

testcase12\_parquet\_parquet\_mismatch\_manual  
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/traversedtestprotoc ol.xlsx
Testcase Name	:testcase12_parquet_parquet_mismatch_manual
Testcase Type	:schema
Test Environment	:SIT
Start Time	:08-Apr-2025 10:40:57 UTC
End Time	:08-Apr-2025 10:42:52 UTC
Run Time	:0:01:54
Test Result	:Failed
Reason	:Content mismatched

2. Configuration Details

Compare Type	:likeobjectcompare
testquerygenerationmode	:Auto
Testcase Type	:schema
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
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	:50
No. of duplicate rows in Target	:0
No. of matched rows	:0
No. of mismatched rows	:0
No. of rows in Source but not in Target	:50
No. of rows in Target but not in Source	:50

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 8 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=target, columnname= COUNTY
4	tabletype=source, columnname= COUNTY
5	tabletype=source, columnname= RACE
6	tabletype=source, columnname= ETHNICITY
7	tabletype=target, columnname= CITY
8	tabletype=target, columnname= MAIDEN

5.2 Keys in target but not in source

S.No	Key Columns
1	tabletype=rawpatients, columnname= HEALTHCARE_COVERAGE
2	tabletype=rawpatients, columnname= STATE
3	tabletype=DRIVERS, columnname= curatedpatients
4	tabletype=rawpatients, columnname= PREFIX
5	tabletype=rawpatients, columnname= ADDRESS
6	tabletype=rawpatients, columnname= PASSPORT
7	tabletype=COUNTY, columnname= curatedpatients
8	tabletype=PREFIX, columnname= curatedpatients

5.3 Keys having one or more unequal column values

None

6. Columnwise Mismatch Summary

None

7. Columnwise Mismatch Details

None