Running the DQD on a Cohort

Clair Blacketer

2024-01-07

Contents

1 DQD Cohort Functionality

Running the Data Quality Dashboard for a cohort is fairly straightforward. There are two options in the `executeDqChecks` function, `cohortDefinitionId` and `cohortDatabaseSchema`. These options will point the DQD to the schema where the cohort table is located and provide the id of the cohort on which the DQD will be run. By default, the tool assumes that the table being referenced is the standard OHDSI cohort table named **COHORT** with at least the columns **cohort_definition_id** and **subject_id**. For example, if I have a cohort number 123 and the cohort is in the *results* schema of the *IBM_CCAE* database, the `executeDqChecks` function would look like this:

```
DataQualityDashboard::executeDqChecks(connectionDetails = connectionDetails,
                                      cdmDatabaseSchema = cdmDatabaseSchema,
                                      resultsDatabaseSchema = resultsDatabaseSchema,
                                      cdmSourceName = "IBM_CCAE_cohort_123",
                                      cohortDefinitionId = 123,
                                      cohortDatabaseSchema = "IBM_CCAE.results",
                                      cohortTableName = "cohort",
                                      numThreads = numThreads,
                                      sqlOnly = sqlOnly,
                                      outputFolder = outputFolder,
                                      verboseMode = verboseMode,
                                      writeToTable = writeToTable,
                                      writeTableName = "dqdashboard_results_123",
                                      checkLevels = checkLevels,
                                      tablesToExclude = tablesToExclude,
                                      checkNames = checkNames)
```

As a note, it is good practice to have the `cdmSourceName` option and the `writeTableName` option reflect the name of the cohort so that the results don't get confused with those of the entire database.