TEKsystems Global Services

QA Process for Datawarehouse and Business Intelligence Projects

Presented by: Biradar Siddamallappa

Date of Presentation:
21-09-2020

Datawarehouse and Business Intelligence

We Will briefly understand following points…

- What is database?
- What is Datawarehouse?
- Datawarehouse and Database difference.
- ETL and BI Process.
- QA Process for ETL and BI.
- QA Team Member Role.
- Skill Prerequisites.

Datawarehouse and Business Intelligence

- **Database:-**

It is a storage system where the information is recorded and stored for business purpose, it collection of related and non related data.

- **Datawarehouse:-**

It is a information systems where data is collected and stored, it is used to store historical data which is used for analysis.

Datawarehouse Uses:-

- Data warehouse helps business users to access critical data from multiple sources all in one place.
- It provides consistent information on various cross-functional activities
- Data warehouse allows you to stores a large amount of historical data to analyze different periods and trends to make future predictions.
- Enhances the value of operational business applications and customer relationship management systems
- Separates analytics processing from transactional databases, improving the performance of both systems
- Stakeholders and users may be overestimating the quality of data in the source systems. Data warehouse provides more accurate reports.

Datawarehouse and Business Intelligence

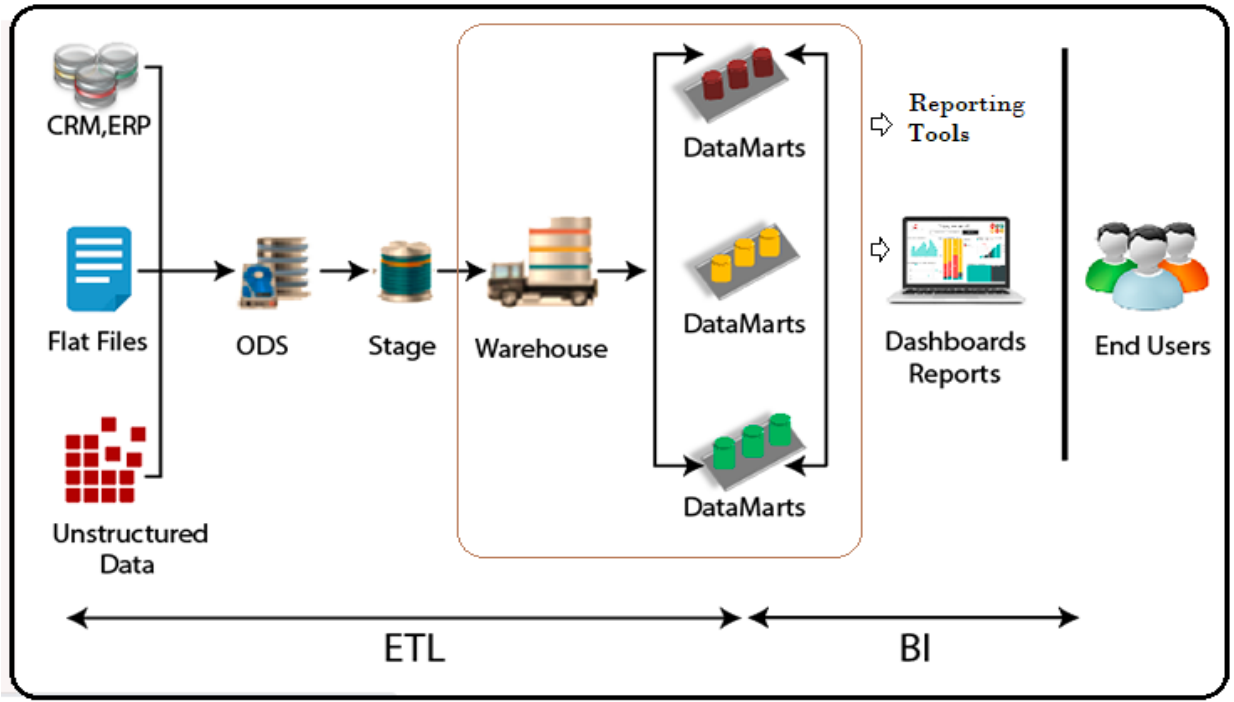
❑ Difference Database and Datawarehouse:-

Database is called Online Transactional Processing (OLTP), Datawarehouse is called Online Analytical Processing (OLAP).

Database	Data Warehouse
An organized collection of data.	A central repository of integrated data from one or more sources.
Usually tied to a single application such as a ticketing system	Usually store data from any number of applications
Primarily insert/write data	Primarily read/retrive data
Data is normalized to allow quick response times.	Data is denormalized for analytical and reporting efficiencies.
Current/Point-in-time data	Historical data
Online Transactional Processing	Online Analytical Processing
Provides a detailed relational view	Provides a summarized multidi-mensional view
For many concurrent transactions	Not for a large amount of concurrent transactions

