Finance Transformation

Focus on Results

VCGAAP Transformation Project

Best Practices From Quality Assurance Standpoint

April 5th, 2011

Author : Dibyendu Saha

Today's Objectives

- Introduction to Data Warehouse
- Why require Best Practices in testing DWH
- Phases in data warehouse testing
- Data warehouse testing goals
- Types of Data Warehouse Testing
- Generic Challenges

Introduction to Data Warehouse

A data warehouse is a:

- ❑ subject oriented
- ❑ Integrated
- ❑ Non-Volatile
- ❑ Time variant

collection of data in support of management's decision.

Why require Best Practices in testing DWH

- How much confident a company can be to implement it's data warehouse in the market without actually testing it thoroughly ?

- How much testing is enough to implement their data warehouse in the market?

Why require Best Practices in testing DWH

- Would it be possible for a company to remain at the top of the business if the bug is detected at the later stage of testing cycles?

"Undoubtedly we need to have in place the best practices in testing a data warehouse application"

Phases in data warehouse testing

- Business Understanding
- Test Plan Creation
- Test case Creation
- Test data file for predictions
- Test predictions creation
- Test case execution
- Deployment in production environment

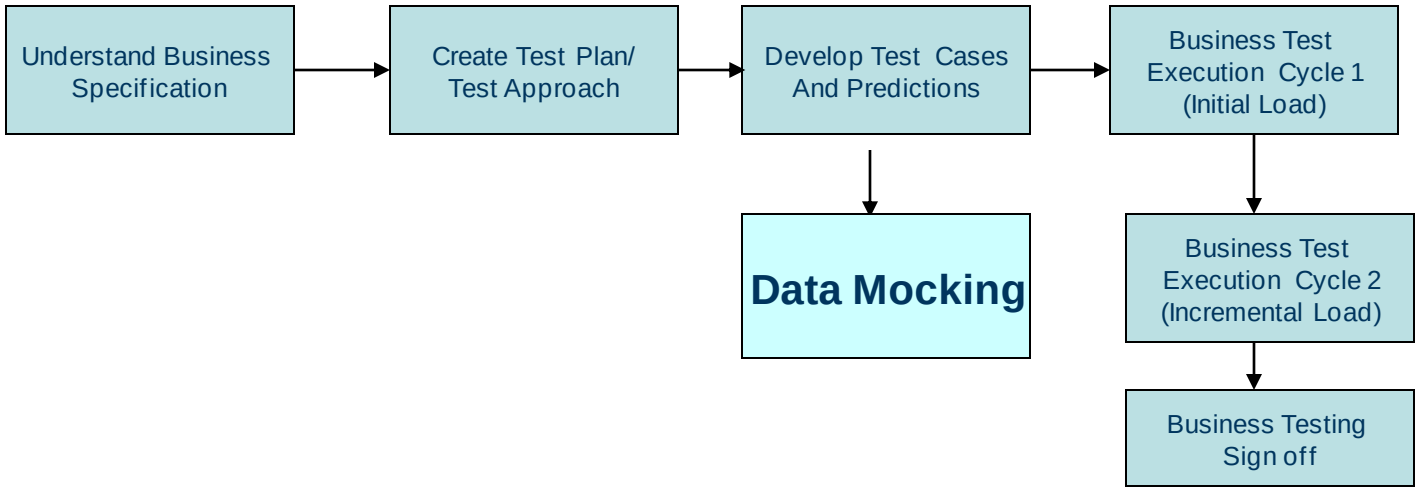
Data Warehouse Testing Goals

- No Data losses
- Correct transformation rules
- Data validation
- Regression Testing
- Oneshot/ retrospective testing
- Prospective testing
- View testing
- Sampling
- Post implementation

