

## testcase27\_csv\_csv\_bigdata\_match duplicate comparison

### 1. Run Summary

Application Name	:sample
Protocol Name	:DBTablesDemo
Protocol File Path	:/app/test/testprotocol/testprotocol.xlsx
Testcase Name	:testcase27_csv_csv_bigdata_match
Testcase Type	:duplicate
Test Environment	:dev
Start Time	:18-Jan-2024 10:28:07 UTC
End Time	:18-Jan-2024 10:28:49 UTC
Run Time	:0:00:42
Test Result	:Failed
Reason	:Duplicates

### 2. Configuration Details

Compare Type	:s2tcompare
testquerygenerationmode	:Auto
Testcase Type	:duplicate
Source Connection Type	:aws-s3
Source Format	:delimitedfile
Source Path	:/app/test/data/source/ar_properties.csv
Source Exclude Columns	:start_date, end_date, operation_type, description
Target Connection Type	:aws-s3
Target Format	:delimitedfile
Target Path	:/app/test/data/target/ar_properties.csv
Target Exclude Columns	:start_date, end_date, operation_type, description
S2T Path	:/app/test/s2t/s2t_27_csv_csv_bigdata_match.xlsx
Primary Keys	:id, created_on

### 3. Duplicate Summary

Test Result	:Failed
No. of rows in Source	:3,380,253
No. of distinct rows in Source	:1,799,578
No. of duplicate rows in Source	:1,580,675
No. of rows in Target	:3,380,253
No. of distinct rows in Target	:1,799,578
No. of duplicate rows in Target	:1,580,675

### 4. SQL Queries

#### 4.1 Source Query

```
SELECT Count(DISTINCT id,created_on) FROM delimitedfile.`/app/test/data/source/ar_properties.csv`
```

#### 4.2 Target Query

```
SELECT Count(DISTINCT id,created_on) FROM delimitedfile.`/app/test/data/target/ar_properties.csv`
```