

testcase12_parquet_parquet_mismatch_manual

fingerprint comparison

1. Run Summary

| | |
|--------------------|---|
| Application Name | :etl_pipeline_goldlayer |
| Protocol Name | :DATFDemo |
| Protocol File Path | :datf_core//test/testprotocol/testselection_template.xlsx |
| Testcase Name | :testcase12_parquet_parquet_mismatch_manual |
| Testcase Type | :fingerprint |
| Test Environment | :SIT |
| Start Time | :10-Mar-2025 09:15:01 UTC |
| End Time | :10-Mar-2025 09:19:28 UTC |
| Run Time | :0:04:26 |
| Test Result | :Failed |
| Reason | :One or more fingerprint mismatched. Please check detail report for more info |

2. Configuration Details

| | |
|-------------------------|------------------------------|
| Compare Type | :likeobjectcompare |
| testquerygenerationmode | :Manual |
| Testcase Type | :fingerprint |
| Source Connection Name | :RawS3Bucket |
| Source Connection Type | :aws-s3 |
| Source Connection Value | :RawS3Bucket |
| Source Format | :parquet |
| Source Name | :patient_data_source_parquet |
| Source Path | :test/data/source |
| Target Connection Name | :CuratedS3Bucket |
| Target Connection Type | :aws-s3 |
| Target Connection Value | :CuratedS3Bucket |
| Target Format | :parquet |
| Target Name | :patient_data_target_parquet |
| Target Path | :test/data/target |
| Primary Keys | :id |

3. Fingerprint Summary

| | |
|------------------------|---------|
| Test Result | :Failed |
| No. of KPIs in Source | :8 |
| No. of KPIs in Target | :8 |
| No. of KPIs matched | :4 |
| No. of KPIs mismatched | :4 |

4. SQL Queries

4.1 Source Query

SELECT ADDRESS,BIRTHDATE,BIRTHPLACE,CITY,COUNTY,DEATHDATE,DRIVERS,ETHNICITY,FIRST,GENDER,HEALTH CARE_COVERAGE,HEALTHCARE_EXPENSES,LAST,LAT,LON,MAIDEN,MARITAL,PASSPORT,PREFIX,RACE,SSN, STATE,SUFFIX,ZIP,id FROM rawpatients

4.2 Target Query

SELECT ADDRESS,BIRTHDATE,BIRTHPLACE,CITY,COUNTY,DEATHDATE,DRIVERS,ETHNICITY,FIRST,GENDER,HEALTH CARE_COVERAGE,HEALTHCARE_EXPENSES,LAST,LAT,LON,MAIDEN,MARITAL,PASSPORT,PREFIX,RACE,SSN, STATE,SUFFIX,ZIP,id FROM curatedpatients

5. Fingerprint Comparison Report

5.1. KPI (1) - Row Count

| | |
|------------------|----------|
| Source Row Count | :1001000 |
| Target Row Count | :1001000 |

KPI Status :Passed
Reason :Source and Target row count is matching

5.2. KPI (2) - Column Count

Source Column Count :25
Target Column Count :25
KPI Status :Passed
Reason :Source column count and target column count is matching

5.3. KPI (3) - Column Level Null Check

List of Columns has null in source :['DEATHDATE', 'DRIVERS', 'MAIDEN', 'PASSPORT', 'SUFFIX']
List of Columns has null in target :['DEATHDATE', 'DRIVERS', 'MAIDEN', 'PASSPORT', 'SUFFIX']
KPI Status :Failed
Reason :Source or/and Target has nulls in some columns

5.4. KPI (4) - Row Level Null Check

No of null rows in source :0
No of null rows in target :0
KPI Status :Passed
Reason :No rows are null

5.5. KPI (5) - Column Level empty Check

List of Columns has empty in source :[]
List of Columns has empty in target :[]
KPI Status :Passed
Reason :Source or/and Target has no empty values in any columns

5.6. KPI (6) - Column Level Duplicate Check

List of Columns has duplicates in source:['ADDRESS', 'BIRTHDATE', 'BIRTHPLACE', 'CITY', 'COUNTY', 'DEATHDATE', 'DRIVERS', 'ETHNICITY', 'FIRST', 'GENDER', 'HEALTHCARE_COVERAGE', 'HEALTHCARE_EXPENSES', 'LAST', 'LAT', 'LON', 'MAIDEN', 'MARITAL', 'PASSPORT', 'PREFIX', 'RACE', 'SSN', 'STATE', 'SUFFIX', 'ZIP', 'id']
List of Columns has duplicates in target :['ADDRESS', 'BIRTHDATE', 'BIRTHPLACE', 'CITY', 'COUNTY', 'DEATHDATE', 'DRIVERS', 'ETHNICITY', 'FIRST', 'GENDER', 'HEALTHCARE_COVERAGE', 'HEALTHCARE_EXPENSES', 'LAST', 'LAT', 'LON', 'MAIDEN', 'MARITAL', 'PASSPORT', 'PREFIX', 'RACE', 'SSN', 'STATE', 'SUFFIX', 'ZIP', 'id']
KPI Status :Failed
Reason :Source or/and Target has duplicate values in some columns

5.7. KPI (7) - Row Level Duplicate Check

No of duplicate rows in source :1000
No of duplicate rows in target :1000
KPI Status :Failed
Reason :One or more rows are duplicated

5.8. KPI (8) - CheckSum Row Level

KPI Status :Failed
Reason :Data is not matching between source and target

5.8.1 Addition Rows in Source

| S.No | hashval |
|------|----------------------------------|
| 1 | 0220751618574bf3bfcf495b633aa9cb |
| 2 | 029d99347cdae3900473637cdfa37d8b |
| 3 | 0c72e1ecff51fe83051d9a796faa1064 |
| 4 | 0ea71e926316da934cc6819d694aca93 |
| 5 | 128fa3c130b9bdcf2202cf13184c41d4 |
| 6 | 158479ce1377e81e11d472ad03054532 |
| 7 | 1701d81614415691269ad7fad5078f71 |

| | |
|---|----------------------------------|
| 8 | 1dd798b65bcd9f8e24c3f761a255e773 |
|---|----------------------------------|

5.8.2 Addition Rows in Target

| S.No | hashval |
|------|----------------------------------|
| 1 | 0eb9776ede72734c20e921c051097b39 |
| 2 | 14eeac6c522417bc9aae184fad623bf3 |
| 3 | 17d888c5d0075c2282eee0a389f45e48 |
| 4 | 1f74d2756a97f6b7cd14a9d75debffaf |
| 5 | 1fa8fc575bf1e75ee81120815f6c2dc4 |
| 6 | 2ad15535510d79645e9d89332a1a6b1c |
| 7 | 2c014e21bc959a29a66c2466602d7f1f |
| 8 | 316d98f1083d0349edc4c11338cc066b |