IBM InfoSphere DataStage – Introduction Training

Course Summary

Description
The Introduction to DataStage course is designed to introduce the student to the world of ETL job development leveraging the high-performance nature of the DataStage environment. Our course is designed around a number of concepts and scenarios. This is achieved by the following:

- Learning in a “Open Workshop” environment. Every concept is explained, demonstrated by the instructor, and then the student demonstrates the concept
- A logical organization of discussion topics centered around “Concepts and Scenarios”
- Enhanced lab exercises based on real-world work scenarios
- Group exercises meant

Objectives
At the end of this course, students will be able to:

- Describe the methods for success for developing an ETL job
- Develop basic ETL jobs
- Debug, Test, and Monitor ETL jobs in development and a production capacity
- Understand the basics about building ETL jobs and using the correct tools to accomplish objectives

Topics
- Introduction to InfoSphere DataStage
- Parallel Job Development Methods
- Transforming Data in DataStage
- Debugging ETL jobs

Audience
This class is designed for the beginning developer who wants to how to use IBM InfoSphere DataStage to perform Extraction, Transformation, and Loading techniques via programmatic logic.

Prerequisites
Basic SQL and Programming knowledge

Duration
Four days
IBM InfoSphere DataStage – Introduction Training

Course Outline

I. Introduction to InfoSphere DataStage
   A. Understanding the needs and uses of ETL jobs

II. Parallel Job Development Methods
   A. Reading data from File Sources in DataStage
   B. Reading Data from Database Sources

III. Transforming Data in DataStage
   A. Building Lookups and Joins to Merge disparate data
   B. Building a Sequence Tree

IV. Debugging ETL jobs
   A. Using the DataStage Director to monitor job execution
   B. Group Project: Populating a Warehouse using ETL Jobs