



ScHARe

ScHARe Repository Introduction

October 16, 2024

Deborah Duran, PhD • NIMHD

Luca Calzoni, MD MS PhD Cand. • NIMHD

Elif Dede Yildirim, PhD • NIMHD



ScHARe

CDE Mapping

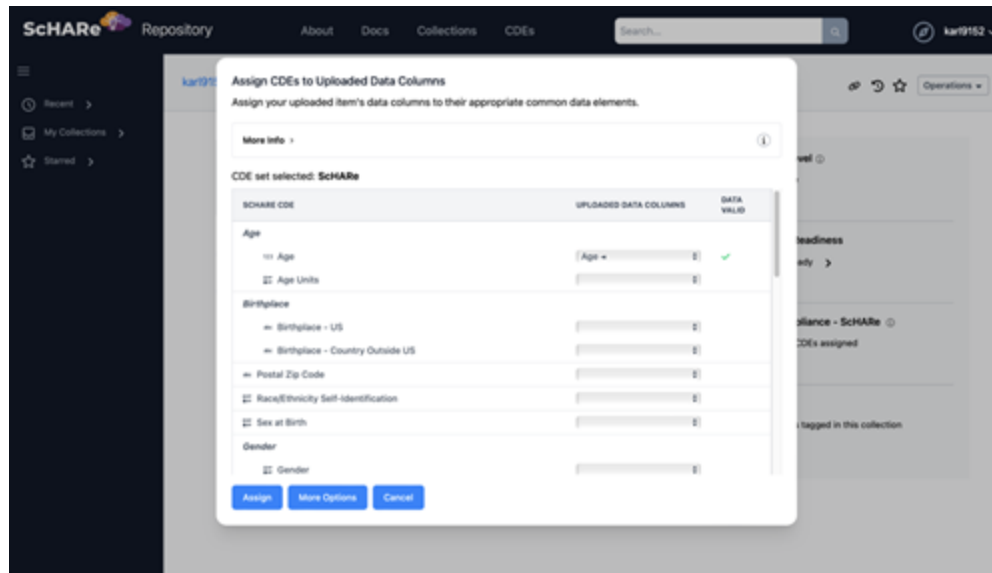


BE A PART OF THE FUTURE
OF KNOWLEDGE GENERATION

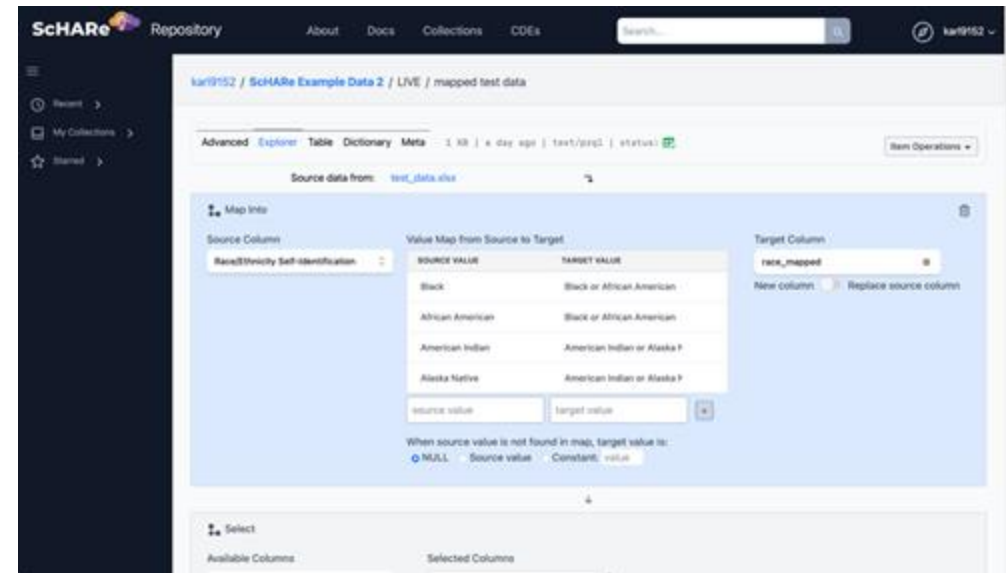
Getting Started with Mapping CDEs

- The Repository does not *enforce* CDEs, but it does *encourage* using CDEs
- Start with the data you have, and use the tools the repository offers to map as much of your data to CDEs as possible

Assign CDEs at Upload



Map CDEs using Dataviews



Assigning CDEs at Upload

File Upload Complete

The ScHARe Repository uses the ScHARe common data elements to enable data discovery, linking, and interoperability. To make your data more interoperable and ready for publication, we will guide you through the process of assigning common data elements to your data column headings. Click OK to continue.

TIP: You can cancel now and complete the process later using the "Configure CDEs" tool from the collection Operations menu.

OK

Cancel

- After you upload a tabular data file, the system will recommend that you begin the process of assigning CDEs to your data column headings.

Match CDEs with column headings

The first step is to automatically detect alignments between your data and the ScHARe CDE set. Unless otherwise indicated, please select ScHARe below.

Select the CDEs you want to assign:

ScHARe (19 CDEs)

CDE matching method:

Using column headers

Next

Cancel

Assigning CDEs at Upload

Some of the assignments will be correct -- double-check the validation and the column name matches what you expect

Some of the assignments will be wrong or missing -- add anything that should be added, and double-check the validation

Complete or cancel the process, or use "More Options" to go to a detailed tool for CDE assignment

Assign CDEs to Uploaded Data Columns

Assign your uploaded item's data columns to their appropriate common data elements.

More Info >

CDE set selected: SchARE

SCHARE CDE	UPLOADED DATA COLUMNS	DATA VALID
Age		
123 Age	Age	✓
Age Units	Age Units	✓
Birthplace		
Birthplace - US		
Birthplace - Country Outside US		
Postal Zip Code		
Race/Ethnicity Self-Identification	Race/Ethnicity Self-Identi	✗
Gender		
Gender		
Gender - Select Other		
Gender - Specify		
Sexual Orientation		

Row 1 "American Indian" is not one of ["American I...
Row 2 "Alaska Native" is not one of ["American Indi...
Row 3 "Asian" is not one of ["American Indian or AI...
Row 4 "Asian American" is not one of ["American I...
Row 5 "Black" is not one of ["American Indian or AI...
Row 6 "African American" is not one of ["American ...
Row 7 "Hispanic" is not one of ["American Indian o...
Row 8 "Latino" is not one of ["American Indian or A...
Row 9 "Spanish" is not one of ["American Indian or...
Row 10 "Native Hawaiian" is not one of ["American ...

Assign More Options Cancel

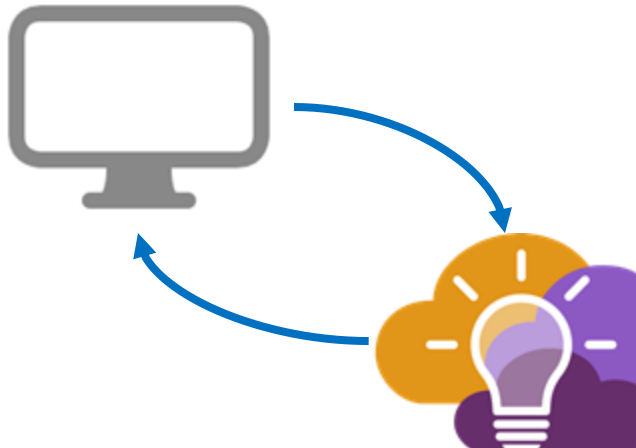
When an assigned CDE fails validation, hover over the red X to get details on why

NOTE: CDE validation errors can be corrected using dataviews (coming up later)

Using Dataviews to Map CDEs

If your data doesn't conform to the CDEs as it was uploaded, you have two options:

