# Exercise 4

## Advanced SDMX Converter

## Part 1 - Transcoding

In this part of the exercise, you will map Antarctica’s SDG dataset to the customized DSD you created in Exercise 2 using transcoding.

1. Open spreadsheet **Exercise 4.xlsx**. Observe that this is the same dataset used in Exercise 2, but a **Parameters** worksheet has been added. Note that the spreadsheet does not use the SDG DSD codes.
2. Open the **Parameters** worksheet. Note that it’s been prefilled with SDG concepts but the mappings are empty.
3. Map the concepts to the spreadsheet.
	* You can use the Global Registry, as described in Exercise 2, to find codes for **SERIES**, **REF\_AREA**, **AGE**, **SEX**, etc.
	* Map **REPORTING\_TYPE** to fixed value **N**, and **NATURE** to fixed value **C**. This should always be the case for national data.
	* Map **UNIT\_MEASURE** to **PT**, **UNIT\_MULT** to **0**, and **OBS\_STATUS** to **A**.
	* Map **AGE** and **SEX** to their respective columns in the data worksheet, even though they do not contain valid SDG codes. Use valid codes for all the other dimensions as appropriate.
	* As appropriate, you can use MIXED mapping for empty dimension cells, e.g.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| SEX | DIM | MIXED | COLUMN | H | FIX | \_T |

This will tell the Converter to use code **\_T** (no breakdown) when the corresponding cell is empty. Alternatively, you can use a simple COLUMN mapping and fill in the empty cells as appropriate.

1. Add a worksheet named **Transcoding**. We will use it to map codes used in the spreadsheet to valid codes from the SDG DSD.
2. Open the worksheet **Transcoding** and type headings **Text**, **Dimension**, and **Value** into the first row as follows:

The **Text** column contains the original value, **Dimension** is ID of the concept being mapped, and **Value** is the destination code from the SDMX DSD.

1. Note that the Data worksheet uses **Female** where the date relates to the female sex. The corresponding SDG DSD code for the **SEX** dimension is **F**. Map **Female** to **F** as follows:
2. Complete mappings for remaining **SEX** and **AGE** codes. Note: if you use the SDMX Global Registry to find the codes, remember that the age group you added in Exercise 2 is not in the global DSD, and therefore is not in the Global Registry. You may wish to consult the Exercise 2 manual to find the code, or open Antarctica’s DSD you created in the DSD Constructor.
3. Open the Parameters worksheet. Add a reference to the Transcoding worksheet as follows:
4. Open SDMX Converter, and convert the spreadsheet to Structure-Specific format using Antarctica’s DSD you created in Exercise 2 (file **LocalRegistry.xml** in the folder **AntarcticaRegistry**). Use **Antarctica.xml** as the output file name.

If any errors are thrown, try to diagnose and resolve the errors.
5. Once you have resolved the errors, click **Download** on the final screen and save the output file.
6. Congratulations! You mapped and converted an Excel spreadsheet using transcoding.

--- continue to Part 2 ---

## Part 2 – External parameters

In this part of the exercise, you will use an external file to store parameters and transcoding.

1. Open spreadsheet **Exercise 4.xlsx**, which you worked on in the first part of the exercise.
2. In Excel, create a new blank workbook and save it as **Parameters.xlsx**.
3. Return to **Exercise 4.xlsx** and:
	* Right-click on the worksheet **Parameters** and select **Move or Copy…**
	* In the **To Book…** dropdown, select **Parameters.xlsx**
	* Click **OK**.
4. Return to **Exercise4.xlsx** using the same procedure as above, move the Transcoding worksheet to **Parameters.xlsx**.
5. Open **Parameters.xlsx** and ensure that it has two worksheets, **Parameters** and **Transcoding**. If one of these is missing, return to **Exercise4.xlsx** and move it from there. If there are any worksheets in addition to these (e.g. Sheet1), delete them.
6. Open **Exercise4.xlsx** and ensure it only has one worksheet named Data.
7. Close Excel.
8. Open SDMX Converter, load spreadsheet **Exercise4.xlsx** and click Next.
9. Load Antarctica’s DSD (**LocalRegistry.xml**) and click Next.
10. In the next screen:
	* In the **Header configuration** section, check **Manual Config**
	* In the **External Parameter File** section, check **External Excel File that contains only Parameter Sheets, Mapping Sheet and/or a Transcoding Sheet**.
	* Click **Select>** and select file **Parameters.xlsx**.
	* Do not check the box Manual Config (Manually map data Sheets and parameter Sheets, the data will be collected in a consecutive screen)
	* Click **Next**.
11. In the screen that opens, click **Save Header** and save the file as **header.prop** into the Exercise 4 folder.
	* You will use this file in the next part of the exercise.



