

# Course Outline:

## 1. Accessing Databases

- Connector Stage Overview
- Connector Stage Functionality
- Error Handling in Connector Stages
- Multiple Input Links
- File Connector Stage
- Handling Database Error
- Parallel jobs with Multiple Connector Input Links
- Using the File Connector Stage to Read and Write HDFS Files

## 2. Processing Unstructured Data

- Using the Unstructured Data Stage in DataStage Jobs
- Extract Data from an Excel Spreadsheet
- Specify a Data Range for Data Extraction in an Unstructured Data Stage
- Specify Document Properties for Data Extraction
- Processing Unstructured Data

## 3. Data Masking

- Using the Data Masking stage in DataStage jobs
- Data Masking Techniques
- Data Masking Policies
- Applying Policies for Masquerading Context-Aware Data Types
- Applying Policies for Masquerading Generic Data Types
- Repeatable Replacement
- Using Reference Tables
- Creating Custom Reference Tables
- Data Masking

## 4. Using Data Rules

- Introduction to Data Rules
- Using the Data Rules Editor

- Selecting Data Rules
- Binding Data Rule Variables
- Output Link Constraints
- Adding Statistics and Attributes to the Output Information
- Use the Data Rules Stage to Valid Foreign Key References in Source Data
- Create Custom Data Rules
- Using Data Rules

## 5. Processing XML Data

- Introduction to the Hierarchical Stage
- Hierarchical Stage Assembly Editor
- Use the Schema Library Manager to Import and Manage XML Schemas
- Composing XML Data
- Using the HJoin Step to Create Parent-Child Relationships between Input Lists
- Using the Composer Step
- Writing Hierarchical Data to a Relational Table
- Using the Regroup Step
- Consuming XML Data
- Using the XML Parser Step
- Propagating Columns
- Transforming XML Data
- Using the Aggregate Step
- Using the Sort Step
- Using the Switch Step
- Using the H-Pivot Step
- Importing XML Schemas
- Compose Hierarchical Data
- Consume Hierarchical Data
- Transform Hierarchical Data

## 6. Updating a Star Schema Database

- Surrogate Keys

- Design a Job that Creates and Updates a Surrogate Key Source Key File from a Dimension Table
- Slowly Changing Dimensions (SCD) Stage
- Star Schema Databases
- SCD Stage Fast Path Pages
- Specifying Purpose Codes
- Dimension Update specification
- Design a Job that Processes a Star Schema Database with Type 1 and Type 2 Slowly Changing Dimensions
- Build a Parallel Job that Updates a Star Schema Database with Two Dimensions