Use your own tools to adjust the data and re-upload



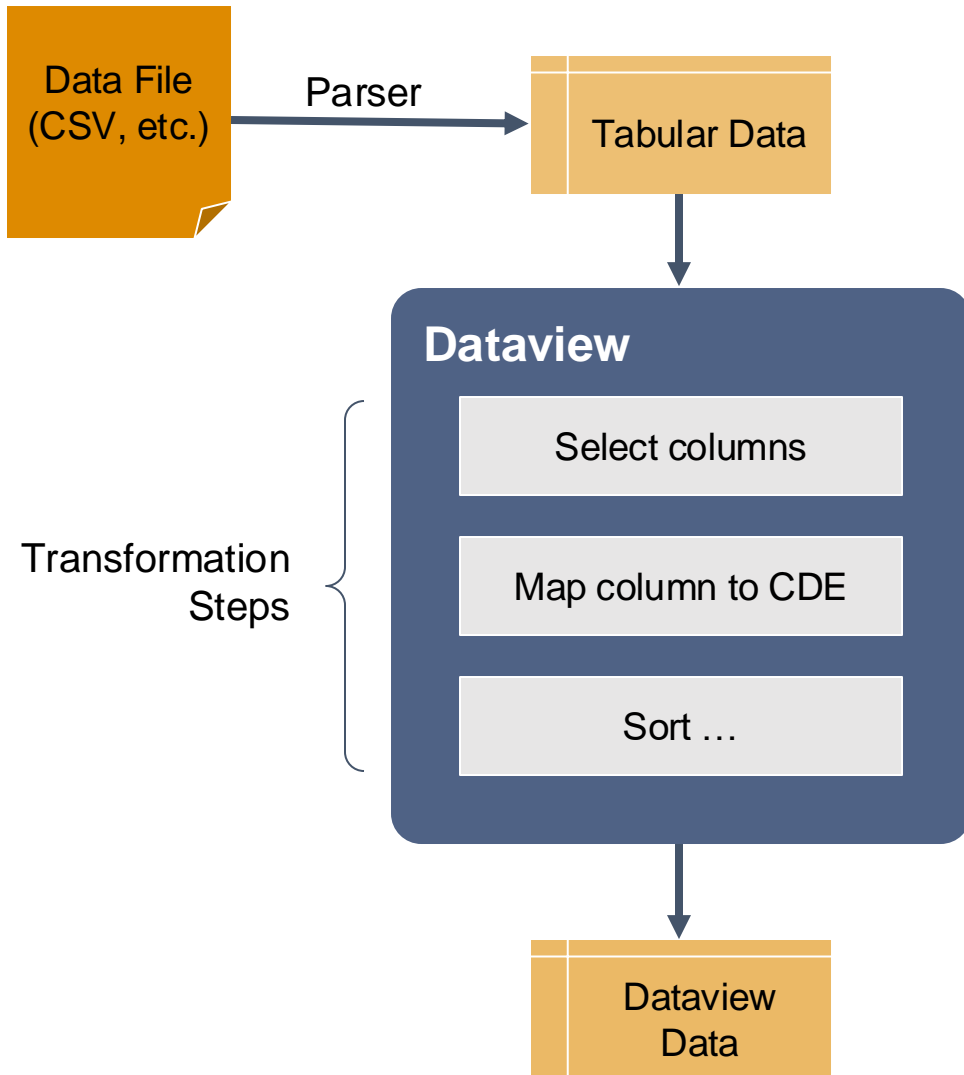
Use a Dataview to Map to CDEs within the Repository

The screenshot shows the SchARE Repository interface. The main content area displays a configuration for a Dataview. The source data is identified as 'test_data.xlsx'. The 'Map Into' section shows a 'Value Map from Source to Target' table with the following entries:

SOURCE VALUE	TARGET VALUE
Black	Black or African American
African American	Black or African American
American Indian	American Indian or Alaska P
Alaska Native	American Indian or Alaska P

Below the table, there are input fields for 'source value' and 'target value'. The 'When source value is not found in map, target value is:' section has radio buttons for 'NULL', 'Source value', and 'Constant: value'. The 'Target Column' is set to 'race_mapped'. The 'New column' and 'Replace source column' options are also visible.

What is a Dataview?

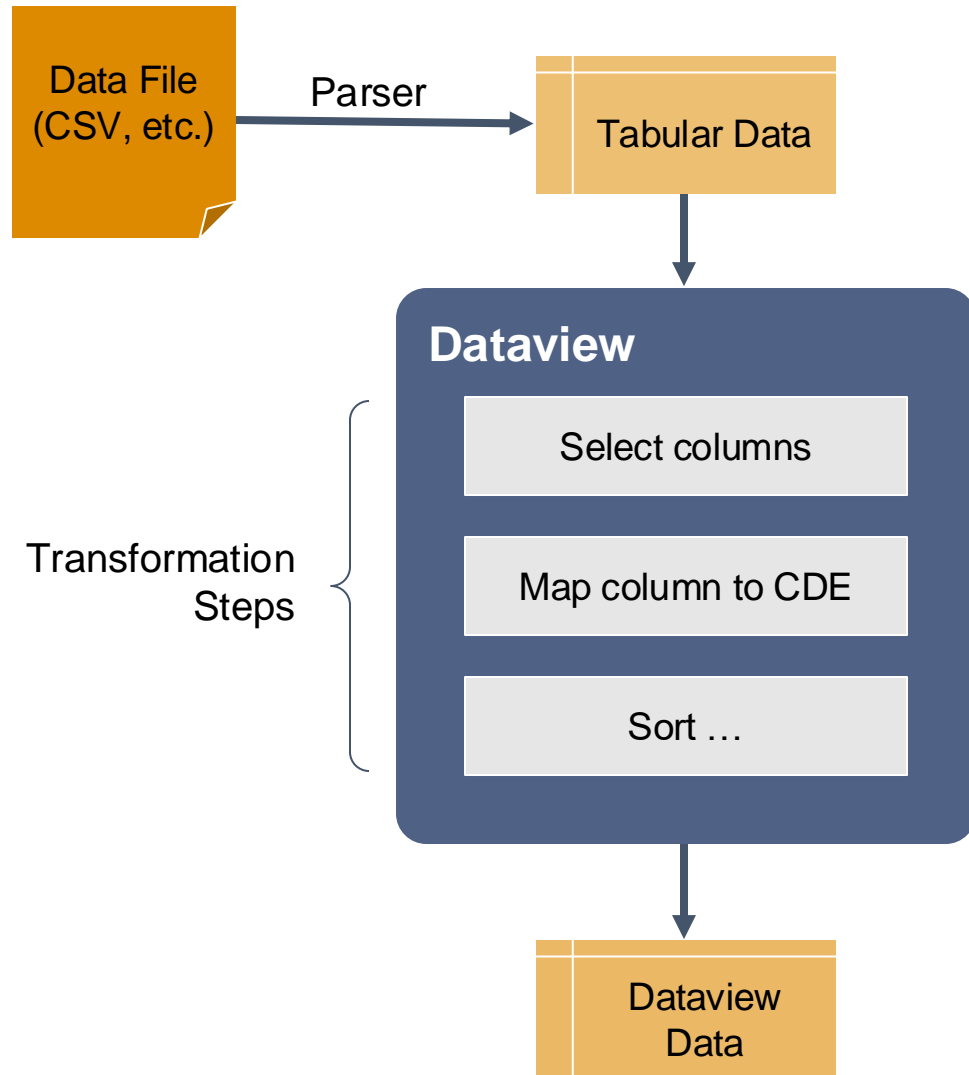


Dataviews take data from one or more sources, apply a series of transformation steps to that data (*filtering, sorting, mapping, etc.*) resulting in a new table of data as output.

Uses of Dataviews:

- Creating subsets of data
- Hiding PHI/PII for publishing
- Summarizing individual-level data into subsets and estimates
- Joining multiple datasets together
- **Mapping to CDEs**
- ... many others!

