User Guide

A Data Harmonization Portal

Athens University of Economics and Bussiness
Charoula Velissarakou
8/5/2017

Table of Contents

Introduction

Data harmonization takes an approach to data quality that involves both machine analytics and human control. It learns which past decisions made to data are most trustworthy and relevant and then uses that intelligence to help present users effectively work with data.

Users may use this web application to manage medical variables in a way that data reliability is guaranteed. They will be able to match their own variables with other user's data so as to create a more trustworthy structure.

This document guides a user through the necessary steps to manage ones data and combine them with other harmonized data in a coherent way.

Account

This is the application's login page.

Note: The application can be used only by registered users.

Register

An unregistered user, who wants to have access to the application, needs to insert a valid username, mail and a password to the proper area at the home page and then click on the "Register" button.

If the username or the e-mail, that have been inserted, already exist, the registration will not be successful and an alert message will be shown. Thus, the user needs to insert a different username or a different e-mail address and try to sign up again.

Log in

A registered user who needs to log in, will have to insert their data to the proper area and then click on the "Log In" button. If the data are valid then the user will be logged in and the home page will be shown.

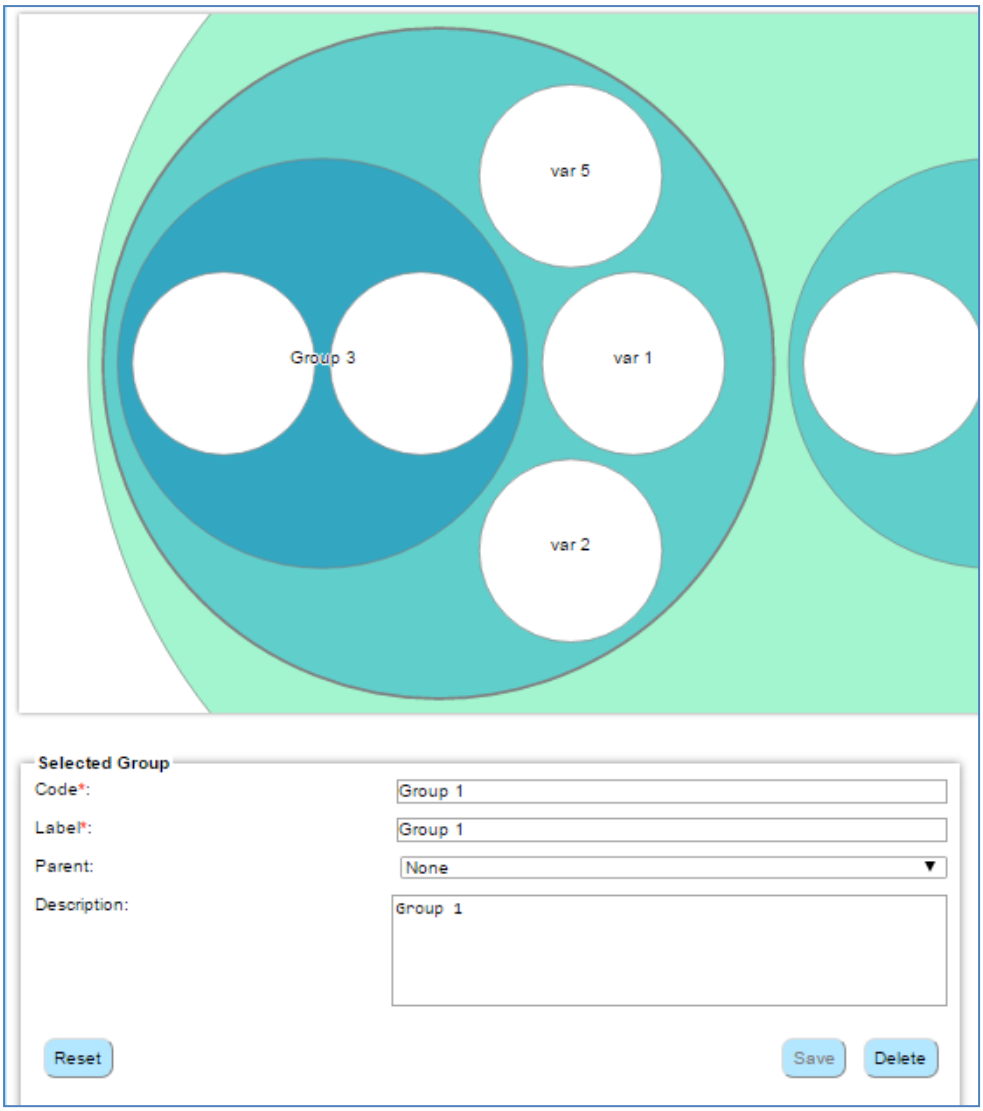
Log out

If the logged user wishes to log out, he/she will have to click on the "Log out" link that exists in side menu of the page.

Group Managing

1. Add New Group

Users who want to add a new group need to fill the blank fields at the "Selected Group" form and press the "Save" button to submit it.



Subsequently, the new group will be plotted on the first panel on the left, according to the hierarchy that has been created so far.