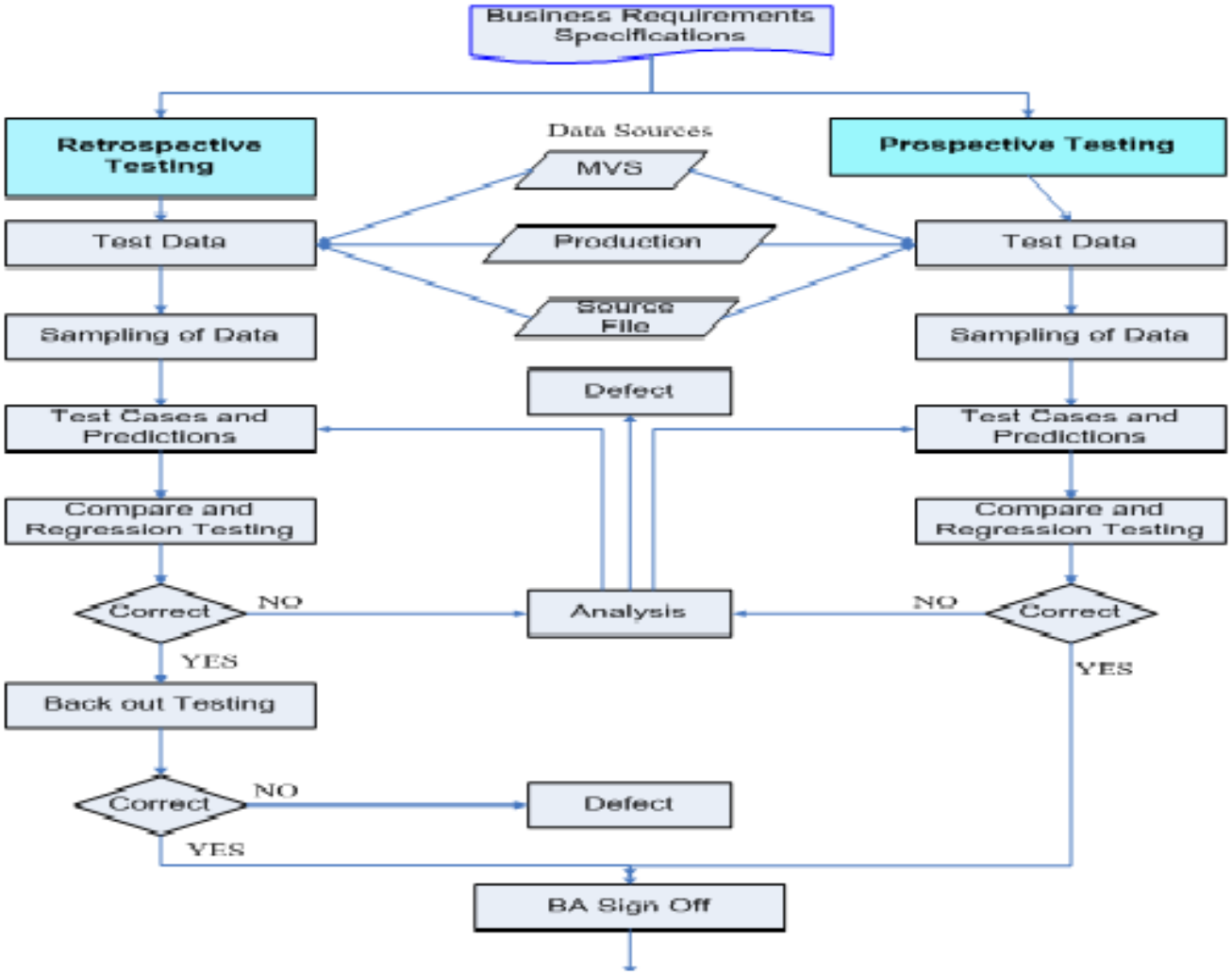
Types Of Data Warehouse Testing

- Business Flow
- One-Shot / Prospective
- Data Quality
- Incremental Load
- Performance
- Version
- Production Checkout

