Abstract

Testing ETL (Extract, Transform, and Load) procedures is an important and vital phase during testing Data warehouse (DW); it's almost the most complex phase, because it directly affects the quality of data. It has been proved that automated testing is valuable tool to improve the quality of DW systems while the manual testing process is time consuming and not accurate so automating tests improves Data Quality (DQ) in less time, cost and attaining good data quality. In this paper the author’s propose testing framework to automate testing data quality at the stage of ETL process. Different datasets with different volumes (stared from 10,000 records till 50,000 records) are used to evaluate the effectiveness of the proposed automated ETL testing. The conducted experimental results showed that the proposed testing framework is effective in detecting errors with the different data volumes.


Index Terms

Computer Science Information Sciences

Keywords
Automated ETL Testing, Data Quality, Data Warehouse, Data Quality checking Routines.