



Enterprise Computing Solutions - Education Services

OFERTA FORMATIVA

Detalles de contacto

Avda Europa 21, 28108 Alcobendas

Email: formacion.ecs.es@arrow.com

Phone: +34 91 761 21 51



IBM InfoSphere Advanced DataStage - Advanced Data Processing V9.1

| | | |
|----------------|------------------|----------------|
| CÓDIGO: | DURACIÓN: | Precio: |
| 1M422G | 16 Hours | €300.00 |

Description

This course is designed to introduce you to advanced parallel job data processing techniques in DataStage V9.1. In this course you will develop data techniques for processing different types of complex data resources including relational data, unstructured data (Excel spreadsheets), Hadoop HDFS ("big data") files, and XML data. In addition, you will learn advanced techniques for processing data, including techniques for masking data and techniques for validating data using data rules. Finally, you will learn techniques for updating data in a star schema data warehouse using the DataStage SCD (Slowly Changing Dimensions) stage. Even if you are not working with all of these specific types of data, you will benefit from this course by learning advanced DataStage job design techniques, techniques that go beyond those utilized in the DataStage Essentials course. If you are enrolling in a Self Paced Virtual Classroom or Web Based Training course, before you enroll, please review the **Self-Paced Virtual Classes and Web-Based Training Classes** on our Terms and Conditions page, as well as the system requirements, to ensure that your system meets the minimum requirements for this course. <http://www.ibm.com/training/terms>

Objetivos

- Use Connector stages to read from and write to database tables
- Handle SQL errors in Connector stages
- Use the Unstructured Data stage to extract data from Excel spreadsheets
- Use the Big Data stage to read from and write to Hadoop HDFS files
- Use the Data Masking stage to mask sensitive data processed within a DataStage job
- Use the XML stage to parse, compose, and transform XML data
- Use the Schema Library Manager to import and manage XML schemas
- Use the Data Rules stage to validate fields of data within a DataStage job
- Create custom data rules for validating data
- Design a job that processes a star schema data warehouse with Type 1 and Type 2 slowly changing dimensions
- Use the Surrogate Key Generator stage to generate surrogate keys

Público

This advanced course is for experienced DataStage developers seeking training in more advanced DataStage job techniques and who seek techniques for working with complex types of data resources.

Requisitos Previos

You should have completed DataStage Essentials course **or** equivalent **and** at least one year of experience developing parallel jobs using DataStage.

Programa

- Unit 1: Accessing Databases
- Unit 2: Processing Unstructured Data
- Unit 3: Processing Big Data
- Unit 4: Data Masking
- Unit 5: Processing XML Data
- Unit 6: Using Data Rules

- Unit 7: Updating a Star Schema Database

Fechas Programadas

A petición. Gracias por [contactarnos](#).

Información Adicional

Esta formación también está disponible en modalidad presencial. Por favor contáctenos para más información.