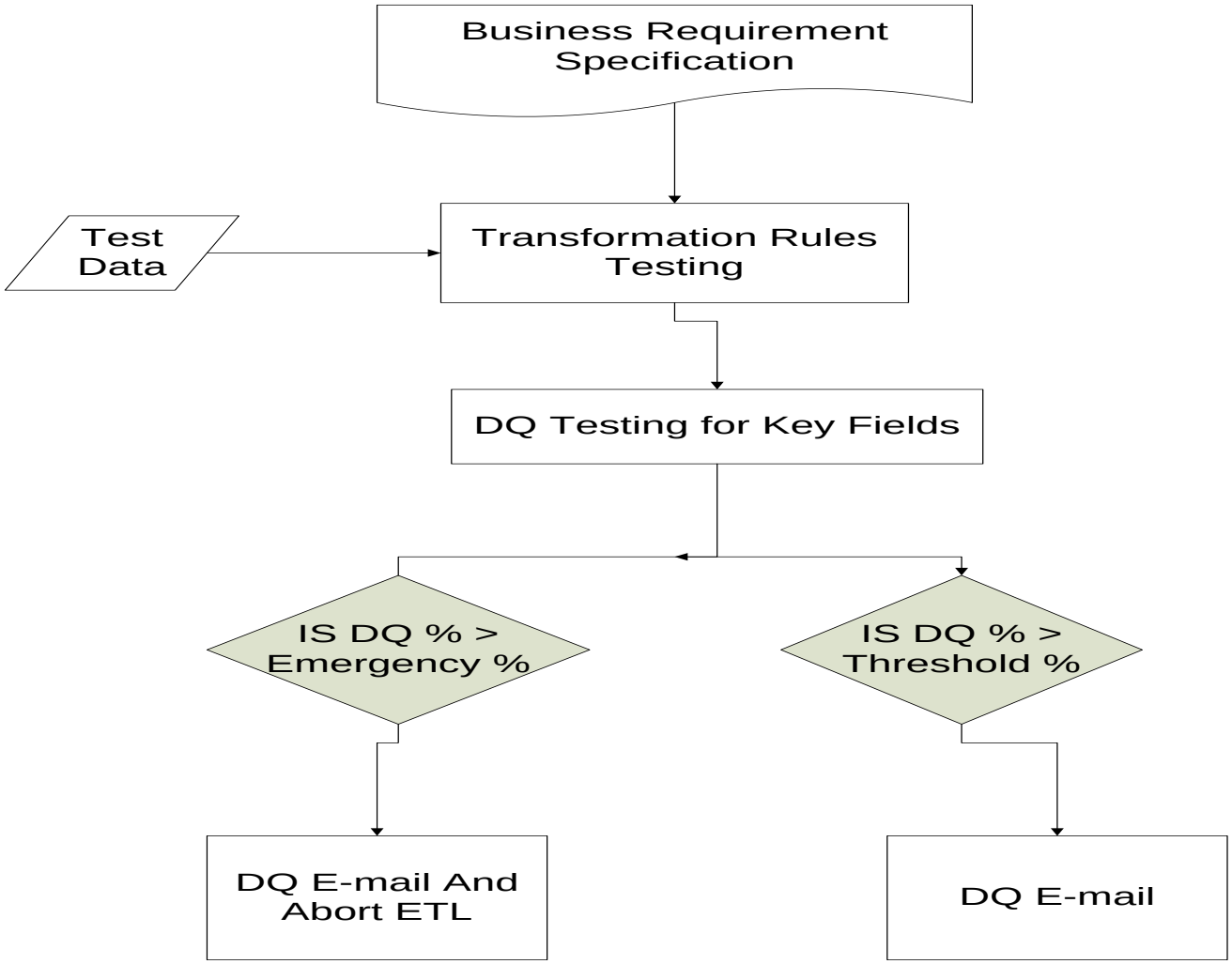
“Business Flow” in Data Warehouse Testing