Mapping CDEs via Dataviews



Next, we'll walk through the process together:

1. Create a new dataview item
2. Add mapping steps to the dataview that map your data values to align with the CDEs
3. Label the newly mapped data columns with the appropriate CDEs and check that it passes validation

Mapping CDEs via Dataviews

Step 1: Create a new Dataview

1. Select **Create Dataview...** from the item's menu
2. Give it a name
3. It will open the Explorer, where you can begin building

Item Operations ▾

- Rename...
- Copy To...
- Link To...
- Move To Folder...
- Configure Table
- Assign Data Dictionary...
- Create Dataview...**
- Import as REDCap Data Dictionary
- Export Table Data
- Download
- Delete

Create Dataview

mapped test data

Create Dataview Cancel

ScHARe Repository

About Docs Collections CDEs Search... kari9152 ▾

kari9152 / ScHARe Example Data 2 / LIVE / mapped test data

Advanced Explorer Table Dictionary Meta 437 bytes | a few seconds ago | text/prql | status: ⌵ Item Operations ▾

Source data from: test_data.xlsx ↗

Take

Number of rows Range of rows Limit output to Number of Rows

10

+ Add Step ▾ Libraries: ⓘ Add Library Add Data Elements Clear Dataview Save Dataview

Results > ✓ Data available ✓ 0 parsing errors ✓ 0 validation errors

Transformation Preview

Participant ID	Age ⓘ	Age Units ⓘ	Birthplace - US	Birthplace - Out...	Postal Zip Code	Race/Ethnicity S...	Sex at Birth ⓘ	Gender	Get
0001	64	Years	AL		20009	American Indian	Female	Man	
0002	47	Years	NC		01581	Alaska Native	Male	Woman	

Mapping CDEs via Dataviews

A Quick Tour of the Dataview Explorer screen

Source data comes in at the top

Add more steps or add a library of additional steps

The result is previewed at the bottom

The screenshot shows the SchARE Dataview Explorer interface. At the top, the source data is identified as 'test_data.xlsx'. A 'Take' step is configured with 'Number of rows' set to 10. Below the step, there are buttons for '+ Add Step', 'Libraries', 'Add Library', and 'Add Data Elements'. The 'Results' section shows 'Data available', '0 parsing errors', and '0 validation errors'. A 'Transformation Preview' table is displayed at the bottom.

Participant ID	Age	Age Units	Birthplace - US	Birthplace - Out...	Postal Zip Code	Race/Ethnicity S...	Sex at Birth	Gender	Ge...
0001	64	Years	AL		20009	American Indian	Female	Man	
0002	47	Years	NC		01581	Alaska Native	Male	Woman	

Data flows vertically down the pipeline of transformation steps

Save your changes using the Save Dataview button

Once you're done, use the **Table** tab to view the full results

Mapping CDEs via Dataviews

Step 2: Add Mapping Steps to your Dataview

1. Using the **Add Step** menu, add a **Map Column** step

+ Add Step ▾

- Select Columns
- Filter Rows
- Sort
- Shuffle
- Join
- Take Rows
- Rename Column
- Map Column**
- Aggregate Rows

2.

Fill in the step fields (source column, target column, value map)

3.

Review the results, adjust as needed

4.

Repeat for all columns that need to be mapped

Map Into

Source Column: Race/Ethnicity Self-Identification

Value Map from Source to Target

SOURCE VALUE	TARGET VALUE
Black	Black or African American
African American	Black or African American
American Indian	American Indian or Alaska Native
Alaska Native	American Indian or Alaska Native

Target Column: race_mapped

New column Replace source column


When source value is not found in map, target value is:
 NULL Source value Constant: value

Race/Ethnicity S...	race_mapped
American Indian	American Indian or Alaska Native
Alaska Native	American Indian or Alaska Native
Black	Black or African American
African American	Black or African American

Mapping CDEs via Dataviews

Step 3: Assign CDEs to Data Columns

1. Click **Add Data Elements**
2. Select the column and data element, click **Add**
3. Verify the data element is on the column and the data validates successfully

 Add Data Elements



Add Data Elements

You can enhance your dataview by adding new data elements or modifying existing ones that were derived from its source data.

COLUMN	DATA ELEMENT	REPRESENTS PV	
Age	Age	N/A	<input checked="" type="checkbox"/>
Age Units	Age Units	N/A	<input checked="" type="checkbox"/>
Sex at Birth	Sex at Birth	N/A	<input checked="" type="checkbox"/>


Add new data element annotation:

From:


CDE set selected: ScHARe

OPTIONAL: Select a PV...




 0 validation errors




race_mapped 

American Indian Alaska Native Race/Ethnicity Self-Identification (string)

Concepts 

- C17049: Race *dataElement*
- C16564: Ethnic Group *dataElement*
- C74528: Self-Report *dataElement*

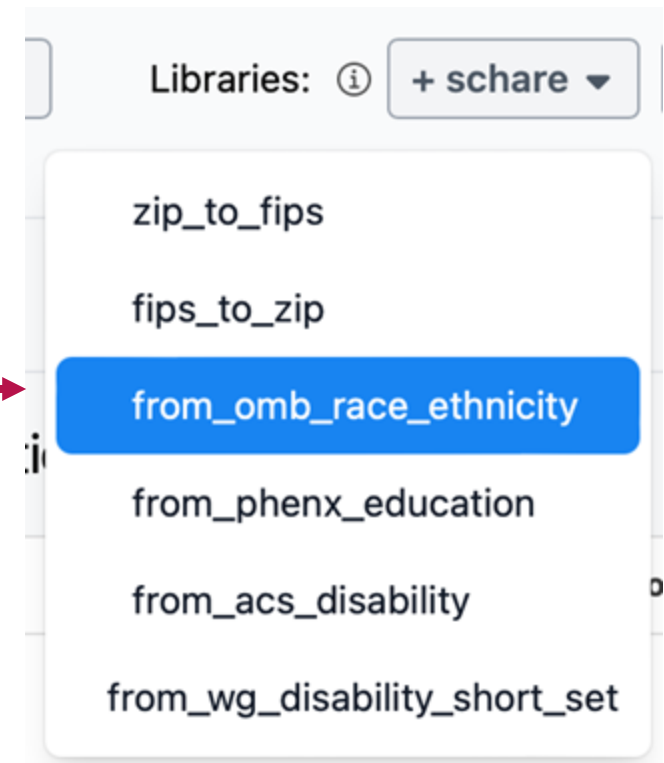
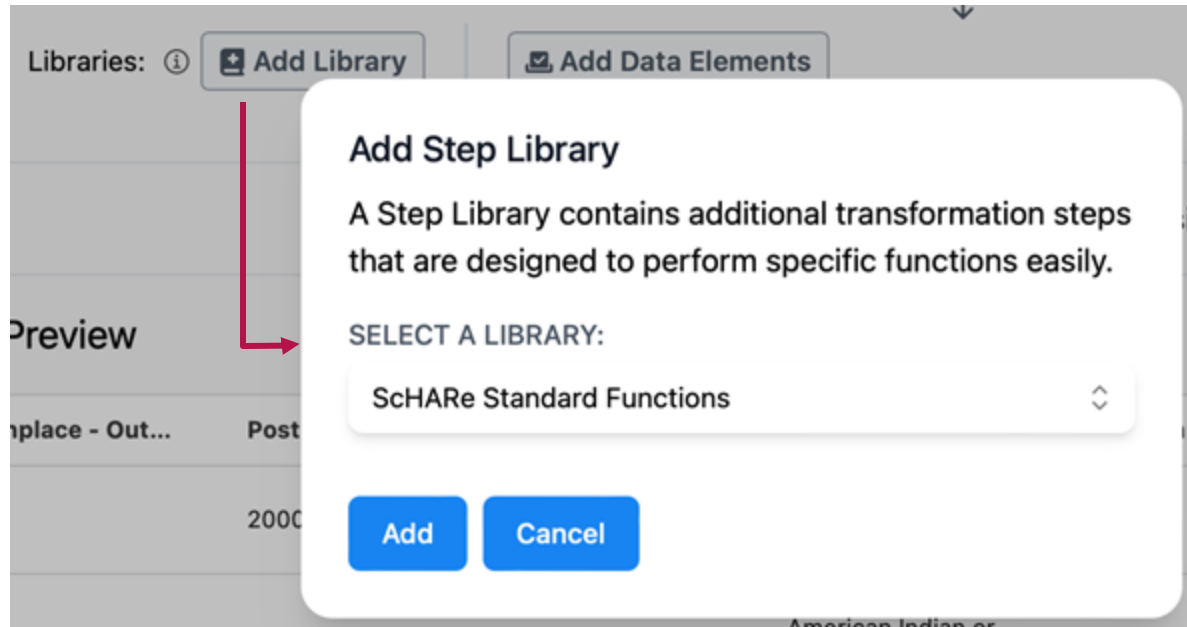
Definitions 

