

Data Stage Course Contents with Data Warehousing & ETL Concepts

1 Database Design

	Normalization
	De-Normalization
	Examples

2 Data-warehouse Introduction

	Define Data-warehouse
	Data Warehouse Architecture
	The Top-Down Approach
	The Bottom-up Approach
	Differences between OLAP & OLTP
	OLAP
	OLAP Architecture
	OLAP Servers
	- R OLAP Server
	- M OLAP Server
	- H OLAP Server
	OLAP Operations
	- Roll up Operation
	- Drill down Operation
	- Slice Operation
	- Dice Operation
	- Pivot Operation
	Types of Data Warehouse
	- Enterprise Data Warehouse
	- Data Marts
	- Dependent Data Mart
	- In-dependent Data Mart
	- ODS (Operational Data Store)

3 Data Models

	Conceptual Data Model
	Logical Data Model
	Physical Data Model
	Comparisons of CDM, LDM & PDM
	Pictorial Representation of these models

4	Dimensional Modeling
	Dimension & Facts
	Attributes, Hierarchy
	Star Schema Architecture
	Snow-Flake Schema
	Pictorial Representation of Star & Snow Flake schemas
	Types of Dimensions
	Types of Facts
	Types of Fact Tables
	Slowly Changing Dimensions [SCD] & their types

5	Data Acquisition
	Define Data Acquisition [ETL]
	Steps in ETL
	ETL Architecture
	ETL tools
	Sources v/s Targets

Data Stage Course Contents

1	DataStage Architecture
	Introduction to Data Stage
	Administrator
	Designer
	Director
	DataStage Workflow
2	Platform Architecture
	Pipeline and Parallel Processing
	Partitioning and Collecting algorithms
	Configurations Files
3	Development Guidelines Overview & Components
	Active stage and Passive Stage
	Creating Parallel jobs
	Reading and Writing Data
	Combining Data
	Sorting and Aggregating Data
	Transforming Data
	Working with Relational Data
	Repository Functions
4	Reading and Writing Data
	Sequential file Stage
	Data Set Stage
	Lookup file Set Stage
	file Set Stage
5	Sorting and Aggregating Data
	Copy Stage
	Sort Stage
	Remove Duplicate Stage

6	Data Stage Utilities
	Job properties
	Parameter Sets
	Import and Export jobs
	Shared Container, Local container
	Multiple Instance
	RCP (Run time column Propagation)
	Shared container
	Routines
7	Combining Data
	Join stage
	Lookup Stage
	Merge Stage
	Funnel Stage
	Filter stage
	Switch Stage
8	Transforming Data
	Transformer Stage
	Modify Stage
	Change Data Capture
	Pivot Stage
9	Working with Relational Data
	Database Connection
	Accessing the Data from Table
	DML operation on Tables
10	Debugging Components
	Peek Stage
	Row Generator Stage
	Column Generator Stage
	Annotations Stage

11 Job Control	
	Job Sequence
	Use Sequencer triggers and stages to control the conditions under which jobs run
	Pass information in job parameters from the master controlling job to the controlled jobs
	Define user variables
	Enable restart

Additional	
	Explanation of a sample project for practicing and better understanding of Concepts
	Assignments
	Resume building
	Mock Interviews