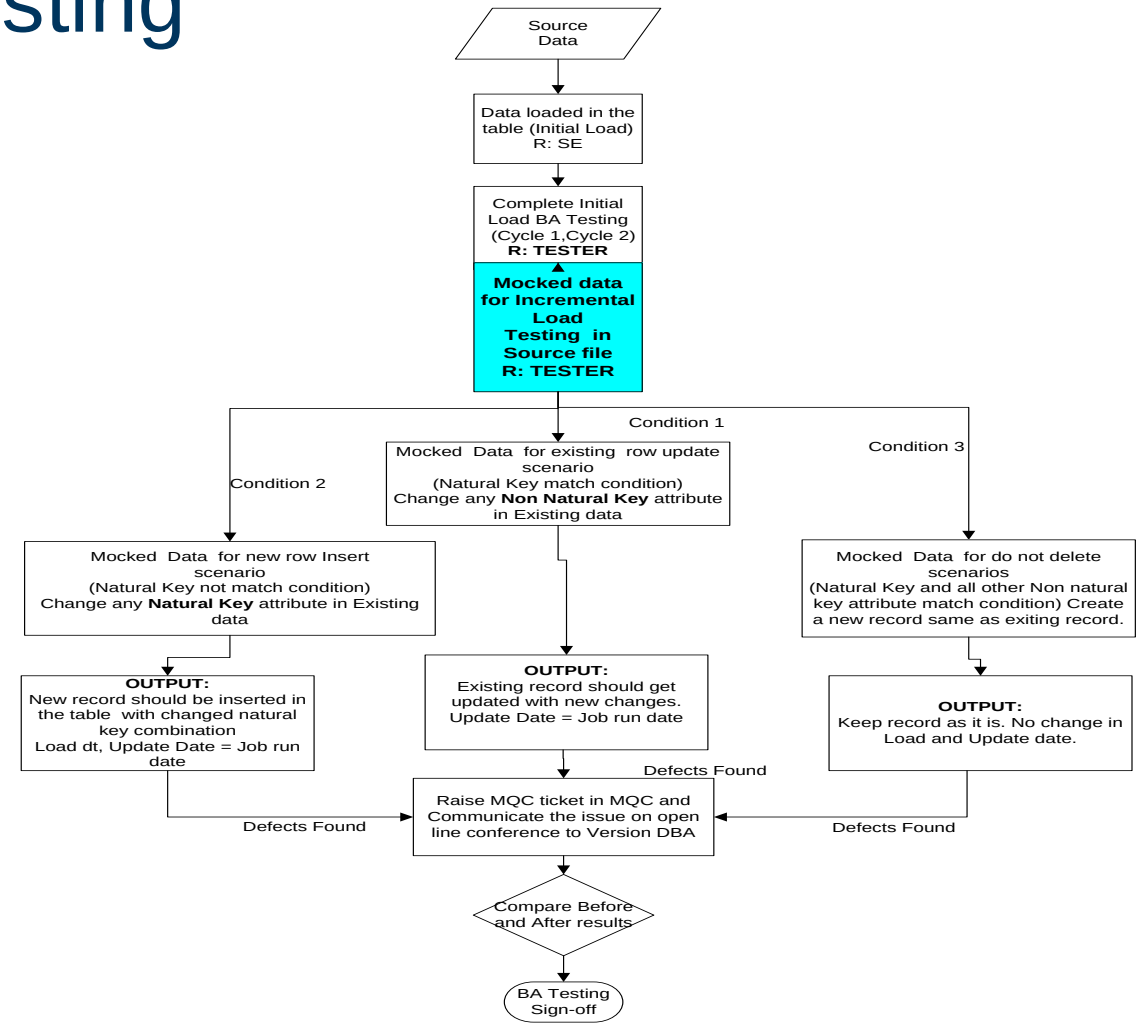
“One-Shot/Prospective” in Data Warehouse Testing



