



SDMX Converter

Workshop on Data and Metadata Sharing
Bangkok, 10-14 December 2018

Abdulla Gozalov, UNSD



SDMX Converter

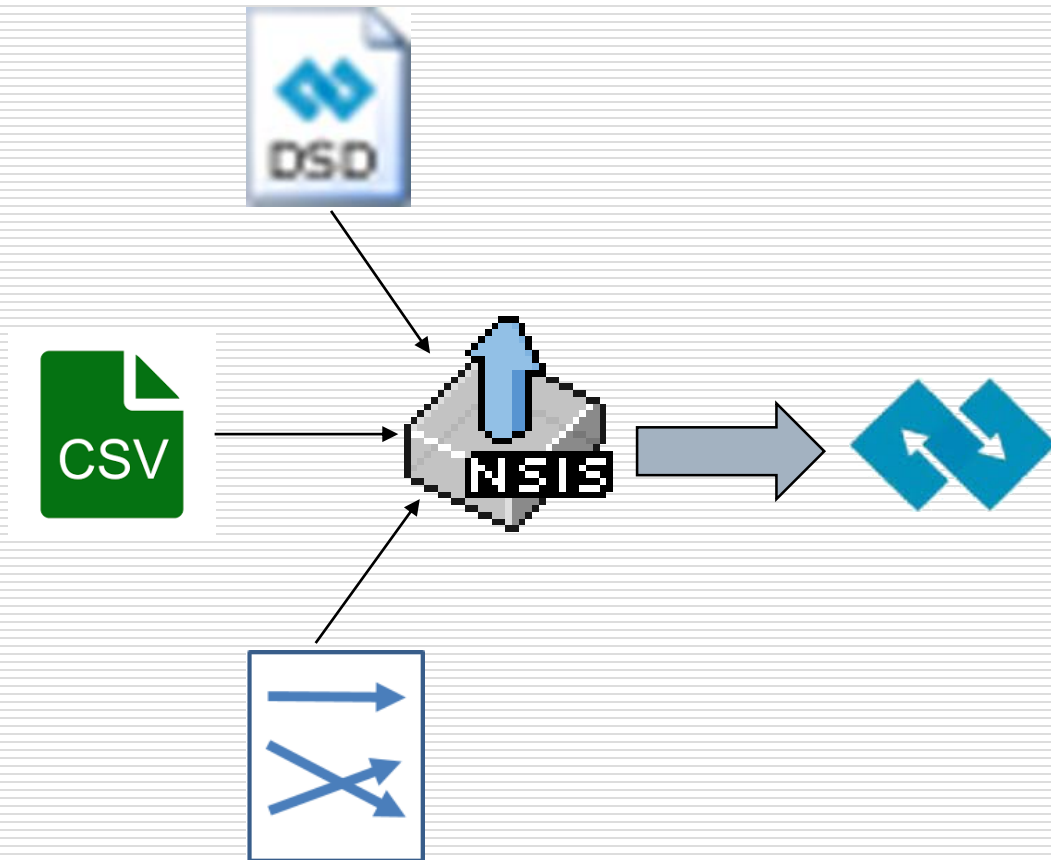
- Software developed by Eurostat
- Can be used to convert data from a variety of format into SDMX and vice versa
- Supports CSV, Excel, DSPL, and others
- Can be used to transform....
 - Non-SDMX data to SDMX
 - SDMX to non-SDMX
 - SDMX format to another SDMX format



SDMX Converter: Applications

- SDMX Converter is available as
 - Desktop application with a Graphical User Interface
 - Command-line application
 - Web service
 - Java library
- Download from <https://ec.europa.eu/eurostat/web/sdmx-infospace/sdmx-it-tools/sdmx-converter>

Converting data to SDMX



- To transform data to SDMX using the SDMX Converter, you need
 - Source data as CSV, DSPL, Excel, etc.
 - A Data Structure Definition (DSD) according to which the SDMX dataset will be structured
 - Mappings that show how the source data maps to the concepts of the Data Structure Definition
 - As always in setting up SDMX exchange, configuring mappings takes the most time and effort



Using SDMX Converter with Excel

- Data and mappings can be placed into the same spreadsheet
- Additional information can be added to facilitate data entry
 - E.g. code lists for validation and display of descriptions



Worksheet names

- Worksheet named **Parameters** contains mappings
 - Only one can be used at a time
 - Shows how cells, rows, and columns map to the DSD dimensions and attributes
- Worksheet names starting with **Val** are ignored
 - Can be used to store code lists or other ancillary information
- All other worksheets are considered to contain data and will be processed



