

QRG - HMIS CSV Data Quality Tool

Table of Contents

WHAT IS A QRG	1
WHO THIS GUIDE IS FOR	1
RUN THE HMIS CSV EXPORT IN CASEWORTHY	2
RETRIEVE THE CSV EXPORT IN HMIS	2
UNZIP HMIS CSV EXPORT FILES	3
RUN THE HUD HMIS CSV DATA QUALITY TOOL	3
NAVIGATING THE TABS IN THE TOOL	5
EXPORTING FLAGGED RECORDS	8
INVESTIGATING DATA ERRORS	8
ADDITIONAL RESOURCES	9

WHAT IS A QRG – QRG stands for **Quick Reference Guide**. Quick Reference Guides are short 1-8 page documents that provide concise practical information and advice on specific topics. Sometimes known as “cheat sheets” quick reference guides are designed for users who know the material or task but need something in front of them to remind them of the steps.

WHO THIS GUIDE IS FOR – The HMIS CSV Data Quality (DQ) Tool was developed to identify common data quality issues found in the CSV export. The tool helps reduce the manual work of finding which client record is causing the error. Please note that this tool can be used with any HMIS system that generates CSV reports.

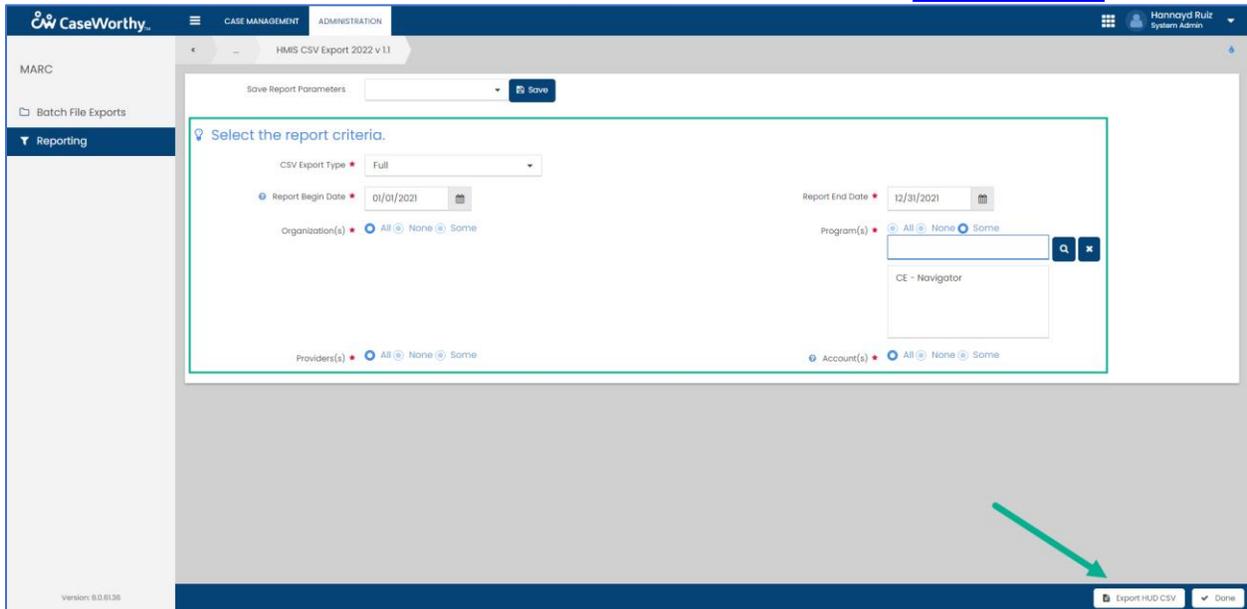
This guide is useful for System Admins who will be reviewing data quality for their programs.

