

testcase12_parquet_parquet_mismatch_manual content comparison

1. Run Summary

| | |
|--------------------|---|
| Application Name | :etl_pipeline_goldlayer |
| Protocol Name | :DATFDemo |
| Protocol File Path | :/Workspace/Shared/QE_ATF_Latest/datf_core/test/testprotocol/traversedtestprotocol.xlsx |
| Testcase Name | :testcase12_parquet_parquet_mismatch_manual |
| Testcase Type | :content |
| Test Environment | :SIT |
| Start Time | :16-Jan-2025 13:42:30 UTC |
| End Time | :16-Jan-2025 13:43:02 UTC |
| Run Time | :0:00:32 |
| Test Result | :Passed |
| Reason | :Content matched |

2. Configuration Details

| | |
|-------------------------|--------------------------|
| Compare Type | :likeobjectcompare |
| testquerygenerationmode | :Manual |
| Testcase Type | :content |
| Source Connection Name | :RawS3Bucket |
| Source Connection Type | :aws-s3 |
| Source Connection Value | :RawS3Bucket |
| Source Format | :parquet |
| Source Name | :patients_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 | :patients_source_parquet |
| Target Path | :/test/data/source |
| Primary Keys | :id |

3. Content Summary

| | |
|---|---------|
| Test Result | :Passed |
| No. of matched columns | :25 |
| No. of columns compared | :24 |
| No. of cols in Source but not in Target | :0 |
| No. of cols in Target but not in Source | :0 |
| No. of rows in Source | :1,171 |
| No. of distinct rows in Source | :1,171 |
| No. of duplicate rows in Source | :0 |
| No. of rows in Target | :1,171 |
| No. of distinct rows in Target | :1,171 |
| No. of duplicate rows in Target | :0 |
| No. of matched rows | :1,171 |
| No. of mismatched rows | :0 |
| No. of rows in Source but not in Target | :0 |
| No. of rows in Target but not in Source | :0 |

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. Sample Mismatches 8 rows

5.1 Keys in source but not in target

None

5.2 Keys in target but not in source

None

5.3 Keys having one or more unequal column values

None

6. Columnwise Mismatch Summary

None

7. Columnwise Mismatch Details

None