

testcase2_parquet_parquet_match
duplicate comparison

1. Run Summary

Application Name	:sample
Protocol Name	:DATFDemo
Protocol File Path	:/Workspace/Repos/sahil.gupta@tigeranalytics.com/QE_ATF/datf_core//test/testprotocol/contenttestprotocol.xlsx
Testcase Name	:testcase2_parquet_parquet_match
Testcase Type	:duplicate
Test Environment	:dev
Start Time	:29-Aug-2024 17:01:14 UTC
End Time	:29-Aug-2024 17:01:22 UTC
Run Time	:0:00:08
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	:/Workspace/Repos/sahil.gupta@tigeranalytics.com/QE_ATF/datf_core/test/data/source/patients_source_parquet
Target Connection Type	:aws-s3
Target Format	:parquet
Target Path	:/Workspace/Repos/sahil.gupta@tigeranalytics.com/QE_ATF/datf_core/test/data/stage/patients_target_parquet_match
S2T Path	:test/s2t/source_parquet_target_parquet_s2t_2.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	:1,171
No. of distinct rows in Target	:1,171
No. of duplicate rows in Target	:0

4. SQL Queries

4.1 Source Query

SELECT Count(DISTINCT id) FROM
parquet.`/Workspace/Repos/sahil.gupta@tigeranalytics.com/QE_ATF/datf_core/test/data/source/patients_source_parquet`

4.2 Target Query

SELECT Count(DISTINCT id) FROM
parquet.`/Workspace/Repos/sahil.gupta@tigeranalytics.com/QE_ATF/datf_core/test/data/stage/patients_target_parquet_match`

testcase12_parquet_parquet_mismatch_manuel

duplicate comparison

1. Run Summary

Application Name	:sample
Protocol Name	:DATFDemo
Protocol File Path	:/Workspace/Repos/sahil.gupta@tigeranalytics.com/QE_ATF/datf_core//test/testprotocol/contenttestprotocol.xlsx
Testcase Name	:testcase12_parquet_parquet_mismatch_manuel
Testcase Type	:duplicate
Test Environment	:dev
Start Time	:29-Aug-2024 17:01:22 UTC
End Time	:29-Aug-2024 17:01:26 UTC
Run Time	:0:00:04
Test Result	:Passed
Reason	:No Duplicates

2. Configuration Details

Compare Type	:s2tcompare
testquerygenerationmode	:Manual
Testcase Type	:duplicate
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/source_parquet_target_parquet_s2t_1.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	:1,171
No. of distinct rows in Target	:1,171
No. of duplicate rows in Target	:0

4. SQL Queries

4.1 Source Query

SELECT Count(DISTINCT id) FROM rawpatients

4.2 Target Query

SELECT Count(DISTINCT id) FROM curatedpatients

testcase14_csv_parquet_mismatch_manuel

duplicate comparison

1. Run Summary

Application Name	:sample
Protocol Name	:DATFDemo
Protocol File Path	:/Workspace/Repos/sahil.gupta@tigeranalytics.com/QE_ATF/datf_core//test/testprotocol/contenttestprotocol.xlsx
Testcase Name	:testcase14_csv_parquet_mismatch_manuel
Testcase Type	:duplicate
Test Environment	:dev
Start Time	:29-Aug-2024 17:01:27 UTC
End Time	:29-Aug-2024 17:01:30 UTC
Run Time	:0:00:03
Test Result	:Passed
Reason	:No Duplicates

2. Configuration Details

Compare Type	:s2tcompare
testquerygenerationmode	:Manual
Testcase Type	:duplicate
Source Connection Name	:RawS3Bucket
Source Connection Type	:aws-s3
Source Connection Value	:RawS3Bucket
Source Format	:delimited
Source Name	:patients_source.csv
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/source_csv_target_parquet_s2t_6_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	:1,171
No. of distinct rows in Target	:1,171
No. of duplicate rows in Target	:0

4. SQL Queries

4.1 Source Query

SELECT Count(DISTINCT id) FROM rawpatients

4.2 Target Query

SELECT Count(DISTINCT id) FROM curatedpatients

testcase16_parquet_parquet_mismatch_likeobject

duplicate comparison

1. Run Summary

Application Name	:sample
Protocol Name	:DATFDemo
Protocol File Path	:/Workspace/Repos/sahil.gupta@tigeranalytics.com/QE_ATF/datf_core//test/testprotocol/contenttestprotocol.xlsx
Testcase Name	:testcase16_parquet_parquet_mismatch_likeobject
Testcase Type	:duplicate
Test Environment	:dev
Start Time	:29-Aug-2024 17:01:31 UTC
End Time	:29-Aug-2024 17:01:35 UTC
Run Time	:0:00:03
Test Result	:Failed
Reason	:Duplicates

2. Configuration Details

Compare Type	:likeobjectcompare
testquerygenerationmode	:Auto
Testcase Type	:duplicate
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_target_parquet_mismatch
Target Path	:test/data/stage
Primary Keys	:id

3. Duplicate Summary

Test Result	:Failed
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 actualpatients

4.2 Target Query

SELECT Count(DISTINCT id) FROM expectedpatients

testcase24_snowflake_snowflake_etljob

duplicate comparison

1. Run Summary

Application Name	:sample
Protocol Name	:DATFDemo
Protocol File Path	:/Workspace/Repos/sahil.gupta@tigeranalytics.com/QE_ATF/datf_core//test/testprotocol/contenttestprotocol.xlsx
Testcase Name	:testcase24_snowflake_snowflake_etljob
Testcase Type	:duplicate
Test Environment	:dev
Start Time	:29-Aug-2024 17:01:35 UTC
End Time	:29-Aug-2024 17:01:44 UTC
Run Time	:0:00:09
Test Result	:Failed
Reason	:Duplicates

2. Configuration Details

Compare Type	:likeobjectcompare
testquerygenerationmode	:Auto
Testcase Type	:duplicate
Source Connection Name	:raw_snowflake_sql_connection
Source Connection Type	:snowflake
Source Connection Value	:raw_snowflake_sql_connection
Source Format	:table
Source Name	:SAMPLE_SRC_PATIENTS
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_TGT_PATIENTS
Target Exclude Columns	:BIRTHDATE,DEATHDATE
Primary Keys	:ID

3. Duplicate Summary

Test Result	:Failed
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 SAMPLE_SRC_PATIENTS

4.2 Target Query

SELECT Count(DISTINCT ID) FROM SAMPLE_TGT_PATIENTS