Sending statistical series to WID.world

Lucas Chancel, Amory Gethin

August 2017

Introduction

This technical note provides information to researchers willing to send statistical series to the WID.world team for integration to the website.

For any question, please contact amory.gethin@psemail.eu or lucas.chancel@wid.world

Material to be sent

Any series sent to WID.world should contain the three following:

1. Final data files to be integrated to the website (see below for detail).
2. A Working paper (see examples here) or Technical note (see examples here) describing the methodology followed. Technical Notes are short papers that only describe the methodology. Working Papers are longer, they often also discuss and analyse the results and can sometimes be turned into research articles.
3. Computer codes and/or datasheets enabling the creation of the final data files.

For further information on concepts and methods used in WID.world, please refer to our Distributional National Accounts Guidelines.

WID.world variables codes

We prefer when series are sent along with their WID.world variable codes, which allow simple identification of the series type (eg. Macro aggregates, averages, shares, etc.), concept (eg. income, wealth, etc.) and the corresponding statistical population (eg. adults, males, etc.).

Variables codes are composed of:

- A prefix: eg. "m" for macroeconomic aggregate.
- A five-letter variable code: eg. "nninc", for national income.
- A 3-number age code: eg. "999", corresponding to "all ages".
- A 1-letter population code: eg. "i", corresponding to individuals.

Variables codes are defined in this document. Please ensure you always have the latest version of the document when using it (ie. by downloading it directly from the above link).

If you have difficulties matching your variables to WID codes, feel free to contact us for any further information

Monetary variables format

- **Unit**: except for indices and ratios, we prefer when variables are expressed in their original unit. For instance, we prefer that population be expressed in number of individuals rather than in thousands or millions.

- **Currency**: We prefer that the series be in local current currency. They are then converted using deflators and exchange rates available on WID.world (unless there is a good reason to modify currently available WID.world deflator or exchange rate series - in this case, a deflator or exchange rate series should also be provided by the researcher). In the database, the series are stored in constant local currency. For information on conversion of WID.world series, read this technical note.

Inequality series format

To facilitate integration to the database, inequality series should be sent in "gpinter" output format, (GPinter is an online statistical tool available on WID.world and R). Ideally, a single Excel file would contain all the series, in the following way:

1. The first sheet summarizes years and income/wealth concepts available.
2. Other sheets contain the series in gpinter format, with one sheet for each year.

The first sheet may look like this, where "fiinc992j", the WID.world variable code refers here to "fiscal income", "adult population", "equal-split of income or wealth between adults".

Note that the variable code does not contain the prefix, which indicates whether the series is a share, threshold or average, since the document contains several variables at the same. Indeed, the following sheets should be in gpinter format as follows:

	A	B	C	D	E	F	G	H	I	J	K
1	year	country	average	p	thr	topsh	bracketsh	topavg	bracketavg	b	
2	1986	CI	485515,625	0	0	1	0,00018451	485515,625	8958,06298	#DIV/0!	
3				0,01	15173,5427	0,99981549	0,0004124	490329,338	20022,8018	32,3147565	
4				0,02	24544,5019	0,99940309	0,00058935	495128,384	28613,7792	20,17268	
5				0,03	32518,741	0,99881374	0,00074485	499937,813	36163,4066	15,3738367	
6				0,04	39702,8293	0,9980689	0,00085925	504768,796	41717,8526	12,7136732	
7				0,05	43592,7013	0,99720965	0,00093511	509643,017	45400,8689	11,6910171	
8				0,06	47166,7042	0,99627454	0,00100634	514581,763	48859,5145	10,909852	
9				0,07	50533,3271	0,9952682	0,00107601	519589,529	52241,9864	10,282116	
10				0,08	54018,0479	0,99419219	0,00115377	524669,394	56017,1945	9,71285364	
11				0,09	58104,7853	0,99303842	0,0012405	529819,418	60228,2719	9,11834396	
12				0,1	62308,6639	0,99179792	0,00132257	535037,097	64212,9824	8,58688124	
13				0,11	65996,0185	0,99047535	0,00139044	540327,256	67508,1614	8,18727051	
14				0,12	68884,4913	0,98908491	0,00144313	545700,2	70066,0919	7,92196022	
15				0,13	71246,0545	0,98764178	0,001496	551167,259	72633,3464	7,73610921	
16				0,14	74197,3336	0,98614577	0,00156853	556731,607	76154,3721	7,50339102	
17				0,15	78289,2225	0,98457725	0,00166044	562385,457	80616,7479	7,18343392	
18				0,16	82913,8267	0,98291681	0,00174879	568120,798	84906,3361	6,85194281	
19				0,17	86776,3643	0,98116803	0,00182524	573942,659	88618,3018	6,61404363	
20				0,18	90524,3658	0,97934278	0,00190756	579861,249	92615,1044	6,40558201	
21				0,19	94790,0607	0,97743522	0,00199784	585876,634	96998,4721	6,18078129	

series | **CI, 1986** | CI, 1987 | CI, 1988 | CI, 1993 | CI, 1998 | CI, 2002 | CI, 2008 | CI, 2014

This format can be obtained simply by running a generalized Pareto interpolation on the website at http://wid.world/gpinter/. Please export all the variables listed above:

- Year, country, average
- Percentiles list
- Thresholds
- Top shares
- Bracket shares
- Top averages
- Bracket averages
- Beta coefficients

Macroeconomic series format

For macroeconomic series, a very standard format is sufficient, with years in rows and variables in column. If possible, please match your variable names (eg. "National Income") to WID codes (eg. "mnninc999i").

	A	B	C	D	E	F
1	Variable	National income	Net domestic product	Gross domestic product	Consumption of fixed capital	Net foreign income
2	WID code	mnninc999i	mndpro999i	mgdpro999i	mconfc999i	mnnfin999i
72	1970	13	13	15	1	0
73	1971	15	15	17	2	0
74	1972	18	18	20	2	0
75	1973	22	22	25	2	0
76	1974	27	27	30	3	0
77	1975	32	32	36	4	0
78	1976	39	39	44	5	0
79	1977	49	49	55	6	0
80	1978	60	60	68	8	-1
81	1979	70	71	80	10	-1
82	1980	82	82	94	12	-1
83	1981	90	92	106	14	-1
84	1982	104	106	123	17	-2
85	1983	119	121	140	19	-2
86	1984	134	137	159	22	-2
87	1985	150	152	176	24	-2
88	1986	174	176	202	26	-2
89	1987	197	199	227	28	-2
90	1988	220	222	254	32	-2
91	1989	249	251	287	36	-2

Sheet1