faoswsStandardization: `pullDataToSUA` plugin

Cristina Muschitiello

Food and Agriculture Organization of the United Nations

21 June 2018

Abstract

This vignette provides a description of the data Pulling Procedure: it is the module that pull the data from different data-sets of different domains, inside the table `sua_unbalanced`, starting point of the overall Standardization and Balancing procedure.

Contents

List of Figures

Disclaimer

The Data flow

This is the first step of the Standardization and Balancing, the step in which data coming from all output data-set are combined in another data-set, which will be the starting point of the following step. It is represented in figure 1
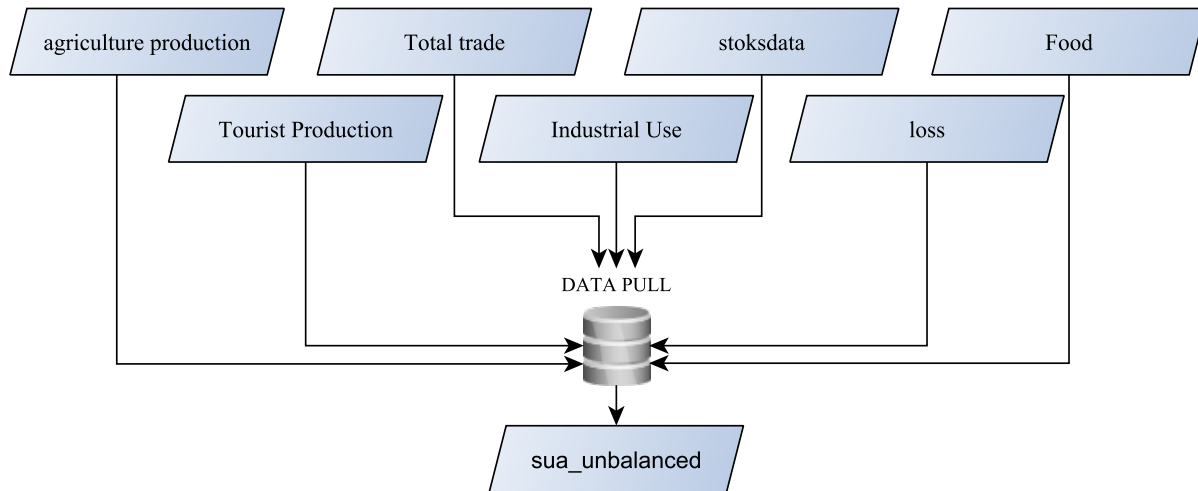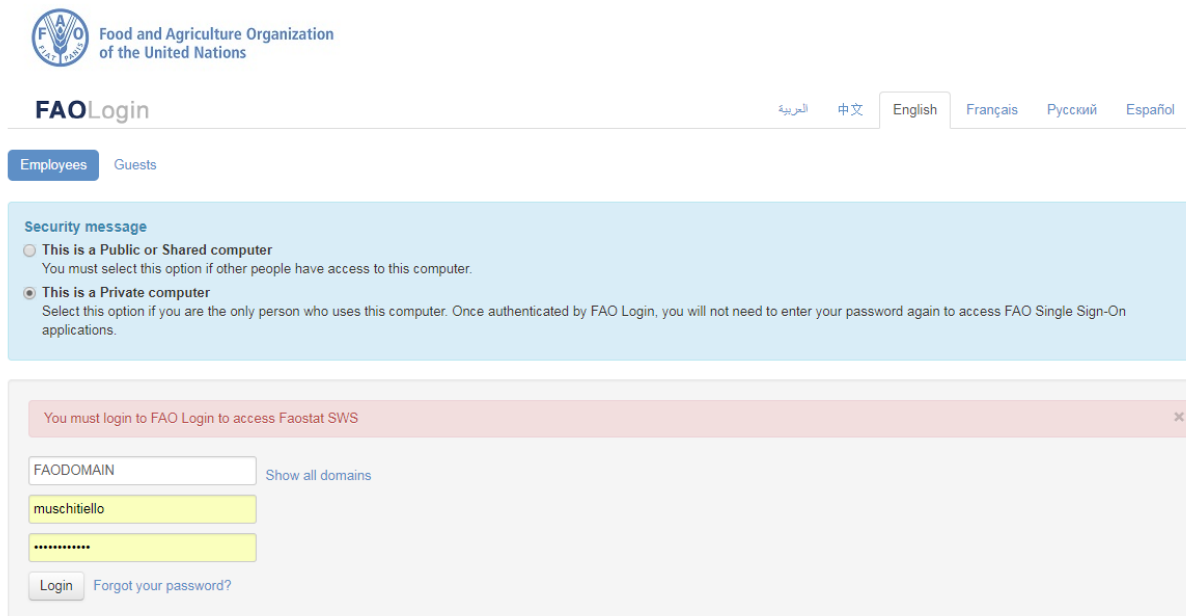

Figure 1: Data pulling

Notice that, from *agriculture production* the following Data are pulled:

- crop production,
- livestock,
- milk and eggs,
- production of derived commodities,
- seed,
- feed.

Plug-in

A general description of all the objects of the *SWS* is given in the document *Food Balance Sheet workflow in the Statistical Working System*. A plug-in, in this framework, is an executable process. In this document, the steps for executing the *pullDataToSUA* plugin are explained:

1   Log-in in the SWS

Figure 2: Log-in in the SWS
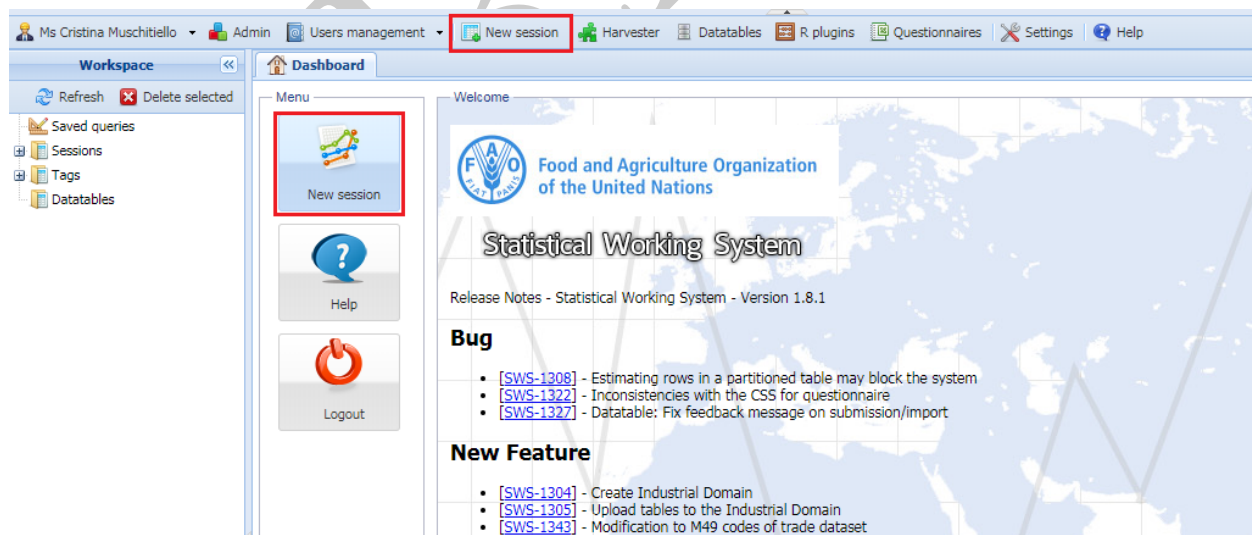
2   Open a new Session

Figure 3: Open a new Session

3   Define dimendions of the session

For the Pulling of the data, a session has to be opened in the *target* data-set, which is the *suafbs:sua_unbalanced*. Therefore *SUA/FBS* domain and *sua_unbalanced* have to be selected from the screen:
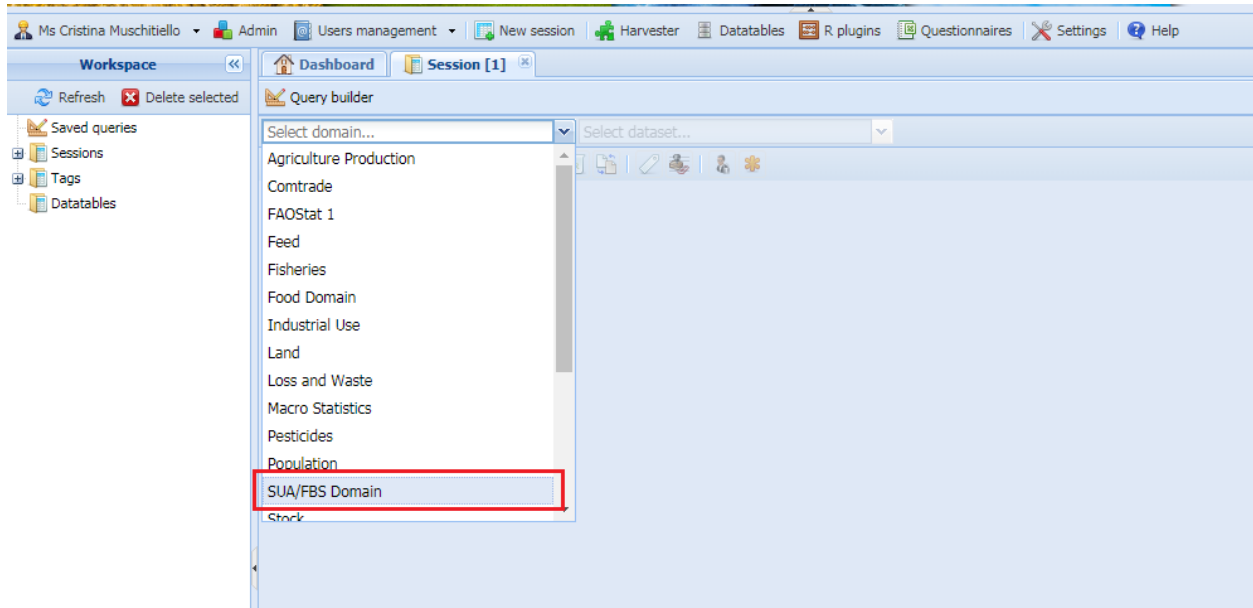
Figure 4: Select Domain
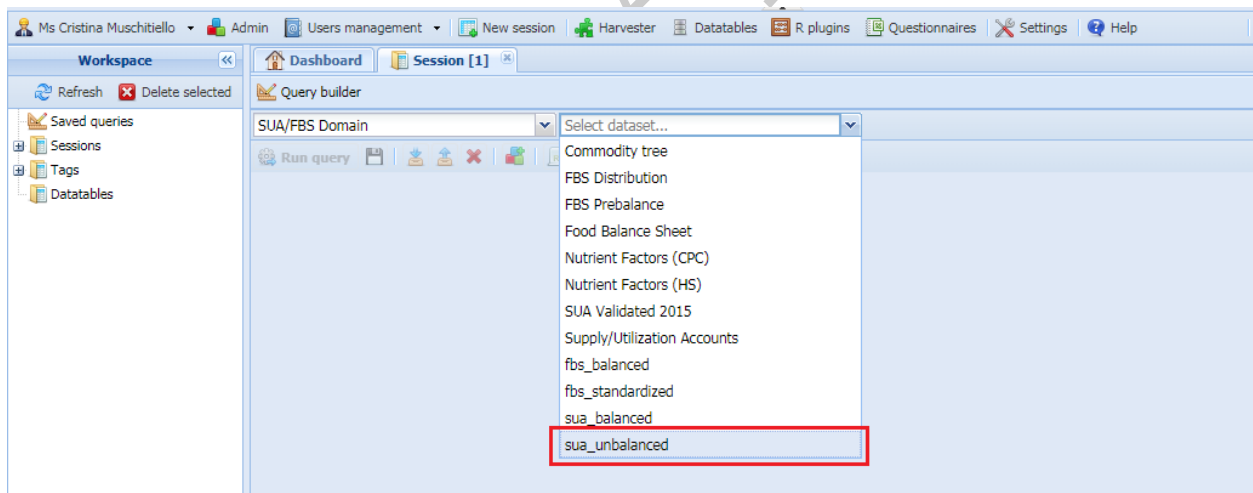
Figure 5: Select Dataset

4 Make an run the query

The query has to be done only on the country for which the Pull data has to be performed. Indeed the plugin could be performed on one of the two following set of countries: *session Countries* or *all countries*. In this example *China, Mainland* is selected.

