

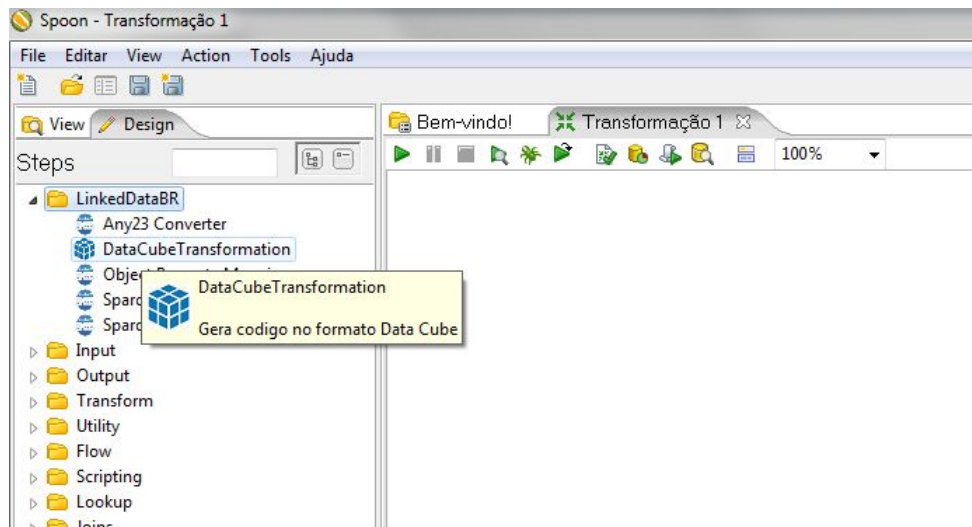
Transformação CSV para Turtle em DataCube Ontology - DataCubeTransformation Kettle Plugin

1. Montagem e execução

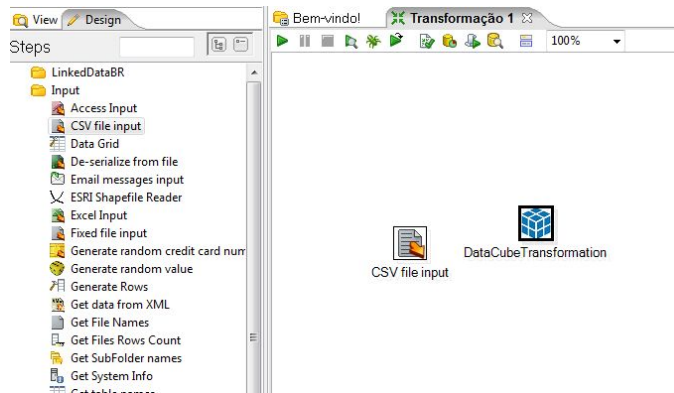
DataCube transformator é um plugin desenvolvido para o Kettle com o objetivo de transformar tabelas CSV (com dimensões e medidas em suas colunas e valores em suas linhas) no formato da ontologia <link>DataCube<link>. Ou seja, a partir de um input de um CSV, receber o código DataCube. Os passos a seguir descrevem como montar e rodar uma transformação.

Para iniciar o Plugin siga os seguintes passos:

