

Documentation for **Life Expectancy at birth (years)** for countries and territories

Gapminder Documentation constitutes work in stepwise progress.
We welcome all sorts of comments, corrections and suggestions through e-mail to the author.

Gapminder Documentation 004

Version 2

Uploaded: 2008-12-22

Published by: The Gapminder Foundation,
Sweden, Stockholm, 2008

Author: Klara Johansson

E-mail: [klara.johansson \(at\) Gapminder.org](mailto:klara.johansson@gapminder.org)

Gapminder is a non-profit foundation
promoting sustainable global development
by increased use and understanding of statistics.
www.gapminder.org



1. Introduction

This is the documentation for the Gapminder compilation of global estimates for average life expectancy at birth.

The Gapminder dataset for life expectancy at birth stretches over 200 years and contains

- a) estimates of average life expectancy from different sources (see section 3)
- b) where no estimates are available, a constructed, simple model for life expectancy (see section 4)

For some countries, there are no reliable estimates before 1950-55, while for others, there is relatively reliable data all the way back into the 19th century. This means that the quality of data varies greatly between countries and years.

The ambition has been to include as many countries and territories as possible and the latest and most updated estimates known to us. A guiding principle has been to use the most transparent and well-documented source whenever possible, however, preference was also given to sources with as wide coverage of countries and territories as possible.

The main purpose of the data is to produce graphical presentations that display the magnitude of difference in life expectancy at birth in the world over time, for the interactive moving graph www.gapminder.org/world. We have not been able to make sure that every single observation is based on the best estimates available. We discourage the use of this dataset for statistical analysis and advice those who require more exact data to investigate the available data more carefully and look for additional sources, when appropriate.

2. Definition

Life expectancy at births is defined as the average number of years a newborn child would live if current mortality patterns were to stay the same (also written as e_0).

3. Estimates from different sources: Data sources used

Below is a short description of each of the datasets used to compile the estimates of average life expectancy at birth for the Gapminder World graph. As a first priority, data from Human Mortality Database (HMD) was used where available. For countries and/or time periods where the HMD did not have data, World Population Prospect was used, if available. The estimates from HMD and WPP are very close. The reason for using HMD in the first hand is that 1) there is data for single years, instead of every 5 years and 2) there is a good, online documentation of sources and methodology, available at www.mortality.org (see individual countries). Where HMD data missing for 2006 and 2007, WPP estimate for 2005-2010 was added (2007).

Where no estimates from either of the above were available, we used estimates from other sources, in the order listed below.

a) Human Mortality database

The Human Mortality Database is a joint project by Department of Demography at the University of California, Berkeley, USA, and at the Max Planck Institute for Demographic Research in Rostock, Germany. The project is described in more detail at www.mortality.org.

The Human Mortality Database only contains data from countries and periods where death registration and census are virtually complete, meaning that the quality of this data is generally very good.

The data we used from the Human Mortality Database is easiest found in the *Life tables 1x1* (1 year of age – 1 calendar year time intervals) for each country, under the heading $e(X)$ which means the expected remaining years to live at exact age X . Thus, data under the column $e(X)$ and the row Age:0 means life expectancy at birth.

The data is available online at www.mortality.org. It is free of charge, but it is necessary to create a username to access it. A detailed description of methodology is also available at this site, as well as the raw data.

b) World Population Prospects: The 2006 Revision

This data is published by the United Nations Population Division and can be found through the "UN data" search engine. <http://data.un.org> or more exactly at <http://data.un.org/Search.aspx?q=life+expectancy+datamart%5bPopDiv%5d>

This source has data for every five years only. In the graph, the estimate has been used for the middle year (e.g. estimates for 1950-55 is used for 1952). The graph automatically interpolates the intervening years.

For estimates of 2005-2010, medium variant projection was used.

c) Publications and files by James C Riley

Historian James Riley has done an extensive review of data and literature on life expectancy. Going through more than 700 different sources in different languages, he also assessed which estimates are the most reliable.

We have had the privilege of getting access to Professor Riley's spreadsheets containing these estimates. These spreadsheets can be accessed online at [URL- address to be added soon].

The spreadsheets contain estimates from in total over 700 different sources, compiled by Riley, and described in <http://www.lifetable.de/RileyBib.htm>. For most countries, one column indicates which estimates are "preferred", i.e. have been estimated to be most reliable. In some cases where there were only estimates which were not preferred, we have used these estimates in the absence of

better estimates. In some few cases, the preferred estimates in the files were probably biased upwards (as noted by Riley 2005b), and were excluded (Australia, Uruguay).

For publications by Riley, see reference list at the end of this document.

d) The Human Life-Table Database

The Human Life-Table Database, HLD, is a database over life-tables from different sources. It is a companion project of the Human Mortality Database, and a joint project of the [Max Planck Institute for Demographic Research](#) (MPIDR) in Rostock, Germany, the [Department of Demography at the University of California at Berkeley](#), USA and the [Institut national d'études démographiques](#) (INED) in Paris, France. The MPIDR is responsible for maintaining the database.

