

# Package ‘wpp2017’

February 10, 2020

**Version** 1.2-3

**Date** 2020-2-7

**Title** World Population Prospects 2017

**Author** United Nations Population Division

**Maintainer** Hana Sevcikova <hanas@uw.edu>

**Depends** R (>= 2.14.2)

**Imports** data.table

**Description** Provides data from the United Nation's World Population Prospects 2017.

**License** GPL (>= 2)

**URL** <http://population.un.org/wpp>

**NeedsCompilation** no

**Repository** CRAN

**Date/Publication** 2020-02-10 06:10:02 UTC

## R topics documented:

|                           |           |
|---------------------------|-----------|
| wpp2017-package . . . . . | 2         |
| e0 . . . . .              | 3         |
| migration . . . . .       | 5         |
| mx . . . . .              | 6         |
| percentASFR . . . . .     | 7         |
| pop . . . . .             | 7         |
| sexRatio . . . . .        | 9         |
| tfr . . . . .             | 10        |
| UNlocations . . . . .     | 11        |
| <b>Index</b>              | <b>13</b> |

## Description

Data from the United Nations World Population Prospects 2017, released on June 21, 2017.

## Details

Package: wpp2017  
Version: 1.2-3  
Date: 2020-2-7  
URL: <http://population.un.org/wpp>

The package contains the following datasets:

- [tfr](#), [tfr\\_supplemental](#), [tfrprojMed](#), [tfrproj80u](#), [tfrproj80l](#), [tfrproj95u](#), [tfrproj95l](#), [tfrprojHigh](#), [tfrprojLow](#): estimates and projections of total fertility rate, including the projected 80% and 95% probability bounds, as well as low and high half child variants.
- [e0F](#), [e0M](#), [e0X\\_supplemental](#), [e0Xproj](#), [e0Xproj80u](#), [e0Xproj80l](#), [e0Xproj95u](#), [e0Xproj95l](#): sex-specific estimates and projections of life expectancy with X="F" and "M", including the projected 80% and 95% probability bounds.
- [pop](#), [popproj](#), [popproj80u](#), [popproj80l](#), [popproj95u](#), [popproj95l](#), [popprojHigh](#), [popprojLow](#): historical estimates of total population counts, as well as the median, probability bounds and the high and low variants of population projections.
- [popFT](#), [popMT](#), [popFTproj](#), [popMTproj](#): historical estimates and projection medians for sex-specific total population.
- [popF](#), [popM](#), [popXprojMed](#), [popXprojHigh](#), [popXprojLow](#): age- and sex-specific population estimates and projections with X="F" and "M", including the high and low variants.
- [mxF](#), [mxM](#): age- and sex-specific mortality rates
- [migration](#): total net migration
- [sexRatio](#): sex ratio at birth as a ratio of female to male
- [percentASFR](#): distribution of age-specific fertility rates
- [UNlocations](#): location dataset

## Note

The package **wppExplorer** offers a shiny user interface to explore these datasets as well as functions for convenient extraction of information from the data, see function `wpp.indicator` in **wppExplorer**, or <https://bayespop.shinyapps.io/wpp2017explorer/>.

**Source**

These datasets are based on estimates and projections of United Nations, Department of Economic and Social Affairs, Population Division (2017).

**References**

World Population Prospects: The 2017 Revision. <http://population.un.org/wpp>.

---

e0

*United Nations Time Series of Life Expectancy*

---

**Description**

Datasets containing the United Nations time series of the life expectancy (e0) for all countries of the world as available in 2017.

**Usage**

data(e0F)

data(e0M)

data(e0F\_supplemental)

data(e0M\_supplemental)

data(e0Fproj)

data(e0Mproj)

data(e0Fproj801)

data(e0Fproj80u)

data(e0Mproj801)

data(e0Mproj80u)

data(e0Fproj951)

data(e0Fproj95u)

data(e0Mproj951)

data(e0Mproj95u)

**Format**

The datasets contain one record per country or region. They contain the following variables:

country\_code Numerical Location Code (3-digit codes following ISO 3166-1 numeric standard)

- see [http://en.wikipedia.org/wiki/ISO\\_3166-1\\_numeric](http://en.wikipedia.org/wiki/ISO_3166-1_numeric).

name Name of country or region (following ISO 3166 official short names in English - see <https://www.iso.org/obp/ui/#search/code/> and United Nations Multilingual Terminology Database

- see <https://unterm.un.org/unterm>).