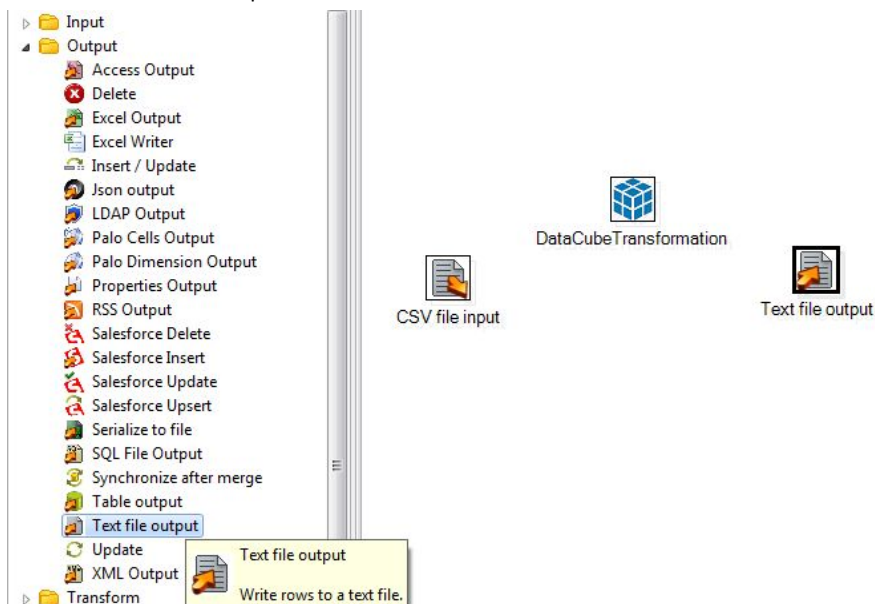
- 1 – Baixe o dataintegration4.1 já com os plugins do LinkedDataBR através do link:
<https://github.com/mayarahmo/Data-Cube-Plugin-Kettle-Transformation.git>
- 2 – Abra o Spoon.bat localizado na pasta data-integration4.1
- 3 – Vá em “file” > novo > Transformação
- 4 - Na aba “Design” vá em “LinkedDataBR” e arraste o “DataCube” para dentro do campo de transformação



- 7 – No “Input” selecione “CSV file input”

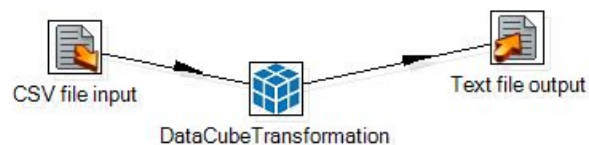


8 – No “Output” selecione “Text file output”

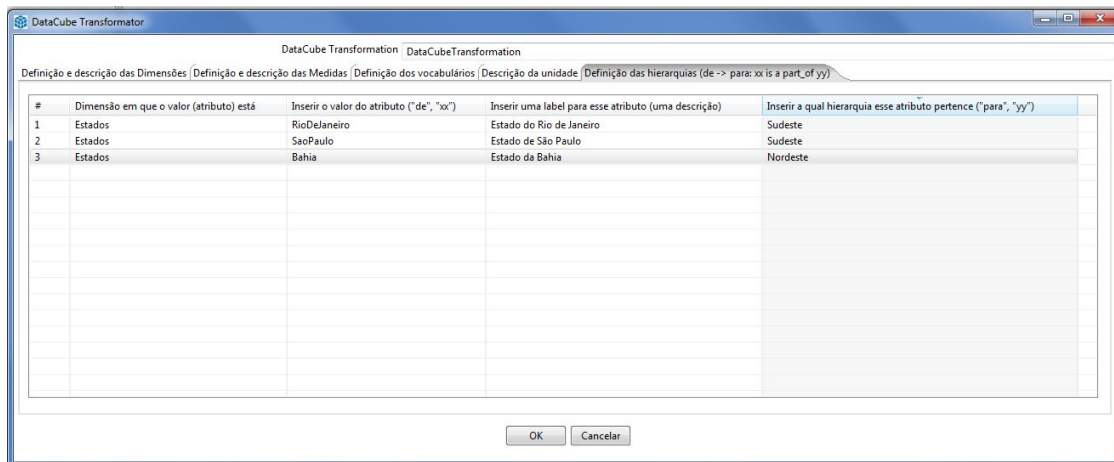


9 – Clique com o “scroll” do mouse em “CSV file input” e depois clique em “DataCubeTransformation”, selecione “Main output of step”

10 – Faça o mesmo caminho de “DataCubeTransformation” para “Text file output”

