Data Quality Analysis Detailed Report

testcase1_parquet_parquet_mismatch

1. Configuration Details
1.1. Source Configuration Details

connectionname	:
connectiontype	:
path	:
format	:
name	:
excludecolumns	:
filter	:
testquerygenerationmode	:Auto
delimiter	:
test_case_name	:testcase1_parquet_parquet_mismatch
autoscripttype	:source
autoscriptpath	:
comparetype	:s2tcompare
filename	:

1.2. Target Configuration Details

connectionname	:
connectiontype	:
path	:
format	:
name	:
excludecolumns	:
filter	:
testquerygenerationmode	:Auto
delimiter	:
test_case_name	:testcase1_parquet_parquet_mismatch
autoscripttype	:target
autoscriptpath	:
comparetype	:s2tcompare
filename	:

2. Testcase Details
2.1. Testcase_ADDRESS_Length_Equal_Validation

Test Status	:Failed
Column Name	:ADDRESS
DQ Validation	:Length
Expected Length	:10
Element Count	:1000
Unexpected Count	:1000
Sample Mismatches	:['427 Balistreri Way Unit 19', '422 Farrell Path Unit 69', '1056 Harris Lane Suite 70', '999 Kuhn Forge', '1050 Lindgren Extension Apt 38', '888 Hickle Ferry Suite 38', '201 Mitchell Lodge Unit 67', '570 Abshire Forge Suite 32', '335 Morar Highlands Apt 75', '1048 Skiles Trailer', '428 Wiza Glen Unit 91', '181 Feest Passage Suite 64', '118 Bailey Orchard Suite 34', '1060 Hansen Overpass Suite 86', '970 Grant Highlands', '585 Greenfelder Camp', '661 Schulist Highlands', '383 Crooks Camp', '1062 Conn Well Suite 4', '290 Hyatt Lane']

2.2. Testcase_ADDRESS_Length_Between_Validation

Test Status	:Failed
Column Name	:ADDRESS
DQ Validation	:Length
Expected Length Range	:20-30
Element Count	:1000
Unexpected Count	:412
Sample Mismatches	:['999 Kuhn Forge', '1048 Skiles Trailer', '970 Grant Highlands', '383 Crooks Camp', '290 Hyatt Lane', '1025 Collier Arcade', '211 Effertz Quay', '547 Ziemann Burg', '560 Herman

Bridge', '687 Howell Frontage road Apt 68', '562 Windler Tunnel', '255 Ondricka Common', '703 Stehr Walk', '257 Jerde Rue', '844 Hammes Byway', '765 Harber Club', '788 Wilderman Vale', '1079 Kuhn Orchard', '314 Torp Byway', '813 Casper Street']

2.3. Testcase_id_Null_Validation

Test Status	:Failed
Column Name	:id
DQ Validation	:Null
Expected Value	:'null'
Element Count	:1000
Unexpected Count	:1000
Sample Mismatches	:['1d604da9-9a81-4ba9-80c2-de3375d59b40', '034e9e3b-2def-4559-bb2a-7850888ae060', '10339b10-3cd1-4ac3-ac13-ec26728cb592', '8d4c4326-e9de-4f45-9a4c-f8c36bff89ae', 'f5dcd418-09fe-4a2f-baa0-3da800bd8c3a', '72c0b9ce-7aa4-430b-aaff-bd0ce7846e55', 'b1e9b0b9-da6e-4f68-b603-bd896a50ca86', '01207ecd-9dff-4754-8887-4652eda231e2', 'b58731cc-2d8b-4c2d-b327-4cab771af3ef', 'cfee79fc-df05-476e-b274-43e09ea345db', 'ad2e9916-4979-40fc-a8c0-68651a0cb5a6', 'bfb6537b-535a-4f31-9a56-073220f96a17', '83719bd7-7a41-4c87-93f9-c5de4db6a14a', '76982e06-f8b8-4509-9ca3-65a99c8650fe', '2ffe9369-24e4-414b-8973-258fad09313a', 'e4f1bd35-93ae-418b-b8cf-d39c3fbb069c', '86b97fc7-ae8f-4e0d-8e66-db68f36e7a76', '1c591f0d-95b1-4289-94a9-9713ab3834e7', 'f1678bde-4814-4d71-bca9-18fdf5282232', 'aac107d8-5187-420f-a90c-448ad49dee41']

2.4. Testcase_id_NotBeNull_Validation

Test Status	:Passed
Column Name	:id
DQ Validation	:NotBeNull
Expected Value	:'not null'
Element Count	:1000
Unexpected Count	:0
Sample Mismatches	:[]

2.5. Testcase_GENDER_Unique_Validation

Test Status	:Failed
Column Name	:GENDER
DQ Validation	:Unique
Expected Value	:to be unique
Element Count	:1000
Unexpected Count	:1000
Sample Duplicate Values	:['F', 'F', 'F', 'F', 'F', 'F', 'F', 'F', 'F', 'F', 'F', 'F', 'F', 'F', 'F', 'F', 'F', 'F', 'F']

2.6. Testcase_GENDER_DistinctSet_Validation

Test Status	:Failed
Column Name	:GENDER
DQ Validation	:DistinctSet
Expected Distinct Value	:['F','M']
Actual Distinct Value	:['F', 'M']

2.7. Testcase_HEALTHCARE_COVERAGE_Sum_Between_Validation

Test Status	:Failed
Column Name	:HEALTHCARE_COVERAGE
DQ Validation	:Sum
Expected Sum Range	:100029-3000039
Actual Sum	:12911883.180000005

2.8. Testcase_ColumnCount_Validation

Test Status	:Failed
DQ Validation	:ColumnCount
Expected Column Count	:40
Actual Column Count	:26

