

DATASTAGE COURSE CONTENT

❖ **UNIT - 1: DATA WAREHOUSE FUNDAMENTALS:**

- An Introduction to Data Warehousing
- Purpose of Data warehouse
- Data Warehouse Architecture
- Operational Data Store
- OLTP Vs Warehouse Applications(OLAP)
- Data Marts
- Data marts Vs Data Warehouses
- Data warehouse Life cycle
- Metadata management.

❖ **UNIT – 2: DATA MODELING:**

- Introduction to Data Modeling
- Entity Relationship model(E-R model)
- Data Modeling for Data Warehouse
- Dimensions and fact tables
- Star Schema and Snowflake Schemas
- Coverage Tables
- Different kind of fact tables

❖ **UNIT – 3: ETL DESIGN PROCESS:**

- Introduction to Extraction, Transformation & Loading
- Types of ETL Tools
- What to look for in ETL Tools
- Key tools in the market
- ETL Trends & New Solution Options

❖ **UNIT - 4: DATA STAGE INSTALLATION:**

- Datastage Installation
- Prerequisites to install Datastage
- Installation process.

❖ **UNIT – 5: INTRODUCTION TO DATA STAGE VERSION 8.0.1:**

- Datastage Introduction
- IBM Information Server architecture
- DataStage within the 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
- Global variable setting
- Environment management
- Runtime Column Propagation(RCP)
- Enable Remote Execution of Parallel jobs

- Add checkpoints for sequencer
- NLS configuration

❖ **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.
- Message Handler Management
- Multiple job compile

❖ **UNIT – 8: DATASTAGE DESIGNER:**

- Introduction to Datastage Designer
- Importance of Parallelism
 - Pipeline Parallelism
- Partition Parallelism
- Partitioning and collecting
- Symmetric Multi Processing (SMP)
- Massively Parallel Processing (MPP)
- Partition techniques
- Datastage Repository
- Palette
- Passive and Active stages
- Job design overview
- Designer work area
- Annotations
- Creating jobs, deleting jobs
- Parameter passing
- Compiling jobs
- Batch compiling
- Validating jobs
- Importing flat file definitions
- Managing the Metadata environment
- Dataset management
- Deletion of Dataset
- Parameter sets

❖ **UNIT – 9: WORKING WITH SERVER JOB STAGES:**

Database stages:

- Oracle
- DB2
- Teradata
- ODBC
- Universe
- Dynamic RDBMS.

File stages:

- Sequential file
- Folder-Hash file

Processing stages:

- Transformer
- Aggregator
- Link Collector
- Link Partitioner
- Merge
- Row splitter
- Row Merger
- Sort

❖ **UNIT – 10: IMPLEMENTATING SCD(SLOWLY CHANGING DIMENSION) IN SEREVR AND PARALLEL JOB:**

- 1)SCD 1.0
- 2)SCD 2.0
- 3)SCD 3.0

❖ **UNIT – 11: WORKING WITH PARALLEL JOB STAGES:**

Database stages

- Oracle
- DB2
- Teradata
- ODBC
- Sql Server stage

File stages

- Sequential file – Dataset

Processing stages

- Copy
- Filter
- Funnel
- Sort
- Remove duplicate
- Aggregator
- Modify
- Pivot stage
- Lookup
- Join
- Merge
- difference between look up, join and merge
- Change capture
- Change apply
- Surrogate key generator
- Transformer.

Debug stages

- Head

- Tail
- Peek
- Column generator
- Row generator
- Restructure stages
- column export--column import

❖ **UNIT – 12: ADVANCED STAGES IN PARALLEL JOBS (VERSION 8.0.1):**

- Range Lookup process
- Surrogate key generator stage
- Slowly changing dimension stage
- Job performance analysis
- Slowly Changing Dimensions implementation
- Performance tuning

❖ **UNIT – 13: JOB SEQUENCERS:**

- Arrange job activities in Sequencer
- Triggers in Sequencer
- Restability
- Recoverability
- Notification activity
- Terminator activity\
- Wait for file activity
- Execute Command activity
- Nested Condition activity
- Routine activity
- Exception handling activity
- User Variable activity
- End Loop activity

❖ **UNIT – 14: MIGRATION FROM ONE VERSION OF DATASTAGE TO ANOTHER VERSION:**

- Creating folder structure
- Taking backup
- Datastage command

❖ **UNIT – 15: WEBSPHERE QUALITY STAGE:**

- About Data Quality
- Datastage Quality stages
- Investigate stage
- Standardize stage
- Unduplicated Match stage

❖ **UNIT – 16: DATASTAGE INTERVIEW PREPARATION:**

- Question asked on Server job
- Question asked on parallel
- Question asked on Datawarehouse
- Question asked on ETL