It will also be added at the "Parent" field's options and the "Group" field's options at the "Selected Variable" form.

Note: Fields (code and label) with a red asterisk are required, for the form to be submitted.

If the user wishes to reset the form, he/she may press the "Reset" button and fill the fields again.

2. Edit Group

Users who want to edit a specified group will have to search in the visualized hierarchy and click the group they need to edit. The details of the target group will be filled in the "Selected Group" form.

If there are any changes, the user should press outside the edited box and the "Save" button will be enabled for them to press. Not only the old details will be edited, but if a different parent is selected, the target group and its variables will be plotted in the new selected parent. Other than that, if the label is changed, the new label will be added at the "Parent" field's options and the "Group" field's options at the "Selected Variable" form and the old one will be deleted. In addition, every variable of the table that had its group's label edited, will also change and they can still be found at the "Grouped" tab.

Note: Fields (code and label) with a red asterisk are required, for the form to be submitted.

If the user doesn't need to edit the group, he/she may press the "Reset" button and the specified group will not be the target anymore.

3. Delete Group

Users who want to delete a specified group will have to search in the visualized hierarchy and click the group they need to delete. The details of the target group will be filled in the "Selected Group" form.
Then, the user should press the "Delete" button if that is the group that needs to be deleted.

Note: If a group is deleted, its variables become ungrouped, so that the user shall specify their group later. Thus, they temporarily disappear from the visualized hierarchy and they can only be found at the variable table at the Ungrouped Tab.

