29 October 2023

Preparing data sets for conversion from “long” to “wide” format

Prepare your input file following this structure:

Column name	Description	Values
unique_id	Unique identifier for each row	You can use your own unique ID. If you do not have one, then just number each row sequentially (1, 2, 3, 4, etc.)
year4	year in which death occurred (expressed in 4 digits)	e.g. 2021
sex	Sex of deceased	**M**: Male **F:** Female: **U:** unknown sex or you can leave it blank or use any other code
age_unit (N, H, D, M, Y) age (see column “Values”)	N = minutes	1-59
	H = hours	1-23
	D = days	1- 27
	M = months	1 -11
	Y = years	1 = 1 year 2= 2 years Up to 120 = 120 years or more 999 = unknown years 888 = unknown days (infant deaths only)
ucod	Underlying cause of death	ICD codes, e.g. A419 (NB : do not put a “.” in the ICD codes)

Do **NOT** modify the column names. Save your file as CSV.

Example of file records using the above structure – your file should have 6 columns in all

Then process the data set in the
Data Formatting Tool

https://icd.who.int/anacod#/conversion

29 October 2023

unique_ID	year4	sex	age_unit	age	ucod
1	2020	M	Y	8	I517
2	2020	F	M	1	W75
3	2020	M	D	10	R95
4	2020	M	M	7	E713
5	2020	M	Y	56	Y20
6	2020	M	D	5	R95
7	2020	M	M	11	A090
8	2020	F	M	8	R95
9	2020	M	H	5	P072
10	2020	M	M	9	Q939
11	2020	F	H	22	P018
12	2020	M	H	4	P072
13	2020	M	Y	3	Q897
14	2020	F	Y	68	I514
15	2020	M	Y	87	R99

The data will be converted from long to wide format and available for download.

However, if your original file contains some errors, a separate file showing the errors will be downloaded as well. Correct these errors and re-upload the entire file for conversion.