

## Oracle Database 11g: Data Warehousing Fundamentals (32 hrs.)

In this course, students learn the basic concepts of a data warehouse and study the issues involved in planning, designing, building, populating, and maintaining a successful data warehouse. Students learn to improve performance or manageability in a data warehouse using various Oracle Database features.

Students also learn the basics about Oracle's Database partitioning architecture and identify the benefits of partitioning.

Students review the benefits of parallel operations to reduce response time for data-intensive operations. Students learn about the extract, transform, and load of data phase (ETL) into an Oracle database warehouse. Students learn the basics about the benefits of using Oracle's materialized views to improve the data warehouse performance. Students also learn at a high level how query rewrite can improve a query's performance. Students review OLAP and Data Mining and identify some data warehouse implementations considerations.

Students briefly use some of the available data warehousing tools such as Oracle Warehouse Builder, Analytic Workspace Manager, and Oracle Application Express.

### Learn To:

Define the terminology and explain basic concepts of data warehousing

Identify the technology and some of the tools from Oracle to implement a successful data warehouse

Describe methods and tools for extracting, transforming, and loading data

Identify some of the tools for accessing and analyzing warehouse data

Describe the benefits of partitioning, parallel operations, materialized views, and query rewrite in a data warehouse

Explain the implementation and organizational issues surrounding a data warehouse project

### Audience:

Application Developers

Data Warehouse Administrator

Data Warehouse Analyst

Data Warehouse Developer

Developer

Functional Implementer

Project Manager

Support Engineer

### Course Objectives:

Define the terminology and explain the basic concepts of data warehousing

Describe methods and tools for extracting, transforming, and loading data

Identify some of the tools for accessing and analyzing warehouse data

Identify the technology and some of the tools from Oracle to implement a successful data warehouse

Define the decision support purpose and end goal of a data warehouse

Describe the benefits of partitioning, parallel operations, materialized views, and query rewrite in a data warehouse

Explain the implementation and organizational issues surrounding a data warehouse project

Use materialized views and query rewrite to improve the data warehouse performance

Develop familiarity with some of the technologies required to implement a data warehouse

## Course Topics:

### Introduction

- Course Objectives
- Course Schedule
- Course Pre-requisites and Suggested Pre-requisites
- The sh and dm Sample Schemas and Appendices Used in the Course
- Class Account Information
- SQL Environments and Data Warehousing Tools Used in this Course
- Oracle 11g Data Warehousing and SQL Documentation and Oracle By Examples

- Continuing Your Education: Recommended Follow-Up Classes
- Data Warehousing, Business Intelligence, OLAP, and Data Mining**
- Data Warehouse Definition and Properties
  - Data Warehouses, Business Intelligence, Data Marts, and OLTP
  - Typical Data Warehouse Components
  - Warehouse Development Approaches
  - Extraction, Transformation, and Loading (ETL)
  - The Dimensional Model and Oracle OLAP
  - Oracle Data Mining
- Defining Data Warehouse Concepts and Terminology**
- Data Warehouse Definition and Properties
  - Data Warehouse Versus OLTP
  - Data Warehouses Versus Data Marts
  - Typical Data Warehouse Components
  - Warehouse Development Approaches
  - Data Warehousing Process Components
  - Strategy Phase Deliverables
  - Introducing the Case Study: Roy Independent School District (RISD)
- Business, Logical, Dimensional, and Physical Modeling**
- Data Warehouse Modeling Issues
  - Defining the Business Model
  - Defining the Logical Model
  - Defining the Dimensional Model
  - Defining the Physical Model: Star, Snowflake, and Third Normal Form
  - Fact and Dimension Tables Characteristics
  - Translating Business Dimensions into Dimension Tables
  - Translating Dimensional Model to Physical Model
- Database Sizing, Storage, Performance, and Security Considerations**
- Database Sizing and Estimating and Validating the Database Size
  - Oracle Database Architectural Advantages
  - Data Partitioning
  - Indexing
  - Optimizing Star Queries: Tuning Star Queries
  - Parallelism
  - Security in Data Warehouses
  - Oracle's Strategy for Data Warehouse Security
- The ETL Process: Extracting Data**
- Extraction, Transformation, and Loading (ETL) Process
  - ETL: Tasks, Importance, and Cost
  - Extracting Data and Examining Data Sources
  - Mapping Data
  - Logical and Physical Extraction Methods
  - Extraction Techniques and Maintaining Extraction Metadata
  - Possible ETL Failures and Maintaining ETL Quality
  - Oracle's ETL Tools: Oracle Warehouse Builder, SQL\*Loader, and Data Pump
- The ETL Process: Transforming Data**
- Transformation
  - Remote and Onsite Staging Models
  - Data Anomalies
  - Transformation Routines
  - Transforming Data: Problems and Solutions

- Quality Data: Importance and Benefits
- Transformation Techniques and Tools
- Maintaining Transformation Metadata

#### **The ETL Process: Loading Data**

- Loading Data into the Warehouse
- Transportation Using Flat Files, Distributed Systems, and Transportable Tablespaces
- Data Refresh Models: Extract Processing Environment
- Building the Loading Process
- Data Granularity
- Loading Techniques Provided by Oracle
- Postprocessing of Loaded Data
- Indexing and Sorting Data and Verifying Data Integrity

#### **Refreshing the Warehouse Data**

- Developing a Refresh Strategy for Capturing Changed Data
- User Requirements and Assistance
- Load Window Requirements
- Planning and Scheduling the Load Window
- Capturing Changed Data for Refresh
- Time- and Date-Stamping, Database triggers, and Database Logs
- Applying the Changes to Data
- Final Tasks

#### **Materialized Views**

- Using Summaries to Improve Performance
- Using Materialized Views for Summary Management
- Types of Materialized Views
- Build Modes and Refresh Modes
- Query Rewrite: Overview
- Cost-Based Query Rewrite Process
- Working with Dimensions and Hierarchies

#### **Leaving a Metadata Trail**

- Defining Warehouse Metadata
- Metadata Users and Types
- Examining Metadata: ETL Metadata
- Extraction, Transformation, and Loading Metadata
- Defining Metadata Goals and Intended Usage
- Identifying Target Metadata Users and Choosing Metadata Tools and Techniques
- Integrating Multiple Sets of Metadata
- Managing Changes to Metadata

#### **Data Warehouse Implementation Considerations**

- Project Management
- Requirements Specification or Definition
- Logical, Dimensional, and Physical Data Models
- Data Warehouse Architecture
- ETL, Reporting, and Security Considerations
- Metadata Management
- Testing the Implementation and Post Implementation Change Management
- Some Useful Resources and White Papers