



## Implementing a Data Warehouse with Microsoft SQL Server

Course 20463D    5 Days    Instructor-led, Hands-on

### Course Description

In this five day instructor-led course, you will learn how to implement a data warehouse platform to support a BI solution. Students will learn how to create a data warehouse with Microsoft SQL Server, implement ETL with SQL Server Integration Services, and validate and cleanse data with SQL Server Data Quality Services and SQL Server Master Data Services.

This course is intended for database professionals who need to create and support a data warehousing solution. Primary responsibilities include:

- Implementing a data warehouse.
- Developing SSIS packages for data extraction, transformation, and loading.
- Enforcing data integrity by using Master Data Services.
- Cleansing data by using Data Quality Services.

### Course Objectives

Upon successful completion of this course, students will be able to:

- Describe data warehouse concepts and architecture considerations.
- Select an appropriate hardware platform for a data warehouse.
- Design and implement a data warehouse.
- Implement Data Flow in an SSIS Package.
- Implement Control Flow in an SSIS Package.
- Debug and Troubleshoot SSIS packages.
- Implement an ETL solution that supports incremental data extraction.
- Implement an ETL solution that supports incremental data loading.
- Implement data cleansing by using Microsoft Data Quality Services.
- Implement Master Data Services to enforce data integrity.
- Extend SSIS with custom scripts and components.
- Deploy and Configure SSIS packages.
- Describe how BI solutions can consume data from the data warehouse.

### Prerequisites

This course requires that you meet the following prerequisites:

- At least 2 years' experience of working with relational databases, including:
  - Designing a normalized database.
  - Creating tables and relationships.
  - Querying with Transact-SQL, Course 20461, or equivalent knowledge and skills.
- Some exposure to basic programming constructs (such as looping and branching).

**Contact ISInc for more information at 916.920.1700 or by visiting our website at <http://www.isinc.com>**



- An awareness of key business priorities such as revenue, profitability, and financial accounting is desirable.
- Administering Microsoft SQL Server Databases, Course 20462 or equivalent knowledge and skills

## **Course Materials**

The student kit includes a comprehensive workbook and other required materials for this class.

## **Course Content**

### **Module 1: Introduction to Data Warehousing**

This module provides an introduction to the key components of a data warehousing solution and the high-level considerations you must take into account when you embark on a data warehousing project.

#### **Lessons**

- Overview of Data Warehousing
- Considerations for a Data Warehouse Solution

#### **Lab : Exploring a Data Warehousing Solution**

After completing this module, you will be able to:

- Describe the key elements of a data warehousing solution
- Describe the key considerations for a data warehousing project

### **Module 2: Data Warehouse Hardware Considerations**

This module discusses considerations for selecting hardware and distributing SQL Server facilities across servers.

#### **Lessons**

- Considerations for Data Warehouse Infrastructure
- Planning Warehouse Hardware Reference Architectures and Appliances

#### **Lab : Planning Data Warehouse Infrastructure**

After completing this module, you will be able to:

- Describe key considerations for BI infrastructure.
- Plan data warehouse infrastructure.

### **Module 3: Designing and Implementing a Data Warehouse**

This module describes the key considerations for the logical design of a data warehouse, and then discusses best practices for its physical implementation.

#### **Lessons**

- Logical Design for a Data Warehouse
- Physical design for a data warehouse

#### **Lab : Implementing a Data Warehouse Schema**

**Contact ISInc for more information at 916.920.1700 or by visiting our website at <http://www.isinc.com>**



After completing this module, you will be able to:

- Describe a process for designing a dimensional model for a data warehouse
- Design dimension tables for a data warehouse
- Design fact tables for a data warehouse
- Design and implement effective physical data structures for a data warehouse

#### **Module 4: Creating an ETL Solution with SSIS**

This module discusses considerations for implementing an ETL process, and then focuses on Microsoft SQL Server Integration Services (SSIS) as a platform for building ETL solutions.

##### **Lessons**

- Introduction to ETL with SSIS
- Exploring Data Sources
- Implementing Data Flow

##### **Lab : Implementing Data Flow in an SSIS Package**

After completing this module, you will be able to:

- Describe the key features of SSIS.
- Explore source data for an ETL solution.
- Implement a data flow by using SSIS

#### **Module 5: Implementing Control Flow in an SSIS Package**

This module describes how to implement ETL solutions that combine multiple tasks and workflow logic.

