

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	:07-Mar-2025 13:59:02 UTC
End Time	:07-Mar-2025 14:00:22 UTC
Run Time	:0:01:19
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	: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. Fingerprint Summary

Test Result	:Failed
No. of KPIs in Source	:8
No. of KPIs in Target	:8
No. of KPIs matched	:6
No. of KPIs mismatched	:2

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	:1171
Target Row Count	:1171

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', 'MARITAL', 'PASSPORT', 'PREFIX', 'SUFFIX', 'ZIP']
List of Columns has null in target	:['DEATHDATE', 'DRIVERS', 'MAIDEN', 'MARITAL', 'PASSPORT', 'PREFIX', 'SUFFIX', 'ZIP']
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:	:['BIRTHDATE', 'BIRTHPLACE', 'CITY', 'COUNTY', 'DEATHDATE', 'DRIVERS', 'ETHNICITY', 'FIRST', 'GENDER', 'HEALTHCARE_COVERAGE', 'LAST', 'MAIDEN', 'MARITAL', 'PASSPORT', 'PREFIX', 'RACE', 'SSN', 'STATE', 'SUFFIX', 'ZIP']
List of Columns has duplicates in target	:['BIRTHDATE', 'BIRTHPLACE', 'CITY', 'COUNTY', 'DEATHDATE', 'DRIVERS', 'ETHNICITY', 'FIRST', 'GENDER', 'HEALTHCARE_COVERAGE', 'LAST', 'MAIDEN', 'MARITAL', 'PASSPORT', 'PREFIX', 'RACE', 'SSN', 'STATE', 'SUFFIX', 'ZIP']
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	:0
No of duplicate rows in target	:0
KPI Status	:Passed
Reason	:No rows are duplicated

5.8. KPI (8) - Check Sum Row Level

KPI Status	:Passed
Reason	:Data is matching between source and target

5.5.1 Addition Rows in Source

None

5.5.2 Addition Rows in Target

None

testcase23_parquet_snowflake_mismatch

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	:testcase23_parquet_snowflake_mismatch
Testcase Type	:fingerprint
Test Environment	:SIT
Start Time	:07-Mar-2025 14:00:22 UTC
End Time	:07-Mar-2025 14:02:05 UTC
Run Time	:0:01:43
Test Result	:Failed
Reason	:One or more fingerprint mismatched. Please check detail report for more info

2. Configuration Details

Compare Type	:likeobjectcompare
testquerygenerationmode	:Auto
Testcase Type	:fingerprint
Source Connection Name	:RawS3Parquet
Source Connection Type	:aws-s3
Source Connection Value	:RawS3Parquet
Source Format	:parquet
Source Name	:patients_source_parquet
Source Path	:test/data/source
Source Exclude Columns	:BIRTHDATE,DEATHDATE
Target Connection Name	:raw_snowflake_sql_connection
Target Connection Type	:snowflake
Target Connection Value	:raw_snowflake_sql_connection
Target Format	:table
Target Name	:SAMPLE_SRC_PATIENTS
Target Exclude Columns	:BIRTHDATE,DEATHDATE
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	:5
No. of KPIs mismatched	:3

4. SQL Queries

4.1 Source Query

```
SELECT
ADDRESS,BIRTHPLACE,CITY,COUNTY,DRIVERS,ETHNICITY,FIRST,GENDER,HEALTHCARE_COVERAGE,HEALTHCARE_EXPENSES,LAST,LAT,LON,MAIDEN,MARITAL,PASSPORT,PREFIX,RACE,SSN,STATE,SUFFIX,ZIP,id
FROM actualpatients
```

4.2 Target Query

```
SELECT
ADDRESS,BIRTHPLACE,CITY,COUNTY,DRIVERS,ETHNICITY,FIRST,GENDER,HEALTHCARE_COVERAGE,HEALTHCARE_EXPENSES,ID,LAST,LAT,LON,MAIDEN,MARITAL,PASSPORT,PREFIX,RACE,SSN,STATE,SUFFIX,ZIP
FROM SAMPLE_SRC_PATIENTS
```

5. Fingerprint Comparison Report

5.1. KPI (1) - Row Count

Source Row Count	:1171
------------------	-------

Target Row Count :1171
KPI Status :Passed
Reason :Source and Target row count is matching

5.2. KPI (2) - Column Count

Source Column Count :23
Target Column Count :23
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 :['DRIVERS', 'MAIDEN', 'MARITAL', 'PASSPORT', 'PREFIX', 'SUFFIX', 'ZIP']
List of Columns has null in target :['DRIVERS', 'MAIDEN', 'MARITAL', 'PASSPORT', 'PREFIX', 'SUFFIX', 'ZIP']
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:['BIRTHPLACE', 'CITY', 'COUNTY', 'DRIVERS', 'ETHNICITY', 'FIRST', 'GENDER', 'HEALTHCARE_COVERAGE', 'LAST', 'MAIDEN', 'MARITAL', 'PASSPORT', 'PREFIX', 'RACE', 'SSN', 'STATE', 'SUFFIX', 'ZIP']
List of Columns has duplicates in target :['BIRTHPLACE', 'CITY', 'COUNTY', 'DRIVERS', 'ETHNICITY', 'FIRST', 'GENDER', 'HEALTHCARE_COVERAGE', 'LAST', 'MAIDEN', 'MARITAL', 'PASSPORT', 'PREFIX', 'RACE', 'SSN', 'STATE', 'SUFFIX', 'ZIP']
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 :0
No of duplicate rows in target :0
KPI Status :Passed
Reason :No rows are duplicated

5.8. KPI (8) - Check Sum Row Level

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

5.5.1 Addition Rows in Source

S.No	hashval
1	1c4ece9e29b7d106ffe4d007f66c17b4
2	74fd5f19fe2a2fd451efa6df6a9ca503
3	8098c11876b5d955286db58da1898c4b
4	9c100a686cc7743495347f016ae35e51
5	ae178818e86205f6b9ffec34ca96a468
6	8beabd0ac8f2619b1c24dbe1558364ad
7	0aca74313f07890a32b04343815d55f4
8	b9a090fb867a141fb3a65575f3f886b4

5.5.2 Addition Rows in Target

S.No	hashval
1	03c4524cce4f16ca934d3a1e6a187231
2	3c94322a36da0797149f078e4d64bbb8
3	428a213a499a24ef4a4bd3ada047bfaa
4	50939aae8222f2f217ee973aa8dc9e3f
5	73fb287354f9051681386e5d37875f1e
6	67af836f16abd907d4b3873f9245695d
7	7dd9b7dd036cf4358f30d766f8e6aa3d
8	01619e9c3ff6723127d177587112a0c2