Black or African American

Mapping CDEs via Dataviews

The system also has a number of specific mapping functions available, for cases when the mapping is standardized or when the mapping is more complex than can be handled by the generic mapping functions.

Please contact our support team if you need any help with custom mapping functions!



You're done uploading your first data set!

What we've accomplished:

- Created a private Collection to hold all of your project data in one place
- Uploaded a data dictionary, documentation, linked to papers, etc.
- Uploaded your tabular data and viewed it
- Assigned and mapped your data to CDEs

Questions so far?

ScHARe

Sharing Data



BE A PART OF THE FUTURE
OF KNOWLEDGE GENERATION

Viewing your Collection

ScHARe Repository About Docs Collections CDEs Search...

karl9152 / ScHARe Example Data 2 / LIVE

ScHARe Example Data 2

Abstract
For purposes of de... according to ...

Levels of Influence
Individual
Community

Access Level
Private

Analysis Readiness
Ready >

CDE Compliance - ScHARe
2 / 19 CDEs assigned

Tags
Topics tagged in this collection
Co-Morbidities Disability

Links and Documents

Data Items

STATUS	NAME	CREATED	SIZE
	test_data.xlsx	an hour ago	14 KB

Page 1 of 1

Drag and Drop or [Browse Files](#) to Upload

All data is initially private (only you can access it).

See whether your data is well-annotated or what you should improve

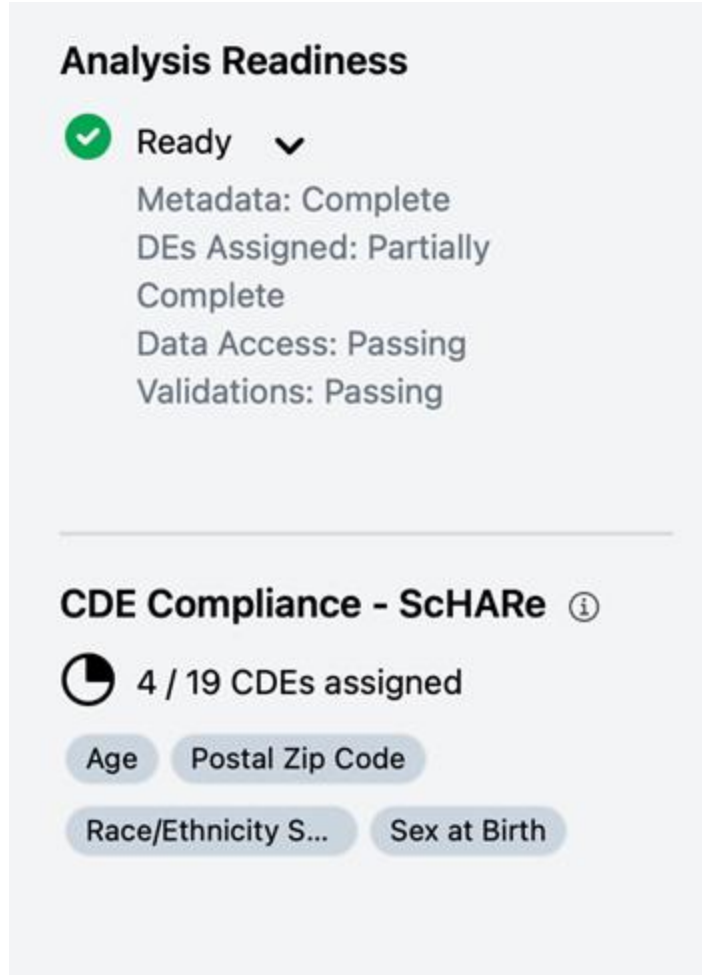
The item status shows all good or any problems in storage, parsing, or validation

View the CDEs that you have assigned

CDE Compliance and Analysis Readiness

Analysis Readiness - a simple metric on whether your data is ready for downstream use

1. Did you assign **metadata** to your collection (tags, project-level CDEs)?
2. Have you partially or fully **assigned CDEs** to your data?
3. Is the data **accessible** to the system (no broken links)?
4. Does the data pass **validation** according to the assigned CDEs?



The screenshot shows a user interface for data analysis readiness and compliance. At the top, under the heading "Analysis Readiness", there is a green checkmark icon followed by the word "Ready" and a dropdown arrow. Below this, four status indicators are listed: "Metadata: Complete", "DEs Assigned: Partially Complete", "Data Access: Passing", and "Validations: Passing". A horizontal line separates this section from the "CDE Compliance - ScHARe" section below. This section features a circular progress indicator showing 4 out of 19 CDEs assigned. Below the progress indicator are four buttons representing assigned CDEs: "Age", "Postal Zip Code", "Race/Ethnicity S...", and "Sex at Birth".

CDE Compliance - Showing how many CDEs have been assigned across the data in the collection



- Click on the metric to pop up the list of CDEs assigned
- Click on an individual CDE to find more information about that CDE

Access Levels and Sharing Data

You have control over how your data is shared on the ScHARe Repository. By default, all collections start out as **Private**.

Share Collection

Users, groups, and collections with access:

ID	ROLE	
 Karl Gutwin (karl9152)	ADMIN	

Share with:

This collection's access level is currently set to **Private**.
To share this collection with others, you must first set the access level to **Confidential**.

Access Levels

The access level of a collection defines the maximum permissions that can be used to share it with others. The following access levels are supported:

- **Private:** Only the collection's owner can access
- **Confidential:** The collection can be shared with named users
- **Controlled:** The collection can be shared with members of a controlled access group, as well as named users
- **Public:** The collection can be read by any user, including those not logged in; it can also be shared with named users

Access Levels and Sharing Data

After changing your collection's access level to Confidential, you can share it with other users, groups, and collections.