Often, data is only available separately for males and females. Estimates for total population have been calculated by Gapminder, assuming a sex ratio of 0.512 males for 0.488 females.

Most of the HLD life tables are life tables for national populations, officially published by national statistical offices and added to the database. All data is available for free. The quality of data can be assumed to be lower than in the Human Mortality Database (see above).

a) Data from national statistical agencies

Some few estimates have also been obtained through the statistical agencies of national governments, such as Statistics and research Åland, homepage <http://www.asub.ax/start.con?iLan=2>

For Taiwan in 2006, we have received a data file from the Directorate General of Budget, Accounting & Statistics (DGBAS) of Taiwan.

4. Gapminder model for life expectancy at birth

In order to show a general outline of the world's development since 1800, Gapminder has constructed a very simple, rough model to show life expectancy for a number countries and periods where we have no estimates. This model to a large extent relies on the online bibliography and published research of Professor James C. Riley (see reference list). Any problems or faults with the model are, however, entirely the responsibility of Gapminder.

PLEASE NOTE THAT: The Gapminder Model for historical life expectancy is extremely rough, and should not be used for research or similar purposes, nor taken as a statement of fact.

The model will be updated and improved continually. Please feel free to send us your comments and suggestions.

The model is based on two basic, very simple assumptions.

Basic assumptions:

1. That all countries start from historical low levels of average life expectancy at birth, with periodic variations.
2. That all countries go through a "health transition" of sustained gains in health, implying lowered crude death rates, improved child survival and increased average life expectancy.
 - exception: improvement in health can be halted, stagnated or reversed due to circumstances such as war, famine, pestilence etc.

CAVEAT: these basic assumptions present a very simplified description, for the purpose of constructing this model. But it is true that all countries seem to have gone through (and many are still going through) sustained periods of decreasing mortality and improving health during the past 200 years.

The Gapminder Model for life expectancy consists of

1. Modelling a crude baseline of life expectancy prior to the "health transition"
2. Modelling the improvements in health during the health transition

CAVEAT:

1. A modeled baseline can never show the historical reality of health and mortality, which can be assumed to have varied greatly on a year-to-year basis due to vulnerability to disease, war and economic circumstances etc.

2. The "health transition" was probably seldom as smooth and straightforward as presented in this model. In addition, the periodic variations would have been big enough to confuse the picture.

How the baseline is estimated:

1. In the case where there are estimates available before the estimated onset of sustained gains in average life expectancy, the earliest estimate is used as baseline.
 - a) Except when these estimates are, for reasons not explained, far outside the reasonable range for average pre-transition life expectancy (which Gapminder for the purpose of this model has set to 25-40 years)
2. In the case where there are no pre-transition estimates, or these estimates are outside the reasonable range, continent averages have been used, from Riley (2005b, table 1).
 - a) For the purpose of the graphic presentation, the countries have been spread tight around the continent averages, based on their relative position within their continent in 1950-55 (the first UN estimate).¹ This was done to avoid the situation where they are placed on top of each other, or on a line (depending on the variable on the other axis).
 - b) For Fiji, Papua New Guinea and Solomon Islands, where the available continent average (from Riley 200a) was lower than 25, the baseline was set to 25.
3. Based on Riley (2005b) 3 countries were assumed to have already begun sustained gains in life expectancy by 1800: Denmark, France, Sweden. For Sweden, there is civil registration data back to the 1750s. For France and Denmark, an interpolation has been made between the estimated level at the beginning of health transition (before 1800) and the earliest available estimate.
4. United Kingdom (for which the earliest estimates only apply to England & Wales) is assumed to have begun sustained gains in health by the earliest years of the 19th century.

Exceptions:

In some cases, the average life expectancy “baseline” can be assumed to have improved (or deteriorated) but without continued gains, for example due to the abolishment of slavery, demographic changes or similar. In a few special cases, we have chosen to illustrate this. More information about special cases can be found in the spreadsheet with detailed metadata (URL to be added soon).

CAVEAT: these baseline estimations should not be taken as a statement regarding the life expectancy of the individual countries during historic periods, nor regarding the relative position of countries. Most likely, the actual life expectancy varied strongly year-to-year due

¹ This was done very roughly, and will hopefully be improved soon. Each country on the continent was given a ranking, based on their relative positions in 1950-55. This was done whether or not the continent average was used for this country. The median country was set at the continent average, and then countries were evenly spread around the continent average, with difference of 0.1 years between each, irrespective of the magnitude of relative differences in 1950-55. Even though all countries on the continent were included in calculating the model, the estimates obtained were only used for countries lacking an estimate that could be used as a baseline.

to circumstances, and these rough baselines can be assumed to be within the reasonable range of such variations for the country in question.

