

testcase1\_parquet\_parquet\_mismatch  
schema comparison

1. Run Summary

|                    |   |
|--------------------|---|
| Application Name   | :etl_pipeline_goldlayer   |
| Protocol Name      | :DATFDemo   |
| Protocol File Path | :/Workspace/Shared/QE_ATF_Enhanced_03/datf_core/test/testprotocol/traversedtestprotocol_schema_loman.xlsx |
| Testcase Name      | :testcase1_parquet_parquet_mismatch   |
| Testcase Type      | :schema   |
| Test Environment   | :SIT  |
| Start Time         | :10-Apr-2025 09:24:12 UTC   |
| End Time           | :10-Apr-2025 09:26:56 UTC   |
| Run Time           | :0:02:44  |
| Test Result        | :Failed   |
| Reason             | :Content mismatched   |

2. Configuration Details

|                         |  |
|-------------------------|--|
| Compare Type            | :s2tcompare  |
| testquerygenerationmode | :Auto  |
| Testcase Type           | :schema  |
| Source Connection Type  | :aws-s3  |
| Source Format           | :parquet   |
| Source Path             | :file:/Workspace/Shared/QE_ATF_Enhanced_03/datf_core/test/data/source/patients_source_parquet          |
| Target Connection Type  | :aws-s3  |
| Target Format           | :parquet   |
| Target Path             | :file:/Workspace/Shared/QE_ATF_Enhanced_03/datf_core/test/data/target/patients_target_parquet_mismatch |
| S2T Path                | :test/s2t/s2t_1_parquet_parquet_mismatch.xlsx  |
| Primary Keys            | :id  |

3. Schema Summary

|   |         |
|---|---------|
| Test Result                             | :Failed |
| No. of matched columns                  | :12     |
| No. of columns compared                 | :10     |
| 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                   | :50     |
| No. of distinct rows in Source          | :50     |
| No. of duplicate rows in Source         | :0      |
| No. of rows in Target                   | :50     |
| No. of distinct rows in Target          | :25     |
| No. of duplicate rows in Target         | :25     |
| No. of matched rows                     | :0      |
| No. of mismatched rows                  | :25     |
| No. of rows in Source but not in Target | :25     |
| No. of rows in Target but not in Source | :0      |

4. SQL Queries

4.1 Source Query

select \* from s2tmappingsheet.schema where tabletype in ('source','target')

4.2 Target Query

DESCRIBE parquetview union DESCRIBE parquetview

5. Sample Mismatches 5 rows

5.1 Keys in source but not in target

| S.No | Key Columns                             |
|------|---|
| 1    | tabletype=source, columnname= id        |
| 2    | tabletype=source, columnname= LAST      |
| 3    | tabletype=source, columnname= COUNTY    |
| 4    | tabletype=source, columnname= RACE      |
| 5    | tabletype=source, columnname= ETHNICITY |

### 5.2 Keys in target but not in source

None

### 5.3 Keys having one or more unequal column values

| S.No | Key Columns                              |
|------|--|
| 1    | tabletype=target, columnname= ADDRESS    |
| 2    | tabletype=target, columnname= BIRTHDATE  |
| 3    | tabletype=target, columnname= BIRTHPLACE |
| 4    | tabletype=target, columnname= CITY       |
| 5    | tabletype=target, columnname= COUNTY     |

### 6. Columnwise Mismatch Summary

No. of rows in Source :50  
No. of rows in Target :50

| S.No | Source Column Name | Target Column Name | Rows in Source | Rows in Target | Rows with Common Keys | Rows Matched | Rows Mismatch |
|------|--------------------|--------------------|----------------|----------------|-----------------------|--------------|---------------|
| 1    | tablename          | tablename          | 50             | 50             | 0                     | 0            | 25            |
| 2    | columnindex        | columnindex        | 50             | 50             | 0                     | 0            | 25            |
| 3    | datatype           | datatype           | 50             | 50             | 0                     | 23           | 2             |
| 4    | primarykey         | primarykey         | 50             | 50             | 0                     | 24           | 1             |

### 7. Columnwise Mismatch Details

#### 7.1 tablename

| S.No | Key Columns                          | Source value | Target value |
|------|--------------------------------------|--------------|--------------|
| 1    | tabletype=target, columnname= COUNTY | targetdata   | dataview     |
| 2    | tabletype=target, columnname= CITY   | targetdata   | dataview     |
| 3    | tabletype=target, columnname= MAIDEN | targetdata   | dataview     |
| 4    | tabletype=target, columnname= LAST   | targetdata   | dataview     |
| 5    | tabletype=target, columnname= LON    | targetdata   | dataview     |

#### 7.2 columnindex

| S.No | Key Columns                          | Source value | Target value |
|------|--------------------------------------|--------------|--------------|
| 1    | tabletype=target, columnname= COUNTY | 20           |              |
| 2    | tabletype=target, columnname= CITY   | 18           |              |
| 3    | tabletype=target, columnname= MAIDEN | 11           |              |
| 4    | tabletype=target, columnname= LAST   | 9            |              |
| 5    | tabletype=target, columnname= LON    | 23           |              |

#### 7.3 datatype

| S.No | Key Columns | Source value | Target value |
|------|-------------|--------------|--------------|
|------|-------------|--------------|--------------|

|   |   |          |      |
|---|---|----------|------|
| 1 | tabletype=target, columnname= DEATHDATE | datetime | date |
| 2 | tabletype=target, columnname= BIRTHDATE | datetime | date |

7.4 primarykey

| S.No | Key Columns                      | Source value | Target value |
|------|----------------------------------|--------------|--------------|
| 1    | tabletype=target, columnname= id | Y            |              |