| **DWH Test Result Summary** |
| --- |
| DQ Strategy |



**Table of contents**

[**Summary**](#_heading=h.n1nqzkrztphl) **3**

[**Test Team**](#_heading=h.561wamsa5y7u) **3**

[**Testing Process Description**](#_heading=h.80i1ch998fwj) **3**

[Pre-requisites](#_heading=h.np7lufi1h1ll) 3

[Testing Approach](#_heading=h.rx0g096zi5xs) 3

[Smoke Testing](#_heading=h.2x4y3yin48sc) 4

[Critical Path Testing( using data reconciliation)](#_heading=h.qfhqhxpozaj3) 4

[Extended Path Testing ( lowest priority checks)](#_heading=h.yvh3mca1zq7x) 4

[**New Defects Statistics**](#_heading=h.qbf6kmb1nekh) **4**

[**New Defects List**](#_heading=h.n9sh8i7s5m25) **5**

[**Recommendations**](#_heading=h.wijd9vosnl98) **8**

[**Attachments / Visualizations**](#_heading=h.yi7wnve4tobv) **8**

[Total Anomalies per Table – Bar Chart](#_heading=h.jhmag8h7nbqw) 8

[Reconciliation Results by Table – Stacked Bar Chart](#_heading=h.cvqoml2o0vpv) 8

[Reconciliation Status Distribution – Bar Chart](#_heading=h.1boiok8phtil) 9

[Reconciliation Status Distribution – Pie Chart](#_heading=h.fesoqdutfcs2) 9

[Bug Severity Distribution (Pie Chart, from Jira Export)](#_heading=h.2qd6ms670k5j) 10

[Bug Priority Distribution (Bar Chart, from Jira Export)](#_heading=h.z4sb5vy2ckgl) 11

# Summary

This report summarizes the results of DWH testing conducted on **dwh\_src\_hw\_db** (Source) and **dwh\_hw\_db** (Target) databases, including landing, DWH, and dashboard (DM) layers. The focus was on:

* Data reconciliation across all layers
* Smoke testing for row-level integrity, key uniqueness, schema, referential integrity, and value sanity
* Extended testing to identify edge cases and potential data quality issues
* Bug logging and tracking in Jira (prefix **DWH\_**)

Testing ensures completeness, consistency, and accuracy of data in the DWH and dashboards.

# Test Team

| TEST TEAM | | | |
| --- | --- | --- | --- |
| # | Role | Name | Responsability |
|
| 1 | QA Lead | Artsem Nikulin | Oversight of DWH testing, reconciliation, bug logging |
| 2 | QA Analyst | Gabriela Cretu | SQL/Python reconciliation, smoke testing, extended path testing |
| 3 | BI Developer | Gabriela Cretu | Dashboard validation, data mapping |

# Testing Process Description

## Pre-requisites

* Executed scripts to create source, landing, DWH, and DM layers:  
  + 1\_create\_sources\_postgres.sql
  + 2\_create\_landing\_postgres.sql
  + 3\_create\_DWH\_layer\_postgres.sql
  + 4\_create\_DM\_layer\_postgres.sql
* Verified table creation and initial data population.

## Testing Approach

### Smoke Testing

* Row-level integrity
* Uniqueness of keys
* Null values in key columns
* Schema validation
* Referential integrity
* Value sanity checks
* Data type and length checks
* Code references: 6\_smoke\_tests.sql, 7\_smoke\_tests\_source.sql

### Critical Path Testing( using data reconciliation)

* SQL-based reconciliation (FULL OUTER JOIN, lnd.reconciliation\_results)
* Python-based reconciliation for flexibility and performance (Python\_reconciliation.ipynb)
* Layers checked:  
  + S1 & S2 → Landing
  + Landing → DWH
  + DWH → Dashboard (DM)

### Extended Path Testing ( lowest priority checks)

* S1 Clients – middle\_name nulls
* S2 Clients – valid\_from / valid\_to consistency
* S1 Sales – units > 0
* S2 Client Sales – future/missing sale dates
* S1 Products – missing product\_name/cost
* S1 Channels – channel\_location mapping
* Code reference: 8\_extended\_path.sql

# New Defects Statistics

| New defects statsistics | | | | |
| --- | --- | --- | --- | --- |
| # | Severity | | Priority | |
| Type | Count | Type | Count |
| 1 | Critical | 5 | Highest | 0 |
| 2 | Major | 8 | High | 9 |
| 3 | Minor | 14 | Medium | 18 |
| 4 | Trivial | 0 | Low | 0 |
| Total | 27 | | | |