##### **Lessons**

- Introduction to Control Flow
- Creating Dynamic Packages
- Using Containers
- Managing Consistency

##### **Lab : Implementing Control Flow in an SSIS Package**

##### **Lab : Using Transactions and Checkpoints**

- Using Transactions
- Using Checkpoints

After completing this module, you will be able to:

- Implement control flow with tasks and precedence constraints
- Create dynamic packages that include variables and parameters
- Use containers in a package control flow
- Enforce consistency with transactions and checkpoints

#### **Module 6: Debugging and Troubleshooting SSIS Packages**

This module describes how you can debug packages to find the cause of errors that occur during execution. It then discusses the logging functionality built into SSIS that you can use to log events for troubleshooting purposes. Finally, the module describes common approaches for handling errors in control flow and data flow.

**Contact ISInc for more information at 916.920.1700 or by visiting our website at <http://www.isinc.com>**



#### Lessons

- Debugging an SSIS Package
- Logging SSIS Package Events
- Handling Errors in an SSIS Package

#### Lab : Debugging and Troubleshooting an SSIS Package

After completing this module, you will be able to:

- Debug an SSIS package
- Implement logging for an SSIS package
- Handle errors in an SSIS package

### **Module 7: Implementing an Incremental ETL Process**

This module describes the techniques you can use to implement an incremental data warehouse refresh process.

#### Lessons

- Introduction to Incremental ETL
- Extracting Modified Data
- Loading Modified data

#### Lab : Extracting Modified Data

After completing this module, you will be able to:

- Plan data extraction
- Extract modified data

### **Module 8: Enforcing Data Quality**

This module introduces Microsoft SQL Server Data Quality Services (DQS), and describes how you can use it to cleanse and deduplicate data

#### Lessons

- Introduction to Data Quality
- Using Data Quality Services to Cleanse Data
- Using Data Quality Services to Match data

#### Lab : Cleansing Data

#### Lab: De-duplicating Data

After completing this module, you will be able to:

- Describe how Data Quality Services can help you manage data quality
- Use Data Quality Services to cleanse your data
- Use Data Quality Services to match data

### **Module 9: Using Master Data Services**

Master Data Services provides a way for organizations to standardize data and improve the quality, consistency, and reliability of the data that guides key business decisions. This module introduces Master Data Services and explains the benefits of using it.

#### Lessons

**Contact ISInc for more information at 916.920.1700 or by visiting our website at <http://www.isinc.com>**



- Master Data Services Concepts
- Implementing a Master Data Services Model
- Managing Master Data
- Creating a Master Data Hub

**Lab : Implementing Master Data Services**

After completing this module, you will be able to:

- Describe key Master Data Services concepts
- Implement a Master Data Services model
- Use Master Data Services tools to manage master data
- Use Master Data Services tools to create a master data hub

## **Module 10: Extending SQL Server Integration Services**

This module describes the techniques you can use to extend SSIS. The module is not designed to be a comprehensive guide to developing custom SSIS solutions, but to provide an awareness of the fundamental steps required to use custom components and scripts in an ETL process that is based on SSIS.

**Lessons**

- Using Scripts in SSIS
- Using Custom Components in SSIS

**Lab : Using Custom Components and Scripts**

After completing this module, you will be able to:

- Include custom scripts in an SSIS package
- Describe how custom components can be used to extend SSIS

## **Module 11: Deploying and Configuring SSIS Packages**

In this module, students will learn how to deploy packages and their dependencies to a server, and how to manage and monitor the execution of deployed packages.

**Lessons**

- Overview of SSIS Deployment
- Deploying SSIS Projects
- Planning SSIS Package Execution

**Lab : Deploying and Configuring SSIS Packages**

After completing this module, you will be able to:

- Describe considerations for SSIS deployment.
- Deploy SSIS projects.
- Plan SSIS package execution.

## **Module 12: Consuming Data in a Data Warehouse**

This module introduces business intelligence (BI) solutions and describes how you can use a data warehouse as the basis for enterprise and self-service BI.

**Lessons**

- Introduction to Business Intelligence
- Introduction to Reporting

**Contact ISInc for more information at 916.920.1700 or by visiting our website at <http://www.isinc.com>**



- An Introduction to Data Analysis

**Lab : Using Business Intelligence Tools**

After completing this module, you will be able to:

- Describe BI and common BI scenarios
- Describe how a data warehouse can be used in enterprise BI scenarios
- Describe how a data warehouse can be used in self-service BI scenarios