

Datastage 8-1 Course Content: 35hours

Course Outline

Unit -1 : Data Warehouse Fundamentals

- An introduction to Data Warehousing
- Purpose of Data Warehouse
- Data Warehouse Architecture
- Operational Data Store
- OLTP Vs Warehouse Applications
- Data Marts Vs Data Warehouses
- Data Warehouse Life cycle

Unit -2 : Data Modelling

- Introduction to Data Modeling
- Entity Relationship model (E-R model)
- Data Modeling for Data Warehouse
- Dimensions and fact tables
- Star Schema and Snowflake Schemas.

Unit -3 : ETL Design Process

- Introduction to Extraction, Transformation & Loading
- Types of ETL Tools
- Key tools in the market .

Unit -4 : Data Stage Installation

- Data stage Installation Prerequisites to install Data stage Installation process.

Unit - 5 : Introduction to Datastage Version 8.1

- Datastage introduction IBM information Server architecture DataStage components DataStage main functions Client components.

Unit - 6 : Datastage Administrator

- Datastage project Administration - Editing projects and Adding Projects
- Deleting projects Cleansing up project files
- Global variable setting - Environement management
- Auto purging - Rutime Column Propagation(RCP)
- Add checkpoints for sequencer - NLS configuration
- Generated OSH (Orchestra Engine) - System formats like data,
- Timestamp - Project protect - Version details.

Unit - 7 : Datastage Director

- Introduction to Datastage Director - Validating Datastage Jobs
- Executing Datastage Jobs - Job execution status
- Monitoring a Job - Job log view - Job scheduling
- Creating Batches - Scheduling batches.

Unit - 8 : Datastage Designer

- Introduction to Datastage Designer - Importance of Parallelism
- Pipeline Parallelism - Partition Parallelism
- Partitioning and collecting - Partition techniques
- Datastage Repository Palette - Passive and Active stages
- Job design overview - Designer work area
- Annotations - Creating jobs - Importing flat file definitions
- Dataset Overview

Unit - 9 : Working with Parallel Job Stages

- Database Stages
- Oracle - Teradata - ODBC - dynamic RDBMS
- **File Stages**
- Sequential file - Dataset - File set - Lookup file set.
- **Processing Stages**
- Copy - Filter - Funnel - Sort Remove duplicate
- Aggregator - Modify - Compress - Expand - Decode
- Encode - Switch - Pivot stage - Lookup
- Join - Merge - difference between look up
- join and merge - change capture - Change apply
- Compare - Difference - Surrogate key generator
- Transformer.
- **Debug Stages**
- Head - Tail - Peek - Column generator
- Row generator
- **Real Time Stages**
- XML input - XML output, Java packs,
- Webservice packs
- **Local and Shared Containers**

Unit - 10 Advanced Stages in Parallel Jobs (Version 8.1)

- Range Look process - Surrogate key generator stage
- Slowly changing dimension stage - iway stage
- FTP stage - Java plug in - Job performance analysis
- Resource estimation
- Slowly Changing Dimensions implementation
- Performance Tuning

Unit - 11 Job Sequencers

- Arrange job activities in Sequencer
- Triggers in Sequencer - Restability
- Recoverability - Notification activity
- Terminator activity
- Wait for file activity Start Look activity
- Execute Command activity - Nested Condition activity
- Exception handling activity - User Variable activity

- End Loop activity - Adding Checkpoints

Unit - 12 Information Analyzer

- IBM WebSphere Information Analyzer Overview
- Data Profiling process - Column analysis
- Primary key analysis - Foreign key analysis
- Cross-domain analysis - Baseline analysis
- Analysis result publication - Deleting statistics reports
- Baseline analysis reports
- Cross-domain analysis summary statistics reports
- Baseline analysis reports
- Cross-domain analysis reports - Primary key reports
- Foreign key analysis reports.

Unit - 13 WebSphere Quality Stage

- About Data Quality - Datastage quality stages
- Investigate stage
- Standardize stage Match Frequency stage
- Un-duplicate Match stage - Reference Match Stage
- Survive stage.

Unit - 14 IBM Information Server Administration Guide

- IBM WebSphere DataStage Administration
- Opening the IBM Information Server Web console
- Creating users in the console
- Assigning security roles to users and groups
- Managing licenses - Managing active sessions
- Managing logs - Managing schedules
- Backing up and restoring IBM Information Server.