1950-1955, 1955-1960, ... Life expectancy in various five-year time intervals (i.e., from 1 July in year  $t$  to 1 July in year  $t+5$  such as the period 1950-1955 refers to the period 1950.5-1955.5 and the mid of the period is 1953.0). The `e0*proj` datasets start at 2015-2020. The `e0*_supplemental` datasets start at 1750-1755. Missing data have NA values.

## Details

Datasets `e0F` and `e0F_supplemental` contain estimates for female historical `e0`; `e0M` and `e0M_supplemental` contain estimates for male historical `e0`. The `*_supplemental` datasets contain a subset of countries for which data prior 1950 are available. Datasets `e0Mproj` and `e0Fproj` contain projections of male and female `e0`, respectively. Datasets `*80l`, `*95l` are the lower bounds of 80 and 95% probability intervals, `*80u`, `*95u` are the corresponding upper bounds.

The historical dataset (`e0F_supplemental.txt` and `e0M_supplemental.txt` for female and male respectively) for 29 countries or areas covers the period 1750-1950 (including 20 countries with data since at least 1900) is based on a series for 5-year periods from the following sources: (1) University of California at Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). (2012). Human Mortality Database Available at <https://www.mortality.org>. Data downloaded on 9 Jan. 2012; (2) University of California at Berkeley (USA), Max Planck Institute for Demographic Research (Germany), and Institut National d'Etudes Demographiques (France). Human Life-Table Database (2011). Available at <https://www.lifetable.de>. Data downloaded on 29 Dec. 2011; (3) Statistics Finland (2006). Statistical Yearbook of Finland 2006; (4) Hungarian Central Statistical Office (2006). Hungary Demographic Yearbook 2005; (5) Japan Ministry of Internal Affairs and Communication (2012). Historical Statistics of Japan (<https://www.e-stat.go.jp/en>); (6) Andreev E.M. et al. (1998). Demographic History of Russia 1927-1959. Informatika, Moscow.

## Source

These datasets are based on estimates and projections of United Nations, Department of Economic and Social Affairs, Population Division (2017).

## References

World Population Prospects: The 2017 Revision. <http://population.un.org/wpp>

## Examples

```
data(e0M)
head(e0M)

data(e0Fproj)
str(e0Fproj)
```

---

|           |                             |
|-----------|-----------------------------|
| migration | <i>Dataset on Migration</i> |
|-----------|-----------------------------|

---

### Description

Estimates and projections of total net migration.

### Usage

```
data(migration)
```

### Format

Data frame with one row per country. It contains the following variables:

country\_code Numerical Location Code (3-digit codes following ISO 3166-1 numeric standard)  
- see [http://en.wikipedia.org/wiki/ISO\\_3166-1\\_numeric](http://en.wikipedia.org/wiki/ISO_3166-1_numeric).

name Country name.

1950-1955, 1955-1960, ... Net migration (in thousand) for the specific five-year time period (i.e., from 1 July in year t to 1 July in year t+5 such as the period 1950-1955 refers to the period 1950.5-1955.5 and the mid of the period is 1953.0).

### Source

These datasets are based on estimates and projections of United Nations, Department of Economic and Social Affairs, Population Division (2017).

### References

World Population Prospects: The 2017 Revision. <http://population.un.org/wpp>.

### Examples

```
data(migration)  
str(migration)
```

---

`mx`*Age-specific Mortality Data*

---

**Description**

Age-specific data on mortality rates for male (mxM) and female (mxF).

**Usage**

```
data(mxM)
data(mxF)
```

**Format**

Data frames with one row per country and age group. They contain the following variables:

`country_code` Numerical Location Code (3-digit codes following ISO 3166-1 numeric standard)  
- see [http://en.wikipedia.org/wiki/ISO\\_3166-1\\_numeric](http://en.wikipedia.org/wiki/ISO_3166-1_numeric).

`name` Country name.

`age` A character string representing an age interval (given by the starting age of the interval).

1950-1955, 1955-1960, ... `mx` for the given five-year time period (i.e., from 1 July in year  $t$  to 1 July in year  $t+5$  such as the period 1950-1955 refers to the period 1950.5-1955.5 and the mid of the period is 1953.0). Not available data are represented by an empty string.

**Note**

Data for ages 85-100+ are not the official UN data. While the published UN mortality datasets contain data only up to 85+, data for ages 85-100+ in this dataset were derived from UN published life table quantities.

**Source**

These datasets are based on estimates and projections of United Nations, Department of Economic and Social Affairs, Population Division (2017).