“Data Quality” in Data Warehouse Testing

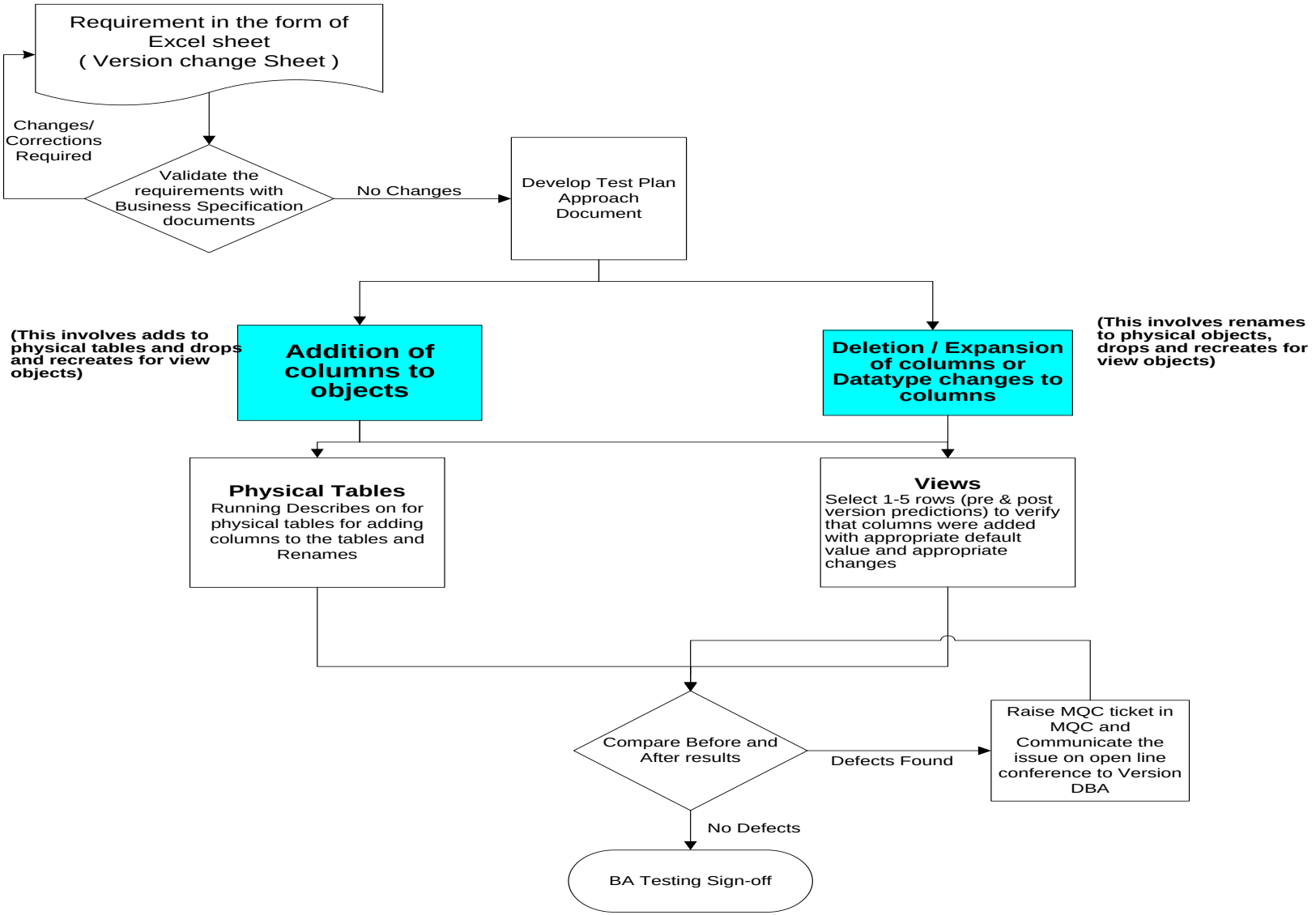


“Incremantal Load” in Data Warehouse Testing

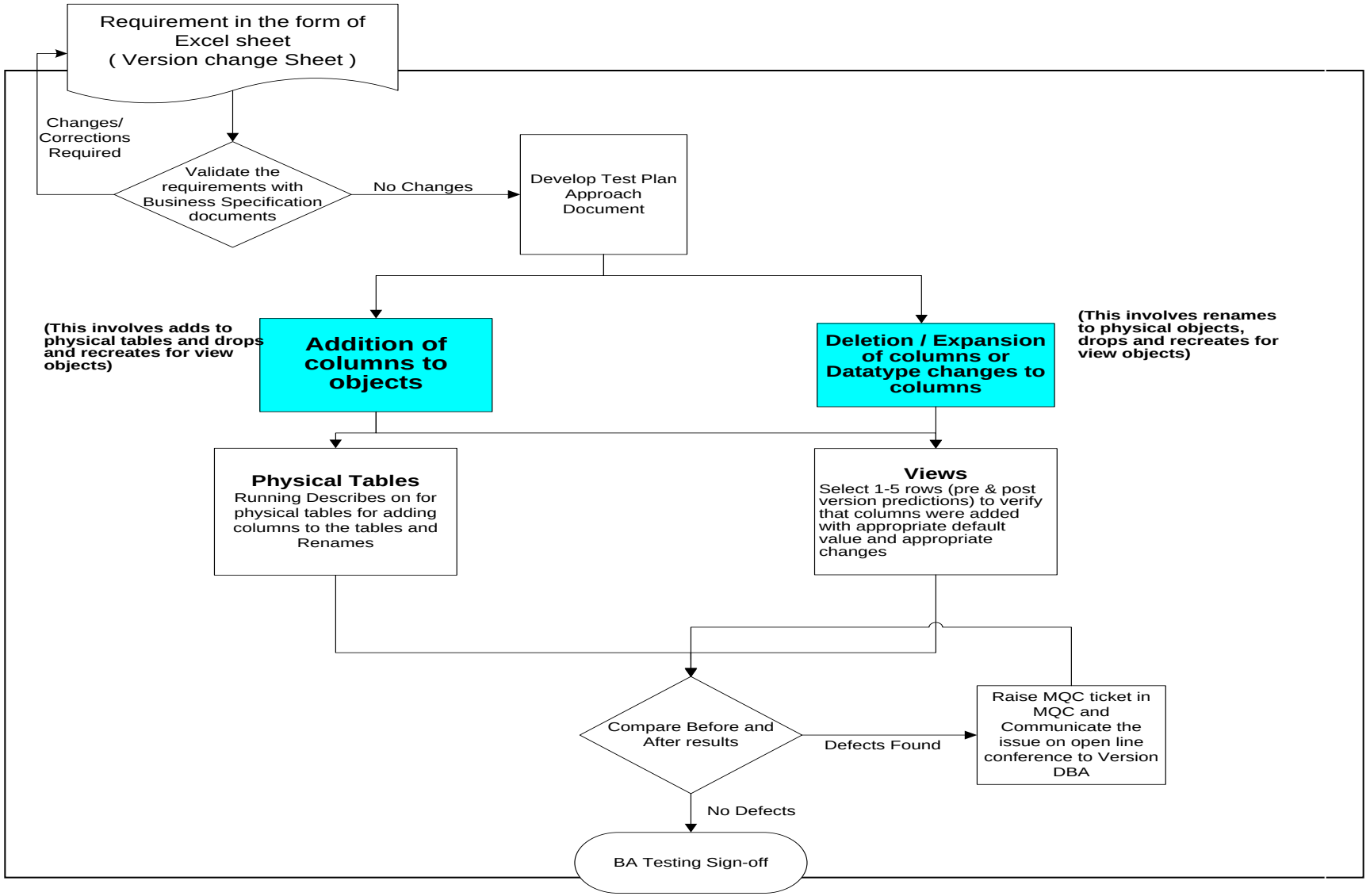


NON Timeline Tables

“Version” in Data Warehouse Testing



“Version” in Data Warehouse Testing



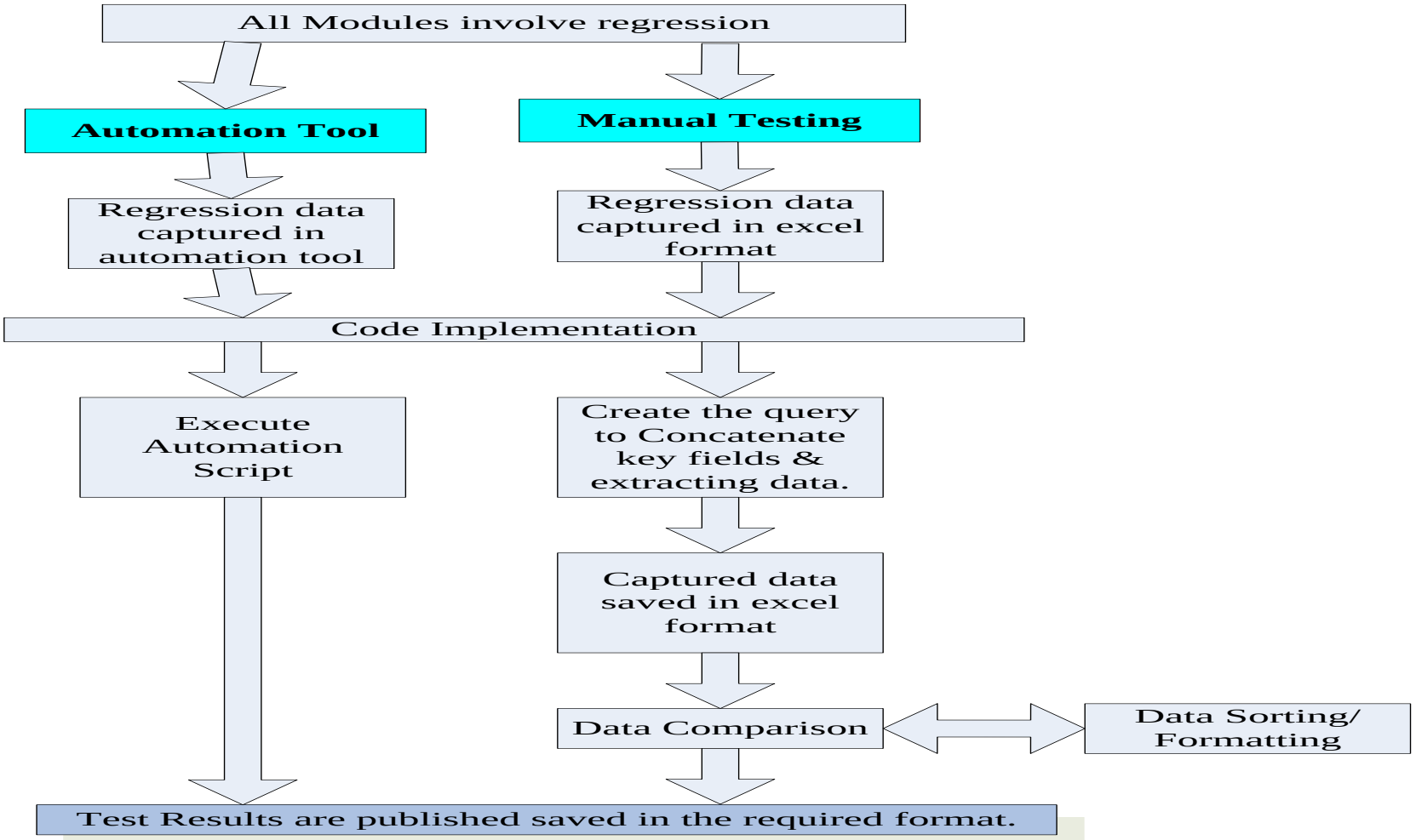
“Production Testing” in Data Warehouse Testing

- Testing only valid values get populated into table (as per BRD’s)
- Testing default values cross threshold % (for given attributes)
- Validates data is not corrupted (environmental factors)
- Verification vs. predictions

Data Warehouse Testing “Challenges”

- Testing only valid values get populated into table (as per BRD’s)
- Testing default values cross threshold % (for given attributes)
- Validates data is not corrupted (environmental factors)
- Verification vs. predictions

Comparison of Flow - Manual Vs Automated



Q & A Session

Any Questions??

Thank You