

|  |
| --- |
| Data quality |
| DW/BI Test Result Report |

**Test Result Report**

**Project:** DWH  
**Test Period:** 09.04.2025 – 11.04.2025  
**Author:** Oleg Arslanov

**1. Summary**

During the testing period, the QA team conducted validation of ETL pipelines and the Data Warehouse. Testing focused on verifying the accuracy of data transformations, correctness of aggregations, and consistency of data across the source, landing, DWH, and Data Mart layers.

Out of approximately 20 potential issues, around half have been identified and registered in Jira (11 bugs reported, with 9 confirmed by the mentor). 8 test cases were designed, and one of them led to a bug discovery. The current bug detection rate suggests there is room for improvement in test coverage and test case effectiveness.

**2. Test Team**

| **Name** | **Role** | **Responsibility** |
| --- | --- | --- |
| Oleg Arslanov | Student | Documentation creation, test – cases execution, participation in code-review |
| Oksana Krangach | Data Quality Engineer (Mentor) | Participation in requirements  testing and code review |

**3. Testing Process Description**

Testing was conducted as part of a single educational project. Test cases were created manually and documented in TestRail, while defects were logged in Jira (view board).

In total, 8 test cases were designed and executed. Additionally, ad hoc testing was performed to supplement planned scenarios and increase coverage. All testing activities were conducted manually, with no automation tools used.

As part of a personal initiative and for educational purposes, an extended smoke test was developed to validate the correctness and integrity of the ETL pipeline. The test included row count checks, data quality validations (NULLs, formats, ranges), and basic aggregation logic in the dm\_main\_dashboard data mart layer.

As a result, 11 bugs were identified, of which 9 were confirmed by the mentor. This represents approximately 50% of the potential issues, considering both structured test cases and ad hoc testing activities. The current bug detection rate suggests that there is room for improvement in test coverage and test case effectiveness.Tools used:

* DBeaver Community Version Version 25.0.2.202504061727
* PostgreSQL version: PostgreSQL 11.13
* Jira (bug tracking)
* Test Rail (test cases)
* Via Docker container Version 4.40.0 (187762)

**4. Timetable**

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Date** | **Activity** | **Duration, hours** |
| Oleg Arslanov | 09.04.2025 | Test cases creation | 8 |
| Oleg Arslanov | 11.04.2025 | Pair testing | 8 |
| Oleg Arslanov | 21.04.2025 | Extended Smoke test (manual) | 2 |
| Oleg Arslanov | 21.04.2025 | Defect reporting | 2 |
| Oksana Krangach | ? .04.2025 | Code review | ? |
| Oksana Krangach | ? .05.2015 | Pair testing | ? |

1. **New Defects Statistics**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | **Severity** | | | |
| **Status** | **Quantity** | **Low** | **Medium** | **Major** | **Critical** |
| Submitted | 11 | 2 | 2 | 4 | 3 |
| Fixed | 0 | 0 | 0 | 0 | 0 |
| Verified | 0 | 0 | 0 | 0 | 0 |
| Extended Smoke Test | 1 | 0 | 0 | 0 | 0 |
| Declined | 2 | 2 | 0 | 0 | 0 |

**6. New Defects List**

| **ID (Jira)** | **Title** | **Severity** | **Status** |
| --- | --- | --- | --- |
| SCRUM-1595 | DWH\_Mismatch between sum of all product sales in DWH layer and DM layer | Major | In Progress |
| SCRUM-1609 | DWH\_Mismatch between sum of all product sales in DWH layer and DM layer | Major | In Progress |
| SCRUM-1631 | DWH\_column valid\_to default value not properly set in DWH\_CLIENTS from S1 Source System | Critical | In Progress |
| SCRUM-1781 | DWH\_extra records in lnd\_s1\_clients not found in s1\_clients | Critical | In Progress |
| SCRUM-1782 | DWH\_data mismatch between s1\_sales and lnd\_s1\_sales | Critical | In Progress |
| SCRUM-1797 | DWH\_Missing rows during ingestion from s2\_client\_sales to lnd\_s2\_client\_sales | Medium | In Progress |
| SCRUM-1818 | DWH\_duplicate rows found in lnd\_s2\_locations after ingestion from s2\_locations | Major | In Progress |
| SCRUM-1832 | Foreign Keys in dwh\_sales table are incorrectly set as NULLABLE | Major | In Progress |
| SCRUM-1839 | DWH\_incorrect location\_src\_id values in dwh\_locations table | Medium | In Progress |

**7. Recommendations**

 **Implement automated data monitoring:**

* Set up monitoring at critical ETL stages to detect anomalies, such as missing data, invalid values, or errors during loading. This will help quickly identify and resolve issues.