2.9. Testcase_ColumnOrder_Validation

Test Status	:Failed
DQ Validation	:ColumnOrder
Expected Column Order	:['id', 'BIRTHDATE', 'DEATHDATE', 'SSN', 'DRIVERS', 'PASSPORT', 'PREFIX', 'FIRST', 'LAST', 'SUFFIX', 'MAIDEN', 'MARITAL', 'RACE', 'ETHNICITY', 'GENDER', 'BIRTHPLACE', 'ADDRESS', 'CITY', 'STATE', 'COUNTY', 'ZIP', 'LAT', 'LON', 'HEALTHCARE_EXPENSES', 'DName', 'HEALTHCARE_COVERAGE']
Actual Column Order	:['Unnamed: 0', 'ADDRESS', 'BIRTHDATE', 'BIRTHPLACE', 'CITY', 'COUNTY', 'DEATHDATE', 'DRIVERS', 'ETHNICITY', 'FIRST', 'GENDER', 'HEALTHCARE_COVERAGE', 'HEALTHCARE_EXPENSES', 'LAST', 'LAT', 'LON', 'MAIDEN', 'MARITAL', 'PASSPORT', 'PREFIX', 'RACE', 'SSN', 'STATE', 'SUFFIX', 'ZIP', 'id']
Column Order Mismatch	:[{'Expected Column Position': 0, 'Expected': 'id', 'Found': 'Unnamed: 0'}, {'Expected Column Position': 1, 'Expected': 'BIRTHDATE', 'Found': 'ADDRESS'}, {'Expected Column Position': 2, 'Expected': 'DEATHDATE', 'Found': 'BIRTHDATE'}, {'Expected Column Position': 3, 'Expected': 'SSN', 'Found': 'BIRTHPLACE'}, {'Expected Column Position': 4, 'Expected': 'DRIVERS', 'Found': 'CITY'}, {'Expected Column Position': 5, 'Expected': 'PASSPORT', 'Found': 'COUNTY'}, {'Expected Column Position': 6, 'Expected': 'PREFIX', 'Found': 'DEATHDATE'}, {'Expected Column Position': 7, 'Expected': 'FIRST', 'Found': 'DRIVERS'}, {'Expected Column Position': 8, 'Expected': 'LAST', 'Found': 'ETHNICITY'}, {'Expected Column Position': 9, 'Expected': 'SUFFIX', 'Found': 'FIRST'}, {'Expected Column Position': 10, 'Expected': 'MAIDEN', 'Found': 'GENDER'}, {'Expected Column Position': 11, 'Expected': 'MARITAL', 'Found': 'HEALTHCARE_COVERAGE'}, {'Expected Column Position': 12, 'Expected': 'RACE', 'Found': 'HEALTHCARE_EXPENSES'}, {'Expected Column Position': 13, 'Expected': 'ETHNICITY', 'Found': 'LAST'}, {'Expected Column Position': 14, 'Expected': 'GENDER', 'Found': 'LAT'}, {'Expected Column Position': 15, 'Expected': 'BIRTHPLACE', 'Found': 'LON'}, {'Expected Column Position': 16, 'Expected': 'ADDRESS', 'Found': 'MAIDEN'}, {'Expected Column Position': 17, 'Expected': 'CITY', 'Found': 'MARITAL'}, {'Expected Column Position': 18, 'Expected': 'STATE', 'Found': 'PASSPORT'}, {'Expected Column Position': 19, 'Expected': 'COUNTY', 'Found': 'PREFIX'}, {'Expected Column Position': 20, 'Expected': 'ZIP', 'Found': 'RACE'}, {'Expected Column Position': 21, 'Expected': 'LAT', 'Found': 'SSN'}, {'Expected Column Position': 22, 'Expected': 'LON', 'Found': 'STATE'}, {'Expected Column Position': 23, 'Expected': 'HEALTHCARE_EXPENSES', 'Found': 'SUFFIX'}, {'Expected Column Position': 24, 'Expected': 'DName', 'Found': 'ZIP'}, {'Expected Column Position': 25, 'Expected': 'HEALTHCARE_COVERAGE', 'Found': 'id'}]

2.10. Testcase_BIRTHDATE_Regexp_Validation

Test Status	:Failed
column Name	:BIRTHDATE
DQ Validation	:Regexp
Expected Pattern	:'^\d{4}-\d{2}-\d{2}$'
Element Count	:1000
Unexpected Count	:1000
Sample Mismatches	:[datetime.datetime(1989, 5, 25, 0, 0), datetime.datetime(1983, 11, 14, 0, 0), datetime.datetime(1992, 6, 2, 0, 0), datetime.datetime(1978, 5, 27, 0, 0), datetime.datetime(1996, 10, 18, 0, 0), datetime.datetime(2017, 7, 27, 0, 0), datetime.datetime(2003, 12, 13, 0, 0), datetime.datetime(2019, 5, 15, 0, 0), datetime.datetime(1970, 5, 16, 0, 0), datetime.datetime(2016, 7, 4, 0, 0), datetime.datetime(2004, 12, 19, 0, 0), datetime.datetime(1991, 7, 3, 0, 0), datetime.datetime(1989, 6, 7, 0, 0), datetime.datetime(1982, 9, 1, 0, 0), datetime.datetime(1958, 7, 1, 0, 0), datetime.datetime(1957, 2, 25, 0, 0), datetime.datetime(1959, 5, 26, 0, 0), datetime.datetime(1954, 4, 4, 0, 0), datetime.datetime(1958, 4, 11, 0, 0), datetime.datetime(1983, 9, 2, 0, 0)]