RUN THE HMIS CSV EXPORT IN CASEWORTHY

Role: System Admin

Navigation Path: Admin > Reporting > Compliance > 2022 HUD HMIS CSV

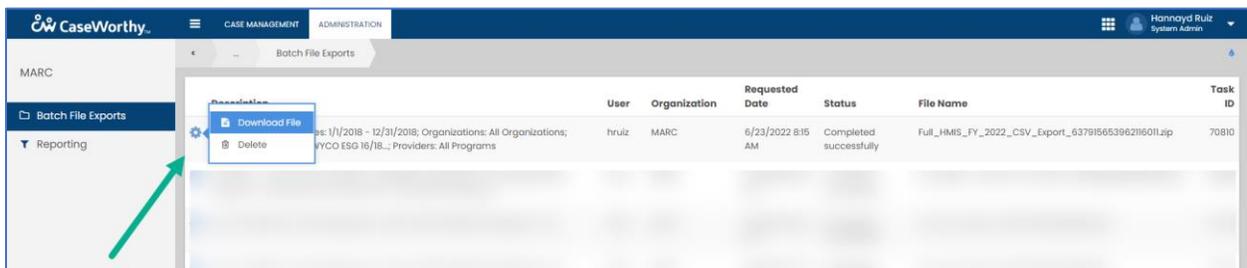
Select CSV Export Type: **Full** along with other reporting parameters and click **Export HUD CSV**



Save the CSV file to a folder you can easily locate.

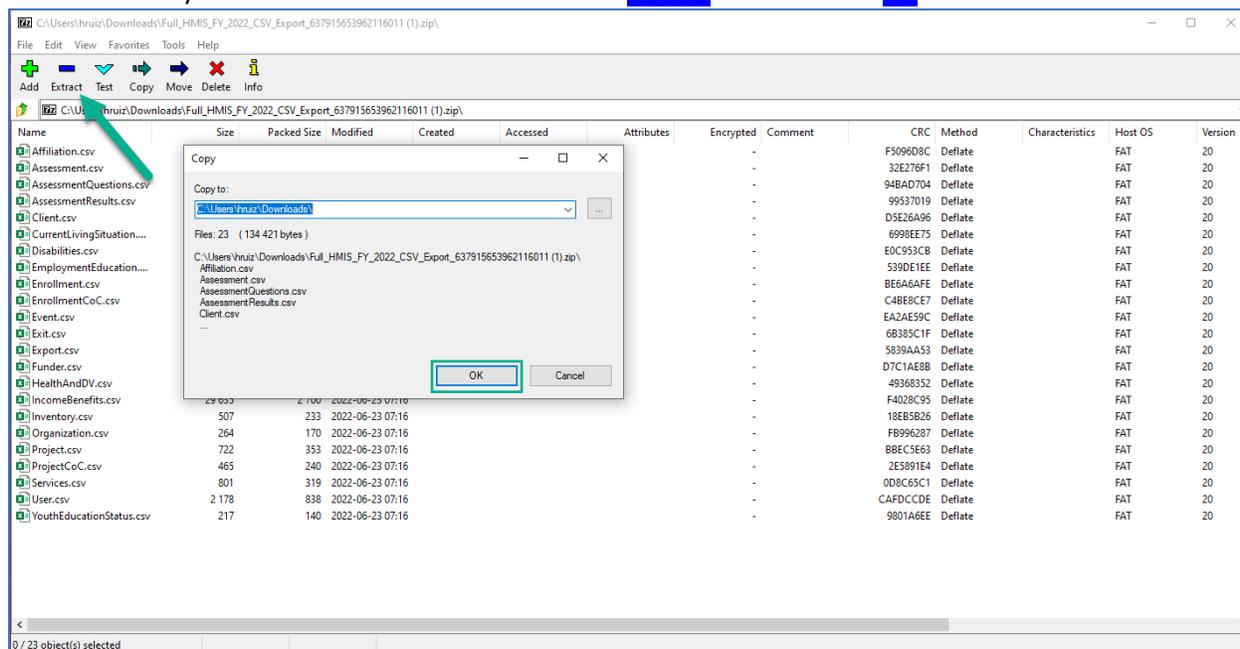
RETRIEVE THE CSV EXPORT IN HMIS

Navigation Path: Admin > Batch File Exports > Gear > Download File



UNZIP HMIS CSV EXPORT FILES

Go to the newly downloaded CSV file and Select the **Extract** icon and click **OK**



RUN THE HUD HMIS CSV DATA QUALITY TOOL

Click on the link to download and open the Data Quality Tool on your computer:

https://bitbucket.org/abtassociates/hmis-csv-dq-tool/downloads/HMIS_CSV_Data_Quality_Tool_v1.4.2.xlsb