Share Collection

Users, groups, and collections with access:

ID	ROLE
Karl Gutwin (karl9152)	ADMIN

Share with: **Users** Groups Collections

John

Role: Reader

Share

John Jacquay (john)

Done

Access Levels

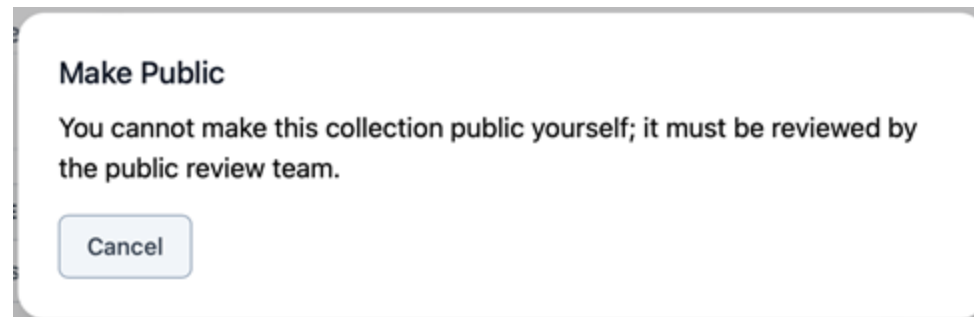
The access level of a collection defines the maximum permissions that can be used to share it with others. The following access levels are supported:

- **Private:** Only the collection's owner can access
- **Confidential:** The collection can be shared with named users
- **Controlled:** The collection can be shared with members of a controlled access group, as well as named users
- **Public:** The collection can be read by any user, including those not logged in; it can also be shared with named users

Publishing Data

We are still working on the final version of the process; however, it will essentially be:

1. Prepare your data according to the published guidelines
2. Add the Public Review group as Admin on your collection
 - a. This is necessary for them to be able to alter your collection's Access Level on your behalf
3. Notify the Public Review group by email
 - a. They will review your data for compliance with the data sharing guidelines
 - b. If any changes are necessary, you can work with the reviewer
 - c. Once complete, they will create a static version of your collection and set its Access Level to Public



ScHARe

Data Aggregation
and Analysis

BE A PART OF THE FUTURE
OF KNOWLEDGE GENERATION



How do I analyze or aggregate data from the ScHARe Repository?

Three quick options:

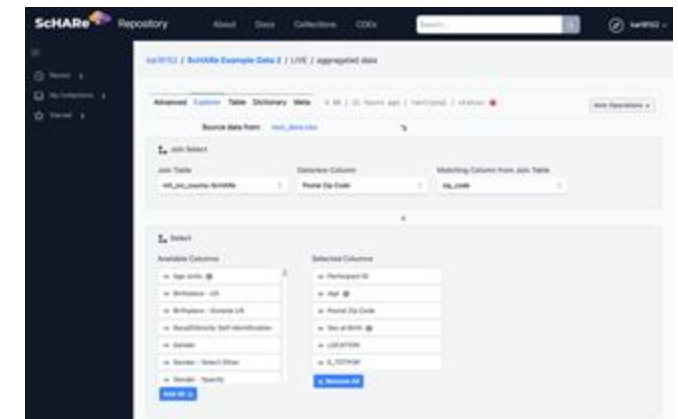
Connect to a Terra Jupyter notebook



Download to your own computer



Use Dataviews on the Repository



Look for a future Think-a-Thon where we will share more about how to link your data to Terra and how to use the ScHARe Repository to create aggregated data sets!

Analyzing Data on Terra (Jupyter Notebooks)

The screenshot shows a Terra workspace interface with a Jupyter notebook. The workspace is named 'schare-gde-demo.ipynb' and is located under 'Workspaces > SchARE-dev/SchARE KG > analyses'. The notebook content is as follows:

```
In [1]: ## Installation of the pypigeon library, do this once
#
# import sys
# ![sys.executable] -m pip install pypigeon

Documentation for the PyPigeon client can be found here:
https://bioteam.github.io/project-pigeon/pypigeon\_api.html

In [2]: from pypigeon import login
client = login('test-schare.nimhd.nih.gov')

To activate your session, visit the URL below:
https://test-schare.nimhd.nih.gov/login/activate/1w4HqehJJkHX0jHQW7DHRw.jnRx5pzandz47A7C4LiKeUHDQ3g

Waiting for session activation...

In [3]: collection = client.get_collection_by_name('Example NMHSS Analysis')

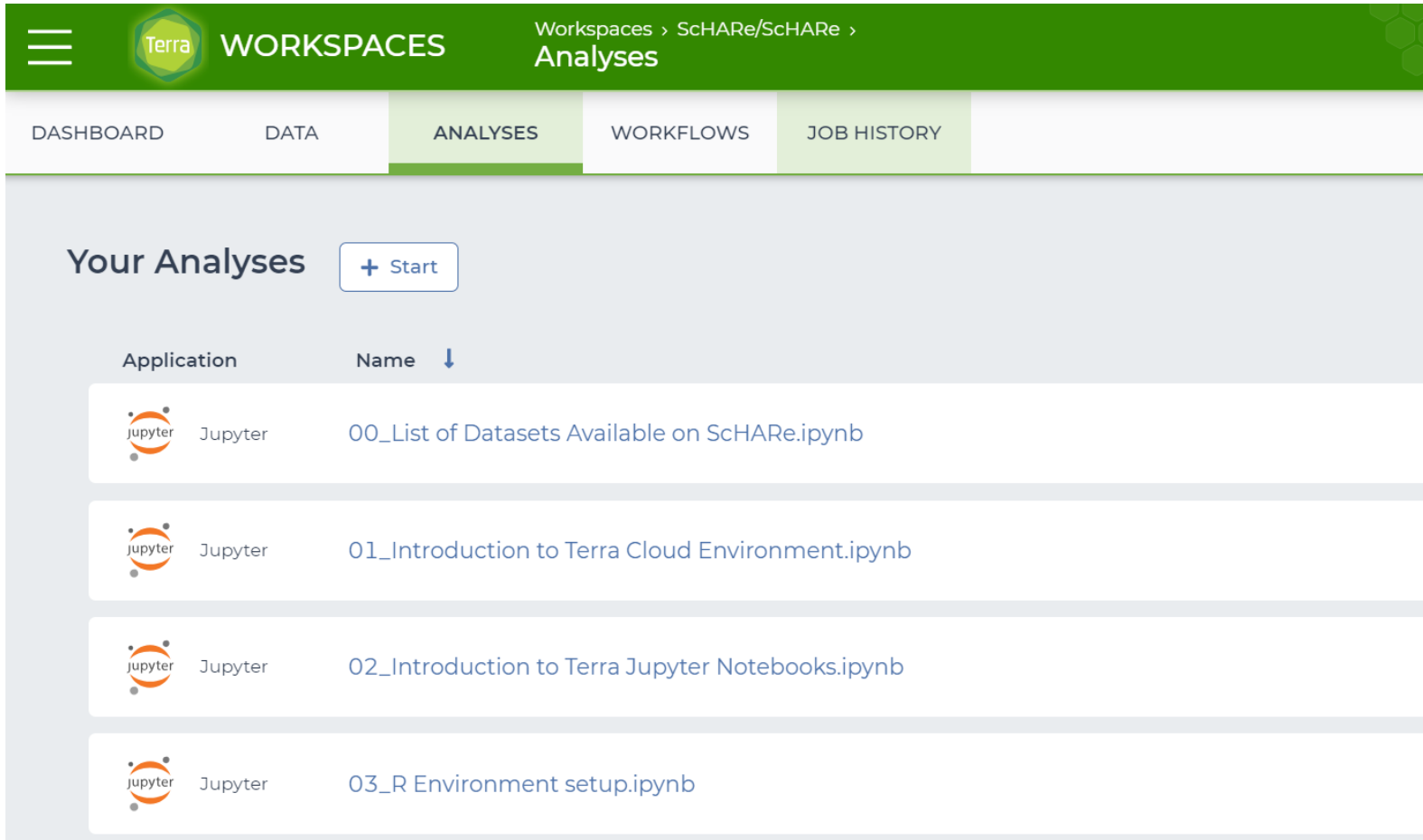
In [4]: nmhss = collection.get_table('nmhss-cbt-facilities')
Loading nmhss-cbt-facilities: 0it [00:00, ?it/s]

In [5]: nmhss
```

The output of the notebook shows a table with 17 columns: CASEID, LST, MHINTAKE, OWNERSHP, PUBLICAGENCY, TREATCOGTHRPY, SENIORS, ALZHDementia, STATE, E_TOTPOP, E_HH, E_POV, E_UNEMP, E_PCI, and E_NOHSDP. The table contains 780 rows of data, with the first few rows showing details for various states like AK, AL, and PR.