Variable Managing

1. Add New Variable

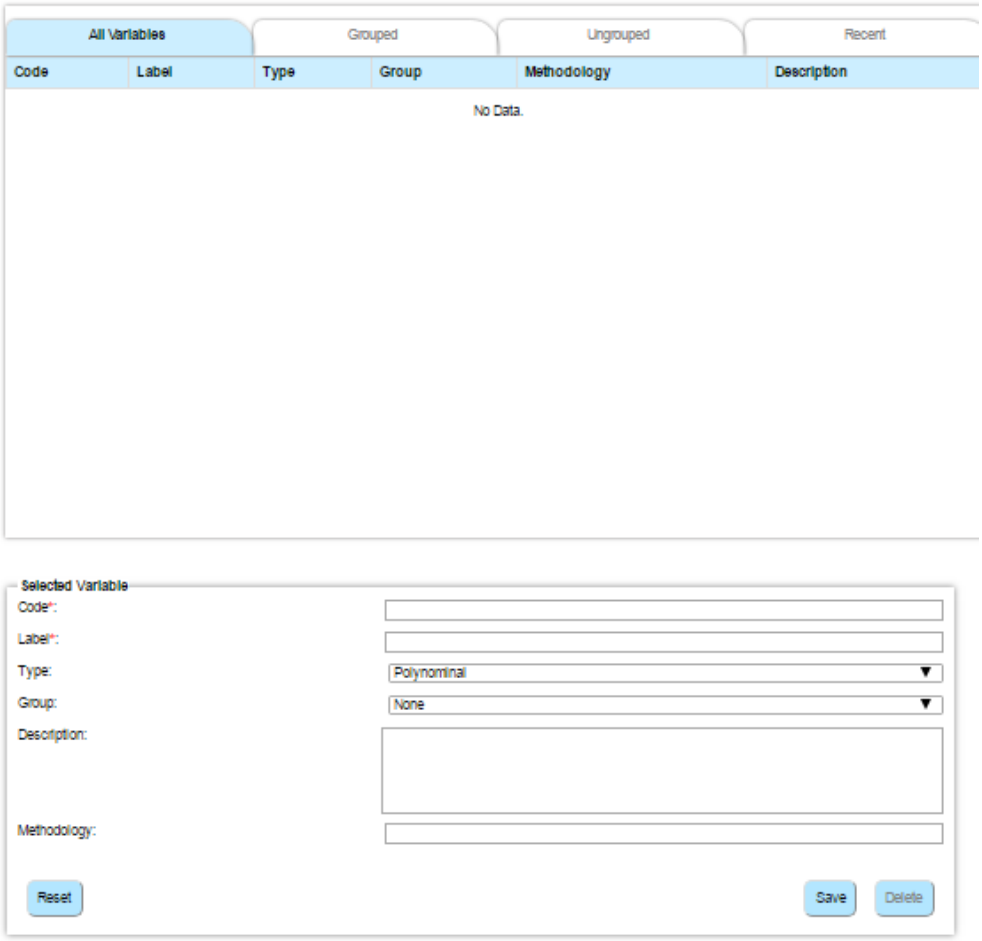
Users who want to add a new variable need to fill the blank fields at the "Selected Variable" form and press the "Save" button to submit it. Subsequently, the new variable will be added to the list, which is shown on the first panel on the right, as a new table row. This variable can also be found at the "Recent" tab that exists above the variable table.

It will also be plotted according to the hierarchy that has been created so far, inside the group that has been selected from the "Group" menu.

Note: Fields (code and label) with a red asterisk are required, for the form to be submitted.

If the user wishes to reset the form, he/she may press the "Reset" button and fill the fields again.

The variable table and the Selected Variable form is shown below.



2. Edit Variable

Users who want to edit a specified variable can either select the variable from the table, or click on the variable at the visualized hierarchy. The selected variable will be highlighted and is details will be filled in the "Selected Variable" form.
If there are any changes, the "Save" button will be enabled for the user to press. Not only the old details will be edited, but if a different group is selected, the highlighted variable will be plotted in the new selected group and will be deleted from the old one. The edited variables will also be shown at the Recent Tab.

Note: Fields (code and label) with a red asterisk are required, for the form to be submitted.

If the user doesn't need to edit the variable, he/she may press the "Reset" button and the specified variable will not be the target anymore.

3. Delete Variable

Users who want to delete a specified variable can either select the variable from the table, or click on the variable at the visualized hierarchy. The selected variable will be highlighted and is details will be filled in the "Selected Variable" form.
Then, the user should press the "Delete" button and the variable will be deleted from the plotting and the variable table.

Upload meta-data

Other functionalities that can be found at the side menu are:
Note: Meta-data can be either JSON files or CSV files.

Upload JSON meta-data

Users who want to upload a file from their computer with a .json extension, need to click on the "Upload meta-data (JSON)" button, at the side menu.
This JSON file should include a valid structure of Grouped Variables or else the hierarchy will not be plotted.