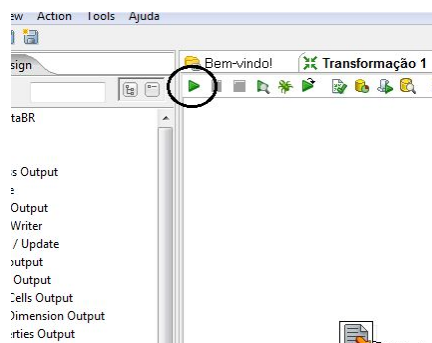


Uma das possíveis hierarquias seria dizer quais estados pertencem a cada região. Rio de Janeiro e São Paulo pertencem ao Sudeste, Bahia ao Nordeste.



20 – Clique em “OK”

21 – Rode o plugin clicando na seta verde



22 – Abra o .txt no local escolhido no output

23 – Exemplos de entrada

	A	B	C	D	E	F
1	Nome	Idade	Sexo	CPF		
2	Gabriel	22	Masculino	xxx.xxx.xxx		
3	Larissa	30	Feminino	xxx.xxx.xxx		
4	Anderson	60	Masculino	xxx.xxx.xxx		
5						
6						
7						
8						
9						
10						
11						
12						

24 – Exemplos de cada aba no plugin

The screenshot shows a software window titled "DataCube Transformator". Inside, there's a tabbed interface with five tabs: "Definição e descrição das Dimensões", "Definição e descrição das Medidas", "Definição dos vocabulários", "Descrição da unidade", and "Definição das hierarquias (de -> para: xx is a part o...)". The first tab is selected. It contains a table with columns: "#", "Dimensões", "Label da Dimensão (Descrição)", "URI da Dimensão", and "URI do tipo da Dimensão (string, float, int...)". There are four rows of data:

#	Dimensões	Label da Dimensão (Descrição)	URI da Dimensão	URI do tipo da Dimensão (string, float, int...)
1	Nome	Nome próprio	www.ExemploDeURIDE.Nome.com	http://www.w3.org/2000/01/rdf-schema#string
2	CPF	CPF de uma pessoa	www.ExemploDeURIDeCPF.com	http://www.w3.org/2000/01/rdf-schema#string
3	Sexo	Gênero	www.ExemploDeURIDeGenero.com	http://www.w3.org/2000/01/rdf-schema#string
4				

At the bottom right of the dialog are two buttons labeled "OK" and "Cancelar".

DataCube Transformator

DataCube Transformation DataCubeTransformation

Definição e descrição das Dimen Definição e descrição das Medid Definição dos vocabulários Descrição da unidade Definição das hierarquias (de ->

@base	http://example.cubeviz.org/datacube/
@prefix owl:	http://www.w3.org/2002/07/owl#
@prefix rdf:	http://www.w3.org/1999/02/22-rdf-syntax-ns#
@prefix rdfs:	http://www.w3.org/2000/01/rdf-schema#
@prefix dc:	http://purl.org/dc/elements/1.1/
@prefix skos:	http://www.w3.org/2004/02/skos/core#
@prefix obo:	http://www.w3.org/2002/07/owl#
@prefix sdmx_code:	http://purl.org/linked-data/sdmx/2009/code#
@prefix sdmx_dimension:	http://purl.org/linked-data/sdmx/2009/dimension#
@prefix cube:	http://purl.org/linked-data/cube#
@prefix ex:	http://meu.exemplo/datacube/
@prefix exProp:	http://meu.exemplo/datacube/properties/