1. Click **Next** and complete the conversion.
2. Congratulations! You completed the conversion of an Excel file to SDMX using an external Parameter file.

--- continue to Part 3 ---

## Part 3 – Using the SDMX Converter command-line application

In this part of the exercise, you will use the SDMX Converter command-line interface to convert an Excel spreadsheet to SDMX.

1. If you haven’t done so, install SDMX Converter Command-Line Interface (CLI):
	* Download file [converter-cli-app\_v9.5.0\_2022.07.25.zip](https://circabc.europa.eu/ui/group/088149e5-0472-405b-839b-57d5970052cc/library/7fe02de8-60ea-4d2b-87e3-06afdedc2ed7/details) from the EU web site.
	* Open the zip file you downloaded and navigate to
	**converter-cli-app\_v9.5.0\_2022.07.25**->**app**
	* Locate the folder **ConverterCLIApp** and copy it to the root of the drive C:, i.e. to
	**C:\ ConverterCLIApp**
		+ If you do not have the permissions to copy the folder to the root of drive C:, copy it to another folder, e.g. C:\Temp.
2. Open the Search box in the lower left corner of the screen, type **cmd** and click on **Command Prompt**.
	* This procedure may slightly vary depending on the version of Windows you are running.
3. Type the following commands:
	* **c:**
	* **cd \ConverterCLIApp**

NB: if you extracted Converter CLI to a different folder, you need to type its address here.

1. Type **dir** and press enter. You will see a list of files in the folder. In Windows, we use batch file **converter-cli.bat** to start the command line application.
2. Type the following command:
	* **converter-cli.bat -help**

You will see a screen with a description of the various options and functions supported by the
SDMX Converter command-line app.

1. In File Explorer, open the folder **Exercise 4**. Make sure that it has the files you previously created, **Exercise 4.xlsx** and **Parameters.xlsx**. If you saved the data file in this folder (e.g. **Antarctica.xml**), delete it. Do not delete the DSD file if you have it there.
2. Open Notepad or another text editor. Copy and paste into Notepad the following command:

**converter-cli.bat -dsd\_file "****<path\_to\_AntarcticaRegistry\_folder>\LocalRegistry.xml" -df true -df\_agency "AQ1" -df\_id "DF\_SDG\_AQ" -df\_version "1.0" -from EXCEL -inputFile "<path\_to\_Exercise4\_folder>\Exercise 4.xlsx" -parameter\_file "<path\_to\_Exercise4\_folder>\Parameters.xlsx" -to STRUCTURE\_SPECIFIC\_DATA\_2\_1 -validation true -header\_file "<path\_to\_Exercise4\_folder>\header.prop" -outputFile "<path\_to\_Exercise4\_folder>\Antarctica.xml"**

1. Replace **<path\_to\_AntarcticaRegistry\_folder>** with the path to the AntarcticaRegistry folder you created in Exercise 2.
	* Tip: you can copy the path from File Explorer.
2. Replace **<path\_to\_Exercise4\_folder>** with the path to the folder where Exercise 4 is saved.
3. Copy the updated command from Notepad and paste it into the Command Prompt.
	* Tip: right-click in the Command Prompt to paste the command.
4. Run the command. You should see the output for successful conversion:
5. For reference, the following parameters were used with the Command Line application:

**-dsd\_file "<path\_to\_Exercise4\_folder>\AQ\_SDG\_DSD.xml"** – path to the Structure file with the DSD
**-df true** – specifies that a dataflow must be used for conversion
**-df\_id "DF\_SDG\_AQ"** – specifies ID of the dataflow
**-df\_agency "AQ1"** - specifies maintenance agency of the dataflow
**-df\_version "1.0"** – specifies version of the dataflow
**-from EXCEL** – specifies source format (must be provided in CLI if input file is not XML)
**-inputFile "<path\_to\_Exercise4\_folder>\Exercise 4.xlsx"** – path to the input data file
**-parameter\_file "<path\_to\_Exercise4\_folder>\Parameters.xlsx"** – path to Excel parameter file
**-to STRUCTURE\_SPECIFIC\_DATA\_2\_1** – specifies output format
**-validation true** – specifies that validation must be performed
**-header\_file "<path\_to\_Exercise4\_folder>\header.prop"** – path to the XML header file
**-outputFile "<path\_to\_Exercise4\_folder>\Antarctica.xml"** – path to the output file.

12. Congratulations! You converted a mapped Excel file using the SDMX Converter command-line application.