**References**

World Population Prospects: The 2017 Revision. <http://population.un.org/wpp>

**Examples**

```
data(mxF)
head(mxF)
```

---

percentASFR

*Datasets on Age-specific Distribution of Fertility Rates*

---

### **Description**

Datasets giving the percentage of fertility rates over ages 15-50.

### **Usage**

```
data(percentASFR)
```

### **Format**

A data frame with one row per country and age group. For each country there are seven age groups. It contains columns `country_code`, `name`, `age` and one column per five-year time interval (i.e., from 1 July in year  $t$  to 1 July in year  $t+5$  such as the period 1950-1955 refers to the period 1950.5-1955.5 and the mid of the period is 1953.0).

### **Source**

This dataset is based on estimates and projections of United Nations, Department of Economic and Social Affairs, Population Division (2017).

### **References**

World Population Prospects: The 2017 Revision. <http://population.un.org/wpp>.

### **Examples**

```
data(percentASFR)
str(percentASFR)
```

---

pop

*Estimates and Projections of Population Counts*

---

### **Description**

Datasets with historical population estimates and projections.

**Usage**

```

data(pop)
data(popMT)
data(popFT)
data(popM)
data(popF)

data(popproj)
data(popproj80l)
data(popproj80u)
data(popproj95l)
data(popproj95u)
data(popprojHigh)
data(popprojLow)

data(popMTproj)
data(popFTproj)

data(popMprojMed)
data(popFprojMed)
data(popMprojHigh)
data(popFprojHigh)
data(popMprojLow)
data(popFprojLow)

```

**Format**

Datasets that start with popM or popF and do not have “T” in their names, are age-specific and are organized as data frames with one row per country and age group. For each country there are 21 age groups. It contains the following variables:

country\_code Numerical Location Code (3-digit codes following ISO 3166-1 numeric standard)  
- see [http://en.wikipedia.org/wiki/ISO\\_3166-1\\_numeric](http://en.wikipedia.org/wiki/ISO_3166-1_numeric).

name Country name.

age A character string representing an age interval. For each country there are 21 values: “0-4”, “5-9”, “10-14”, “15-19”, “20-24”, “25-29”, “30-34”, “35-39”, “40-44”, “45-49”, “50-54”, “55-59”, “60-64”, “65-69”, “70-74”, “75-79”, “80-84”, “85-89”, “90-94”, “95-99”, and “100+” in that order. If there are NAs for ages 85 and older, the “80-84” category is to be treated as “80+”.

1950, 1955, ... Population estimate or projection (in thousand) for the given time (mid-year).

The remaining datasets, i.e. those that do not have “M” or “F”, or have “T” in their names, contain one row per country.

**Details**

Dataset pop provides estimates of historical total population counts.



Datasets popMT and popFT provide estimates of total counts of male and female population, respectively.

Datasets popM (popF) contain age-specific estimates of the historical population counts for male (female).

Dataset popproj provides median projection of total population counts, i.e. aggregated over sex and age. Datasets popproj80l, popproj80u, popproj95l, and popproj95u are the lower (l) and upper (u) bounds of the 80 and 95% probability intervals of the total population. Datasets popprojHigh and popprojLow contain the upper and lower variant of total population defined as  $\pm 1/2$  child.

Datasets popMTproj and popFTproj provide median projection of total counts of male and female population, respectively.

Datasets popXprojMed, popXprojHigh and popXprojLow contain median, high and low variants of age-specific projections, respectively, with X=M for male and X=F for female.

All values are in thousands.

### Source

These datasets are based on estimates and projections of United Nations, Department of Economic and Social Affairs, Population Division (2017).

### References

World Population Prospects: The 2017 Revision. <http://population.un.org/wpp>.

### Examples

```
data(popM)
str(popM)
```

---

|          |                           |
|----------|---------------------------|
| sexRatio | <i>Sex Ratio at Birth</i> |
|----------|---------------------------|

---

### Description

Estimates and projections of the sex ratio at birth derived as the number of female divided by the number of male.

### Usage

```
data(sexRatio)
```

### Format

A data frame with one record per country. It contains columns country\_code, name, and one column per five-year time interval (i.e., from 1 July in year t to 1 July in year t+5 such as the period 1950-1955 refers to the period 1950.5-1955.5 and the mid of the period is 1953.0).

**Source**

This dataset is based on estimates and projections of United Nations, Department of Economic and Social Affairs, Population Division (2017).

**References**

World Population Prospects: The 2017 Revision. <http://population.un.org/wpp>.