Figure 6: Select Country/ies

All elements here have to be selected (figure 7) and all items(8). The years to be selected depend on the interest of the user.

Figure 7: Select all Elements

Figure 8: Select items and years

When all Variables have been defined, the query can be run:

Figure 9: Run query

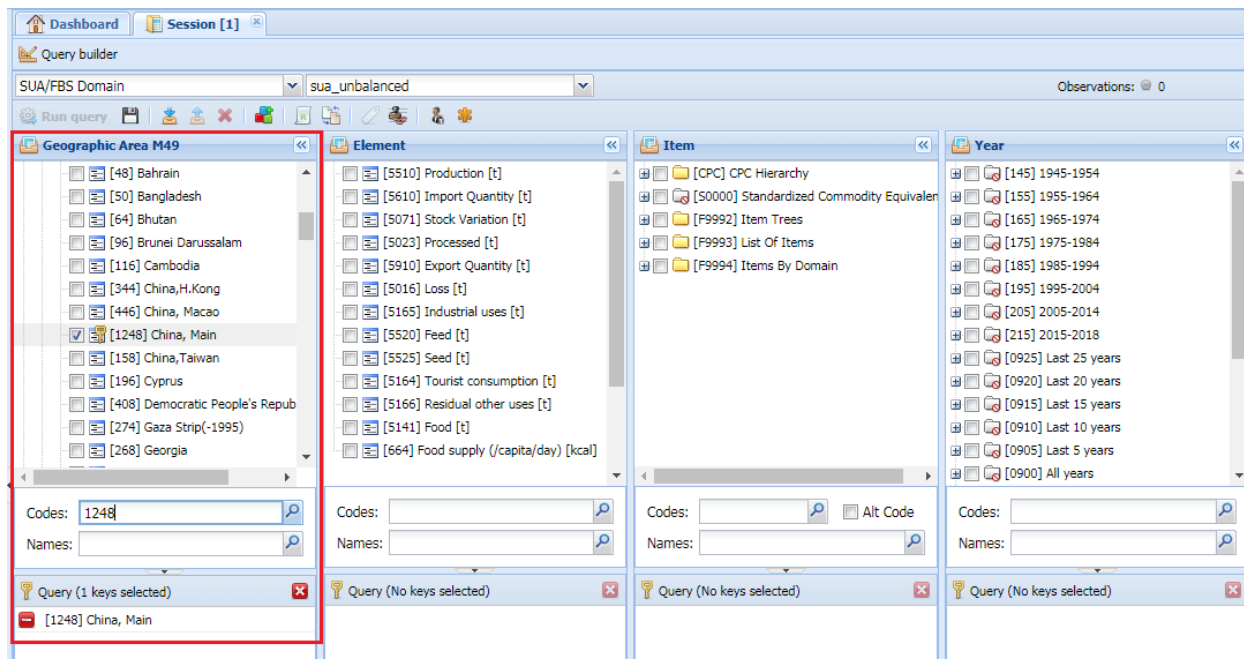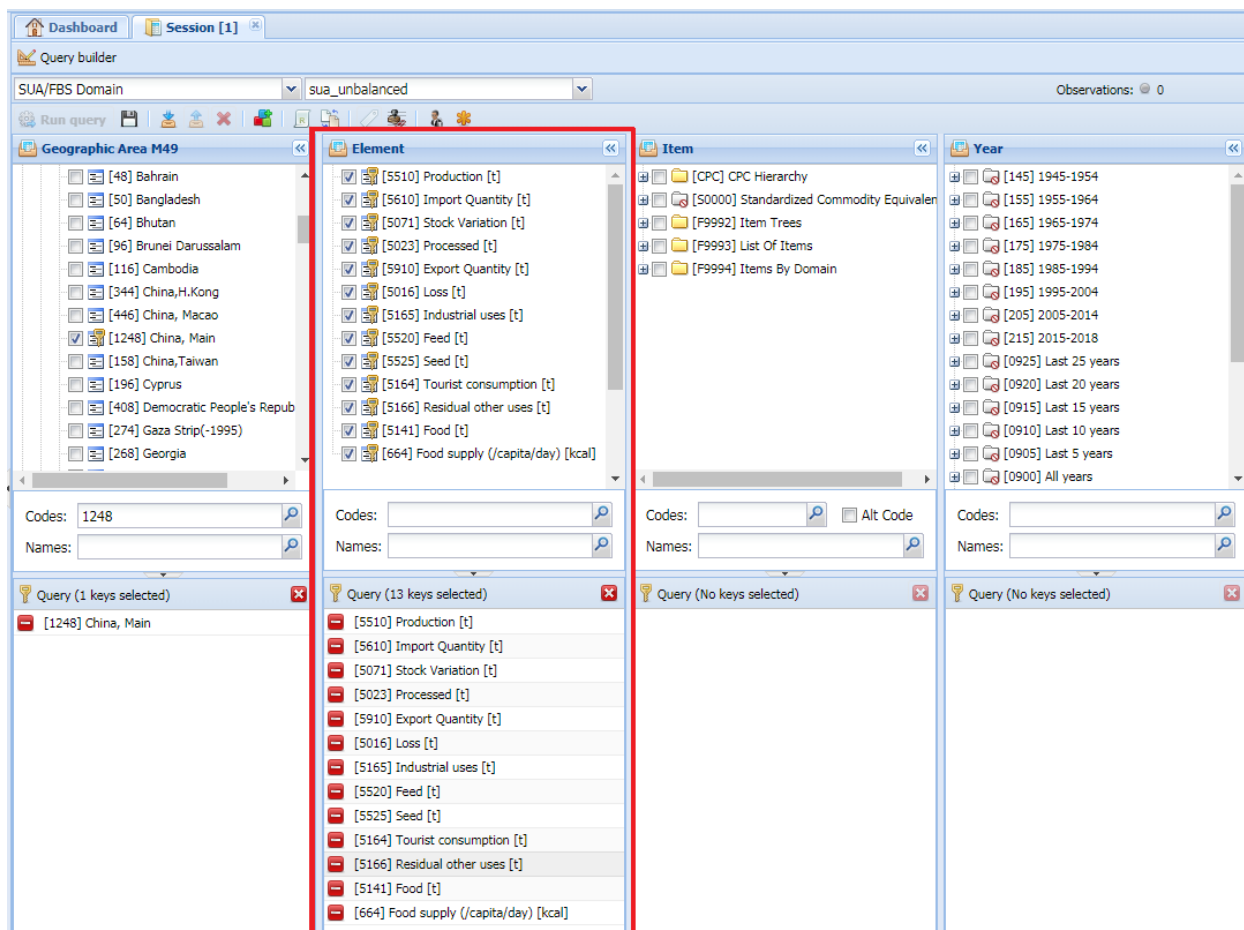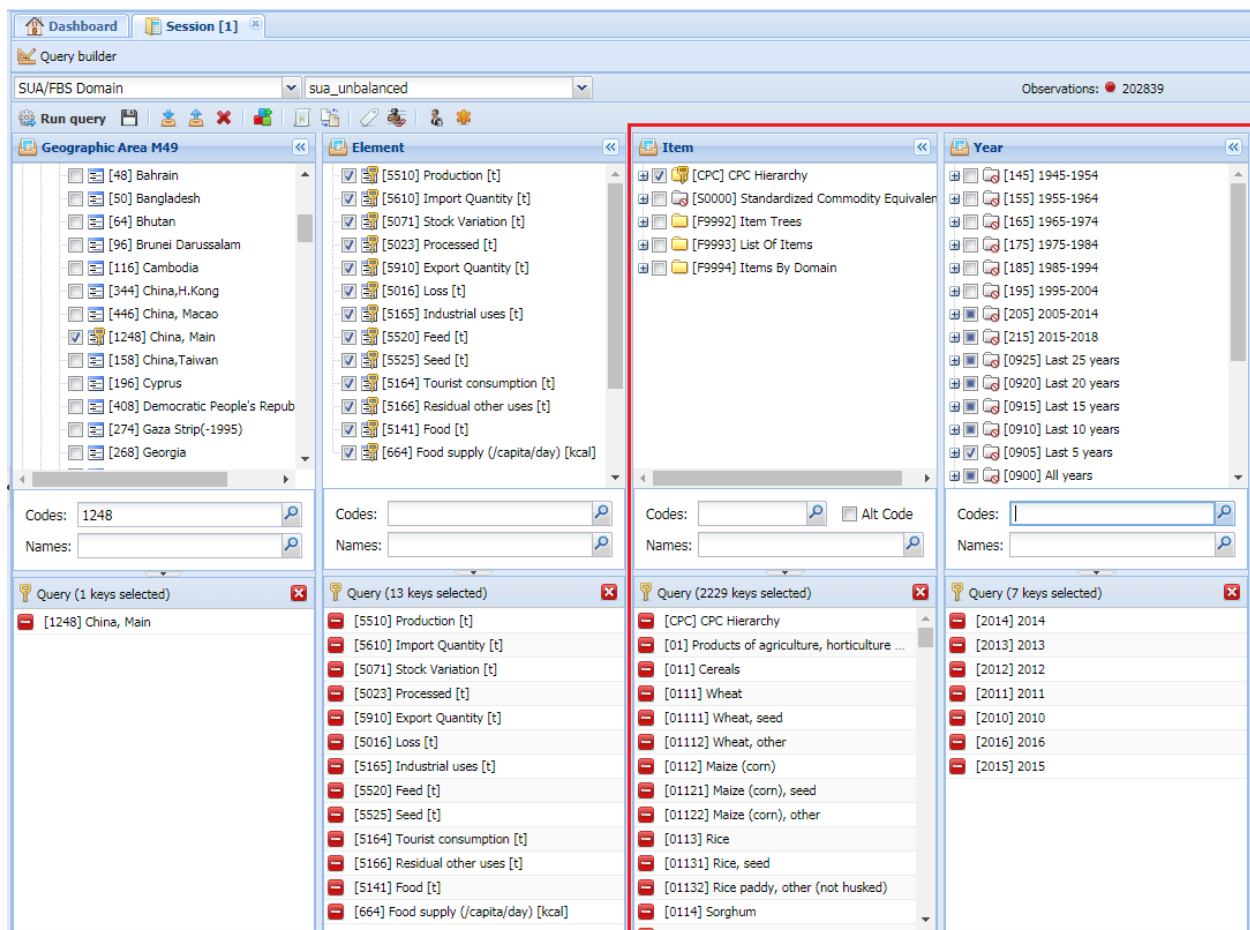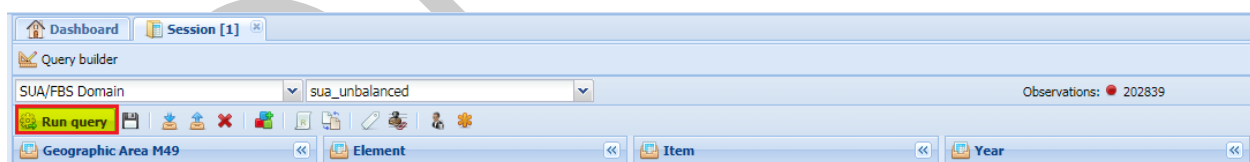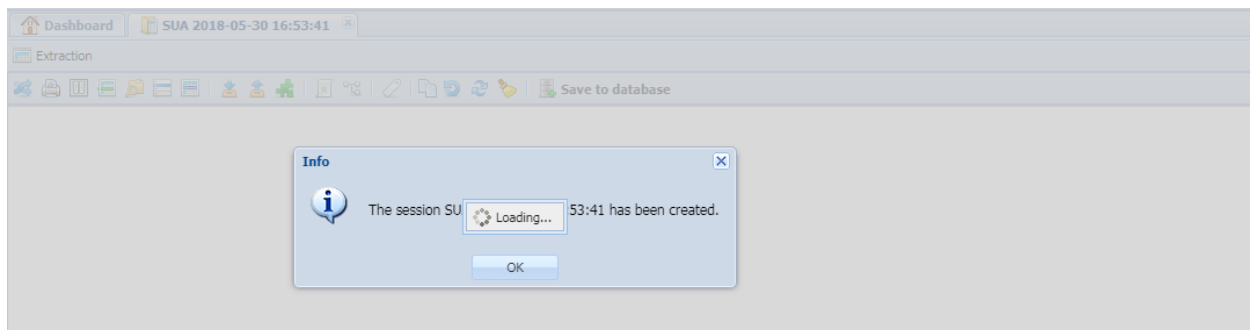Figure 10: Execution run

