Excent – Enrich Datavalidation Tool

By

Muthu veera

Dsdar ExcentDBA

## Data Origin

### 1.1 Flat files

Following table describes the source files (CSV files) having data to import into enrich database (into DATAVALIDATION Schema).

|  |  |
| --- | --- |
| Source file name | Description |
| District | This file have the all the district names and district codes of Aurora under Colorado state. |
| Goal | This file has IEP reference id, sequence, Goal area state assigned code, Goal statement. |
| Iep | This file has student's special education records, their Iep start date, Iep end date, next review dates.. etc |
| Selectlists | This file has the details about lookup's table LRE, Exit, ServLoc..etc and their Codes |
| Objective |  |
| School | This file has the school code, name, DistrictRefId, MinutesPerWeek. |
| Service | This file has the Service related date, Service Type (SPED, Related), IEPRefID, Service Definition code, Begindate, End date..etc |
| SpedStaffMember | This file has Special education staff records stafflocalID, name, email etc |
| Student | This file has student records student localId, stateId, name ,gender, birthdate, ethnicity code , Disabiltycode, special education status.etc |
| TeamMember | This file has the records SpedStaffRefID, SpedStudentRefID, ISCASEMANAGER (?). |
| StaffSchool | This file has the records SpedStaffEmailID, SchoolCode. |

### 1.2 Source Data from Database

Following table describes the source view having data to import data into enrich database (into DATAVALIDATION Schema). The view should be in dbo schema in Source database (for ex: Excent Online database).

|  |  |
| --- | --- |
| Source view name | Description |
| dbo.District\_EO | This view have the all the district names and district codes. |
| dbo.Goal\_EO | This view has IEP reference id, sequence, Goal area state assigned code, Goal statement. |
| dbo.IEP\_EO | This view has student's special education records, their Iep start date, Iep end date, next review dates.. etc |
| dbo. Selectlists\_EO | This view has the details about lookup's table LRE, Exit, ServLoc.etc and their Codes |
| dbo.Objective\_EO | This view has the details about ObjectiveStatement, Goal Reference ID, sequence. |
| dbo.School\_EO | This view has the school code, name, DistrictRefId, MinutesPerWeek. |
| dbo.Service\_EO | This view has the Service related date, Service Type (SPED, Related), IEPRefID, Service Definition code, Begindate, End date..etc |
| dbo.SpedStaffMember\_EO | This view has Special education staff records stafflocalID, name, email etc |
| dbo.Student\_EO | This view has student records student localId, stateId, name ,gender, birthdate, ethnicity code , Disabiltycode, special education status.etc |
| dbo.TeamMember\_EO | This view has the records SpedStaffRefID, SpedStudentRefID, ISCASEMANAGER (?). |
| dbo.StaffSchool\_EO | This view has the records SpedStaffEmailID, SchoolCode. |

### 2.1 Flow Chart for Data validation process with Flat file:

Flat files (Source Data Files)

Ex: District.csv

Number of fields in Header with other rows

Validation Summary Report

Validated Data Ex: Datavalidation.student (in Enrich Database)

5

4

3

2

1

Fail

Success

If not match

If match

Validation Report

Checking with Enrich specifications

Validation Report History

Raw Data (Before validate data with Enrich specification) Ex: Datavalidation.Student\_local

### 2.2 Flow Chart for Data validation process with Database:

Fail

Success

Validation Report History

Validation Report

Validation Summary Report

Validated Data Ex: Datavalidation.student (in Enrich Database)

Checking with Enrich specifications

Source Database (Database views) ex: dbo.Student\_EO

Ex DB: ExcentOnline

Raw Data (Before validate data with Enrich specification) Ex: Datavalidation.Student\_local

|  |  |
| --- | --- |
| Table Name | Description |
| DATAVALIDATION.ValidationReport | To store validation details for that iteration (storing single iteration) |
| DATAVALIDATION .ValidationSummaryReport | To store validation summary of that iteration (storing single iteration) |
| DATAVALIDATION .ValidationReportHistory | To store validation details history for last 60 days. |
| DATAVALIDATION .ValidationRules | To store validation rules (such as data length, data type, referential integrity, etc., ). |

### 3. Datavalidation Schema

And DATAVALIDATION.SelectLists\_LOCAL has the records before the data validation (same for all data files). And DATAVALIDATION.SelectLists has the validated records after the data validation (same for all data files).

#### 3.1 Data Flow:

1. Clean up the data validation report of previous iteration and clean up the validation report history of older records (for ex: more than 60 days).
2. Import data (from Flat file or Data from Database) into the \_Local (Ex: Datavalidation.Student\_LOCAL) table and before that checking the number of fields in header row of the file with number of fields in other rows and if not those records will be stored in Validation report. And if record does not satisfy the condition, those records will be stored in validation report table.
3. Check the records with Enrich Specifications. And if record does not satisfy the enrich specification, those records will be stored in data validation report with respective issue details and Create the validation summary report for that iteration.
4. Import the validated data into the main table (Ex: Datavalidation.Student), if records are satisfied with enrich specifications.
5. Create the validation report history for that iteration.

**Notes:**

In validation rules, for checking the composite referential key, checking the date range issues like, we are created SQL queries directly instead of creating dynamic SQL.

### 3.2 ValidationRules:

This table (DATAVALIDATION.ValidationRules) has the details of each column of the table (like Datatype, Datalength, IsRequired, IsUniqueColumn, IsUniqueField, ParentTable, ParentColumn, etc.,) . If the data specifications change, we will change the specification checking process by changing the records in this table.

For some rules, we are created the dynamic SQL to check (for ex: Checking the Composite Unique field).

### 4. Steps to Integrate the Data validation tool with Enrich:

### 4.1 For Flat file data validation

1. In setupETL\_Faltfile.sql file in VC3ETL.ExtractDatabase, change the database name as Enrich database (in which database we are going to import the data).
2. In setupETL\_Faltfile.sql file in VC3ETL.LoadTable, change the data file location.

### 4.2 For validating data from Database:

1. In setupETL\_EO.sql file in VC3ETL. ExtractDatabase, change the database name as Source database name (in which database we are going to get the data).

### 5. Steps to do Data validation thorough UI:

1. Upgrade the Enrich database by using upgrade\_db\_script.bat.
2. Keep the data file in data file location.
3. Keep the ValidationReport.bat & ValidationReport.sql in ValidationReportfiles directory and Change the database name & Validationreport files location in Validationreport.sql.
4. Run the Data validation through Enrich UI.
5. Run the ValidationReport.bat to prepare the data validation report files.