**Examples**

```
data(sexRatio)
str(sexRatio)
```

---

|     |   |
|-----|---|
| tfr | <i>United Nations Time Series of Total Fertility Rate</i> |
|-----|---|

---

**Description**

Datasets containing the United Nations time series of the total fertility rate (TFR) for all countries of the world as available in 2017.

**Usage**

```
data(tfr)
data(tfr_supplemental)

data(tfrprojMed)
data(tfrproj80l)
data(tfrproj80u)

data(tfrproj95l)
data(tfrproj95u)

data(tfrprojHigh)
data(tfrprojLow)
```

**Format**

The datasets contain one record per country or region. It contains the following variables:

country\_code Numerical Location Code (3-digit codes following ISO 3166-1 numeric standard)  
- see [http://en.wikipedia.org/wiki/ISO\\_3166-1\\_numeric](http://en.wikipedia.org/wiki/ISO_3166-1_numeric).

name Name of country or region (following ISO 3166 official short names in English - see <https://www.iso.org/obp/ui/#search/code/> and United Nations Multilingual Terminology Database - see <https://unterm.un.org/unterm>).

1950-1955, 1955-1960, ... TFR in various five-year time intervals (i.e., from 1 July in year t to 1 July in year t+5 such as the period 1950-1955 refers to the period 1950.5-1955.5 and the mid of the period is 1953.0). The tfrproj\* datasets start at 2015-2020. The tfr\_supplemental datasets start at 1740-1745. Missing data have NA values.

## Details

Dataset `tfr` contains estimates of the historical TFR starting with 1950; `tfr_supplemental` contains a subset of countries for which data prior 1950 are available. Datasets `tfrprojMed` contain the median projections. Datasets `tfrproj80l`, `tfrproj80u`, `tfrproj95l`, and `tfrproj95u` are the lower (l) and upper (u) bounds of the 80 and 95% probability intervals, respectively. Datasets `tfrprojHigh` and `tfrprojLow` contain high and low variants, respectively, defined as  $\pm 1/2$  child.

The historical dataset is contained in (`tfr_supplemental.txt`) for 103 countries or areas covering the period 1740-1950 (including 24 countries with data before 1850), and is based on series for five-year periods from the following sources: (1) Max Planck Institute for Demographic Research (Germany) and Vienna Institute of Demography (Austria). (2012). Human Fertility Database (HFD). Available at <https://www.humanfertility.org>. Data downloaded on 13 May 2012; (2) Festy, P. (1979). *La fecondite des pays occidentaux de 1870 a 1970*. Paris: Presses universitaires de France; (3) Chesnais, J.C. (1992). *The demographic transition: stages, patterns, and economic implications: a longitudinal study of sixty-seven countries covering the period 1720-1984*. Oxford ; New York: Clarendon Press; (4) Bhat, P.N.M. (1989). "Mortality and fertility in India, 1881-1961: a reassessment." pp. 73-118 in *India's historical demography: studies in famine, disease and society*, edited by T. Dyson. London and Riverdale, Md: Curzon and Riverdale Co.; (5) Hofsten, E.A.G.v. and H. Lundstrom. (1976). *Swedish population history: Main trends from 1750 to 1970*. Stockholm: Statistiska centralbyran: LiberForlag; (6) Ajus, F. and M. Lindgren. (2012). *Gapminder fertility dataset, 2010 (including documentation for Children per Woman (Total Fertility Rate) for countries and territories, Version 2*. The Gapminder Foundation. Sweden, Stockholm. <http://www.gapminder.org/data/documentation/gd008/>. Data downloaded on 8 April 2012.

## Source

These datasets are based on estimates and projections of United Nations, Department of Economic and Social Affairs, Population Division (2017).

## References

World Population Prospects: The 2017 Revision. <http://population.un.org/wpp>.

## Examples

```
data(tfr)
head(tfr)

data(tfrprojMed)
str(tfrprojMed)
```

---

UNlocations

*United Nations Table of Locations*

---

## Description

United Nations table of locations, including regions, for statistical purposes as available in 2017.