5 The session

The new session is reported in figure 11. Items are in the "code-list order", which means in the way they are stored in the SWS. This means that they have a numerical order but there might be some codes in a position not consistent with their number, just because they have been inserted in a different moment in the code-list.

Geographic Area M49, Item	[2010] 2010	[2011] 2011	[2012] 2012	[2013] 2013	[2014] 2014	[2015] 2015	[2016
[1248] China, Main, [23511.02] Cane sugar, non-centrifugal							
[5510] Production [t]	432,000 I i	436,000 I i	440,000 I i	440,000 I i	387,084 I e	348,470 I e	32
[5520] Feed [t]							
[5141] Food [t]	432,000 I b	436,000 I b	440,000 I b	440,000 I b	387,084 I i	348,470 I i	32
[5164] Tourist consumption [t]					-99.76 I e	-161 I e	-
[664] Food supply (/capita/day) [kcal]	3.116 I i	3.125 I i	3.134 I i	3.115 I i			
[1248] China, Main, [0111] Wheat							
[5510] Production [t]	115,181,000 -	117,410,000 -	120,580,000 -	121,926,400 -	126,208,400 p	130,185,000 q	128,84
[5610] Import Quantity [t]	1,218,722 -	1,248,822 -	3,688,617 -	5,506,712 -	2,971,249 s	2,971,794 s	3,37
[5071] Stock Variation [t]	6,089,710 I b	-5,990,972 I b	135,823 I b	1,782,592 I b	1,120,565 I i	614,435 I i	-14
[5910] Export Quantity [t]	12 T -	39,794 -	39,794 E t	2,520 -	957.5 s	5,296 s	1
[5520] Feed [t]	13,500,000 T -	26,000,000 T -	24,500,000 E f	25,500,000 E f	29,181,617 I e	29,776,854 I e	28,66
[5525] Seed [t]	4,690,000 T -	4,690,000 T -	4,580,000 T -	4,600,000 E f	4,277,567 I e	4,286,805 I e	4,29
[5016] Loss [t]	2,585,000 T -	2,635,000 T -	2,663,000 T -	2,678,000 T -	2,713,000 E f	2,900,000 E f	5,86
[5023] Processed [t]	86,800,000 E f	88,500,000 E f	89,500,000 E f	90,000,000 E f			
[5165] Industrial uses [t]	2,735,000 T -	2,785,000 T -	2,850,000 T -	2,870,000 T -	2,985,279 E e	3,077,089 E e	3,05
[1248] China, Main, [0112] Maize (corn)							
[5510] Production [t]	177,425,000 -	192,781,000 -	208,130,000 -	218,489,000 -	215,646,300 q	224,630,000 q	219,55
[5610] Import Quantity [t]	1,572,394 -	1,752,825 -	5,207,111 -	3,264,886 -	2,598,461 s	4,728,587 s	3,16
[5071] Stock Variation [t]	8,378,079 I b	8,395,828 I b	10,852,848 I b	8,509,260 I b	-37,733 I i	1,716,991 I i	-1,02
[5910] Export Quantity [t]	127,315 -	135,997 -	257,263 -	77,626 -	20,006 s	11,067 s	1
[5520] Feed [t]	114,500,000 E f	125,000,000 E f	140,000,000 E f	150,000,000 E f	143,788,614 E e	151,117,420 E e	146,74

Data | Metadata | Block metadata | History | Live helper | Validation

Data inspector — Data value — Status — Method — Revert — Apply

Figure 11: The new Session

*At the moment this dataset is filled with data coming from the old system (dataset "suaValidated2015"), from 2000 to 2013 for **all countries** and New data from 2014 onward. On these set of data, some changes are*

made from the users when the FBS are validated. If a data pull is performed in this time range for these countries, the data taken from old system would be overwritten. Is very important to look at the history of data and ask for clarification to the last persons who saved data.