Excel Mappings

Mapping type

Cell where data starts

Concept

Number of obs. columns

Concept Role

	A	B	C	D	E	F	G
1	Element	Type	PosType	Position		DataStart	K12
2	FREQ	DIM	FIX	A		NumColumns	1
3	REPORTING TYPE	DIM	CELL	B4			
4	SERIES	DIM	CELL	B3			
5	REF_AREA	DIM	COLUMN	2			
6	TIME_PERIOD	DIM	COLUMN	3			
7	SEX	DIM	COLUMN	4			
8	AGE	DIM	COLUMN	5			
9	URBANISATION	DIM	COLUMN	6			
10	INCOME WEALTH_QUANTILE	DIM	COLUMN	7			
11	EDUCATION_LEV	DIM	COLUMN	8			
12	OCCUPATION	DIM	COLUMN	9			
13	CUST_BREAKDOWN	DIM	FIX	_T			
14	COMPOSITE_BREAKDOWN	DIM	COLUMN	10			
15	DISABILITY_STATUS	DIM	COLUMN	11			
16	OBS_VALUE	DIM	OBS_LEVEL	12			
17	NATURE	ATT	COLUMN	13			
18	SOURCE_DETAIL	ATT	COLUMN	14			

Position or value



Excel mappings worksheet

- **Element:** name of the DSD concept
- **Type:** role of the concept
 - **DIM:** Dimension
 - **ATT:** attribute
- **DataStart:** the first cell containing an observation
- **NumColumns:** number of observations per row



Column PosType: mapping or position type

- The following mapping/position types are supported:
 - **CELL**
 - **ROW**
 - **COLUMN**
 - **FIX**
 - **OBS_LEVEL**
 - **MIXED**
 - **SKIP**



Mapping type: CELL

- The value for the entire dataset is provided in the cell provided in the column **Position**
- E.g. if the spreadsheet is expected to only contain data for a single country, its code can be provided in a cell.

	A	B	C	D	E	F	G	H
1	Percentage of ever-married women who have experienced emotional, physical or sexual violence by any husband in the past 12 months, according to background characteristics, Nepal DHS 2016							
2	Background characteristic	Emotional violence	Physical violence	Sexual violence	Physical and sexual	Physical and sexual and emotional	Physical or sexual	Physical or sexual or emotional
3	Age							
4	15-19	8.3	14	5.7	2.6	1.4	17	18.1
5	20-24	6	10.3	5.8	3.7	2.2	12.4	13.4
6	25-29	7.9	10.9	4.1	3	2	12	13.9
7	30-39	8.5	10.5	3.5	2.8	2.3	11.2	14.2
8	40-49	7.4	7.2	2.9	1.9	1.5	8.2	10.9
9								
10	sources: Nepal DHS 2016							



Mapping type: ROW

- Values for the concept are stored in the row specified in column **Position**

	A	B	C	D	E	F	G	H
1	Percentage of ever-married women who have experienced emotional, physical or sexual violence by any husband in the past 12 months, according to background characteristics, Nepal DHS 2016							
2	Background characteristic	Emotional violence	Physical violence	Sexual violence	Physical and sexual	Physical and sexual and emotional	Physical or sexual	Physical or sexual or emotional
3	Age							
4	15-19	8.3	14	5.7	2.6	1.4	17	18.1
5	20-24	6	10.3	5.8	3.7	2.2	12.4	13.4
6	25-29	7.9	10.9	4.1	3	2	12	13.9
7	30-39	8.5	10.5	3.5	2.8	2.3	11.2	14.2
8	40-49	7.4	7.2	2.9	1.9	1.5	8.2	10.9
9								
10	sources: Nepal DHS 2016							



Mappings type: COLUMN

- Values for the concept are stored in the column specified in column **Position**

	A	B	C	D	E	F	G	H
1	Percentage of ever-married women who have experienced emotional, physical or sexual violence by any husband in the past 12 months, according to background characteristics, Nepal DHS 2016							
2	Background characteristic	Emotional violence	Physical violence	Sexual violence	Physical and sexual	Physical and sexual and emotional	Physical or sexual	Physical or sexual or emotional
3	Age							
4	15-19	8.3	14	5.7	2.6	1.4	17	18.1
5	20-24	6	10.3	5.8	3.7	2.2	12.4	13.4
6	25-29	7.9	10.9	4.1	3	2	12	13.9
7	30-39	8.5	10.5	3.5	2.8	2.3	11.2	14.2
8	40-49	7.4	7.2	2.9	1.9	1.5	8.2	10.9
9								
10	sources: Nepal DHS 2016							