**Usage**

```
data(UNlocations)
```

**Format**

A data frame with one observations per country or region. It contains the following variables:

name Name of country or region (following ISO 3166 official short names in English - see <https://www.iso.org/obp/ui/#search/code/> and United Nations Multilingual Terminology Database - see <https://unterm.un.org/unterm>).

country\_code Numerical Location Code (3-digit codes following ISO 3166-1 numeric standard) - see [http://en.wikipedia.org/wiki/ISO\\_3166-1\\_numeric](http://en.wikipedia.org/wiki/ISO_3166-1_numeric).

reg\_code Code of the regions.

reg\_name Name of the regions.

area\_code Area code.

area\_name Area names, such as Africa, Asia, Europe Latin America and the Caribbean, Northern America, Oceania, World.

location\_type Code giving the type of the observation: 0=World, 2=Major Area, 3=Region, 4=Country/Area, 5=Development group, 12=Special groupings. Other numbers are allowed and they can be used for aggregation, see below.

agcode\_1500, agcode\_1501, agcode\_1502, agcode\_1503, agcode\_1517, agcode\_901, agcode\_902, agcode\_921, agcode\_...  
Optional columns that can be used for aggregations. To aggregate a region with country\_code= $x$ , get the value of its location\_type, say  $y$ . Then look for the column agcode\_ $y$  and locate all records with agcode\_ $y$ = $x$  that have location\_type=4, see Example below.

**Source**

Data provided by the United Nations Population Division.

The designations employed in this dataset do not imply the expression of any opinion whatsoever on the part of the Secretariat of the United Nations concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

**Examples**

```
data(UNlocations)
# Find high income countries in Africa (based on World Bank groups)
grouprec <- subset(UNlocations, name == "High-income countries")
# grouprec$location_type is 1503, thus look for column agcode_1503
subset(UNlocations, agcode_1503 == grouprec$country_code &
      location_type == 4 & area_name == "Africa")
```

# Index

## \*Topic **datasets**

- e0, 3
- migration, 5
- mx, 6
- percentASFR, 7
- pop, 7
- sexRatio, 9
- tfr, 10
- UNlocations, 11

## \*Topic **package**

- wpp2017-package, 2

- e0, 3
- e0\_supplemental (e0), 3
- e0F, 2
- e0F (e0), 3
- e0F\_supplemental (e0), 3
- e0Fproj (e0), 3
- e0Fproj801 (e0), 3
- e0Fproj80u (e0), 3
- e0Fproj951 (e0), 3
- e0Fproj95u (e0), 3
- e0M, 2
- e0M (e0), 3
- e0M\_supplemental (e0), 3
- e0Mproj (e0), 3
- e0Mproj801 (e0), 3
- e0Mproj80u (e0), 3
- e0Mproj951 (e0), 3
- e0Mproj95u (e0), 3
- e0X\_supplemental, 2
- e0Xproj, 2
- e0Xproj801, 2
- e0Xproj80u, 2
- e0Xproj951, 2
- e0Xproj95u, 2
- migration, 2, 5
- mx, 6
- mxF, 2

- mxM (mx), 6
- mxM, 2
- mxM (mx), 6
- percentASFR, 2, 7
- pop, 2, 7
- popF, 2
- popF (pop), 7
- popFprojHigh (pop), 7
- popFprojLow (pop), 7
- popFprojMed (pop), 7
- popFT, 2
- popFT (pop), 7
- popFTproj, 2
- popFTproj (pop), 7
- popM, 2
- popM (pop), 7
- popMprojHigh (pop), 7
- popMprojLow (pop), 7
- popMprojMed (pop), 7
- popMT, 2
- popMT (pop), 7
- popMTproj, 2
- popMTproj (pop), 7
- popproj, 2
- popproj (pop), 7
- popproj801, 2
- popproj801 (pop), 7
- popproj80u, 2
- popproj80u (pop), 7
- popproj951, 2
- popproj951 (pop), 7
- popproj95u, 2
- popproj95u (pop), 7
- popprojHigh, 2
- popprojHigh (pop), 7
- popprojLow, 2
- popprojLow (pop), 7
- popXprojHigh, 2
- popXprojLow, 2

popXprojMed, 2

sexRatio, 2, 9

tfr, 2, 10

tfr\_supplemental, 2

tfr\_supplemental (tfr), 10

tfrproj80l, 2

tfrproj80l (tfr), 10

tfrproj80u, 2

tfrproj80u (tfr), 10

tfrproj95l, 2

tfrproj95l (tfr), 10

tfrproj95u, 2

tfrproj95u (tfr), 10

tfrprojHigh, 2

tfrprojHigh (tfr), 10

tfrprojLow, 2

tfrprojLow (tfr), 10

tfrprojMed, 2

tfrprojMed (tfr), 10

UNlocations, 2, 11

wpp2017 (wpp2017-package), 2

wpp2017-package, 2