OK Cancelar

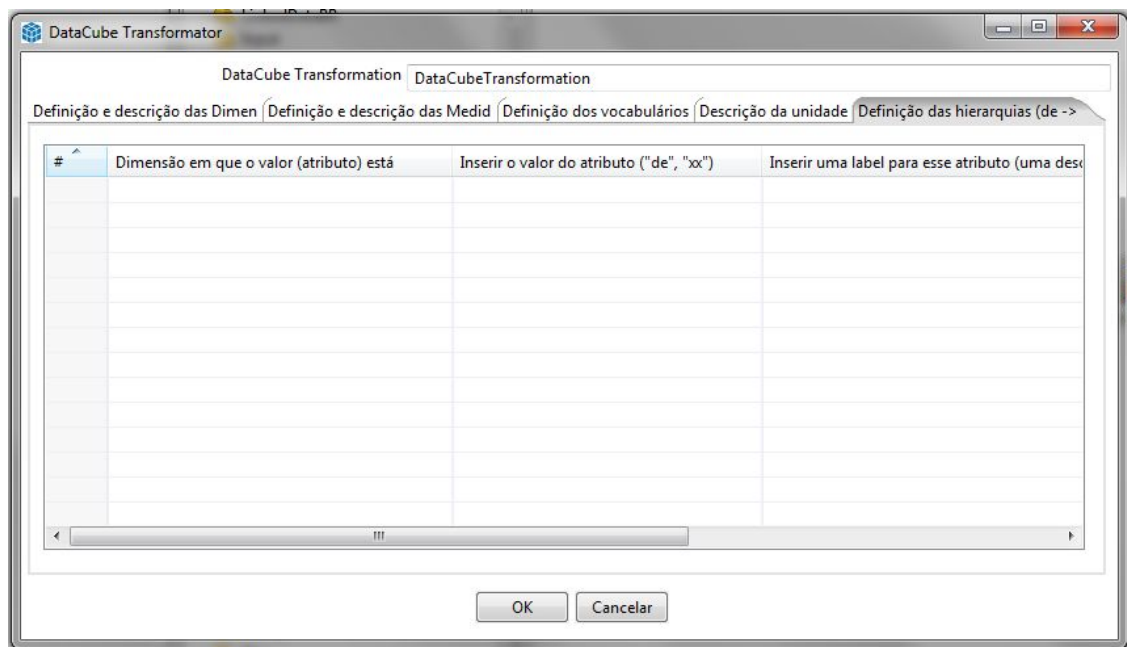
DataCube Transformator

DataCube Transformation DataCubeTransformation

Definição e descrição das Dim Definição e descrição das Med Definição dos vocabulários Descrição da unidade Definição das hierarquias (de ->

Descrição da unidade do DataSet (descrever sobre Idade de uma população por nome, CPF e gênero

OK Cancelar



Obs: o campo hierarquia fica em branco pois não há nenhuma hierarquia definida para esse exemplo. Este campo serve para complementar e dar valor semântico quando necessário. Ele não é obrigatório.

Exemplos de arquivos de saída podem ser vistos [aqui](#) e [aqui](#).

2. Um exemplo

No github <https://github.com/mayarahmo/Data-Cube-Plugin-Kettle-Transformation.git> baixe a transformação DataCubeTransformation.ktr e importe para o Kettle. Altere os caminhos no CSV file input step e no Text File Output.

3. Código fonte do plugin

Os passos a seguir descrevem como importar o código fonte do plugin DataCubeTransformation do pacote LinkedDataBR.

Download os dois projetos neste link

<https://github.com/mayarahmo/Data-Cube-Plugin-Kettle-Transformation/tree/master/Source%20Code/Java%20Source%20Projects> > KettlePluginTemplate e KettlePluginTools.

No Eclipse selecione Import > Existing projects into your workspace > Select the folders of the Projects > Finish. Para exportar File > Export > JAR File > Select the resources to export (both projects > finish

Export into [Data Integration Folder]/plugins/steps/Datacube.

4. Considerações finais

O DataCubeTransformation faz parte do pacote de plugins LinkedDataBR como o N23Converter, NTripleGenerator, ObjectPropertyMapping, SparqlEndpoint e SparqlUpdateInsert. No dataintegration4.1 há também esses outros plugins.