Step 1: Locate and Check CSVs

- Select the **Locate and Check CSVs** button. This button prompts you to locate the folder where the HMIS CSVs are saved. You only need to select the main folder where the CSVs are saved, not the individual CSV files.
- After the CSVs are located and checked, the CSV Check Report tab will appear. Additionally, the *Raw CSV Files Folder*, *Initial Cleaned CSV Folder*, and *CSV Check Report* boxes will appear and populate:
 - The *Raw CSV Files Folder* shows the folder path of the main folder that you selected.
 - The *Initial Cleaned CSV Folder* generates a hyperlink to a folder where alternate versions of the CSV files are saved. These files are a version of the CSV files with a subset of columns and in some cases a subset of rows. They will be used for the validation and data quality checks. Do **not** alter or delete these files.
 - The **“Go to Report”** hyperlink navigates you to the second tab in the workbook – the **CSV Check Report**.
 - After reviewing the CSV Checks, return to the **Controller tab** by clicking *‘Controller’* at the bottom left of the screen.

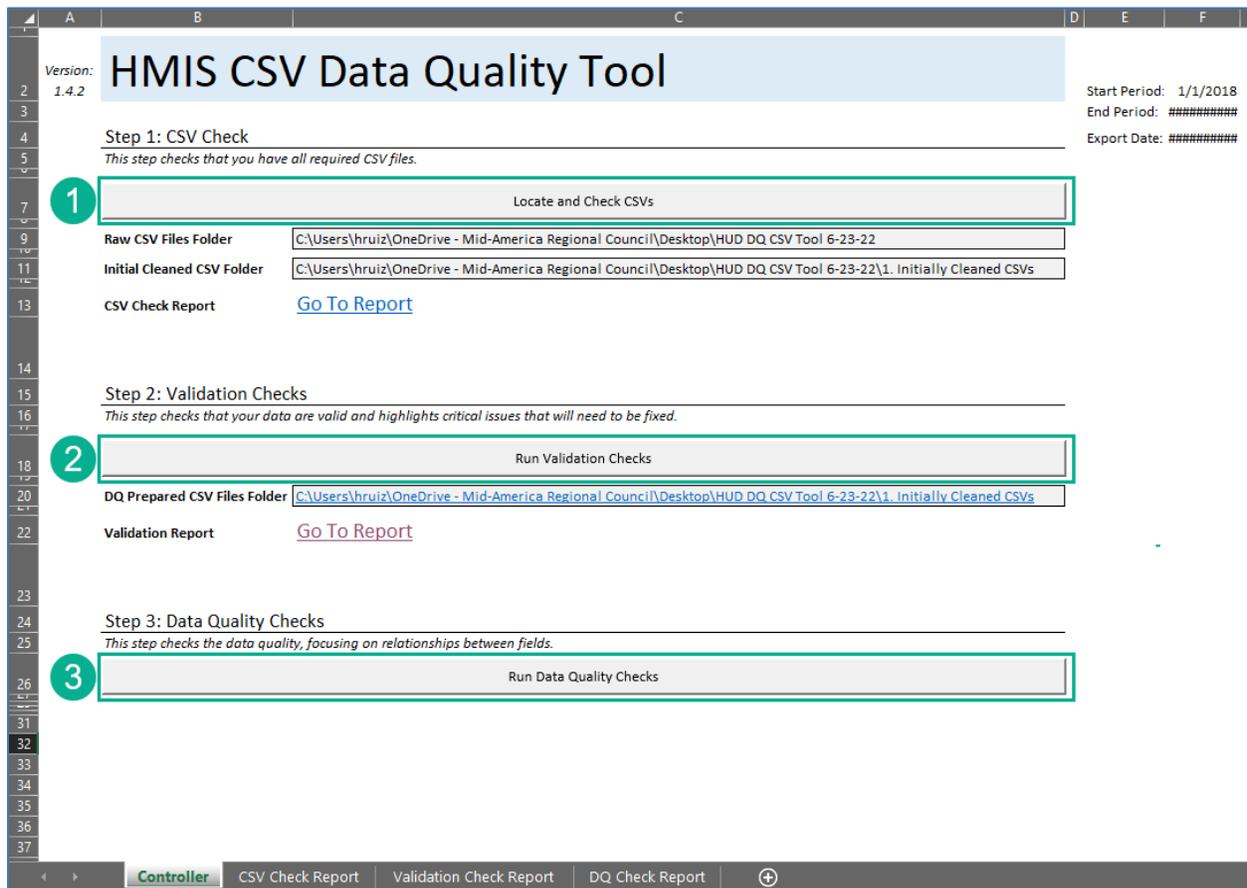
Step 2: Run Validation Checks

- Select the **Run Validation Checks** button. This button prompts you to confirm that they are ready to run the validation checks. To run the checks, select **OK**. Once the checks are complete, the **Validation Check Report tab** will appear.
- The tool lets you know that the data passed all the validation checks OR that you need to address validation issues to move on to the next step – DQ checks.