Datawarehouse and Business Intelligence

- **Datawarehouse:-**

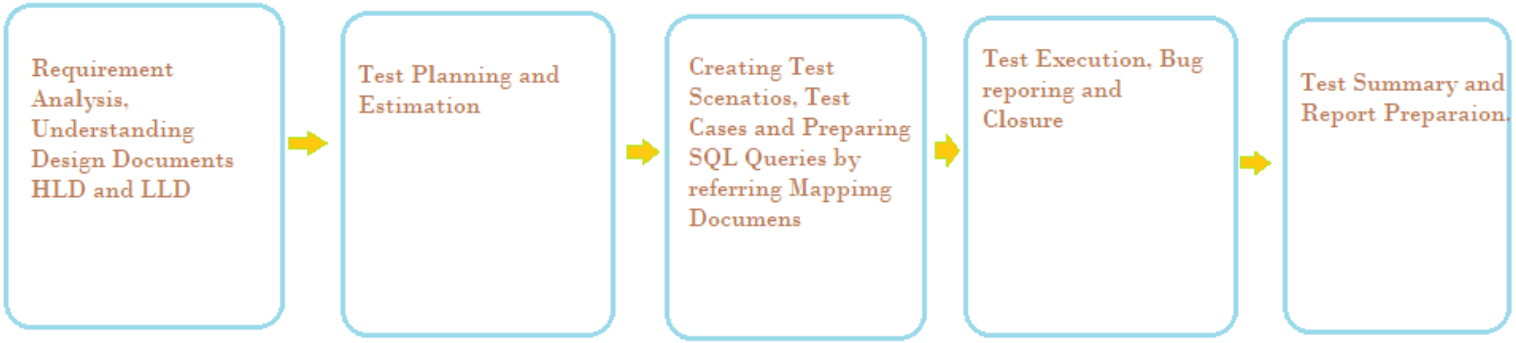


Datawarehouse and Business Intelligence

- **Datawarehouse Testing – Validation Stages:-**

Datawarehouse and Business Intelligence

Datawarehouse Testing Flow



** HLD - High Level Documnets that give Schema amd Tables information
** LLD - Low Level Documnets that gives Columnar level info, Business Logic and Transformation Information

Datawarehouse and Business Intelligence

- **Datawarehouse Testing - Test Scenario and Test Cases**

- Mapping Document Validation – Verify if all information in mapping document is accurate all Target tables and columns have source information.
- Table Structure in Target – Verify all Tables are present in Target and all Target Tables are created as per Mapping requirements including Column names, data types, data length, etc,.
- Constraint Validation – Validate all tables have required constraints in place like Primary Key, not nulls, unique constraints, etc,.
- Data Validation from source to staging – Data Extraction
- Data Validation from staging to target – Data Transformation and Data Loading
- Data Validation from data warehouse to reports.

Datawarehouse and Business Intelligence

Datawarehouse Testing – Data Validation of Target database

- Source to Target Count Testing
- Field to Field Testing
- Duplicate Check Testing
- ETL Transformation and Business Logic Testing
- Dimension Tables Testing – SCDs
- Fact Tables Testing
- Initial Loading and Incremental Loading Testing

Datawarehouse and Business Intelligence

Datawarehouse Testing – BI Report Testing

- Look and Feel of the report, dashboard titles, alignment, data tables, graphs, etc
- Field displays like decimals, rounding, currencies, date and time, etc
- Validate report data against data warehouse tables
- Check all aggregates like count, sum, etc
- Check with different browsers
- Verify any links if present

Datawarehouse and Business Intelligence

Test Data Management

Subset of PRODUCTION data catered for Testing Requirements

Suppose if we are working in an Enhancement projects or maintenance project then we can avail the existing data that is already loaded into warehouse.

In some instance we can't use the data that's pulled directly from Production as it's sensitive like Bank's data, Insurance Data, Healthcare Data and its Customers details.

Mask the data as per regulatory and data privacy requirements

All sensitive data need to be masked or de-identified before it is used for testing purpose.

Create/Generate test Data in case of unavailability

If source system cannot provide the data for Testing Scenarios. Lack of Test data is the major issue in the Data warehouse projects. A tester should analyze the system and need to create the test data in the form of Insert Scripts. The test data should be extracted correctly by the query before its loaded into the Target tables.

Datawarehouse and Business Intelligence

Test Automation

Automation tools:-

There are tools available which perform automation of specific task

Tools examples:-

Query Surge, ETL Validator, Informatica Data Validator, Datagaps, etc,.

Automation using programming and scripting:-

Data validation can be automated by developing validation programmes using python, vb, java, etc

These programmes are built such that they take sql queries and connect to data bases pull the data, compare the data and show the discrepancies, these are cost effective as they can be built on our own.

Datawarehouse and Business Intelligence

QA Member Role

- Understand the data warehouse and reporting needs
- Analyze and understand the design and data flow
- Understand the business logic and transformation logic and identify the data to be reported
- ETL Transformation and Business Logic Testing
- Prepare Test plan and estimates
- Create, design and execute test cases, test plans and test harness.
- Creating required SQL queries for data validation.
- Report, retest and close defects.
- Prepare sign-off, test summary document.

Skills of a ETL/BI Tester:- Analysis, Communication, QA process knowledge, Good in database SQL, ETL and DW knowledge, ETL Tools, Reporting tools, scripting knowledge.

TEKsystems
Own change

THANK YOU