testcase21_mysql_csv_match

content comparison

1. Run Summary

Application Name	:sample
Protocol Name	:DBTablesDemo
Protocol File Path	:/app/test/testprotocol/testprotocol.xlsx
Testcase Name	:testcase21_mysql_csv_match
Testcase Type	:content
Test Environment	:dev
Start Time	:22-Dec-2023 13:45:04 UTC
End Time	:22-Dec-2023 13:51:21 UTC
Run Time	:0:06:17
Test Result	:Failed
Reason	:Content mismatched

2. Configuration Details

Compare Type	:s2tcompare
testquerygenerationmode	:Auto
Testcase Type	:content
Source Connection Name	:raw_mysql_database_connection
Source Connection Type	:mysql
Source Connection Value	:raw_mysql_database_connection
Source Format	:table
Source Name	:patients
Source Path	:patients/patients_db
Source Exclude Columns	:BIRTHDATE,DEATHDATE
Target Connection Name	:raw_mysql_database_connection
Target Connection Type	:aws-s3
Target Format	:delimitedfile
Target Name	:patients_target.csv
Target Path	:/app/test/data/target/patients_target.csv
Target Exclude Columns	:BIRTHDATE,DEATHDATE
S2T Path	:/app/test/s2t/s2t_21_mysql_csv_match.xlsx
Primary Keys	:id

3. Content Summary

Test Result	:Failed
No. of matched columns	:23
No. of columns compared	:22
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	: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
No. of matched rows	:0
No. of mismatched rows	:1,171
No. of rows in Source but not in Target	:0
No. of rows in Target but not in Source	:0

4. SQL Queries

4.1 Source Query

spark.sql("SELECT src.ADDRESS as ADDRESS, src.BIRTHPLACE as BIRTHPLACE, src.CITY as CITY, src.COUNTY as COUNTY, src.DRIVERS as DRIVERS, src.ETHNICITY as ETHNICITY, src.FIRST as FIRST, src.GENDER as GENDER, src.HEALTHCARE_COVERAGE as HEALTHCARE_COVERAGE, src.HEALTHCARE_EXPENSES as HEALTHCARE_EXPENSES, src.LAST as LAST, src.LAT as LAT, src.LON as LON, src.MAIDEN as MAIDEN, src.MARITAL as MARITAL, src.PASSPORT as PASSPORT, src.PREFIX as PREFIX, src.RACE as RACE, src.SSN as SSN,

src.STATE as STATE, src.SUFFIX as SUFFIX, src.ZIP as ZIP, src.id as id FROM patients_db.patients src ")

4.2 Target Query

readdatadf=spark.read.format('delimitedfile').option('delimiter',,).option('header','true').load('/app/test/data/target/patients_target.csv')

readdatadf.createOrReplaceTempView('dataview')

spark.sql("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 dataview tgt ")

5. Sample Mismatches 8 rows

5.1 Keys in source but not in target

None

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	id=00185faa-2760-4218-9bf5-db301acf8274
2	id=0042862c-9889-4a2e-b782-fac1e540ecb4
3	id=0047123f-12e7-486c-82df-53b3a450e365
4	id=010d4a3a-2316-45ed-ae15-16f01c611674
5	id=01207ecd-9dff-4754-8887-4652eda231e2
6	id=0149d553-f571-4e99-867e-fcb9625d07c2
7	id=01e1f394-7219-4189-bceb-3cbd90cff90b
8	id=023a7d29-32b3-4db5-89c8-b88bd7582ec0

6. Columnwise Mismatch Summary

None

7. Columnwise Mismatch Details

None