# New Defects List

| New defects list | | | | | |
| --- | --- | --- | --- | --- | --- |
| # | BUG\_ID | Issue | Severity | Priority | Link |
|
| 1 | SCRUM-1750 | The column location\_name in table dwh.dwh\_locations has length character varying(256) but should be character varying(100). | Major | Medium | [Atlassian](https://dq-team-1-mimlefokko.atlassian.net/browse/SCRUM-1750?atlOrigin=eyJpIjoiZmU0Y2FlNjNkZjFiNGIwYzllODdjY2MzZTg1ZWUxNDkiLCJwIjoiaiJ9) |
| 2 | SCRUM-1749 | The column is\_valid in table dwh.dwh\_clients is defined as character(1) but should be character varying(1). This may cause ETL or validation failures. | Major | High | [Atlassian](https://dq-team-1-mimlefokko.atlassian.net/browse/SCRUM-1749?atlOrigin=eyJpIjoiMTIyZjRlOWRiYWI1NGQxODg0MmIwMzEyMmM3MjRkNzYiLCJwIjoiaiJ9) |
| 3 | SCRUM-1748 | The column channel\_name in table dwh.dwh\_channels has incorrect length. Expected character varying(30) but found character varying(50). | Major | Medium | [Atlassian](https://dq-team-1-mimlefokko.atlassian.net/browse/SCRUM-1748?atlOrigin=eyJpIjoiMGU0ZmM0ZmNhMDcwNDA4YTg1MmQ5MTY4MjQyMmEzMGIiLCJwIjoiaiJ9) |
| 4 | SCRUM-1747 | The column total\_cost in table dm.dm\_main\_dashboard has an incorrect numeric precision. Expected numeric(18,4) but found numeric(16,4). This may cause rounding errors or ETL failures. | Critical | High | [Atlassian](https://dq-team-1-mimlefokko.atlassian.net/browse/SCRUM-1747?atlOrigin=eyJpIjoiYWM2Mzg3MWVlZmQyNGNiYWIyNDEyODM1ZDM0NmIzNWYiLCJwIjoiaiJ9) |
| 5 | SCRUM-1746 | During reconciliation between the S2 landing lnd\_s2\_client\_sales table and the source s2.s2\_client\_sales table, 5 sales exist in the source table but are missing in the landing table. | Minor | Medium | [Atlassian](https://dq-team-1-mimlefokko.atlassian.net/browse/SCRUM-1746?atlOrigin=eyJpIjoiMTRhYWU4NDVhZmQ3NDNiYmFkYzQ5YTFiYTZlYmRlOTkiLCJwIjoiaiJ9) |
| 6 | SCRUM-1745 | During reconciliation between the S2 landing lnd\_s2\_locations table and the source s2.s2\_locations table, 4 locations exist only in the source table. | Minor | Medium | [Atlassian](https://dq-team-1-mimlefokko.atlassian.net/browse/SCRUM-1745?atlOrigin=eyJpIjoiZDAxZjhmOThkZmZlNGU1MjgxNTkyNjA5ZjY3MGE5MTIiLCJwIjoiaiJ9) |
| 7 | SCRUM-1744 | During reconciliation between the S2 landing lnd\_s2\_locations table and the source s2.s2\_locations table, 1 location exists only in the landing table. | Minor | Medium | [Atlassian](https://dq-team-1-mimlefokko.atlassian.net/browse/SCRUM-1744?atlOrigin=eyJpIjoiMWVjMTQwMWI3OTVkNGE3ZWEwOGZlYTBkMWY3YzQwYzIiLCJwIjoiaiJ9) |
| 8 | SCRUM-1743 | During reconciliation between the S1 landing lnd\_s1\_sales table and the source s1.s1\_sales table, 176 sales records exist only in the source table. | Major | High | [Atlassian](https://dq-team-1-mimlefokko.atlassian.net/browse/SCRUM-1743?atlOrigin=eyJpIjoiZTk5OGVlOWU3MDdlNGFiMDk0YjNkMTk5MjM5NGRiMTQiLCJwIjoiaiJ9) |
| 9 | SCRUM-1742 | During reconciliation between the S1 landing lnd\_s1\_products table and the source s1.s1\_products table, 1 product record exists only in the landing table. | Minor | Medium | [Atlassian](https://dq-team-1-mimlefokko.atlassian.net/browse/SCRUM-1742?atlOrigin=eyJpIjoiMmFiMTNiY2NiNmNhNGU1OGE5NTU2ZGFlNDdhMWJhY2QiLCJwIjoiaiJ9) |