- The “Go to Report” hyperlink navigates you to the third tab in the workbook – Validation Check Report.

Step 3: Run Data Quality Checks

- You will select the Run Data Quality Checks button (shown in the screenshot below). This button will prompt you to confirm that they are ready to run the data quality checks. To run the checks, select **OK**. Once the checks are complete, the **DQ Check Report tab** will appear.
- The tool will either let you know that the data has passed all the data quality checks OR that you needs to address a data error or warning displayed on the **DQ Check Report tab**.
- The “Go to Report” hyperlink navigates you to the fourth tab in the workbook – **DQ Check Report**.



NAVIGATING THE TABS IN THE TOOL

CSV Check Report Tab— This tab displays the results of the CSV checks and allows you to access each file through a hyperlink.

CSV Check Report

The Standard HMIS CSV export contains 23 CSV files. Only the files referenced by the data quality checks are shown below. Not all files are required to have data depending on the HMIS implementation. Please see the user guide for definitions of each column

Table 1: File Checks

File Name	File in Folder?	# Records	Missing Column	Required?
Client	Yes	24	No	TRUE
Enrollment	Yes	38	No	TRUE
Exit	Yes	19	No	TRUE
Services	Yes	4	No	FALSE
CurrentLivingSituation	Yes	0	No	FALSE
Project	Yes	2	No	TRUE
Inventory	Yes	1	No	TRUE
EnrollmentCoC	Yes	25	No	TRUE
Organization	Yes	1	No	TRUE

You have all required CSVs.
 However, some non-essential files do not have data. If that is expected, please proceed. If not, please inquire with your vendor.

- a) File Name**— The file name and a hyperlink that opens the CSV file.
- a) File in Folder?**— If the CSV file exists in the main folder.
 - o If all the necessary CSV files are present, then green text to the right of the table will say “You have all required CSVs!”
 - o If the main folder is missing a file, then the text to the right of the table will say, “You are missing some CSVs!”
- b) # Records**— The number of records in each CSV file.
- c) Missing Columns?**— If the CSV file is missing any columns that should be present in the file according to the FY 2022 HMIS Data Standards.
- d) Required?**— If the CSV file is required for the validation or data quality checks.

Validation Check Report Tab—This tab displays critical errors (invalid values, and duplicate primary IDs) in select data fields across the files. It also allows you to investigate potential issues in your data by displaying the relevant IDs associated with each data error.

Validation Check Report

Export All Flagged Records

This report is ensuring that there are no invalid values (values not listed in the 'Valid Values' column) for the fields listed below. Please see the user guide for examples of invalid values. Additionally, this report looks for duplication of the primary key for each file. Any duplication can be examined by double clicking the corresponding cell in the Duplicate Primary ID column

