**Data Quality & User Acceptance Test Specifications**

Version: 1.3

Prepared by: Thando Calana & Thembani Faleni

# Table of Contents

1. Overview

2. Test Categories

2.1 One-to-One Tests

2.2 Type 2 SCD Tests

2.3 Fact Table Tests

2.4 Business Logic Tests

3. Execution Flow

4. Result Logging

5. Acceptance Certificate

# Document Control

## Revision History

|  |  |  |  |
| --- | --- | --- | --- |
| **Revision Number** | **Date** | **Name** | **Revision** |
| V 1.0 | 30/06/25 | T. Calana | Initial version |
| V 1.1 | 03/07/25 | T. Faleni | Update to V2 |
| V 1.2 | 21/07/25 | T. Calana | Addition of tests |
| V 1.3 | 30/07/25 | T. Calana | Addition of overview |
|  |  |  |  |
|  |  |  |  |

## 

## Reviewers

|  |  |  |
| --- | --- | --- |
| **Reviewer** | **Date** | **Feedback** |
|  |  |  |
|  |  |  |
|  |  |  |

## 1. Overview

This document serves as a formal specification for the automated data quality testing framework implemented across the client’s data pipeline. It outlines the testing scope, objectives, data movement strategy, and types of validations conducted to ensure that the delivered data platform meets quality and functional standards.  
  
The goal is to provide the client with clear evidence that the ingestion, transformation, and storage layers function as expected, preserving data accuracy, consistency, and usability across different layers of the architecture.  
  
**Testing Strategy**  
The framework adopts a multi-layer testing approach aligned with the modern data lakehouse model, covering:  
  
- **Staging Layer** (RAW / SRC): Ensures data from the source systems is completely and accurately captured without loss or corruption.  
- **Lake Layer** (Bronze → Silver → Gold): Tracks data refinement and transformation steps; focuses on consistency, type integrity, and content validation.  
- **Warehouse Layer** (DIM / FCT / SCD2): Emphasizes conformance to slowly changing dimension rules, surrogate key uniqueness, historical integrity, and fact table structure.  
  
**Types of Tests**  
- **One-to-One Tests**: Validates row-level and column-level parity between source and target tables across RAW, Landing, and Lake layers.  
- **Type 2 Slowly Changing Dimension (SCD2) Tests**: Ensures that SCD2 tables in the warehouse adhere to dimensional modeling standards.  
- **Fact Table Tests**: Similar to SCD2 tests but applied to fact tables.  
- **Business Logic Tests** (Planned Phase): These will validate client-specific rules and KPIs.

## 1.2 Data Orchestration Architecture

Data flows through a layered architecture:  
- Bronze: Raw ingested data  
- Silver: Cleansed and structured data with minor transformations  
- Gold: Final analytical model with conformed dimensions and standardized measures  
- Warehouse/Data Mart: Subject-specific views (DIMs, FCTs) for reporting

# 2. Test Cases

**2.1 One-to-One tests: Source to Target**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Test\_ID** | **Test Name** | **Description** | **Applicable Layer(s)** | **Frequency** | **Defect (Y/N)** | **Slipstream**  **Signature** | **Customer Signature** | **Comments** |
| 2.1.1 | row\_count\_match | Compares row counts between source and target entities to verify full coverage | Landing, Lake | Daily |  |  |  |  |
| 2.1.2 | column\_presence\_and\_types | Checks column presence and type consistency between landing and lake tables | Lake, Warehouse | Daily |  |  |  |  |
| 2.1.3 | unique\_key | Validates that natural/primary key is unique in both source and target tables | Landing, Lake, Warehouse | Daily |  |  |  |  |
| 2.1.4 | not\_null\_key | Checks that key columns are not null in both source and target tables | Landing, Lake | Daily |  |  |  |  |
| 2.1.5 | hash\_match | Verifies row-level content match between source and target using hash comparison | Landing, Lake | Daily |  |  |  |  |

**2.2 Type 2 Slowly Changing Dimensions**

**2.2.1 Initial Load Tests**

The purpose of these tests is to ensure that when the data is initially sent to the data warehouse, it has the correct initial configurations (i.e. effective dates and expiry dates are set).