	CASEID	LST	MHINTAKE	OWNERSHP	PUBLICAGENCY	TREATCOGTHRPY	SENIORS	ALZHDementia	STATE	E_TOTPOP	...	E_HH	E_POV	E_UNEMP	E_PCI	E_NOHSDP
0	201800025	AK	1	2.0	-2.0	1.0	1.0	0.0	ALASKA	738516.0	...	253462.0	77865.0	28067.0	32531.206897	34760.0
1	201800093	AL	1	2.0	-2.0	1.0	1.0	1.0	ALABAMA	4864680.0	...	1860269.0	829400.0	147898.0	23072.835821	470043.0
2	201800099	AL	1	1.0	-2.0	1.0	1.0	1.0	ALABAMA	4864680.0	...	1860269.0	829400.0	147898.0	23072.835821	470043.0
3	201800104	AL	1	1.0	-2.0	1.0	1.0	1.0	ALABAMA	4864680.0	...	1860269.0	829400.0	147898.0	23072.835821	470043.0
4	201800109	AL	1	2.0	-2.0	1.0	1.0	0.0	ALABAMA	4864680.0	...	1860269.0	829400.0	147898.0	23072.835821	470043.0
...
779	201809433	PR	1	1.0	-2.0	1.0	0.0	0.0	None	NaN	...	NaN	NaN	NaN	NaN	NaN
780	201809435	PR	1	1.0	-2.0	1.0	1.0	1.0	None	NaN	...	NaN	NaN	NaN	NaN	NaN

Analyzing Data on Terra (Jupyter Notebooks)

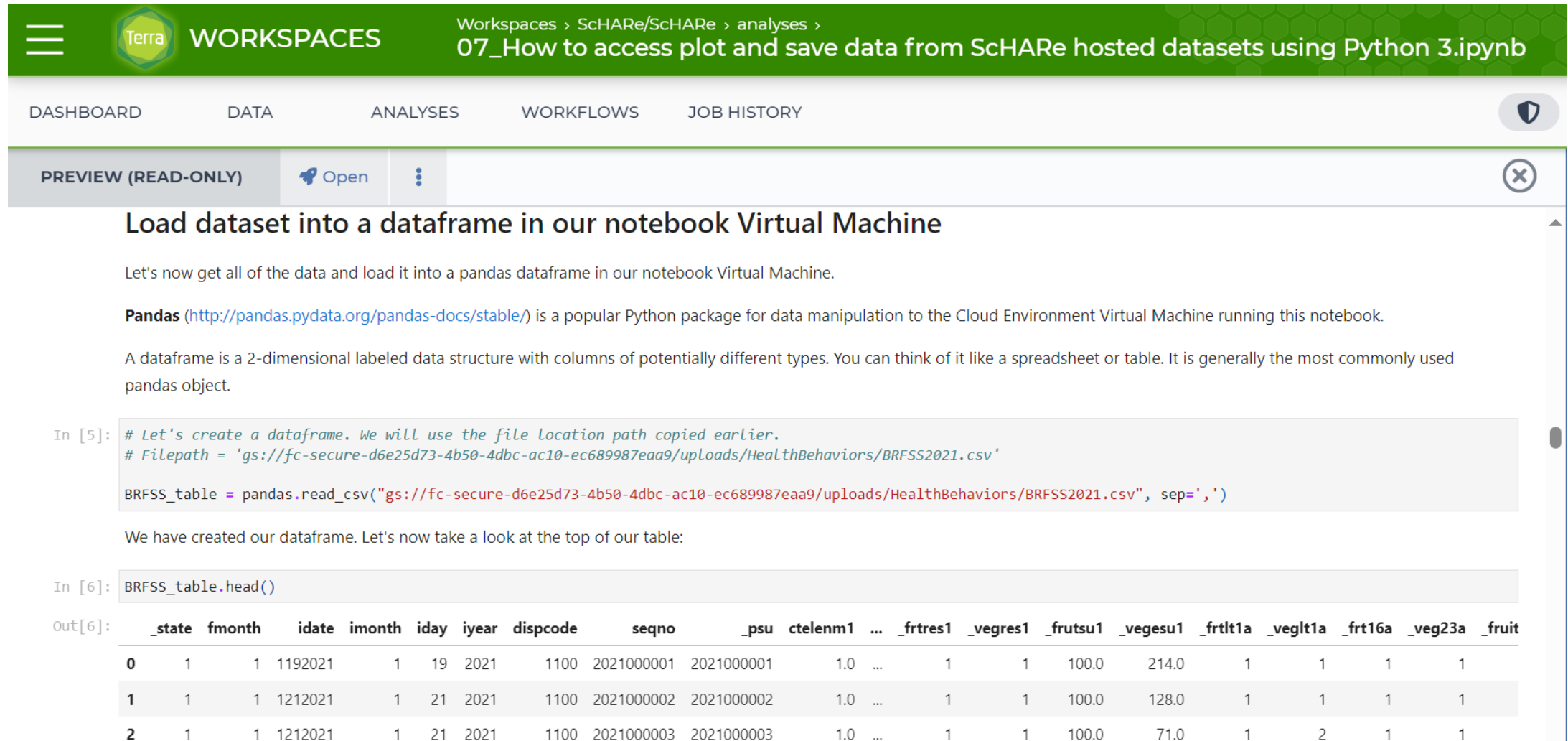


The screenshot displays the Terra WORKSPACES interface. The top navigation bar is green and contains the Terra logo, the word "WORKSPACES", and a breadcrumb trail: "Workspaces > ScHARe/ScHARe > Analyses". Below this is a horizontal menu with tabs for "DASHBOARD", "DATA", "ANALYSES" (which is highlighted), "WORKFLOWS", and "JOB HISTORY".

The main content area is titled "Your Analyses" and includes a "+ Start" button. Below this is a table listing the analyses:

Application	Name ↓
 Jupyter	00_List of Datasets Available on ScHARe.ipynb
 Jupyter	01_Introduction to Terra Cloud Environment.ipynb
 Jupyter	02_Introduction to Terra Jupyter Notebooks.ipynb
 Jupyter	03_R Environment setup.ipynb

Analyzing Data on Terra (Jupyter Notebooks)



Terra WORKSPACES Workspaces > ScHARe/ScHARe > analyses > 07_How to access plot and save data from ScHARe hosted datasets using Python 3.ipynb

DASHBOARD DATA ANALYSES WORKFLOWS JOB HISTORY

PREVIEW (READ-ONLY) Open

Load dataset into a dataframe in our notebook Virtual Machine

Let's now get all of the data and load it into a pandas dataframe in our notebook Virtual Machine.

Pandas (<http://pandas.pydata.org/pandas-docs/stable/>) is a popular Python package for data manipulation to the Cloud Environment Virtual Machine running this notebook.

A dataframe is a 2-dimensional labeled data structure with columns of potentially different types. You can think of it like a spreadsheet or table. It is generally the most commonly used pandas object.

```
In [5]: # Let's create a dataframe. We will use the file location path copied earlier.
# Filepath = 'gs://fc-secure-d6e25d73-4b50-4dbc-ac10-ec689987eaa9/uploads/HealthBehaviors/BRFSS2021.csv'

BRFSS_table = pandas.read_csv("gs://fc-secure-d6e25d73-4b50-4dbc-ac10-ec689987eaa9/uploads/HealthBehaviors/BRFSS2021.csv", sep=',')
```

We have created our dataframe. Let's now take a look at the top of our table:

```
In [6]: BRFSS_table.head()
```

```
Out[6]:
```

	_state	fmonth	idate	imonth	iday	iyear	dispcode	seqno	_psu	ctelenm1	...	_ftres1	_vegres1	_frutsu1	_vegesu1	_frlt1a	_vegl1a	_frt16a	_veg23a	_fruit
0	1	1	1192021	1	19	2021	1100	2021000001	2021000001	1.0	...	1	1	100.0	214.0	1	1	1	1	1
1	1	1	1212021	1	21	2021	1100	2021000002	2021000002	1.0	...	1	1	100.0	128.0	1	1	1	1	1
2	1	1	1212021	1	21	2021	1100	2021000003	2021000003	1.0	...	1	1	100.0	71.0	1	2	1	1	1

