Informatica Data Quality Analyst/Developer 9x Combined

Course Description

Delivery Methods
- Public Instructor-led training
- Onsite Instructor-led training

Course Duration
- Five Days of instructor-led training
- 60% lecture, 40% hands on lab

Target Audience
- New users of Informatica PowerCenter 9x, including Business and Data Developers

Prerequisites
- Windows GUI
- A theoretical understanding of process development and design

Pricing
- R 15,000 excluding VAT

Course Overview

This five-day, instructor led course introduces students to Informatica Analyst and Developer through lecture and hands-on exercises. The course is designed primarily for individuals that will perform the dual role of Data Quality Analyst and Developer and who are new to Informatica Data Quality.

Through lecture and hands on labs this course will introduce attendees to working with both Informatica Data Quality applications, Informatica Analyst and Developer.

In the Analyst application participants will be taught to use the Web-based application to perform tasks such as creating projects and objects, analyzing data using Column and Rule Profiling and identifying anomalies. Additionally, this course covers how to build and maintain Reference Tables, Scorecards, as well as learn how to collaborate between Analysts and Developers using Rules, Comments, and Tags.

In the Informatica Developer application the participants fix the data quality problems and anomalies discovered using the Informatica Data Analyst tool by developing cleansing and standardization mappings. Additionally, building mapping for address validation using postal reference data is taught as well as how to use different matching methods (Classic and Identity) to identify duplicate and/or related records. This course covers both automatic consolidation strategies and manual consolidation and data cleansing, using the Data Quality Assistant in the Informatica Analyst. Integration of IDQ mappings with PowerCenter mappings is also covered in the Developer Application.
Objectives

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

- Be able to navigate through the Analyst and Developer Tools
- Cover Metadata and Data access/import
- Perform Column and Rule Profiling
  - Apply prebuilt and custom built Rules
  - Apply Filters, Tags and Comments to objects in the project
- Learn how to manage reference tables (RTM)
- Use Project Collaboration techniques such as shared Tags, Comments, Profiles, reference tables to share information between Analysts and Developers across projects
- Use a variety of profiling methods to profile data, with both Analyst and Developer.
- Perform Data Quality Scorecarding
- Use Developer to develop mappings and mapplets to cleanse and standardize
- Perform Address Validation
- Build Matching Mappings to identify duplicate or related records
- Automatically consolidate duplicate records in a mapping and manual consolidation/cleansing with the Data Quality Assistant
- Integrate IDQ mappings/mapplets with PowerCenter
- Use Data Quality Mappings in an Excel Spreadsheet

Course Agenda

1. Introduction to Informatica Data Quality 9
   - An introduction into the process and theory of Data Quality Management.

2. Data Objects
   - Create projects and understand the difference between shared and unshared projects.
   - Understand the types of Data Objects that can exist in Informatica 9.
   - Be able to import flat files and relational table as physical data objects.
   - Data access and preview

3. Profiling
   - Column and Rule Profiling
     - Interpret Profiling results
     - Apply Filters to profiles and drilldowns
     - Comment and Tag profiling results
     - Pre-built and Custom built Rules.
   - Collaborate with Data Quality Developers

4. Reference Table Management
   - Understand what reference tables are
   - Create Reference Tables using the Reference Table Editor and the import flat files option
     - Create Managed and Unmanaged Reference Tables
     - Edit and update reference tables
     - Auditing of changes

5. Data Quality Scorecarding
   - Build and execute Scorecards using a variety of methods including:
     - Valid values per column
     - Applying Pre Built Rules
     - Using Custom Rules

6. An Overview of Informatica Developer 9
   - Overview of Data Quality and Informatica Developer covering areas such as the GUI, Mappings, Mapplets, Transformations, Content Sets, Data Objects and Tips for working with Developer.
Course Agenda (Continues...)

7. Analyst Collaboration
   - Review Information from the Analyst
     o Profiles/Comments/Tags
     o Create/Add to Reference Tables
   - Shared projects and team collaboration means no loss of information or duplication of work across projects.

8. Developer Profiling
   - Perform Join Analysis Profiling
   - Perform Column Profiling
     o Multi Object Profiling
   - Build and execute a Mapping using core transformations
     o Configure Transformation Properties

9. Data Standardization
   - Develop data standardization mapplets and mappings
   - Cleanse, transform and parse data
     o Data Quality cleansing and parsing transformations

10. Address Validation
    - Configuring a Reusable AV Transformation
      o AV Transformation Properties
      o AV Inputs and Outputs
    - Define a Reusable AV Mapplet
    - Build and address validation mapping to incorporate your mapplet

11. Matching
    - Grouping Data
      o Key Generator Transformation
    - Key Gen Methods and Properties
    - Analyze Detail Report
    - DQ Matching
      o Perform matching using the Match Transformation
    - Match Properties
    - Matching Algorithms
    - Cluster Analysis Report
    - Create and update matching Mapplets
      o Additional Matching Transformations

12. Identity Matching
    - Build matching mappings using Identity Matching
      o Identity Populations and Strategies

13. Automatic Consolidation
    - Automatically associate and consolidate your match data using a variety of association and consolidation transformations and techniques

14. Manual Consolidation Using the Data Quality Assistant
    - Build mappings to create and populate the DQA tables using the new Exception Transformation
    - Perform Manual consolidation and bad record management using the DQA

15. PowerCenter Integration
    - Execute Data Quality mappings/mapplets in PowerCenter

16. Object Import/Export
    - Learn how to export and import projects using both basic and advanced methods

17. DQ for Excel
    - Run Data Quality mappings on Excel spreadsheets

18. Parameters
    - How to use parameters in Data Quality mappings, transformations and reference tables

19. Content
    - What content is available with IDQ9?
    - Content Management Service
      o How to configure AV information
    - What are Accelerators?
      o What comes with the Core Accelerator?

20. Project
    - Profile, standardize and identify duplicates in product data.
    - Use all that has been covered and apply it to product data
Onsite Training

On-site training courses are designed to meet your specific business needs and requirements. Convenient and cost-effective, on-site training events are delivered at your location to reduce impact on your day-to-day business. Onsite training is ideal for companies who desire the flexibility of training at their location. You set the date and time, and we come to you! We provide a blending approach to learning – instructor-led training with hands on exercises and coaching – that encourages team building and collaboration, all in the convenience of your office.

Extremely cost effective for multiple students
- Your team is trained together
- Your team is accessible
- You pick the date

Public Training & onDemand

Courses are conveniently scheduled around the world. Please visit our website at www.infoflow.co.za to find the class that is right for you.

Learn More:

Visit us at http://www.infoflow.co.za/services.html#training or call +27 (0)11 462 9676

About Infoflow:

Infoflow for the past eight years has focused on Sales, Support and Training of the Informatica Data Integration Platform.