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Test\_ID** | **Test Name** | **Description** | **Applicable Layer(s)** | **Frequency** | **Defect (Y/N)** | **Slipstream**  **Signature** | **Customer Signature** | **Comments** |
| 2.2.1.1 | scd\_initial\_active\_state | Confirms first record per natural key is marked active | Warehouse | Once-off |  |  |  |  |
| 2.2.1.2 | scd\_initial\_effective\_date\_set | Ensures earliest record has a ‘valid\_from’ date set | Warehouse | Once-off |  |  |  |  |
| 2.2.1.3 | scd\_initial\_expiry\_is\_null | Confirms earliest version has NULL expiry | Warehouse | Once-off |  |  |  |  |

**2.2.2 Structure Test**

The purpose of this test is to ensure that the structure of the SCD Type 2 in the data warehouse is correct and has the necessary columns.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Test\_ID** | **Test Name** | **Description** | **Applicable Layer(s)** | **Frequency** | **Defect (Y/N)** | **Slipstream**  **Signature** | **Customer Signature** | **Comments** |
| 2.2.2.1 | scd\_columns\_check | Verifies existence of required SCD2 fields | Warehouse (SCD2 only) | Daily |  |  |  |  |

**2.2.3 Data Validation and quality tests**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Test\_ID** | **Test Name** | **Description** | **Applicable Layer(s)** | **Frequency** | **Defect (Y/N)** | **Slipstream**  **Signature** | **Customer Signature** | **Comments** |
| 2.2.3.1 | scd\_unique\_surrogate\_key | Ensures surrogate key is unique | Warehouse (SCD2 only) | Daily |  |  |  |  |
| 2.2.3.2 | scd\_not\_null\_surrogate\_key | Ensures surrogate key field is not null | Warehouse (SCD2 only) | Daily |  |  |  |  |
| 2.2.3.3 | scd\_effective\_before\_expiry | Ensures valid\_from < valid\_to for all SCD2 records | Warehouse (SCD2 only) | Daily |  |  |  |  |
| 2.2.3.4 | scd\_unique\_active\_nkey | Ensures only one active record exists per natural key | Warehouse (SCD2 only) | Daily |  |  |  |  |
| 2.2.3.5 | scd\_active\_null\_expiry | Validates active records have NULL expiry date | Warehouse (SCD2 only) | Daily |  |  |  |  |
| 2.2.3.6 | scd\_inactive\_historical\_records | Ensures inactive records have is\_active = false | Warehouse (SCD2 only) | Daily |  |  |  |  |
| 2.2.3.7 | scd\_no\_scd2\_valid\_from\_gaps | Validates continuity in valid\_from dates | Warehouse (SCD2 only) | Daily |  |  |  |  |

**2.3 Fact Table Tests**

**2.4 Business Logic Tests**

# 3. Execution Flow

Tests are orchestrated using a centralized macro called run\_tests\_controller, which serves as the execution backbone for the entire data quality testing framework.

**Key Components:**

* **Controller Macro (run\_tests\_controller)**:  
  Dynamically invokes tests based on metadata definitions. This modular macro evaluates each test configuration and dispatches execution logic accordingly, promoting scalability and standardization.
* **Metadata Tables**:
  + TEST\_CONFIG: Holds runtime configurations for each test, including frequency (daily, weekly, once-off), severity, and alerting preferences.
  + TEST\_METADATA: Defines each test’s logic, including SQL template, test type (row count, hash check, etc.), and descriptive metadata.
  + ENTITY\_METADATA: Maps source and target entities, identifies key columns, and links tests to schema/table combinations.

**Test Trigger Conditions:**

* Tests are triggered only when IS\_ACTIVE = TRUE in TEST\_CONFIG.
* Filtering can be applied by schema, layer (RAW, Lake, Warehouse), and frequency.
* The controller batches all active and relevant tests into a single execution pipeline and logs outcomes to TEST\_RESULTS.

# 4. Result Logging

Test results are logged into ‘TEST\_RESULTS’, capturing:

- RESULT\_ID

- TEST\_NAME

- TEST\_ID

- TEST\_CONFIG\_ID,  
- SRC\_ENTITY

- TRG\_ENTITY  
- RESULT (PASS / FAIL)  
- RESULT\_DESCRIPTION

- TEST\_DATE  
- HWM\_FROM

- HWM\_TO (where applicable)

## 5. Acceptance Certificate

Slipstream Consulting (Pty) Ltd  
  
CERTIFICATE OF ACCEPTANCE  
  
Client Name: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  
Project Name: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  
Environment: ☐ Development ☐ QA ☐ Production  
Date: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  
  
We hereby certify that the asset(s) specified below have been successfully tested and accepted by the client:  
  
Assets Covered:  
☐ One-to-One Data Quality Tests  
☐ Type 2 SCD Validation  
☐ Fact Table Conformance  
☐ Execution Workflow & Logging  
☐ Business Logic (if applicable)  
  
Test Summary:  
All critical and high-priority tests passed successfully, with no outstanding defects at the time of acceptance.  
  
Client Representative: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_   
Signature: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_   
Date: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  
  
Slipstream Representative: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_   
Signature: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_   
Date: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_