How the health transition is illustrated

1. Riley (2005b) suggests, for each country, a period when sustained gains in health can be assumed to have begun, and also gives the exact year which he has used for his research. This year is used for each country as the end year of the baseline. Between the end of the baseline and the first estimate, if later, the graph will interpolate (linear).
 - a. Exception: for countries with an estimated onset of health transition in 1950, but which have estimates for 1950-55 (UN) that are higher than the continent average, the health transition is assumed to have begun from the continent average in 1945. (Example: most countries in Africa south of Sahara)
 - b. Exception: for countries where there is an estimate higher than 40 for the suggested year or earlier, in which case a lower baseline is set, and the onset of health transition assumed to be earlier but still within the time period suggested by Riley 2005b, appendix 1. (Example: Thailand)
2. For a number of countries, we have estimates going back to before the estimated onset of the health transition. If there are 10 years or less between the estimates, nothing further has been done to illustrate the health transition: it is illustrated by the existing observations. For countries where there is more than 10 years between the latest pre-transition observation and the estimated time of health transition, we have used the latest pre-transition estimate to set an end year of the baseline (the year given by Riley 2005b, appendix 1) and mark the onset of the health transition.

CAVEAT: this will show a very simplified version of the health transition, which was probably seldom as smooth and straightforward as suggested by the graph. The timing of the health transition is sometimes placed within a period of a few decades, where we have chosen either the middle year or the year assumed to be most likely (Riley 2005b). For more information, consult Riley (2005b) and Riley's online bibliography-

Illustrating disasters and temporary setbacks

Where we have the possibility, we make an effort to illustrate famines, wars and other disasters and setbacks for health. More information about special cases can be found in the spreadsheet with detailed metadata (URL to be added soon).

Example of a special case:

Philippines: we have estimates from Riley's file for the Philippines (not preferred) for a famine in 1902 and also in 1918 (which is also the year of the Spanish flu, which increased adult mortality worldwide). Since these estimates are considerably lower than the baseline, we have chosen to illustrate these disasters by using the estimates, but reset the baseline for the year immediately before and after the disaster – with the effect in the graph of a straight baseline with two strong falls in life expectancy in 1902 and 1918. According to Riley's online bibliography, the improvements in health slowed down during the period 1942-45, but there are no estimates to show this. We have chosen to set the life expectancy for the whole

period of 1942-45 at a steady 42 years (an interpolation between 40 in 1938 and 47.8 in 1950-55).

5. Links to the Gapminder dataset and other links

The Gapminder dataset for life expectancy at birth can be accessed at <http://spreadsheets.google.com/pub?key=phAwcNAVuyj2tPLxKvvnNPA>

A detailed dataset with metadata for each observation can be accessed at:

[URL to be added shortly]

The graph for which the data is used can be accessed through:

www.gapminder.org/world

Please learn more about Gapminder at www.gapminder.org

6. Updates

This dataset and the documentation are being updated continually.

References (incomplete, will be updated soon)

Bell, Felicitie C. & Miller, Michael L. Life Tables for the United States Social Security Area 1900-2100 Actuarial Study No. 116 http://www.lifetable.de/data/MPIDR/USA_1901-1999.pdf

Riley, James C (no date) Bibliography of Works Providing Estimates of Life Expectancy at Birth and Estimates of the Beginning Period of Health Transitions in Countries with a Population in 2000 of at Least 400,000. Compiled by James C. Riley. Available online at <http://www.lifetable.de/RileyBib.htm> [2008-12-10]

Riley, James C (2005a) Estimates of Regional and Global Life Expectancy, 1800–2001. Population and development review 31(3): 537–543 (September 2005)

Riley, James C (2005b) The Timing and Pace of Health Transitions around the World. Population and development review 31(4): 741–764 (December 2005) <http://www.blackwell-synergy.com/doi/abs/10.1111/j.1728-4457.2005.00096.x>

Riley, James C (files on countries) Can be accessed at: [URL to be added soon]

Statistics and research Åland, homepage <http://www.asub.ax/start.con?iLan=2> , life expectancy file: <http://www.asub.ax/files/BE13en.xls>

The Human Mortality Database <http://www.mortality.org/> The Department of Demography at the University of California, Berkeley, USA, and at the Max Planck Institute for Demographic Research in Rostock

The Human Life-Table Database www.lifetable.de The Max Planck Institute for Demographic Research (MPIDR) in Rostock, Germany, the Department of Demography at the University of California at Berkeley, USA and the Institut national d'études démographiques (INED) in Paris, France.

The Human Life-Table Database: Methodology note <http://www.lifetable.de/methodology.pdf>

United Nations Population Division. World Population Prospects: The 2006 Revision. Life expectancy at birth, both sexes. Can be accessed at: <http://data.un.org/Search.aspx?q=life+expectancy+datamart%5bPopDiv%5d>