Prospective Age Dataset Codebook

Maja Založnik

August 16, 2018

Background

This file describes the dataset `2017_prospective-ages.csv` which contains the calculated prospective old age thresholds for 241 countries and regions, 1950—2100, as well as the proportions of the population (male, female and total) over this threshold.

For background see Sanderson and Scherbov (2008).

The data was used in Založnik (n.d.), which is accessible here.

The methods of analysis used to produce the data are available here.

The dataset is deposited on figshare [here](https://figshare.com/articles/Prospective_Age_Dataset/69744140.

The original data used in the calculations UN (2017)

- UN World Population Prospects [Standard Projections]; accessed 13.08.2018
- UN Life Tables—Mortality Indicators; accessed 13.08.2018

Description of the dataset

Rows

Each row is a location-time-group combination e.g.: *Algeria-1956-female*. There are 241 locations and 151 times, and three groups (male, female and total) giving a total of 109,173 rows.

Columns

1. **location**—name of the region, subregion, country or area [string].

- 241 unique names
- Source: UN (2017)

2. **time**—single year for which the data refer to [numeric].

- range: 1950—2100
- Source: UN (2017)

3. **group**—population to which the data refers to [string]

- either male, female or tota
- Source: UN (2017)

4. **threshold**—old age threshold $RLE^{1}5$ [numeric].

- Data for men and women (but not total) randomly cross-checked with Scherbov, Andruchowitz, and Sanderson (2018).
- Source: own calculation using splines from abridged life-tables.

5. **prop.over.65**—proportion of the population aged 65 or over.

- Data randomly cross-checked with World Bank Data[1]
- Source: own calculation using WPP standard projections.

6. **prop.over.t**—proportion of population older than old-age threshold [**threshold** above].

- For three countries was possible to visually cross-check with Scherbov, Andruchowitz, and Sanderson (2018), in all cases theirs are slightly higher.[2]
- Source: own calculation using splines from abridged life-tables and WPP standard projections.

References

Sanderson, Warren, and Sergei Scherbov. 2008. *Rethinking age and aging.* Population Reference Bureau Washington, DC.

Scherbov, S, S Andruchowitz, and W Sanderson. 2018. "Aging Demographic Data Sheet 2018." International Institute for Applied Systems Analysis.

UN. 2017. *World Population Prospects: The 2017 Revision.* New York: Deartment of Economic; Social Affairs, Population Division.

Založnik, Maja. n.d. "Ageing in the Middle East and North Africa–Measuring Population Ageing Using Prospective Instead of Chronological Age." *Population Horizons.*

[1] The huge majority of the checks were correct to 7 significant digits. There were a few discrepancies however, to the tune of 0.1-0.5 percentage point (e.g Malawi females). These are ascribed to WB perhaps using an earlier version of the data. Given the preponderance of the evidence the calculations performed here are correct, however the results are only as reliable as the original data.

[2] The Scherbov, Andruchowitz, and Sanderson (2018) datasheet is the only validation source known to me. They only calculate the proportions over the threshold age for Korea, Italy and the US and these values are plotted, not reported numerically. Despite this it seems by visual inspection that their values for the proportions are slightly higher than mine (perhaps .5 %). Not sure as to the cause, could again have to do with original data. e.g. I used the `.csv` life table files where the life expectancy is to 2 decimal points precision, while there also exist on the WPP site Excel files with one more digit precision that does not seem to be correctly rounded in the .csv file. Similar things may be happening in the population files, which could lead to errors. Since the Scherbov, Andruchowitz, and Sanderson (2018) public methodology, I cannot repoduce their numbers. Given that we match on the thresholds and I match with the World Bank on the proportions over 65, I feel quite confident that my proportions over the threshold are correct as well. [Have another look at the LEQ issue maybe?]