6 Open the plug-in window.

For run the plug-in first the window for the plug-in selection and definition has to be opened (figure 12).
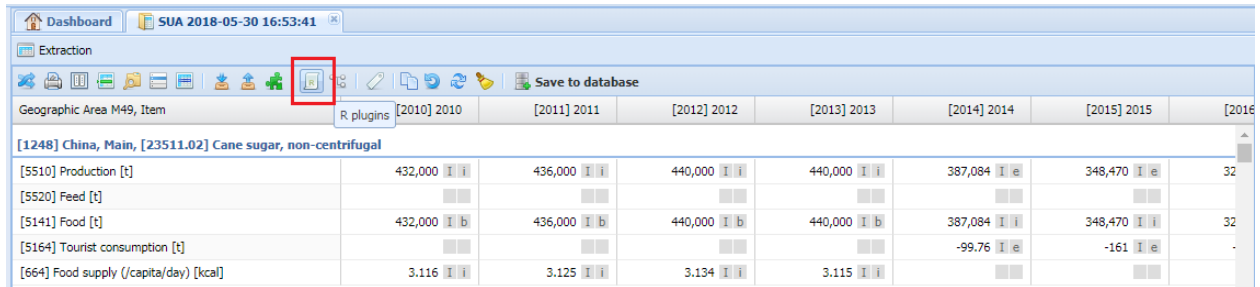

Figure 12: Plug-in botton

In the `script` session, select the plug-in *pullDataToSUA* (figures 13 and 14).
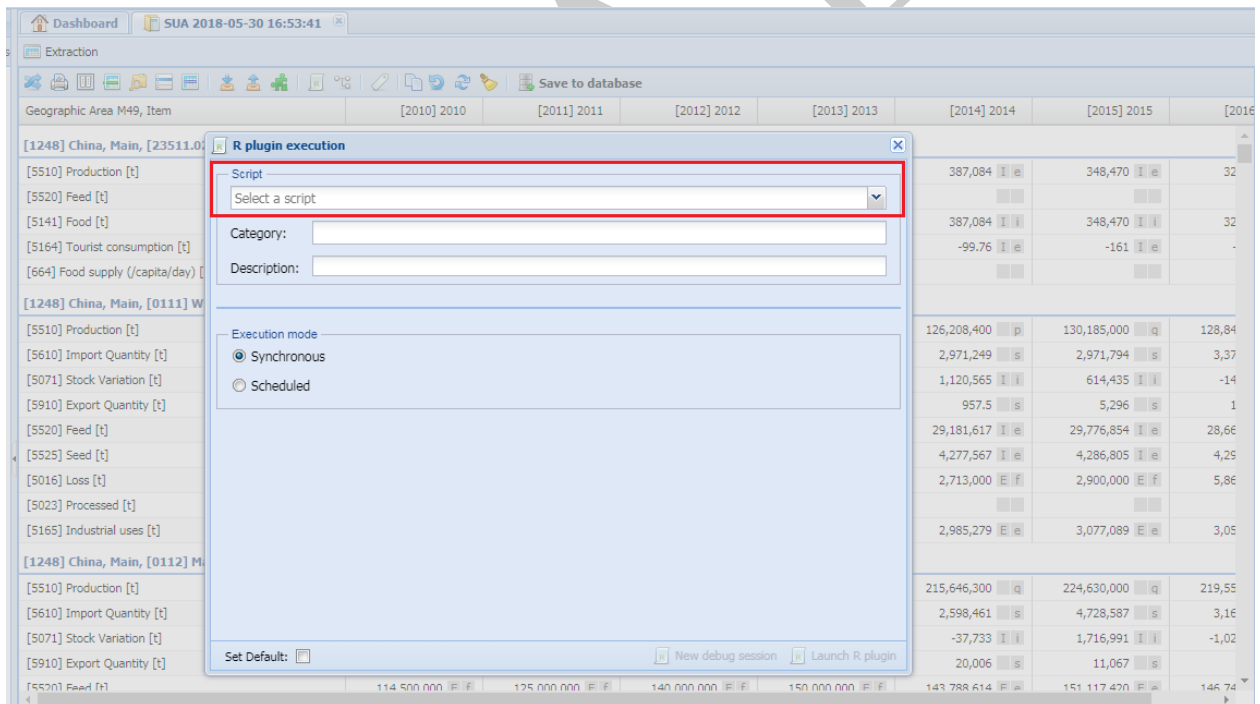
Figure 13: Plug-in window