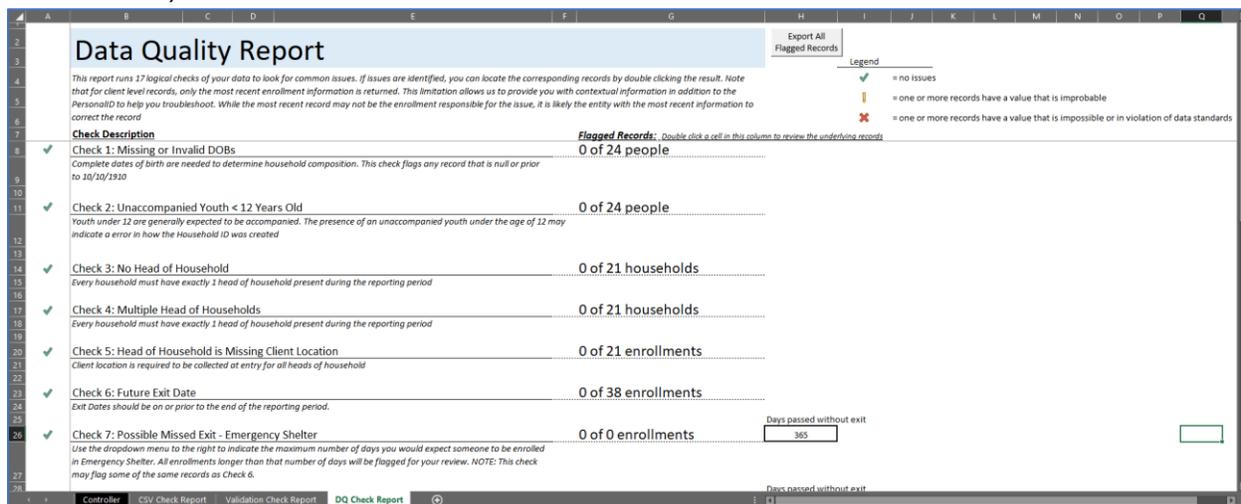
Double click on a cell to see the underlying IDs

Table 2: Validation Flags

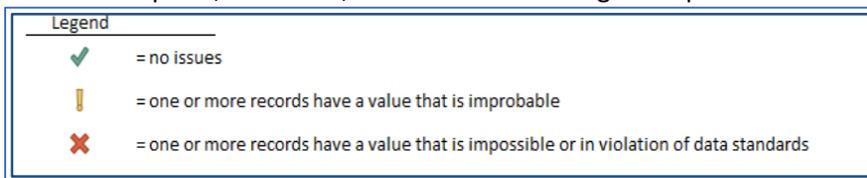
File Name	Field	Valid Values	Invalid Value	Duplicate Primary IDs	Total Flags
Client	Male	0,1,99	0 (0%)	N/A	0
Client	NoSingleGender	0,1,99	0 (0%)	N/A	0
Client	Transgender	0,1,99	0 (0%)	N/A	0
Client	Questioning	0,1,99	0 (0%)	N/A	0
Enrollment	EnrollmentID	ID	N/A	0 (0%)	0
Enrollment	DisablingCondition	0,1,8,9,99	0 (0%)	N/A	0
Enrollment	LivingSituation	1,2,3,4,5,6,7,8,9,10,11,14,15,16,18,19,20,21,25,26,28,29,31,32,33,34,35,36,99,NULL	0 (0%)	N/A	0
Enrollment	RelationshipToHoH	1,2,3,4,5,99	0 (0%)	N/A	0
Exit	ExitDate	D	0 (0%)	N/A	0
Exit	Destination	1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32	0 (0%)	N/A	0
Services	DateProvided	D	0 (0%)	N/A	0
CurrentLivingSituation	CurrentLivingSituation	1,2,3,4,5,6,7,8,9,10,11,14,15,16,17,18,19,20,21,25,28,29,31,32,33,34,35,36,37,99	0 (0%)	N/A	0
CurrentLivingSituation	InformationDate	D	0 (0%)	N/A	0
Project	ProjectID	ID	N/A	0 (0%)	0
Project	ProjectType	1,2,3,4,6,7,8,9,10,11,12,13,14	0 (0%)	N/A	0
Inventory	InventoryID	ID	N/A	0 (0%)	0
EnrollmentCoC	CoCCode	S6	0 (0%)	N/A	0
EnrollmentCoC	EnrollmentCoCID	ID	N/A	0 (0%)	0
Organization	OrganizationID	ID	N/A	0 (0%)	0

- a) **File Name**—The CSV file name associated with the data field.
- b) **Field**—The name of the data field.
- c) **Valid Names**— Valid values associated with that data field.
- d) **Invalid Values**— The number and percent of invalid values.
- e) **Duplicate Primary IDs**— The number and percent of duplicate primary IDs for some data fields.
- f) **Total Flags**— The total number of flags associated with each data field.

DQ Check Report Tab—This tab displays select data quality checks and allows you to investigate data issues by showing the relevant IDs. The data quality checks fall into two categories: errors and warnings. Errors are critical issues that need to be fixed and warnings are potential issues that could be valid, but also could be an error.