Mapping type: COLUMN (2)

- Also used with record-based representation, when row contains one record or observation

	A	B	D	E	F	G	H	I	J	K	L	M	N
1	M49 Code	Series Code	Indicator R	Country	Disaggre	Year	IMR	Age Group	Unit	Nature	Footnote	Source Det	Time Detail
2	4	SH_DYN_IM	3.2.1	Afghanistan	BOTHSEX	2000	90.8	<1Y	PER_1000_LIVE	NA		Source: Uni	2000
3	8	SH_DYN_IM	3.2.1	Albania	BOTHSEX	2000	23.1	<1Y	PER_1000_LIVE	NA		Source: Uni	2000
4	12	SH_DYN_IM	3.2.1	Algeria	BOTHSEX	2000	33.9	<1Y	PER_1000_LIVE	NA		Source: Uni	2000
5	20	SH_DYN_IM	3.2.1	Andorra	BOTHSEX	2000	4	<1Y	PER_1000_LIVE	NA		Source: Uni	2000
6	24	SH_DYN_IM	3.2.1	Angola	BOTHSEX	2000	122.9	<1Y	PER_1000_LIVE	NA		Source: Uni	2000
7	28	SH_DYN_IM	3.2.1	Antigua and	BOTHSEX	2000	13.1	<1Y	PER_1000_LIVE	NA		Source: Uni	2000
8	32	SH_DYN_IM	3.2.1	Argentina	BOTHSEX	2000	17.3	<1Y	PER_1000_LIVE	NA		Source: Uni	2000
9	51	SH_DYN_IM	3.2.1	Armenia	BOTHSEX	2000	26.6	<1Y	PER_1000_LIVE	NA		Source: Uni	2000
10	36	SH_DYN_IM	3.2.1	Australia	BOTHSEX	2000	5.1	<1Y	PER_1000_LIVE	NA		Source: Uni	2000
11	40	SH_DYN_IM	3.2.1	Austria	BOTHSEX	2000	4.6	<1Y	PER_1000_LIVE	NA		Source: Uni	2000
12	2	SH_DYN_IM	3.2.1	Azerbaijan	BOTHSEX	2000	60.8	<1Y	PER_1000_LIVE	NA		Source: Uni	2000



Mapping type: **FIX**

- Fixed value for the entire dataset is stored in the column **Position** and does not appear in the data spreadsheet
 - E.g. if the data is always expected to be annual, frequency can be coded for the entire spreadsheet

	A	B	C	D
1	Element	Type	PosType	Position
2	FREQ	DIM	FIX	A
...



Mapping type: MIXED

- The concept value is conditional
- Can be used to provide a default value

	A	B	C	D	E	F	G
1	Element	Type	PosType	Position		DataStart	K12
2	FREQ	DIM	FIX	A		NumColumns	1
3	REPORTING_TYPE	DIM	CELL	B4			
4	SERIES	DIM	COLUMN	2			
5	REF_AREA	DIM	MIXED	CELL	B3	FIX	TH
6	TIME_PERIOD	DIM	COLUMN	4			

- “Use cell B3 for concept REF_AREA. If the cell is empty, use fixed value TH”



Mapping type: OBS_LEVEL

- Can be used in to specify attributes attached at the observation level relative to the cell containing the observation.
 - E.g. when each row has multiple observations *and their attributes*.

	A	B	C	D	E	F	G
37	OBS_EDP_WBB	ATT	SKIP				
38	OBS_STATUS	ATT	MIXED	OBS_LEVEL 1 CELL			H14
39	REF_PERIOD_DETAIL	ATT	CELL	B4			
40	REF_YEAR_PRICE	ATT	CELL	H6			

- “For attribute OBS_STATUS, use cell that is 1 column to the right of the cell containing the observation value. If that cell is empty, use the value in cell H14.”



Transcoding

- Transcoding refers to code mapping, when internal codes are different from DSD codes.

Internal Age Group		SDG DSD Age Group
15-19	➔	Y15T19
20-24	➔	Y20T24
25-29	➔	Y25T29

- With CSV and DSPL, you can configure transcodings in external files or specify them directly in the SDMX Converter.
- With Excel, SDMX Converter expects to find DSD codes in the spreadsheet.



THANK YOU!