| 10 | SCRUM-1741 | During reconciliation between the S1 landing lnd\_s1\_clients table and the source s1.s1\_clients table, 1 client record has a last name mismatch. | Minor | Medium | [Atlassian](https://dq-team-1-mimlefokko.atlassian.net/browse/SCRUM-1741?atlOrigin=eyJpIjoiNjhhNzNlNGE2MDMwNGI0N2FiNWFjNTc0YTA2OTgzYjEiLCJwIjoiaiJ9) |
| 11 | SCRUM-1740 | During reconciliation between the S1 landing lnd\_s1\_clients table and the source s1.s1\_clients table, 1 client record has a phone number mismatch. | Minor | Medium | [Atlassian](https://dq-team-1-mimlefokko.atlassian.net/browse/SCRUM-1740?atlOrigin=eyJpIjoiNmVhMDQwY2UzYzQ2NDY0NDkwZWY4NDBjMzYyNWY4YzEiLCJwIjoiaiJ9) |
| 12 | SCRUM-1739 | During reconciliation between the S1 landing lnd\_s1\_channels table and the source s1.s1\_channels table, 2 records exist in the source database that are missing from the landing table. | Minor | Medium | [Atlassian](https://dq-team-1-mimlefokko.atlassian.net/browse/SCRUM-1739?atlOrigin=eyJpIjoiYjdmOTRiMzZkYmFhNDQ5Yjg0MTgxZDk4OTVkNDNlYjEiLCJwIjoiaiJ9) |
| 13 | SCRUM-1738 | During reconciliation between the S1 landing lnd\_s1\_channels table and the source s1.s1\_channels table, 1 record was found where the channel\_name in landing does not match the corresponding channel\_name in the source database. | Minor | Medium | [Atlassian](https://dq-team-1-mimlefokko.atlassian.net/browse/SCRUM-1738?atlOrigin=eyJpIjoiODM4NmYwMGJiZmY3NGRjMjhmYzE4OWQzOTA5ODQwZDAiLCJwIjoiaiJ9) |
| 14 | SCRUM-1737 | During reconciliation between the S1 landing lnd\_s1\_channels table and the source s1.s1\_channels table, 5 records were found where the channel\_location in landing does not match the corresponding channel\_location in the source database. | Minor | Medium | [Atlassian](https://dq-team-1-mimlefokko.atlassian.net/browse/SCRUM-1737?atlOrigin=eyJpIjoiZWRlODVhYzY3YWY0NGZmNWFjMzZjNzYzMjcwM2YxMWMiLCJwIjoiaiJ9) |
| 15 | SCRUM-1736 | During reconciliation between the landing sales tables (lnd\_s1\_sales / lnd\_s2\_client\_sales) and DWH dwh\_sales, 181 sales records exist only in DWH and are missing in the landing layer. | Major | High | [Atlassian](https://dq-team-1-mimlefokko.atlassian.net/browse/SCRUM-1736?atlOrigin=eyJpIjoiMzlhY2ExNzhmNGJlNGMxYjhmOTg3ODc0M2ZlNThiNmIiLCJwIjoiaiJ9) |
| 16 | SCRUM-1735 | During reconciliation between the DWH dwh\_clients table and landing clients (lnd\_s1\_clients / lnd\_s2\_clients), 3 client records have mismatched phone numbers between landing and DWH. | Minor | Medium | [Atlassian](https://dq-team-1-mimlefokko.atlassian.net/browse/SCRUM-1735?atlOrigin=eyJpIjoiNjE5NTgwNTk1NThhNDRjODhjZmU0Yjc2MmNhZTA4M2UiLCJwIjoiaiJ9) |
| 17 | SCRUM-1734 | During reconciliation between the DWH dwh\_clients table and landing layer clients (lnd\_s1\_clients / lnd\_s2\_clients), 1 client record has a mismatched last\_name value between landing and DWH. | Minor | Medium | [Atlassian](https://dq-team-1-mimlefokko.atlassian.net/browse/SCRUM-1734?atlOrigin=eyJpIjoiM2U1YTg4NTQzZTBkNDBhYzk2Nzg1YjVkYzA5M2UzOTAiLCJwIjoiaiJ9) |
| 18 | SCRUM-1733 | During reconciliation between the DWH Sales fact and the Dashboard (dm\_main\_dashboard), 7 sales records show mismatched total\_cost values between the dashboard and DWH calculation (quantity \* product\_cost). | Minor | Medium | [Atlassian](https://dq-team-1-mimlefokko.atlassian.net/browse/SCRUM-1733?atlOrigin=eyJpIjoiYzI2MWVkNmQ1NTk3NDRkMTlkNmFkZTcyMTk3YzRhYWEiLCJwIjoiaiJ9) |
