

# Data Quality 9.x: Developer, Level 1

## Course Description

### Delivery Methods

- Instructor-led

### Course Duration

- Four Days of instructor-led training
- 55% lecture, 45% hands on lab

### Target Audience

- New users of Informatica Developer, including Business and Data Developers.
- Users who wish to re-familiarize themselves with the Developer tool

### Prerequisites

- Windows GUI and theoretical understanding of process development and design

### Pricing

- Contact your Informatica Education Representative for pricing information

### More Information

Courses may be conveniently scheduled around the world.

Please visit our website at <http://www.informatica.com/us/services-and-training/training/> to find the class that is right for you.

For onsite class information contact your local Education Sales Specialist.

## Course Overview

This four day, instructor-led course introduces students to Informatica Data Quality 9.x through lecture and hands-on exercises. The course is designed for Developers who are new to Informatica Data Quality.

Informatica Developer is a thick client workbench tool that developers can use to profile, cleanse, standardize, de duplicate and consolidate data in an enterprise.

The course will introduce attendees to working with the Developer Tool, performing tasks such as creating projects and objects, profiling data and identifying anomalies.

It will cover developing mappings and mapplets for use in processes such as data standardization, address validation, identifying duplicate records and consolidation.

This course will provide users with a high-level overview of the Informatica Data Quality Process and will set the foundation for the processes that follow.

## Objectives

After successfully completing this course, students should be able to:

- Navigate the Developer Tool and collaborate on projects with Analysts.
- Perform Column, Rule, Multi object, Comparative and Mid-Stream Profiling.
- Manage reference tables.
- Design and develop Mapplets.
- Create standardization, cleansing and parsing routines.
- Validate addresses.
- Identify duplicate records using Classic and Identity matching.
- Build mappings used to associate and consolidate matched records.
- Create and execute Workflows to populate IDD user inboxes with Exception and Duplicate record tasks.
- Run data quality mappings in PowerCenter.

Run data quality mappings in DQ Standalone.

## Course Agenda

### Introduction to Data Quality Management

#### 1: Working with Informatica Developer 9

- GUI, Mappings, Mapplets, Transformations, Content Sets, Data Objects

#### 2: Analyst Collaboration

- Reviewing projects created by the Analyst
- Review objects, Profiles, Rules, Scorecards Comments and Tags
- Create/add Reference tables

#### 3: Developer Profiling and Logical Data Objects

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

#### 4: Data Standardization

- Cleanse and transform data using Standardization Transformations
- Develop standardization mappings and mapplets

#### 5: Parsing

- Perform parsing using a variety of methods such as:
  - Token Parser
  - Pattern Parser
  - NER/Probabilistic parsing

#### 6: Address Validation

- Create a Reusable AV Transformation
- Become familiar with the range of Address Validation Inputs and Outputs
- Build and execute an Address Validation Mapping

#### 7: Matching

- Become familiar with the concepts involved in Matching
- Grouping data
- Match Performance Analysis
- DQ Matching
- Match Cluster Analysis

#### 8: Identity Matching

- Be able to differentiate between the need for Identity or Classic matching
- Become familiar with the concepts involved in Identity Matching
- Work with Identity Populations and Strategies

#### 9: Automatic Consolidation

- Use the Consolidation Transformation to automatically Consolidate matched records.
- Be able to choose the most affective consolidation strategy for your data; simple, row based or advanced.

#### 10: Manual Exception and Consolidation Management

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

#### 11: Task and Workflow Management

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

#### 12: Informatica Data Director

- Update exception and duplicate records in IDD.

#### 13: PowerCenter Integration

- Export DQ Mappings to PowerCenter as either Mappings or Mapplets.
- Build and execute a workflow in PowerCenter Developer using DQ mapplets.

#### 14: Running DQ in a Standalone environment

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

#### 15: Object Import/Export

- Import Projects using both Basic and Advanced methods.
- Export Projects

#### 16: Content

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

