

Online Training in Data Modeling

By:

Sanjay Sharma

Pragya Consulting Ltd.

October 15, 2025

Training in data modeling

Data modeling is very important task in any database design project. It creates a blue print of the database which ensures that database design supports the business requirements. Developing a database with good design in the beginning of a project saves high cost of changing database design once it is developed and in use. The skill of data modeling is essential for data architects, data engineers, data analysts, data scientists, analytic developers and business people who deal with business data. This training program in data modeling is a customized training specially tailored for the participants based on their current level of data modeling knowledge and skill.

Training participants

1. Santosh Biswal - biswals1984@gmail.com
2. Milind Gawande - milind_gawande@ymail.com
3. Ashutosh Mishra - babu.mishra@gmail.com

Training goal

The goal of this training program is to teach participants fundamentals of data modeling and equip them with skill to develop various types of data model in an enterprise.

Training topics

After discussing their learning goals with the participants, the training has been structured with following topics: have been included in the training program.

1. **Introduction**
2. **What is data modeling and why it is needed?**
3. **Data modeling symbols and notations**
4. **Introduction to data modeling tools.**
5. **Relational data modeling fundamentals**
6. **Data modeling for OLTP systems**
7. **Data modeling for ODS system**
8. **Going from OLTP to ODS**
9. **Data modeling for OLAP system**
10. **Going from ODS to OLAP system**

- 11. Data profiling
- 12. Data lineage
- 13. Source-to-Target-Mapping
- 14. Capstone Project assignment and discussion

It is expected that the above topics can be covered in a 15 hours training program. Based on 1 hour of training each day, the training program has been designed to run for 15 days.

Training schedule

Day 1

- 1. **Introduction:** Time : 15 minutes

Introduction to the training program and how the sessions would be structured.

- 2. **What is data modeling:** Time 30 minutes

In early human history, when exchange of goods started, need for recording the transactions emerged. These recordings were first data records of human history. Medium for recording these records were clay tablets, stones, papyrus etc. Later on paper sheets and printed binders were used for recording the transactions. When computer age came then need for modeling these transactions to save them in computer hard discs emerged. In this section we will study how computer storage system started different types of data models

- A. History of data storage.
- B. Computer age
- C. Digital data storage.
- D. Data modeling needed for digital storage.

Q&A: 15 minutes

Day 2

- 3. Data modeling notations : 45 minutes

Data models are compared with city maps or engineering blue prints. All these artifacts use symbols and notations for representing real world things. In this section we will study what symbols and notations are used in developing data models.

- A. Data modeling symbols and notations
- B. Why symbols and notations are needed ?
- C. Value of data model in organizations

Q&A: 15 minutes.

Day 3:

4. Introduction to Data modeling tools : 45 minutes

Data models can not be developed without help of tools. In this section, we will study various types of. Tools available for data modeling.

- A. Pure data modeling tools.
- B. Data modeling tools with forward engineering support.
- C. Data modeling tools with reverse engineering support.

Q&A: 15 minutes

Day 4:

5. Relational data modeling fundamentals: 45 minutes

Despite emergence of various types of databases, relational databases are still most widely used. They are primary data storage for analytics use case. In this section we will study the relations database models and why they are so powerful and widely used.

- A. Recap of relational model.
- B. Domain, Conceptual, Logical, Physical data models

Q&A: 15 minutes

Day 5:

6. Data modeling for Online Transaction Processing System(OLTP) systems: 45 minutes

Normalization and de-normalization are core to data modeling. This concept needs deep understanding to become a good data modeler. In this section we will cover this important concept

- A. What is normalization and de-normalization ?
- B. What is OLTP system ?
- C. Why normalize in OLTP system ?

Q&A: 15 minutes

Day 6:

7. Data modeling for Operational Data Store(ODS) system: 45 minutes

Beside transactional systems, Operational Data Store also called Data Marts are another important type of database in organizations. In this section we will cover this category of databases.

- A. What is Operational Data Store ?
- B. Going from OLTP to ODS

Q&A: 15 minutes

Day 7:

8. Data modeling for OLAP system: 45 Minutes

Online Analytical Processing System (OLAP) system support all enterprise reports, dashboards, KPIs, advanced analytics etc. Fast query processing is the major requirement for this type of system. In this section we will study how do we model this type of system to run fast queries on massive datasets.

- A. What is OLAP system ?
- B. What is dimensions data modeling ?

Q&A: 15 minutes

Day 8:

9. Deep dive in OLAP system: 45 minutes

There are some special types of schema design which are used for specialized OLAP systems. In this section we will study these schema design.

- A. Star and snowflake schema

Q&A: 15 minutes

Day 9:

10. Data Profiling: 45 minutes

Data Profiling is from where the task of a data analyst starts. In this section we will study what is data profiling and how it is done.

- A. What is data profiling ?
- B. Tools for data profiling ?

Q&A: 15 minutes

Day 10:

11. Data Lineage: 45 minutes

Data Lineage helps to understand how data moves through an enterprise and how it is transformed in its journey from source system to target system. It helps catching any error in analytic reports and understand the components of the data.

- A. What is data lineage ?
- B. Tools for extracting data lineage.

Q&A: 15 minutes

Day 11:

12. Source-to-Target-Mapping: 45 minutes

This task is required before any ETL data pipeline is built. Here we will see the tools available for source to target mapping.

- A. Transaction flow
- B. Tools for Source-To-Target Mapping

Q&A: 15 minutes

Day 12:

13. Data modeling full life cycle review: 45 minutes

In this section we will learn how data model projects are initiated in enterprise and what are its deliverables and outcomes.

- A. Data modeling project lifecycle

Q&A: 15 minutes

Day 13:

14. Capstone project assignment and discussion: 60 minutes

After learning all theoretical concepts, we will start a hands-on practice for data modeling. We will select one business area for example stock trading, ride sharing, food delivery etc. and develop a data model for this sector

- A. Selection of Capstone project
- B. How to work on Capstone project

Q&A: 15 minutes

Day 14:

15. Capstone project review: 60 minutes

In this session we will review our data models for capstone project. We will learn what we have done correctly and what we have missed.

Q&A: 15 minutes

Day 15:

16. Training Review : Lesson learned : 60 minutes

In this concluding session, we will review our full training experience. We will review what we did right and what we did wrong in the training and how we can improve the training material.

- A. What we did right?
- B. What we did wrong?

Training materials

All training materials would be put on a central place on web from where it can be downloaded by the participants.

We would use some open source data modeling tools for the training purpose. The instructions to download and install these tools would be provided during introductory session

Training medium

The training would be provided over zoom with support for session recordings. The recorded sessions would be uploaded to a central place for later viewing by the participants.