| 19 | SCRUM-1732 | During reconciliation between the DWH channels table (dwh\_channels) and the landing channels data, 1 channel record has a name mismatch (Mismatch in channel\_name). | Minor | Medium | [Atlassian](https://dq-team-1-mimlefokko.atlassian.net/browse/SCRUM-1732?atlOrigin=eyJpIjoiOTQ4OWQ2ZjIxNDdjNDFmYjg2ZTgyYjZhMjYwNjI5MzkiLCJwIjoiaiJ9) |
| 20 | SCRUM-1731 | During reconciliation between the DWH Sales fact and the Dashboard (dm\_main\_dashboard), 2 sales records exist in the dashboard but have no corresponding record in the DWH (Only in dashboard). | Minor | Medium | [Atlassian](https://dq-team-1-mimlefokko.atlassian.net/browse/SCRUM-1731?atlOrigin=eyJpIjoiZDI5Mzk2ZTIxZWM2NDgzMmEzMzQ2NDI0NDRhNTUwNmMiLCJwIjoiaiJ9) |
| 21 | SCRUM-1730 | During reconciliation between the DWH Sales fact and the Dashboard (dm\_main\_dashboard), 7 sales records show mismatched total\_cost values between the dashboard and DWH calculation (quantity \* product\_cost). | Major | Medium | [Atlassian](https://dq-team-1-mimlefokko.atlassian.net/browse/SCRUM-1730?atlOrigin=eyJpIjoiZWEzZDAwOTFhNDQ3NGYzNmFhZGIyNjg2OWUyZWYzMzciLCJwIjoiaiJ9) |
| 22 | SCRUM-1729 | 43 sales records have mismatched phone numbers between dm\_main\_dashboard and the DWH client dimension. | Major | High | [Atlassian](https://dq-team-1-mimlefokko.atlassian.net/browse/SCRUM-1729?atlOrigin=eyJpIjoiOWMxYWIyZWU1NTJkNDdiMDhhMDQ3NzU3NTA2NTg1NTQiLCJwIjoiaiJ9) |
| 23 | SCRUM-1728 | During reconciliation between the **DWH Sales fact** and the **Dashboard (dm\_main\_dashboard)** view, we identified **67 sales records** where the client email in the dashboard does not match the email stored in the DWH client dimension. | Major | Medium | [Atlassian](https://dq-team-1-mimlefokko.atlassian.net/browse/SCRUM-1728?atlOrigin=eyJpIjoiYmM1NmJkN2QwMWVhNDBjMzhjZDc3NWFmNTE5NGI3M2MiLCJwIjoiaiJ9) |
| 24 | SCRUM-1727 | Sales transactions in s1\_sales reference product IDs that do not exist in s1\_products. This breaks referential integrity. | Critical | High | [Atlassian](https://dq-team-1-mimlefokko.atlassian.net/browse/SCRUM-1727?atlOrigin=eyJpIjoiZjU0YmM5NzYzMWZhNGMwMDhmMGZjMzE0OWUwMzBmZDYiLCJwIjoiaiJ9) |
| 25 | SCRUM-1726 | The deduplicated warehouse layer (s2\_locations) still contains duplicate IDs, suggesting incorrect ETL logic. | Critical | High | [Atlassian](https://dq-team-1-mimlefokko.atlassian.net/browse/SCRUM-1726?atlOrigin=eyJpIjoiYzYwMzQ2YzhiN2Y3NGI3OThjMTM1Mzg1NDc3ZDdlOTgiLCJwIjoiaiJ9) |
| 26 | SCRUM-1725 | The lnd\_s1\_clients table contains duplicate client\_ids. This breaks uniqueness assumptions for primary keys and affects sales-to-client relationships. | Critical | High | [Atlassian](https://dq-team-1-mimlefokko.atlassian.net/browse/SCRUM-1725?atlOrigin=eyJpIjoiOTg0MWFjNTJjZmM3NGMxYWFiNDc4NWRkNTZlYjNhMWYiLCJwIjoiaiJ9) |
| 27 | SCRUM-1724 | The lnd\_s2\_locations table contains duplicate primary keys (location\_id). This violates uniqueness rules and may cause incorrect joins in downstream transformations. | Critical | High | [Atlassian](https://dq-team-1-mimlefokko.atlassian.net/browse/SCRUM-1724?atlOrigin=eyJpIjoiNTIzM2EwZWU2NDM3NDIyNTk3MmE1ZmFhYmVlODAwNTEiLCJwIjoiaiJ9) |