10
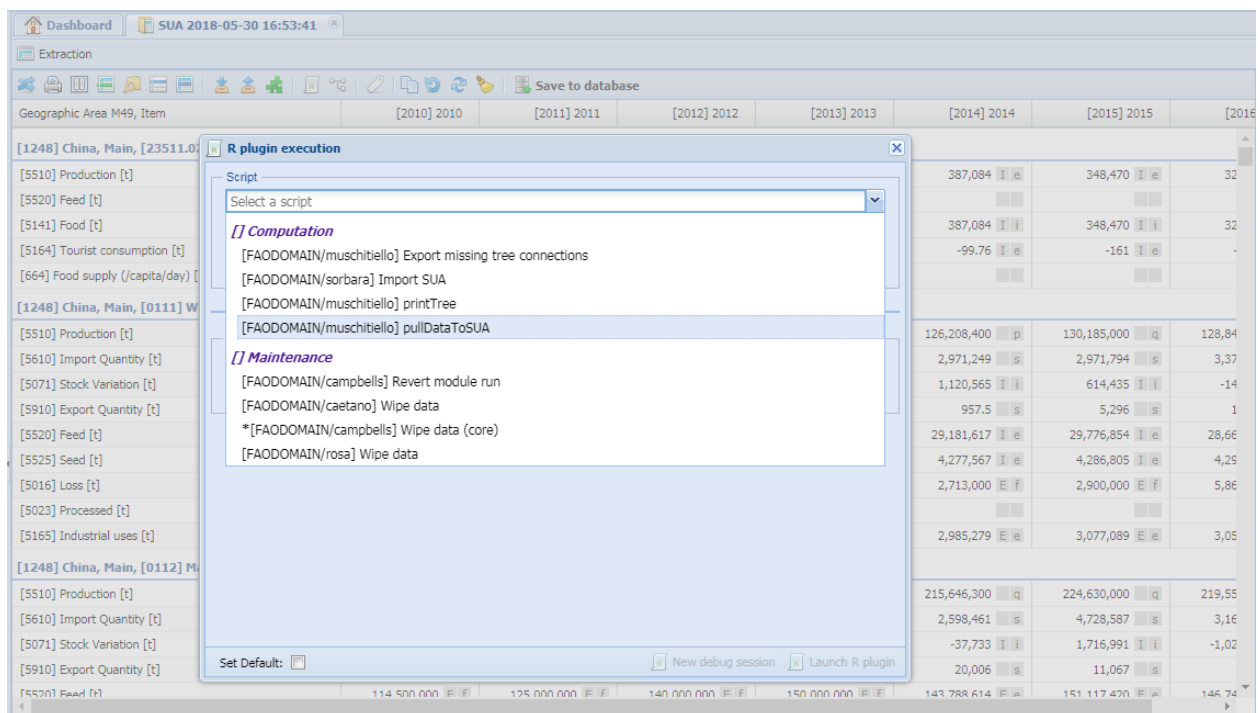
Figure 14: Select Plug-in

This will automatically bring to a sub-window where the other variables of this plugin can be selected (figure 15).
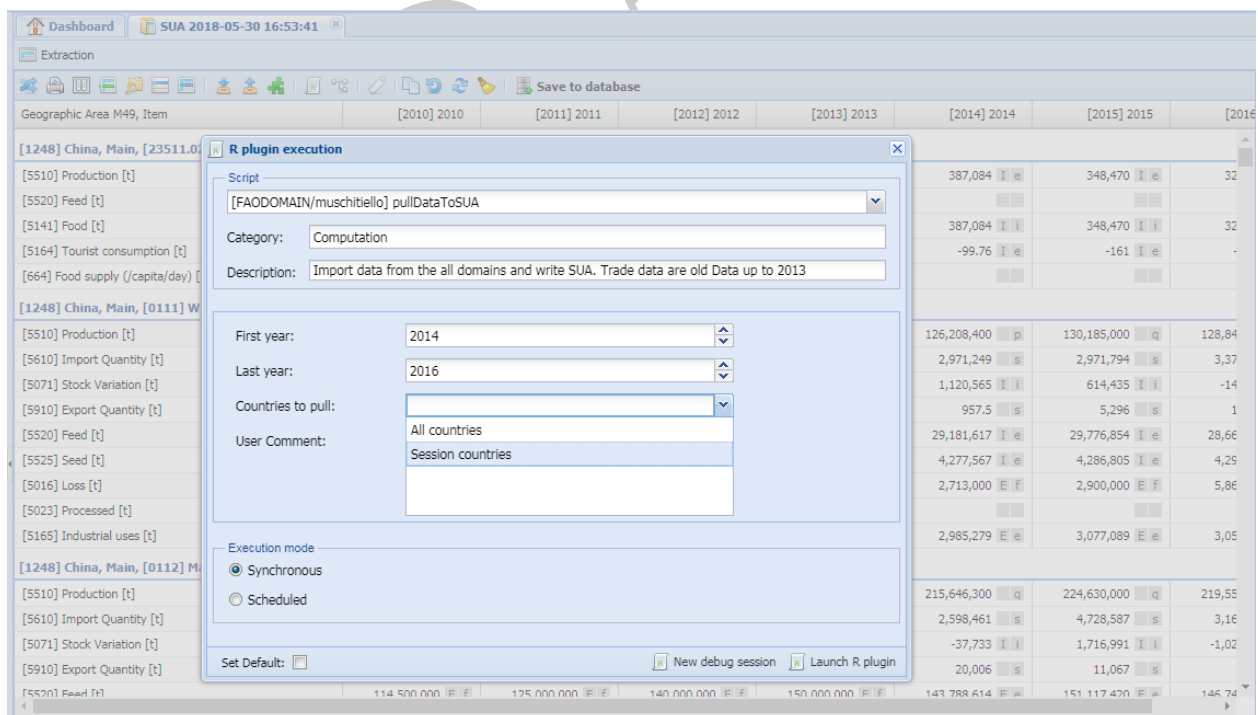
Figure 15: Select other parameters

- **First Year**: the year from which data have to be pulled. As previously said, in this data-set there are Validated and old data up to 2013, therefore in this example 2014 is selected as first year.
- **Last Year**: the last year until which the data have to be pulled.
- **countries to pull**: plugin could be run on "session countries" or "all countries". Time of execution has to be taken into account in this case. To run this plugin on all countries might require almost an hour and generate a session so big that the SWS is not able to handle it.

7 Launch Plug-in.

In the *Execution mode* section of the window, the option *Synchronous* is selected as default. This option imply that, when the *Launch R plugin* button is clicked, the run starts immediately. The other option, *Scheduled*, imply that another window is opened for the selection of the time of execution.