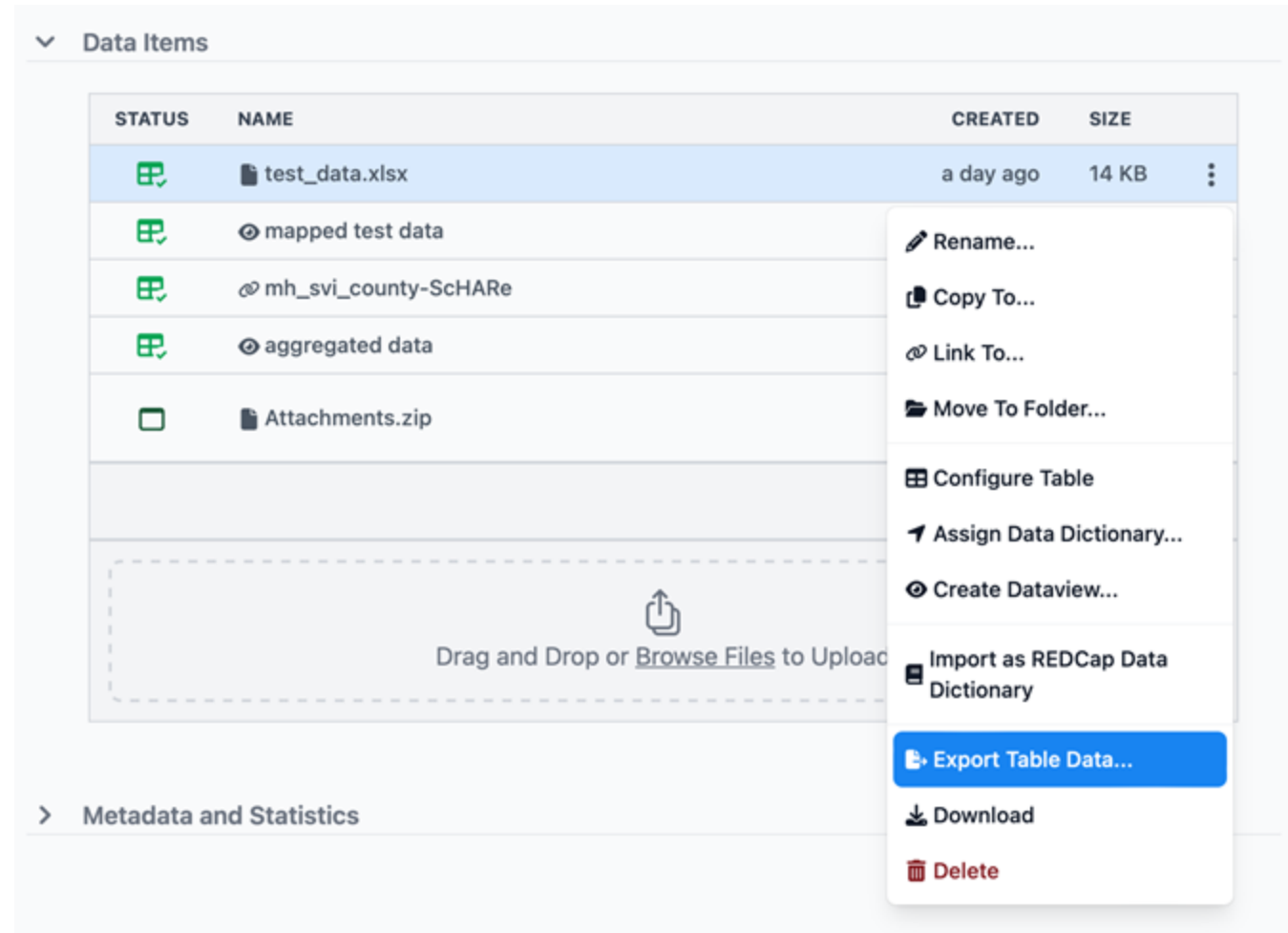
Downloading data to your own computer

Use the item's menu to download the original file

- Best for working directly with raw data

You can also export and download the tabular data in CSV, TSV, or Parquet format

- Best for working with mapped data, summary data, and other processed data



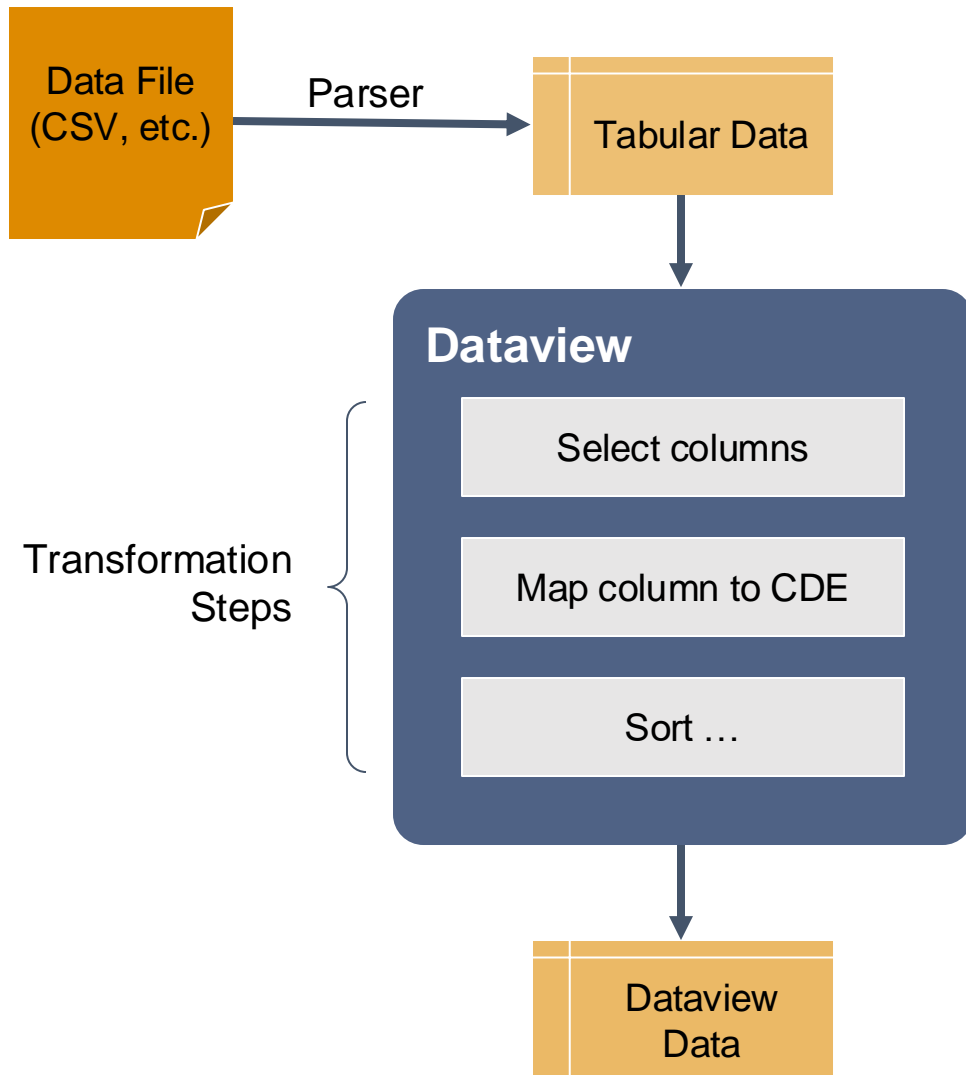
The screenshot displays a 'Data Items' section with a table listing various data files. The table has columns for STATUS, NAME, CREATED, and SIZE. The first row, 'test_data.xlsx', is selected. A context menu is open over this row, offering several actions. The 'Export Table Data...' option is highlighted in blue. Below the table is a dashed box for file uploads with the text 'Drag and Drop or Browse Files to Upload'. At the bottom, there is a 'Metadata and Statistics' section.