 **Ensure versioning and data logging:**

* Keeping logs of data changes and model versions helps track what changes were made and when, making it easier to diagnose problems.

 **Optimize ETL performance:**

* ETL processes can take a long time, especially when working with large data volumes. It is recommended to optimize queries and processes to speed up processing, such as using indexes, parallel processing, or optimizing scripts.

 **Regular Data Quality Checks:**

* Regularly perform checks for data quality, including uniqueness, completeness, adherence to data type expectations, and duplication checks, to maintain high data quality.

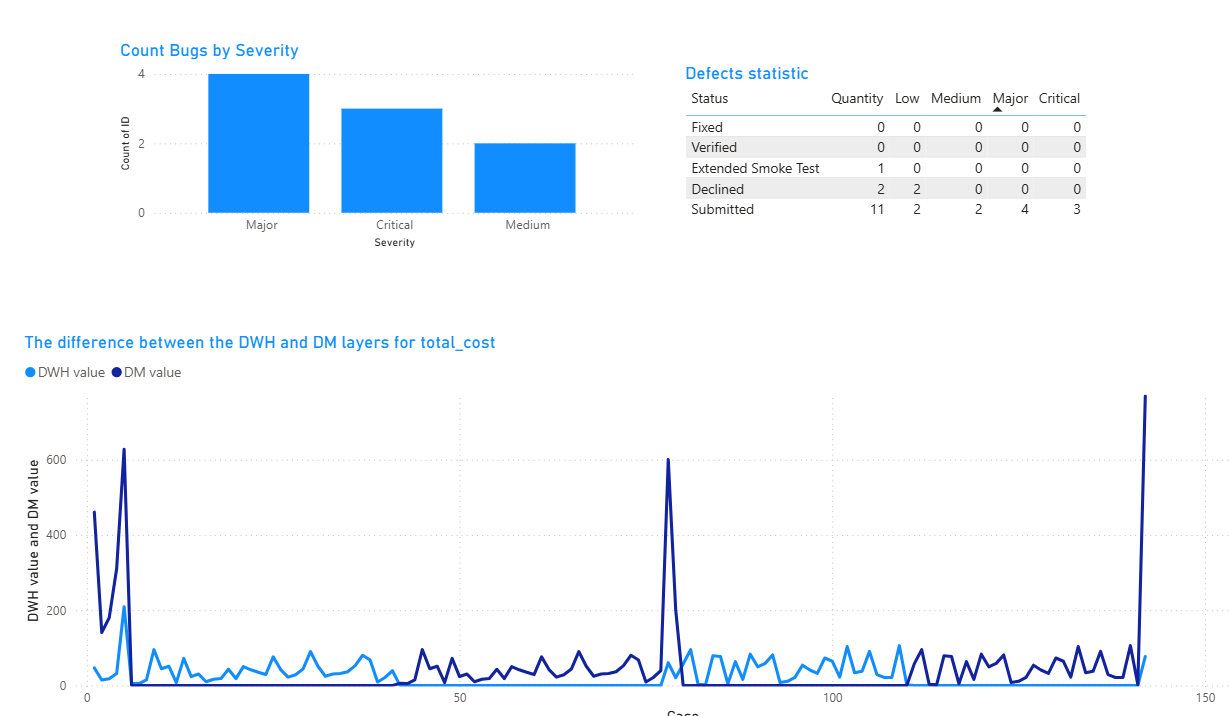
 **Team training and improved documentation:**

* Regular training for staff on data processing standards and improving documentation for data handling processes will also help reduce errors and increase efficiency.

 **Use containerization for the development environment (e.g., Docker):**

* Containerization allows you to create isolated environments for testing and development, making deployment and managing dependencies easier, and reducing the likelihood of errors when transitioning between different environments (development, testing, production).

**8. Attachment - visuals**

****