

ETL TESTING

classmate

Date 08/09/2016
Page

→ INTRODUCTION TO ETL (DWH) TESTING

One of the key elements contributing to the success of a DWH soft² is the ability of the test team to plan, design and execute a set of effective test that will help to identify multiple issues related to data inconsistency, data quality failures in the extract, Transform and Load (ETL) process.

The primary focus of testing should be on the ETL process. This includes validating the loading of all required rows, correct execution of all transformations and successful completion of the cleansing operation.

DWH projects largely concentrate on data and moving data from one stream to another stream to all reporting.

→ The sources of this data can be any one of

- Database of an application
- Flat file which have been processed from other applications.

→ OVERVIEW OF ETL OPERATIONS

ETL : Extract, Transform and Load is used to populate the DWH.

1 EXTRACTION : Extract is where data is pulled from source system.

- * Databases
- * Flat Files
- * XML Files
- * Excel File etc.

Unit 4 L13

2 TRANSFORMATION : Data transforms into new formats according to business rules. Apply business logic in this stage.

NOTE : Transformation can be the most complex part of DWH.

3 LOADING : Load is the where data is loaded into the DWH.

→ TEST STRATEGY FOR ETL TESTING

There will be some standard tests for DWH that should be carried as a part of testing for every DWH project.

* These are strategies for testing ETL applications are identified as below

i) Data Completeness :- Ensures that all expected data is loaded.

ii) Data Transformation :- Ensures that all the data transformed correctly according to business rules.

iii) Data Quality :- Ensures that all the ETL applications are correctly rejects, substitutes default values and reports invalid data.

iv Initial Load / Full load testing

v Incremental load testing

vi Presentation layer (BI / reports) Testing : Testing BI reports in DWH testing.

vii Job dependency Testing.