

Implementing ETL Solutions using SQL Server Integration Services (SSIS) 2022

Day 1 – Data Warehousing Fundamentals & Architecture

Module 1: Introduction to Data Warehousing

Topics:

- Data Warehouse Concepts
- OLTP vs OLAP
- Data Warehouse Architecture
- Data Marts and Enterprise Data Warehouse
- ETL vs ELT Concepts
- Star Schema and Snowflake Schema
- Fact and Dimension Tables

Lab:

- Identify OLTP and OLAP scenarios
 - Design Star Schema for sample business case
 - Create Fact and Dimension model
-

Module 2: Planning Data Warehouse Infrastructure

Topics:

- Infrastructure Planning
- Capacity Planning
- Compute and Storage Design
- SQL Server Components for Data Warehouse
- Scalability and High Availability Considerations
- Security and Governance
- Environment Planning

Lab:

- Design Data Warehouse Infrastructure
 - Estimate Storage and Capacity
 - Create Environment Planning Document
-

Module 3: Designing and Implementing a Data Warehouse

Topics:

- Dimensional Modeling

- Fact Table Design and Fact Table Types
- Dimension Table Design
- Slowly Changing Dimensions (SCD)
- Surrogate Keys
- Historical Data Handling
- Warehouse Loading Strategies

Lab:

- Create Warehouse Database
 - Implement SCD Example
 - Design Loading Strategy
-

Day 2 – Performance Optimization

Module 4: Columnstore Indexes

Topics:

- Columnstore Index Architecture
- Clustered Columnstore Index
- Nonclustered Columnstore Index
- Compression Concepts
- Performance Optimization
- Query Improvements in Analytical System
- Best Practices

Lab:

- Create Columnstore Index
 - Compare Performance Before and After
 - Execute Analytical Queries
-

Module 5: Creating an ETL Solution

Topics:

- ETL Lifecycle
- SSIS Architecture
- Creating SSIS Projects
- Control Flow and Data Flow
- Sources and Destinations
- Data Transformation Concepts
- Complex Data Flow Pipeline

- Package Execution

Lab:

- Create First SSIS Package
- Extract and Load Data
- Apply Multiple Transformations

Module 6: Advanced Transformations and ETL Optimization

Topics:

- Lookup Transformation
- Derived Column
- Conditional Split
- Merge and Merge Join
- Aggregate Transformation
- Performance Optimization
- Parallel Processing

Lab:

- Build Complex Transformation Flow
- Optimize ETL Performance

Day 3 – SSIS Package Development & Data Extraction

Module 7: Implementing Control Flow in an SSIS Package

Topics:

- Execute SQL Task
- Data Flow Task
- File System Task
- Script Task
- Sequence Container
- For Loop Container
- Foreach Loop Container
- Variables and Parameters
- Package Expressions

Lab:

- Create Dynamic SSIS Package

- Implement Loop Execution
 - Configure Variables and Expressions
-

Module 8: Debugging and Troubleshooting SSIS Packages

Topics:

- Breakpoints
- Logging
- Event Handlers
- Error Outputs
- Checkpoints
- Performance Monitoring
- Troubleshooting Package Failures

Lab:

- Configure Logging
 - Debug Package Failures
 - Implement Error Handling
-

Module 9: Incremental ETL and Data Extraction Solutions

Topics:

Lesson 1: Introduction to Incremental ETL

- Incremental ETL Concepts
- Full Load vs Incremental Load
- Incremental Loading Strategies
- ETL Design Considerations

Lab A:

- Configure Incremental ETL Process
 - Identify Changed Data for Processing
-

Lesson 2: Extracting Modified Data

- Change Detection Techniques
- Extracting Modified Records
- Change Data Capture (CDC)

- Change Tracking
- Timestamp-Based Extraction

Lab B:

- Extract Modified Data using CDC
 - Validate Incremental Extraction Results
-

Lesson 3: Loading Modified Data

- Incremental Data Loading
- Insert, Update and Delete Handling
- Merge Operations
- Warehouse Synchronization

Lab C:

- Load Incremental Changes into Target
 - Validate Data Consistency
-

Lesson 4: Temporal Tables

- Introduction to Temporal Tables
- System-Versioned Tables
- Historical Data Management
- Querying Historical Data

Lab D:

- Configure Temporal Tables
- Track and Retrieve Historical Changes

Module 10: Data Source Integration and Extraction

Topics:

- Data Extraction Strategies
- SQL Server Source Extraction
- Flat File Extraction
- Excel Integration
- Multiple Source Integration
- Import and Export Operations

Lab:

- Extract Data from Multiple Sources
 - Configure Source Connections
 - Validate Extracted Data
-

Day 4 – Data Quality & Advanced SSIS

Module 11: Enforcing Data Quality

Topics:

- Data Validation
- Data Cleansing
- Data Profiling
- Duplicate Detection
- Lookup Validation
- Data Quality Best Practices

Lab:

- Implement Data Validation Rules
 - Create Cleansing Workflow
 - Generate Error Records
-
-

Module 12: Extending SQL Server Integration Services

Topics:

- Script Component
- Script Task
- Custom Transformations
- Dynamic Package Configuration
- Package Expressions
- Reusable ETL Framework

Lab:

- Develop Custom Script Component
 - Implement Dynamic Configuration
-

Day 5 – Deployment, Configuration & Enterprise Implementation

Module 13: Deploying and Configuring SSIS Packages

Topics:

- Deployment Models
- Package Deployment
- Project Deployment
- SSIS Catalog (SSISDB)
- Environment Variables
- SQL Server Agent Scheduling
- Monitoring Execution
- Package Security

Lab:

- Deploy Package to SSISDB
 - Configure Environment Variables
 - Schedule Package Execution
 - Validate Data Load
 - Monitor Package Execution
 - Verify ETL Results
-