Every group should have these attributes:

- code
- label
- parent
- description
- an array with children (the array can be empty)

Every variable should have these attributes:

- code
- label
- type
- group
- description
- methodology

An example structure is shown below:

```
[{
  "code": "Group A",
  "label": "Group A",
  "parent": "None",
  "description": "Group A",
  "children": [
    {
    "code": "Group B",
    "label": "Group B",
    "parent": "Group A",
    "description": "Group B",
    "children": [
     {
        "code": "var 2",
        "label": "var 2",
        "type": "Polynominal",
        "group": "Group C",
        "description": "var 2",
        "methodology": "var 2"
     }
    ]
    },
    {
      "code": "var 1",
      "label": "var 1",
      "type": "Polynominal",
      "group": "Group B",
      "description": "var 1",
      "methodology": "var 1"
    }
   ]
 }]
```

Users will be asked to select a proper json file and once they choose it, the meta-data hierarchy (groups and variables) will be plotted on the first panel on the left.
The variables that are included in the selected file, will be also added to the variable table on the first panel on the right. Those variables, can also be found at the "Grouped" tab that exists above the table.

1. **Upload variables from CSV**

   Users who want to upload variables from their computer, using a file with a .csv extension, need to click on the "Upload Variables (CSV)" button, at the side menu.

   Users will be asked to select a valid file and once they choose it, the variables that the file includes, will be uploaded on the first panel on the right.

   Those variables, can also be found at the "Recent" tab that exists above the variable table.

   Note: The CSV file does not include the variable's group information, so the user should edit the variable's groups later.

Download meta-data

Users who want to download the data that have been created during ones session to their computer, need to click on the "Download meta-data (JSON) button, at the side menu.

The filename of the downloaded data will be "meta-data" and it is going to have a json extension.

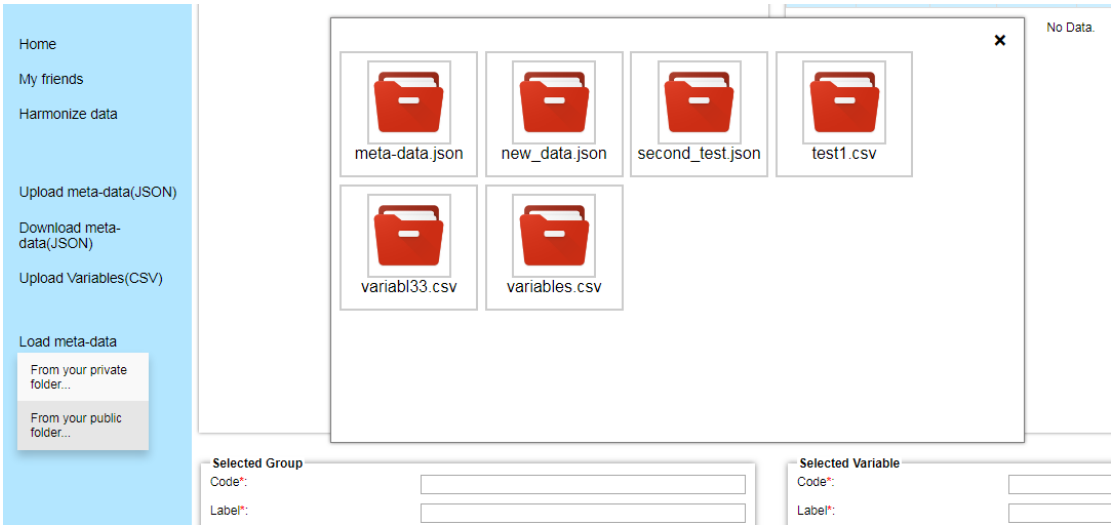
Sharing data

The Data Harmonization tool is not only for managing ones data but also sharing those data with the other members of the community. Thus, its functionalities make possible to:

1. **Load meta-data**

   Users are able to load the data that they have saved to the server by clicking the "Load meta-data" from the side menu. Consequently, users should choose which folder they want to load their data from. There are two folders available:

   - Private: The files that exist in this folder are accessible only by the user who is logged in.
   - Public: The files that exist in this folder are accessible by the user who is logged in and the friends that this user shares these data with.