When the plug-in has finished to run, a window appears on the screen and an email is sent to the user (figure 17).
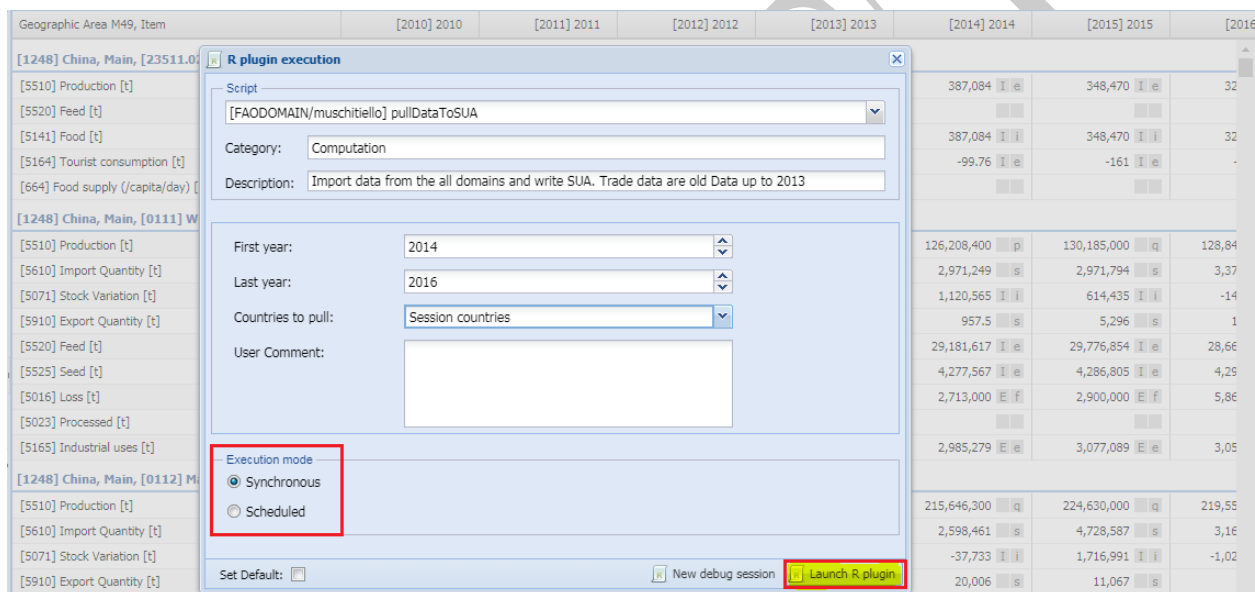
Figure 16: Launch plug-in
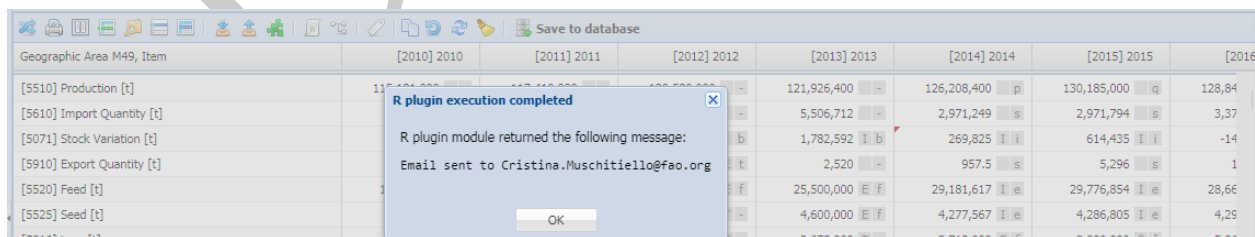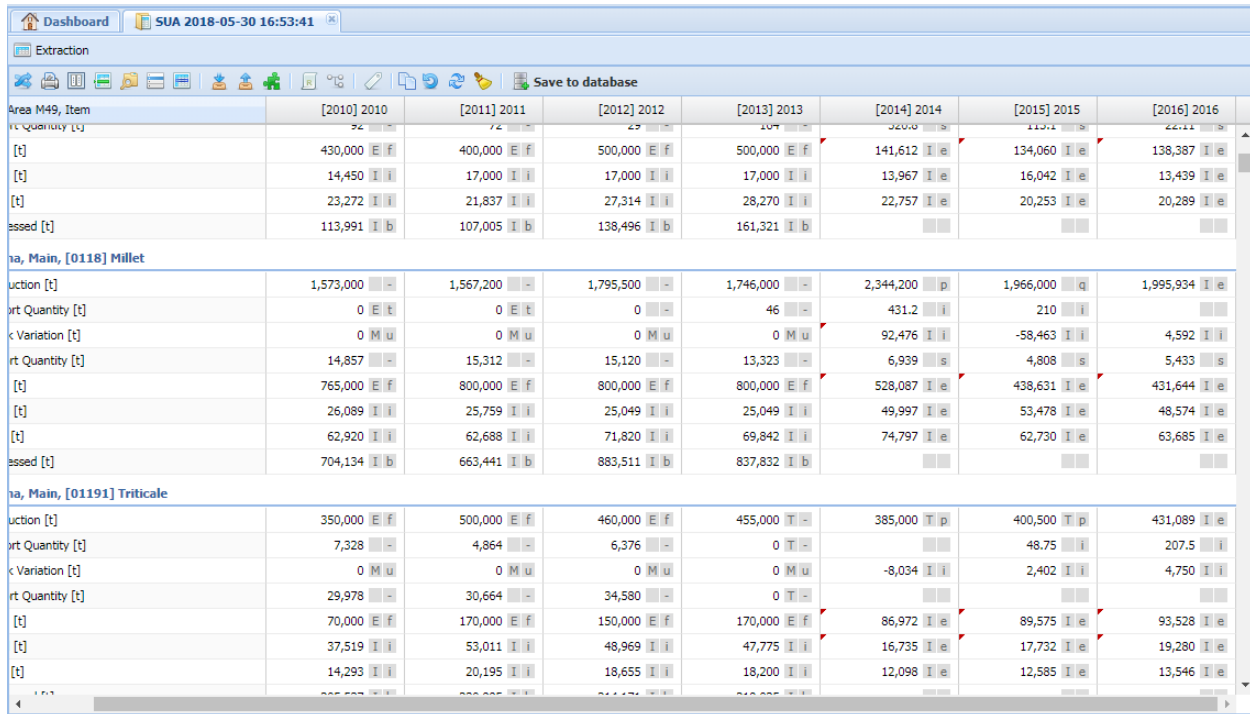
Figure 17: End of plug-in run

12

8 Session updated

In the session, all the figure that have been changed/added from the plug-in have a small red triangle on the top-left side of the figure box (figure 18).
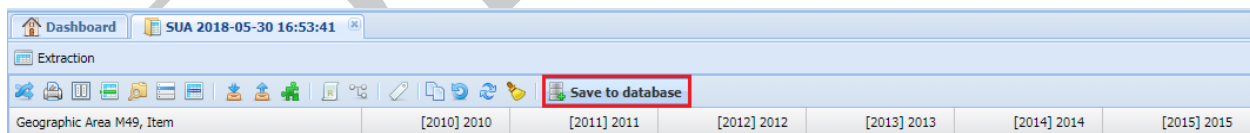

Figure 18: Updated session

9 Save back to the database

For the new figures to be used in following steps, the data have to be saved to the database.

Figure 19: Save to Database