and taken home for additional work.

[Training Courses and Lab-Work:](#)

Genesee Academy offers these courses in person only. Additional lab work and consulting is available on-site for review / assessment of your Data Vault efforts. We have software (RapidACE) which assists you in the Data Vault efforts along the way. At the end of this course students may take the certification test which is included in the fee, and is mandatory.

Day 1:

Business:

Module 1: Foundational Concepts

- Segment 1: Architecture Constructs
- Segment 2: Definition of the Data Vault
- Segment 3: Issues and Evolution in Data Modeling
- Segment 4: Scalability, Agility, Pros and Cons
- Segment 5: Operational Data Warehousing
- Segment 6: Business Case Studies
- Segment 7: Common Business Process Modeling Errors
- Segment 8: Classified Information Systems
- Segment 9: Business Process Agility in I.T.

Technical:

Module 2: Foundational Concepts

- Segment 1: Hubs and Lab on Hub Discovery
 - a) Hub Definition
 - b) Business Keys
 - c) Composite Business Keys
 - d) Real-time Arrival of Keys

- Segment 2: Links and Lab on Link Discovery
 - a) Link Definition
 - b) Link to Link
 - c) Links and Role Playing mistakes

Day 2:

- Segment 3: Satellites and Labs on Satellite Discovery
 - a) Satellite Definition
 - b) Type of Information
 - c) Rate Of Change
 - d) Splitting/Consolidating Satellites

Module 3: Key Structures, Elements in Tables, Common Table Structures

Segment 4: Point in Time Tables, Bridge Tables, User Defined and Record Sources

- a) Point in Time Structures
- b) Bridge Tables
- c) Reference Structures
- d) Report Collections
- e) Net Change and Tracking incoming data across loads
- f) Aggregation Points

Segment 5: User Grouping Sets

- a) Excel Sources
- b) Flat File Sources
- c) Defining User Data and processing into the Data Vault

Segment 6: EDW Processing

- a) Business Processing Overview
- b) Tying the business to the technical data models
- c) The business process of defining a Data Vault

Segment 7: Conversion to A Data Vault

- a) Converting from Normalized format to a Data Vault
- b) Converting a Dimension to a Data Vault
- c) Converting from a Report to a Data Vault

Day 3:

Module 4: Loading and Querying processes within a Data Vault

Segment 8: Golden Rules, Loading Constructs, Mathematics behind the Load Cycles

- a) Golden Rule of the Data Vault
- b) Basic Mathematics behind the loading processes
- c) Paradigm Shift in executing Business Logic
- d) Agility of the Loading Processes
- e) Scalability of through Architecture
- f) Hub, Link, Satellite Load Process Flows

Segment 9: Loading Data Delivery Mechanisms

- a) Loading Dimensions
- b) Loading Facts
- c) Loading Error Marts
- d) Loading Report Collections

Segment 10: Scalability of Loading Processes

- a) Issues with the current loading paradigm
- b) Shifting the paradigm for Compliance reasons
- c) Delta Processing and Set Logic
- d) What to do about “bad data”

Segment 11: What Happens when you Break the Data Vault Rules?

- a) Role Playing Links
- b) Mistakes with the Load Dates

Segment 12: Business Constructs

- a) SLA's needed for changes, and sign-off
- b) Default Values versus Error Codes
- c) Zero Rows populated in the Data Vault
- d) Views in the Data Vault

Segment 13: Strategies for Managing the Data Vault

- a) Backup and Restore
- b) Suggested Topic Areas
- c) Suggested Satellites
- d) Items To Note

CERTIFICATION:

We provide a study guide for the exam on day one of this course; we also provide hints throughout the course on what is on the test. Certification is a closed-book, closed-note test mostly essay in nature. The test is written in English, we do not grade on grammar or spelling constructs – only on ideas and concepts. Two hours are provided in the United States for the test, and three hours are provided in non-English speaking countries.

INSTRUCTOR'S BIO IS AVAILABLE ON THE WEB SITE:

<http://www.GeneseeAcademy.com/instructors.htm>