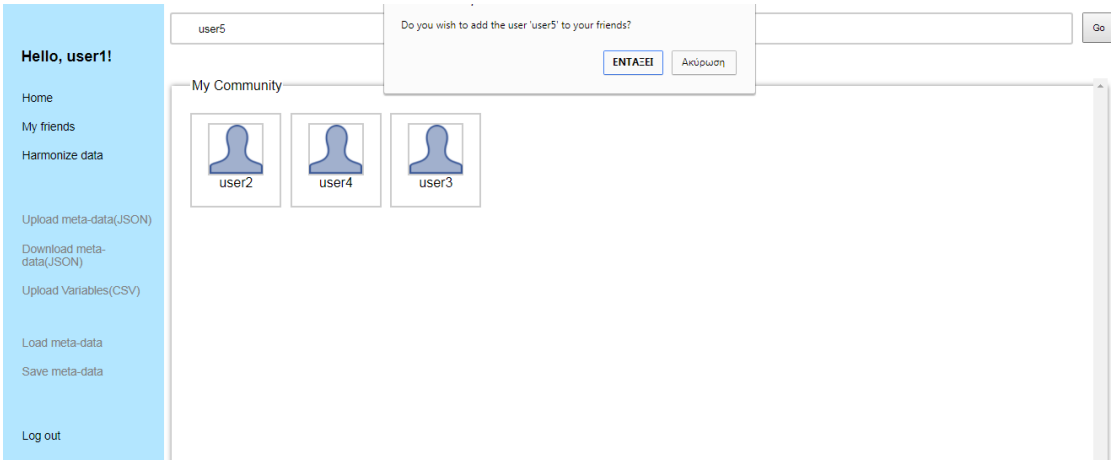
Once the user chooses a file from the pop up window, the meta-data will be loaded and a new hierarchy will be plotted on the first panel on the left. The new and the old variables will be shown on the table for the user to edit them.

2. Share meta-data

Users who want to save the data to the server have to click at the "Share meta-data" from the side menu. The meta-data can be saved in the Public or the Private folder. Once the user chooses a folder, they should enter a valid name for the new file. The new file will be saved in the selected folder in JSON format.

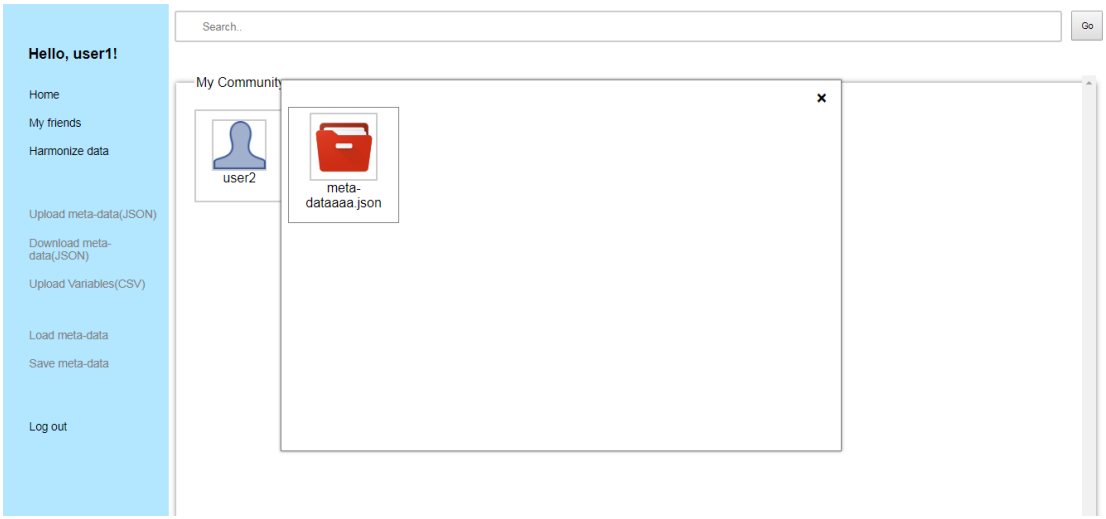
My Friends

Users can send requests to other members of the community so as to be able to share their files and create more trustworthy data.
At the "My Friends" Tab, users can search new friends by entering ones username or email at the search bar.