STATUS	NAME	CREATED	SIZE
	test_data.xlsx	a day ago	14 KB
	mapped test data		
	mh_svi_county-SCHARE		
	aggregated data		
	Attachments.zip		

- Rename...
- Copy To...
- Link To...
- Move To Folder...
- Configure Table
- Assign Data Dictionary...
- Create Dataview...
- Import as REDCap Data Dictionary
- Export Table Data...**
- Download
- Delete

Using Dataviews to Aggregate Data

Dataviews take data from one or more sources, apply a series of transformation steps to that data (*filtering, sorting, mapping, etc.*) resulting in a new table of data as output.



Uses of Dataviews:

- Creating subsets of data
- Hiding PHI/PII for publishing
- Summarizing individual-level data into subsets and estimates
- **Joining multiple datasets together**
- Mapping to CDEs
- ... many others!

Browsing and Searching for Data

The screenshot shows the SchARE Repository interface. At the top, there is a navigation bar with 'About', 'Docs', 'Collections', and 'CDEs' links, a search bar, and a user profile 'karl9152'. The main content area is titled 'Collections' and includes a 'Create New Collection' button. It is divided into three sections: 'My Collections' with one item 'SchARE Example Data 2' (LIVE, 4 hours ago); 'Shared with Me' with a message 'You don't have any collections shared with you right now!'; and 'Public Collections grouped by: Domains of Influence'. The domains shown are 'Biological', 'Health Care Systems and Clinical Care', and 'Sociocultural Environment'. Under 'Health Care Systems and Clinical Care', there are two items: 'Example NMHSS Analysis' (LIVE, 5 months ago) and 'Minority Health SVI' (LIVE, 7 months ago).

The screenshot shows the search results for 'nmhss'. The search bar at the top contains 'nmhss'. The results page is titled 'Search' and features a 'Filters' sidebar on the left with categories: 'metadata' (0 / 90), 'content type' (0 / 51), 'type' (0 / 6), and 'tags' (0 / 7). The main content area displays the top result, 'Example NMHSS Analysis', which is 5 months old. The description states: 'This contains data from the 2018 National Mental Health Services Survey (N-MHSS) and links to Minority Health SVI data, also from 2018.' Below the description are three tags: 'nmhss-puf-2018-csv...', 'nmhss-cbt-facilities', and 'nmhss-dictionary-pa...'. An 'Apply Filters' button is located at the bottom of the filters sidebar.

Look for improvements to the data organization and search experience in 2025!

Linking and Aggregating Data

If you find public data that you would like to combine with your project data, you can link it into your own collection and use it just like a file that you had uploaded yourself.

▼ Data Items

STATUS	NAME	CREATED	SIZE
📄	MH SVI Overview_11.19.2021.pdf	7 months ago	193 KB
📄	MH SVI Fact Sheet_7.15.2021.pdf	7 months ago	276 KB
📄	mh_svi_county_2018.csv	7 months ago	7.5 MB
📄	MinorityHealthSVI_DataDictionary_2018.csv	7 months ago	25 KB
🔗	Ⓞ mh_svi_county-SchARE	7 months ago	4 KB
📄	Ⓞ mh-svi-by-state-2018		
📄	MinorityHealthSVI_DataDictionary_2018.pdd.js...		

Copy To...
Link To...
Export Table Data...
Download

> Metadata and Statistics

Link Item To...

My Collections

- ✓ SchARE Example Data 2

Public Collections

- test unlisted data
- Example NMHSS Analysis
- Minority Health SVI

Link Cancel

SchARE Example Data 2

Abstract
For purposes of demonstration, this project collected data according to ...

Levels of Influence
Individual
Community

Domains of Influence
Biological
Behavioral

> Links and Documents

▼ Data Items

STATUS	NAME	CREATED	SIZE
🔗	test_data.xlsx	3 hours ago	14 KB
🔗	Ⓞ mapped test data	an hour ago	1 KB
🔗	Ⓞ mh_svi_county-SchARE	a few seconds ago	4 KB

<< Page 1 of 1 >>

Linking and Aggregating Data

Once linked into your collection, you can create a dataview that combines the public data with your own project data.

The screenshot displays the ScHARE Repository interface. At the top, the navigation bar includes 'ScHARE Repository', 'About', 'Docs', 'Collections', 'CDEs', a search bar, and a user profile 'kari9152'. The main content area shows a breadcrumb path: 'kari9152 / ScHARE Example Data 2 / LIVE / aggregated data'. Below this, there are tabs for 'Advanced', 'Explorer', 'Table', 'Dictionary', and 'Meta'. The 'Advanced' tab is active, showing a 'Source data from: test_data.xlsx' and an 'Item Operations' dropdown. The 'Join Select' section is configured with 'Join Table' set to 'mh_svl_county-ScHARE', 'Dataview Column' set to 'Postal Zip Code', and 'Matching Column from Join Table' set to 'zip_code'. Below this is the 'Select' section, which has two columns: 'Available Columns' and 'Selected Columns'. The 'Available Columns' list includes 'Age Units', 'Birthplace - US', 'Birthplace - Outside US', 'Race/Ethnicity Self-Identification', 'Gender', 'Gender - Select Other', and 'Gender - Specify'. The 'Selected Columns' list includes 'Participant ID', 'Age', 'Postal Zip Code', 'Sex at Birth', 'LOCATION', and 'E_TOTPOP'. At the bottom of the 'Select' section, there are buttons for 'Add All' and 'Remove All'. Below the 'Select' section, there are buttons for 'Add Step', 'Libraries: Add Library', and 'Add Data Elements'. On the right side of this row, there are buttons for 'Clear Dataview' and 'Save Dataview'. The 'Results' section shows 'Data available', '0 parsing errors', and '0 validation errors'. The 'Transformation Preview' section shows a table with the following data:

Participant ID	Age	Postal Zip Code	Sex at Birth	LOCATION	E_TOTPOP
0001	64	20009	Female	District of Columbia, District of Columbia	684498
0002	47	01581	Male	Worcester County, Massachusetts	822280

SCHARe

Conclusion

BE A PART OF THE FUTURE
OF KNOWLEDGE GENERATION



ScHARe Repository

In summary:

- When getting started, first create a collection, provide metadata, and upload documents.
- Upload your data and use the system to help you map to the ScHARe CDEs.
- View your data, see CDE compliance and analysis readiness.
- Data can be shared with your colleagues, and can be made publicly available after review.
- You can use dataviews to create subsets of your data, join datasets together and more.
- Data can also be analyzed on the ScHARe Terra workspace.

We are here to support you!

We want to hear your questions, issues and comments about the ScHARe Repository!

For any questions regarding how to use the Repository, please reach out to:

 schare@mail.nih.gov

Response time: within 24 hours

For Technical Support inquiries, contact:

 schare-repository-support@bioteam.net

 Office Hours: Wednesdays, 4-5 PM US Eastern

Questions?

Think-a-Thon poll

1. Rate how useful this session was:

- Very useful
- Useful
- Somewhat useful
- Not at all useful

Think-a-Thon poll

2. Rate the pace of the instruction for yourself:

- Too fast
- Adequate for me
- Too slow

Think-a-Thon poll

3. How likely will you participate in the next Think-a-Thon?

- Very interested, will definitely attend
- Interested, likely will attend
- Interested, but not available
- Not interested in attending any others

ScHARe

Thank you



ScHARe

Next Think-a-Thons:



bit.ly/think-a-thons

Register for ScHARe:



bit.ly/join-schare

 schare@mail.nih.gov

