

Importing The Dataset

- You might have your data in .csv files, .xlsx files
- Let's load the excel data file into pandas using the `read_excel()` function. We will need to locate the directory of the excel file at first (it's more efficient to keep the dataset in the same directory as your program).

```
data=pd.read_excel(r"D:\Crude_oil_price_prediction\dataset\Crude Oil Prices Daily.xlsx")
```

- If your dataset is in some other location, Then
`Data= pd.read_excel(r"File_location")`
Note: `r` stands for "raw" and will cause backslashes in the string to be interpreted as actual backslashes rather than special characters.
- If the dataset is in the same directory of your program, you can directly read it, without giving raw as `r`.

You can create reference data sets outside of the catalog in [CSV \(comma-separated value\) format](#), then import or reimport values. You can also [import value-mapping relationships between values in multiple reference data sets](#).

A reference data set includes a number of reference data values. You can use a CSV file import to:

- Import the information about the reference data set itself. Such import does not include reference data values. The format of the CSV file is described in [CSV file format for importing governance artifacts](#).
- Import the reference data values into an existing reference data set.

The CSV import file for reference data values at a minimum consists of the following columns that are defined for the reference data set:

1. Code
2. Value
3. Description (optional)
4. Parent (optional)

For example, the first several rows of a reference data set for NAICS codes look like this:

```
Code,value,description,parent
11,Agriculture, Forestry, Fishing and Hunting
111,Crop Production,Crop Produ,11
```

```
1111,Oilseed and Grain Farming,111
11111,Soybean Farming,Soybean Farming,1111
111110,Soybean Farming,Soybean Farming,1111
11112,Oilseed (except Soybean) Farming,Oilseed (except Soybean)
Farming,1111
111120,Oilseed (except Soybean) Farming,Oilseed (except Soybean)
Farming,1111
```

You can also use an API request to import a ZIP file that includes both, the CSV files for reference data sets and separate CSV files for reference data values of each imported set. Read more about the different import methods in [Import methods for governance artifacts](#).

Importing files with reference data values

In the selected reference data set, click **Upload file** to select a CSV file from which to import values:

- Rows in the file with existing codes update existing rows.
- Rows with new codes are added.

When you import values from the CSV file, you map the columns from the file to any default or custom columns. If the columns do not exist, you can create them.

Custom columns

Values in a reference data set by default have the following columns: code, value, description (optional), and parent (optional). However, you might want to capture additional information related to the code such as translations of the value in different languages or other supportive attribute information relevant to your needs. For example, you might have a reference data set of country codes and want to capture some other attributes like prime minister of the country, language, or alternative names.

You can add custom columns during the creation of a reference data set or during a reimport of values from a CSV file into a reference data set. See the following example of a portion of a CSV file for custom columns.

```
code,value,description,Capital City,National Day,Official
Language,Population
AFG,Afghanistan,The Islamic Republic of
Afghanistan,Kabul,19/08/1919,د افغانستان اسلامي جمهوریت,37200000
ALA,Åland Islands,Åland,Mariehamn,07/05/1920,sve^nska,28007
ALB,Albania,The Republic of Albania,Tirana,28/11/1912,Albanian,2850000
DZA,Algeria,The People's Democratic Republic of
Algeria,Algiers,05/07/1962,الجزائر,42200000
```

```
ASM,American Samoa,The Territory of American Samoa,Pago  
Pago,14/06/1889,English,55465
```

When you import values from the CSV file, you map the columns from the file to any default or custom columns:

1. Click a field under **Target columns**. You can enter a name to search for an existing column. If the column does not exist, type the name for a custom column that you want to create, then click the + sign to add it.
2. If you created a column, choose a type and optionally add a description.
3. Click **Save**.
4. Repeat the steps for each column from the CSV file that you want to map.