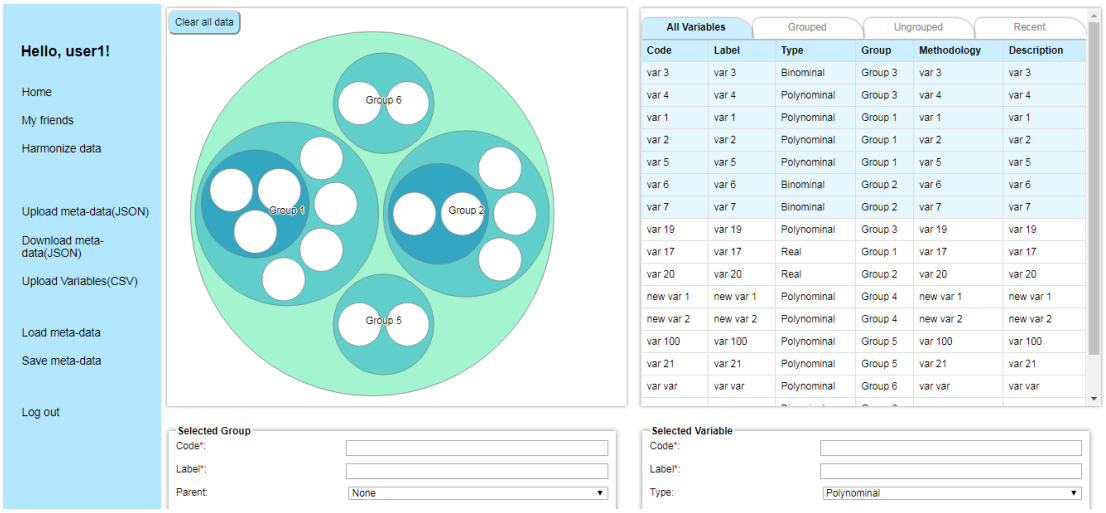


If the user with this username exists then an alert box will ask the user if she or he wants to add that user to their community. If the user sends the request then the other user will have to search in the same way the former one so that they will be both added in "My Community" field.

By selecting a user from the Community Field, one can load meta-data, that this user saved in their public folder. Once a file is selected, the new hierarchy will be plotted or new variables will be added in the tables that exist at the Home Page.



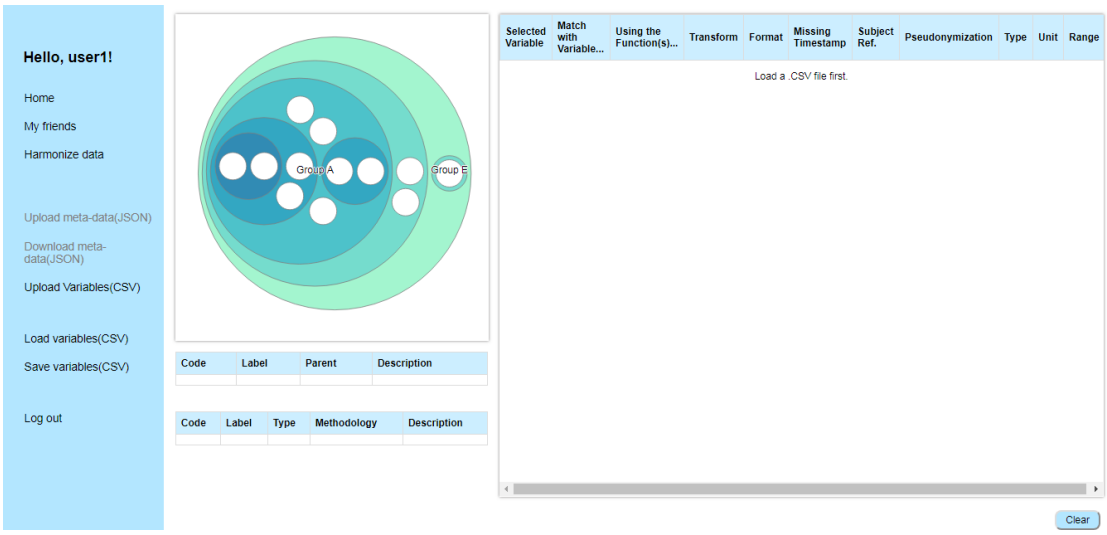
Those variables can be shown with a blue color at the "Grouped" tab if it was a JSON file that was uploaded or at the "Ungrouped" tab if it was a CSV file.



