

ETL Testing Training Outline

DWH: Data Ware Housing Concepts

- What is Data Warehouse?
- Need of Data Warehouse
- Introduction to OLTP, ETL and OLAP Systems
- Difference between OLTP and OLAP
- Data Warehouse Architecture
- Data Marts
- ODS [Operational Data Store]
- Dimensional Modelling
- Difference between relation and dimensional modelling
- Star Schema and Snowflake Schema
- What is fact table
- What is Dimension table
- Normalization and De-Normalization

ETL Testing

- ETL architecture.
- What is ETL and importance of ETL testing
- How DWH ETL Testing is different from the Application Testing
- SDLC/STLC in the ETL Projects (ex: V Model, Water fall model)
- Challenges in DWH ETL Testing compare to other testing
- Incompatible and duplicate data.
- Loss of data during ETL process.
- Testers have no privileges to execute ETL jobs by their own.
- Volume and complexity of data is very huge.
- Fault in business process and procedures.
- Trouble acquiring and building test data.
- ETL Testing Work flow activities involved
- Analyze and interpret business requirements/ workflows to Create
- estimations

- o Approve requirements and prepare the Test plan for the system testing
- o Prepare the test cases with the help of design documents provided by the
- developer team
- o Execute system testing and integration testing.
- o Best practices to Create quality documentations (Test plans, Test Scripts and
- Test closure summaries)
- o How to detect the bugs in the ETL testing
- o How to report the bugs in the ETL testing
- o How to co-ordinate with developer team for resolving the defects
- Types of ETL Testing
- Data completeness.
- Data quality.
- Performance and scalability.
- Integration testing.
- User-acceptance testing.
- SQL Queries for ETL Testing
- Incremental load testing
- Initial Load / Full load testing
- Different ETL tools available in the market
- Informatica
- Ab Initio
- IBM Data stage
- Power Center Components
- Designer
- Repository Manager
- Workflow Manager
- Workflow Monitor
- Power Center Admin Console
- Informatica Concepts and Overview
- Informatica Architecture.

- Sources
- Working with relational Sources
- Working with Flat Files
- Targets
- Working with Relational Targets
- Working with Flat file Targets
- Transformations - Active and Passive Transformations
- Expression
- Lookup --Different types of lookup Caches
- Sequence Generator
- Filter
- Joiner
- Sorter
- Rank
- Router
- Aggregator
- Source Qualifier
- Update Strategy
- Normalizer
- Union
- Stored Procedure
- Slowly Changing Dimension
- SCD Type1
- SCD Type2 -- Date, Flag and Version
- SCD Type3
- Creating Reusable tasks
- Workflows, Worklets & Sessions
- Tasks
- Session
- Decision task
- Control Task
- Event wait task
- Timer task
- Monitoring workflows and debugging errors

- Indirect Loading
- Constraint based load ordering
- Target Load plan
- Worklet ,Mapplet ,Resuable transformation
- Migration –XML migration and Folder Copy.
- Scheduling Workflow
- Parameter and variables
- XML Source, Target and Transformations
- Performance Tuning
- Pipeline Partition
- Dynamic Partition
- Pushdown optimization
- Preparation of Test Cases
- Executing Test case
- Preparing Sample data
- Data validation in Source and target
- Load and performance testing
- Unit testing Procedures.
- Error handling procedures.