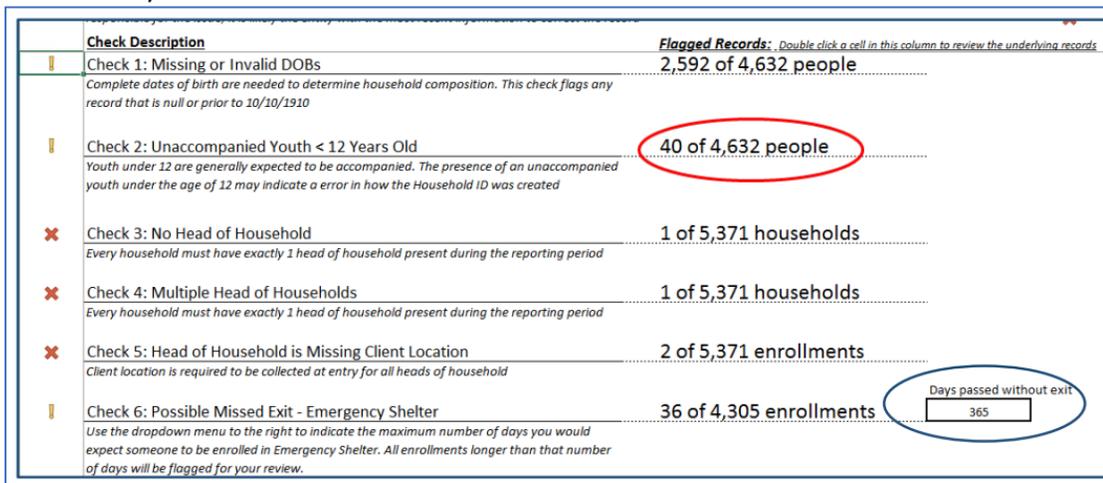
a) **Check Status**— the first column includes the status of the check which is represented by either an exclamation point, an x mark, or a check mark. A legend is provided for the status symbols.



b) **Check Description**— the second column includes the name and a description of the data check.

c) **Flagged Records**— the third column lists the number of errors/warnings by unit for each check. The units (households, enrollments, or people) vary by check.

d) **Time Period**— the fourth column includes a drop-down menu where you can select different time periods which will then apply to the check (circled in blue in the screenshot below). This feature is only available for some checks.



EXPORTING FLAGGED RECORDS

You all the data quality flags by selecting the **Export All Flagged** button in the upper right corner. This will generate a workbook with a tab for each check containing the flagged records.

INVESTIGATING DATA ERRORS

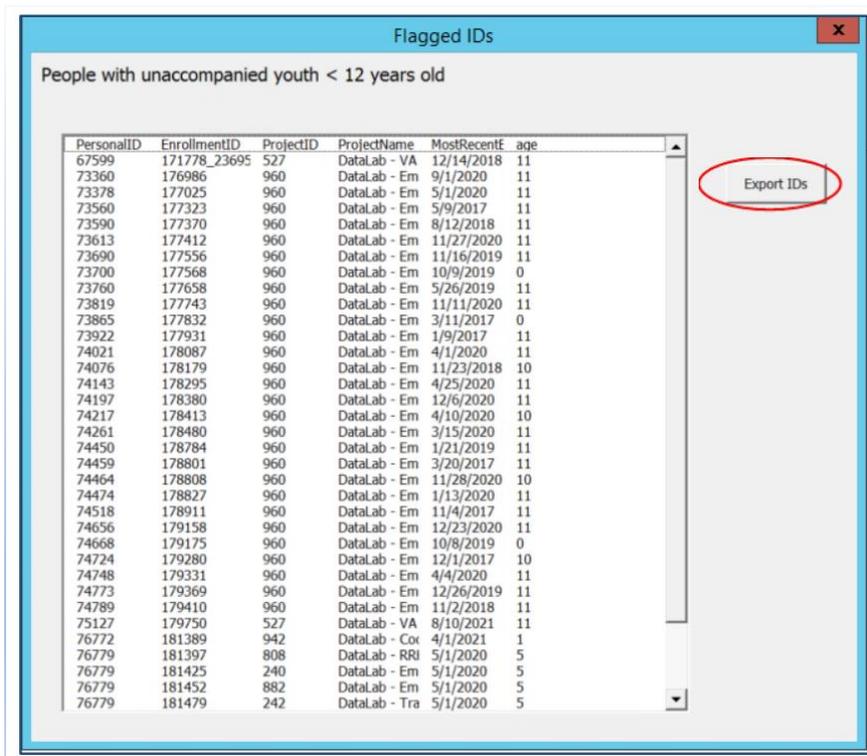
Validation Check Report Tab— The third column of the DQ report displays the number or records with a specific data error or warning. When double clicked each cell shows the underlying IDs associated with that data error/warning in the **Flagged IDs** window (like the **Validation Check Report**). For example, in the screenshot above, the workbook shows that there are 40 people under the age of 12 that are unaccompanied youth. This is possible, but not likely given that youth under the age of 12 are usually accompanied. If you double clicks the cell circled in red then the window in the screenshot below will appear. You can then select the **Export IDs** button to export the information in the window into Excel.