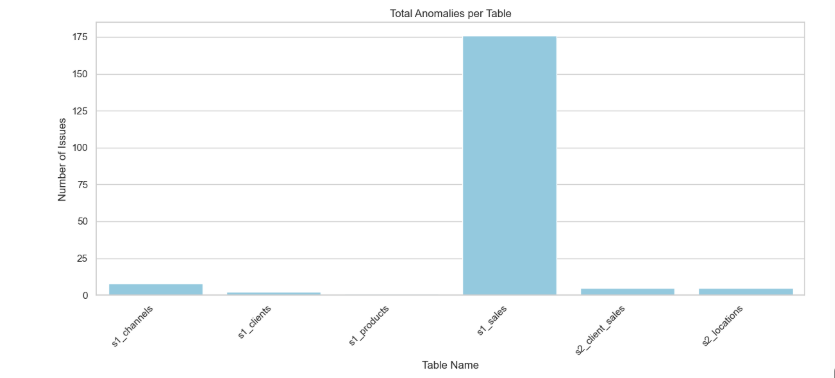
# Recommendations

* Automate reconciliation scripts for future DWH updates
* Implement stricter validation at source ingestion stage
* Review data mappings for dashboards, especially for edge cases
* Periodically monitor key metrics and anomalies in S1/S2 → Landing
* Standardize middle\_name, product\_name, and date validations

# Attachments / Visualizations

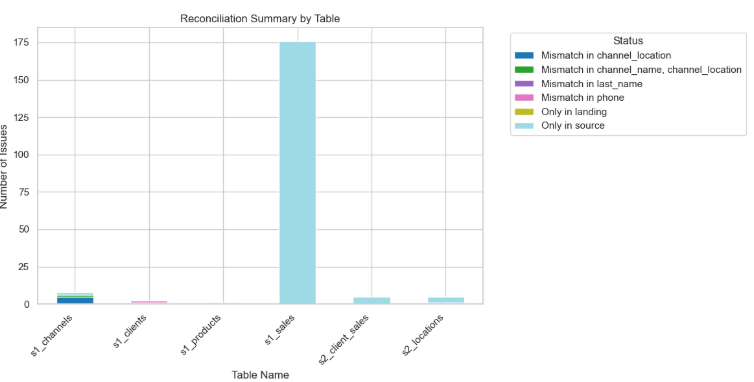
## Total Anomalies per Table – Bar Chart

Shows the **total number of issues detected per table**. Provides a quick overview of which tables contain the most discrepancies during reconciliation.



## Reconciliation Results by Table – Stacked Bar Chart

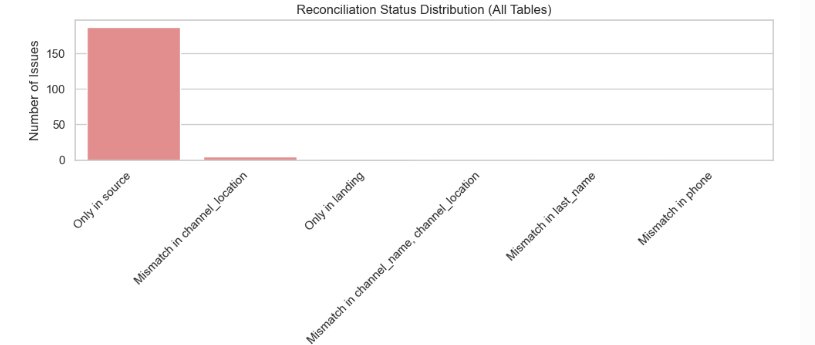
Displays the **breakdown of issue types per table**, categorized by reconciliation status:

* **Only in Source** – Records missing in the target.
* **Only in Target/Landing** – Records missing in the source.
* **Mismatch** – Records that exist in both but have differing column values.  
   Helps identify both the volume and type of data inconsistencies across tables.  
  

