

testcase1_parquet_parquet_mismatch duplicate comparison

1. Run Summary

| | |
|--------------------|---|
| Application Name | :etl_pipeline_goldlayer |
| Protocol Name | :DATFDemo |
| Protocol File Path | :datf_core//test/testprotocol/testselection_template.xlsx |
| Testcase Name | :testcase1_parquet_parquet_mismatch |
| Testcase Type | :duplicate |
| Test Environment | :SIT |
| Start Time | :11-Mar-2025 14:08:53 UTC |
| End Time | :11-Mar-2025 14:09:15 UTC |
| Run Time | :0:00:21 |
| Test Result | :Passed |
| Reason | :No Duplicates |

2. Configuration Details

| | |
|-------------------------|--|
| Compare Type | :s2tcompare |
| testquerygenerationmode | :Auto |
| Testcase Type | :duplicate |
| Source Connection Type | :aws-s3 |
| Source Format | :parquet |
| Source Path | :datf_core/test/data/source/patients_source_parquet |
| Target Connection Type | :aws-s3 |
| Target Format | :parquet |
| Target Path | :datf_core/test/data/target/patients_target_parquet_mismatch |
| S2T Path | :test/s2t/s2t_1_parquet_parquet_mismatch.xlsx |
| Primary Keys | :id |

3. Duplicate Summary

| | |
|---------------------------------|---------|
| Test Result | :Passed |
| 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 | :1,169 |
| No. of distinct rows in Target | :1,169 |
| No. of duplicate rows in Target | :0 |

4. SQL Queries

4.1 Source Query

```
SELECT Count(DISTINCT id) FROM parquet.`datf_core/test/data/source/patients_source_parquet`
```

4.2 Target Query

```
SELECT Count(DISTINCT id) FROM parquet.`datf_core/test/data/target/patients_target_parquet_mismatch`
```