faoswsIndustrial: Data Sources

Bruno Caetano Vidigal
Food and Agriculture Organization of the United Nations

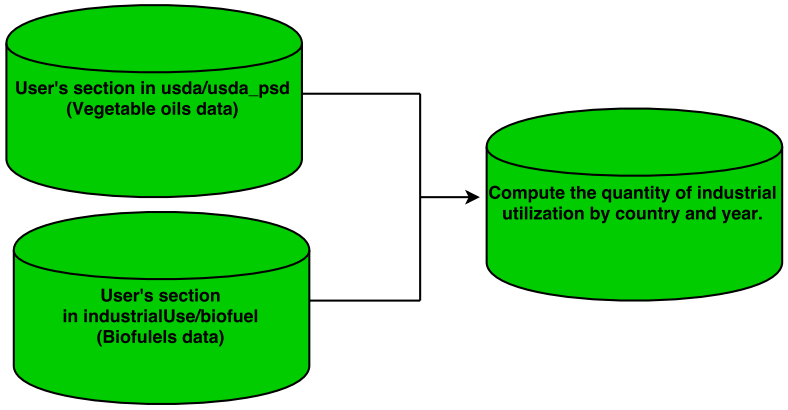
About this document

This vignette is a detailed guide of the various data sources used in the **Industrial Use**.

1 Data Sources

1.1 Flow Chart

Description of how the Industrial Use works



1.2 Example of tables

- usda domain, usda_psd dataset. In this table there are the vegetables oils data by country and year provided by the United States Department of Agriculture.

geographicUsda	measuredElementPsd	measuredItemPsd	timePointYears	Value
CA	140.08	4239100	2015	120
CA	140.08	4239100	2014	108
CA	140.08	4239100	2013	85
CH	140.08	0813100	2015	1000
CH	140.08	0813100	2014	980
CH	140.08	0813100	2013	960

- industrialUse domain, biofuel dataset. In this table there are the biofuels (such as maize) by country and year.

geographicAreaM49	measuredItemCPC	timePointYears	Value	flagObservationStatus_measuredElement_5150
112	01801	2000	39.63288	
112	01801	2001	40.95939	
112	01801	2002	42.25281	
112	01801	2003	43.67861	
112	01801	2004	44.90584	
112	01801	2005	46.82236	

2 Process of Industrial Use

1. to pull the vegetables oils data by country and year provided by the United States Department of Agriculture and stored in **usda/usda_psd**. The values are in 1000 MT. We have to convert the country code provided by USDA to m49 country codes and the item psd to cpc codes as well. There are two datasets mapping these codes and they are stored on SWS Datatables.

geographicUsda	measuredElementPsd	measuredItemPsd	timePointYears	Value
CA	140.08	4239100	2015	120
CA	140.08	4239100	2014	108
CA	140.08	4239100	2013	85
CH	140.08	0813100	2015	1000
CH	140.08	0813100	2014	980
CH	140.08	0813100	2013	960

2. to pull the biofuels data by country and year provided by the **OECD-FAO (Make double check)** and stored in **industrialUse/biofuel**. The values are in **tonnes**.

geographicAreaM49	measuredItemCPC	timePointYears	Value	flagObservationStatus_measuredElement_5150
112	01801	2000	39.63288	
112	01801	2001	40.95939	

geographicAreaM49	measuredItemCPC	timePointYears	Value	flagObservationStatus_measuredElement_5150
112	01801	2002	42.25281	
112	01801	2003	43.67861	
112	01801	2004	44.90584	
112	01801	2005	46.82236	

3. to merge both datasets by country/year/cpc and calculate the quantity of industrial utilization. This quantity is just the sum of both values from each table. It's relevant to say that we have to convert the values from the second dataset to 1000 MT multiplying by 1,000. This dataset was saved in domain = "agriculture", dataset = "aproduction", data = industrialUsesData.

geographicUsda	measuredItemCPC	timePointYears	measuredElement	Value	flagObservationStatus	flagMethod
100	2163	1992	5195	15000	I	e
100	2163	1993	5195	39000	I	e
100	2163	1994	5195	62000	I	e
100	2163	1995	5195	57000	I	e
100	2163	1996	5195	45000	I	e
100	2163	1997	5195	25000	I	e