## Reconciliation Status Distribution – Bar Chart

Aggregates **all tables and shows counts of each reconciliation status**, allowing quick understanding of the overall

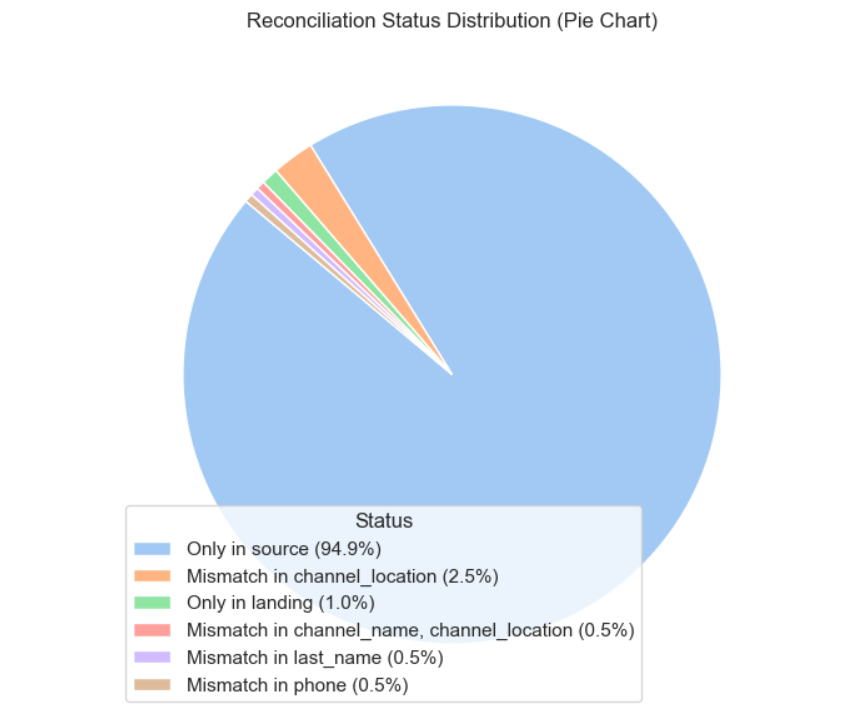
distribution of issue types.



## Reconciliation Status Distribution – Pie Chart

Presents the **proportion of each reconciliation status** across all tables as percentages, providing an easy-to-read

summary of data quality issues.

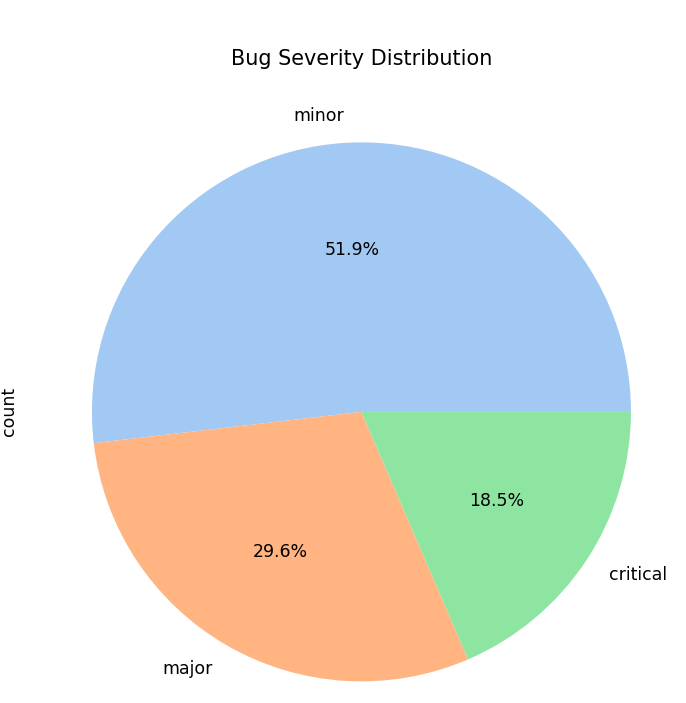


Next, I implemented a bugs profiling to see the type of severity or priority among the bugs distribution

## Bug Severity Distribution (Pie Chart, from Jira Export)

A bar chart showing the number of bugs assigned to each priority level (e.g., Critical, High, Medium, Low).

**Purpose:** Highlights which bugs require immediate attention from the development/BI team and supports resource planning for fixing issues.



## Bug Priority Distribution (Bar Chart, from Jira Export)

If time-stamped data is available, this line chart can show trends of flagged records per table or layer over the testing period.

**Purpose:** Identifies patterns or recurring problems, enabling proactive measures to prevent repeated issues in future DWH loads.

