

IDQ – Informatica Data Quality Course Content:

This course can be delivered through online instructor led format, which is comprised of 17 modules.

Topics covered include

- Profiling, Standardization, Data Cleaning using Labeler & Parser, de-duplication and Address Validation
- Matching and Consolidation Techniques
- Reference Table Management and its usage

Course Objectives

On completion of this course Attendees will be able to:

Navigate the Developer Tool and collaborate on projects with Analysts using the Analyst Tool

- Perform Column Profile, Join Profile, Multi object and Mid-Stream Profiling, Mid-Stream data preview, LDO, Scorecard and working with DQ transformations
- Manage Reference Tables in the Developer & Analyst Tool
- Design Rule & Mapplet, Mapping, Workflows and develop of Applications
- Create standardization, cleansing and parsing routines
- Identify duplicate records
- Build mappings used to associate and consolidate matched records
- Exception Process – Bad Records & Duplicate Records
- Validate addresses

Course Agenda

Introduction to Data Quality Management

Unit 1: Working with Informatica Developer 10X

- GUI, Mappings, Mapplets, Transformations, Content Sets, Data Objects, Reference Tables, LDO

Unit 2: Analyst Collaboration

- Creating Profile and Scorecards
- Adding Comments/Tags
- Reviewing information from the Analyst
- Creating/adding to Reference tables
- Creating Profile, Reference Tables

Unit 3: Developer Profiling and Logical Data Objects

- Perform:
 - Column Profiling
 - Join Profiling
 - Mid-stream profiling
- Create a Logical Data Object
- Create Mappings and work with DQ and Core transformations

Unit 4: Labeler and Data Standardization

- Cleanse and transform data using Labeler and Standardization Transformations
- Develop data standardization mapplets and mappings
- Working with Reference tables

Unit 5: Parsing

- Perform parsing using a variety of methods such as:
 - Token Parser
 - Pattern Parser
 - Working with Reference Tables

Unit 6: Field Matching

- Grouping data
- DQ Matching
- Match Cluster Analysis
- Matching Performance Analysis

Unit 7: Identity Matching

- Build Matching mappings using Identity matching
- Identity Populations and Strategies

Unit 8: Automatic Consolidation & Key Generator

- Associate and Consolidate data

Unit 9: Manual Exception and Consolidation Management

- Build and execute Mappings, using the Exception Transformation, to identify bad records and duplicate records.

Unit 10: Task and Workflow Management

- Build and execute workflows to populate Informatica Data Director user inboxes with exception and duplicate records

Unit 11: Informatica Data Director (Informatica Analyst)

- Update exception and duplicate records in IDD

Unit 12: PowerCenter Integration

- Export DQ Mapping to PowerCenter
- Run DQ Mappings/Mapplets in PowerCenter
- Build and execute a workflow in PowerCenter Developer using DQ mapplets.

Unit 13: Running DQ in a Standalone environment

- Schedule DQ mappings to run in DQ Standalone using Windows Task Manager

Unit 14: Object Import/Export to Informatica PowerCenter

- Import Projects using both Basic and Advanced methods
- Export Projects
- Deploying DQ Jobs to Application

Unit 15: Content

- What content is available with IDQ 10.X?
- Content Management Service
- Accelerators
- Core Accelerator

Unit 16: Parameters and Schedule

- How to use Parameters in Data Quality mappings, transformations and reference tables
- Scheduling Profile, Scorecards and Applications.

Unit 17: Address Validation

- Create a Reusable AV Transformation
- AV Transformation Properties, Inputs and Outputs
- Build and execute an Address Validation Mapping
- Reusable AV Mapplet