Table 2: Validation Flags

Double click on a cell to see the underlying IDs

File Name	Field	Valid Values	Invalid Value	Duplicate Primary IDs	Total Flags
Client	BlackAfAmerican	0,1,99	0 (0%)	N/A	0
Client	NativeHIPacific	0,1,99	0 (0%)	N/A	0
Client	White	0,1,99	0 (0%)	N/A	0
Client	Ethnicity	0,1,8,9,99	0 (0%)	N/A	0
Client	GenderNone	8,9,99,NULL	0		0
Client	Female	0,1,99	0		0
Client	Male	0,1,99	0		0
Client	NoSingleGender	0,1,99	0		0
Client	Transgender	0,1,99	0		0
Client	Questioning	0,1,99	0		0
Enrollment	EnrollmentID	ID	N/A	0	0
Enrollment	DisablingCondition	0,1,8,9,99	0 (0%)	N/A	0
Enrollment	LivingSituation	1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33	289 (9%)	N/A	289
Enrollment	RelationshipToHoH	1,2,3,4,5,99	0 (0%)	N/A	0
Exit	ExitDate	D	0 (0%)	N/A	0
Exit	Destination	1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34	1 (0%)	N/A	1
Services	DateProvided	D	0 (0%)	N/A	0
CurrentLivingSituation	CurrentLivingSituation	1,2,3,4,5,6,7,8,9,10,11,14,15,16,17,18,19,20,21,25,28,29,31,32,33,34,35,36,37,99	0 (0%)	N/A	0
Assessment	AssessmentID	ID	N/A	0 (0%)	0
AssessmentResults	AssessmentResultID	ID	N/A	0	0
Project	ProjectID	ID	N/A	0 (0%)	0
Project	ProjectType	1,2,3,4,6,7,8,9,10,11,12,13,14	0 (0%)	N/A	0
Inventory	InventoryID	ID	N/A	0 (0%)	0
EnrollmentCoC	CoCCode	56	0 (0%)	N/A	0
EnrollmentCoC	EnrollmentCoCID	ID	N/A	0	0
Organization	OrganizationID	ID	N/A	0	0

DQ Check Report—The third column of the DQ report displays the number of records with a specific data error or warning. When double clicked each cell shows the underlying IDs associated with that data error/warning in the **Flagged IDs** window (like the **Validation Check Report**). For example, in the screenshot above, the workbook shows that there are 40 people under the age of 12 that are unaccompanied youth. This is possible, but not likely given that youth under the age of 12 are usually accompanied. If you double click the cell circled in red then the window in the screenshot below will appear. You can then select the Export IDs button to export the information in the window into Excel.



ADDITIONAL RESOURCES

- HUD HMIS CSV Data Quality Tool:
- <https://bitbucket.org/abtassociates/hmis-csv-dq-tool/downloads/>
- HUD HMIS CSV Data Quality Tool XLS: [https://bitbucket.org/abtassociates/hmis-csv-dq-tool/downloads/HMIS CSV Data Quality Tool v1.4.2.xlsb](https://bitbucket.org/abtassociates/hmis-csv-dq-tool/downloads/HMIS_CSV_Data_Quality_Tool_v1.4.2.xlsb)
- HUD HMIS CSV Data Quality Tool User Guide: [https://bitbucket.org/abtassociates/hmis-csv-dq-tool/downloads/User Guide Data Quality Tool v1 508 compliant.pdf](https://bitbucket.org/abtassociates/hmis-csv-dq-tool/downloads/User_Guide_Data_Quality_Tool_v1_508_compliant.pdf)