Note: The variables that are uploaded from "My friends" tab are shown with blue color, when the white ones are the ones that the user uploaded from the local space.

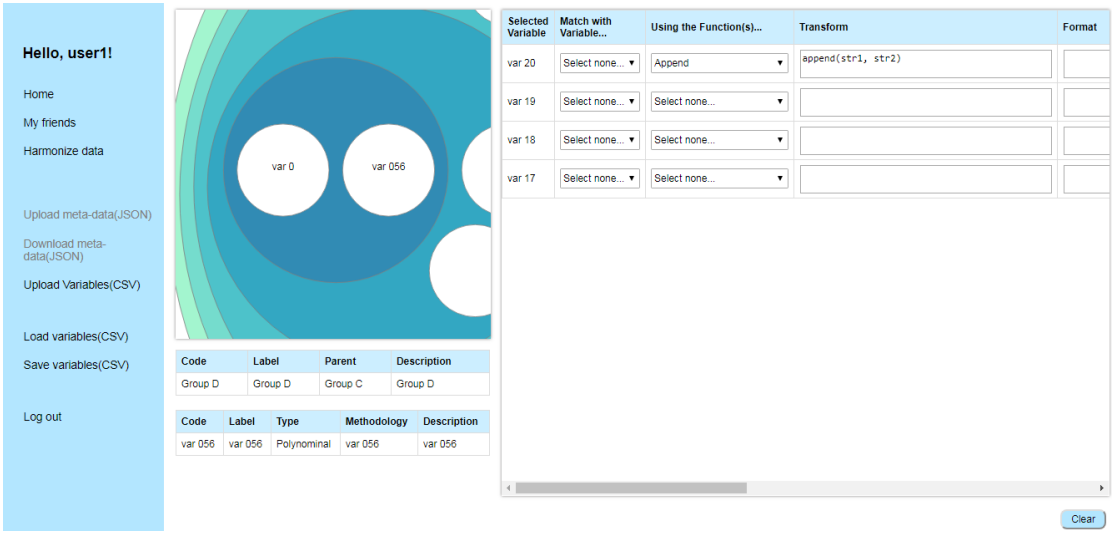
Data Harmonization

Using this tab, the user can match new variables with the old ones that are uploaded at the Home Page or have been created during their session. The Data Harmonization Tab is shown below.



At the left side the hierarchy that has been created so far is shown and the user can view the group's or variable's details below the plot. To be able to match and transform the

variables the user should upload a CSV using the navigation menu on the left. Once a csv file is uploaded the table on the right will change as shown.



Later the user can choose the variable they want to match their new ones using the available functions. The transformation can be saved as a .csv file with the form:

The csv can be saved at the server in the form:
old_variable, new_variable format, missing timestamp, subject ref., pseudonymization, type, unit, range, transformation

where transformation comes from the altered functions of the textarea that is beside each variable.

Note: If one needs to clear the table, they can press the Clear button at the bottom.