

# Reading Human Fertility Database and Human Mortality Database data into R

Tim Riffe

Department of Demography, University of California, Berkeley

May 22, 2015

## Abstract

This is about

## 1 Motivation

The Human Fertility Database and Human Mortality Database are two widely used data sources for the comparative and historical study of fertility and mortality. Both databases offer the option to download data in bulk in a few different formats. After a bulk download, users can then set up local databases in a variety of convenient ways, such as that described by Minton (2015) for R (R Development Core Team 2012) users. Another option for R users is to read data directly into an interactive session from the respective database websites. The `HMDHFDplus` package provides easy direct access to the databases using a simple standard set of arguments. Issues such as authentication and fixing column classes are handled automatically. This report outlines the basic features and provides usage examples.

## 2 Installation

The `HMDHFDplus` package is hosted by and can be installed directly from `github`.<sup>1</sup> Two external dependencies help R handle `html` parsing and database authentication, and these must be installed first in order for `HMDHFDplus` to properly install. These two dependencies are `cURL` and `XML`, and their installation unfortunately depends on one's operating system at this time.

In Linux and similar systems, open the Terminal and run:

---

<sup>1</sup>The package is under the following url: <http://github.com/timriffe/TR1/tree/master/TR1/HMDHFDplus>. This report, as well as a useful `README` document, can be found under <http://github.com/timriffe/TR1>.

```
sudo apt-get install libcurl
sudo apt-get install libxml2-dev
```

Then install HMDHFDplus using:

```
library(devtools)
install_github("timriffe/TR1/TR1/HMDHFDplus")
```

This is the simplest case. For Windows, Mac, or other special cases, some further steps are required. Please consult the `README`<sup>2</sup> file on the `github` repository for more details, and contact the maintainer if necessary.

### 3 Usage and examples

Load the package using:

```
library(HMDHFDplus)
```

The two main functions of interest are `readHMDweb()` and `readHFDweb()`, and both functions have the same essential arguments. These two functions only require the user to supply country codes, date item names, and database authentication parameters. The HFD also allows users to extract data from former updates, which may be useful for strict replication purposes. In order to make use of this feature, users must be aware of the 8-digit date code associated with the specific country series update. By default, `textttreadHMDweb()` extracts the most recent update, and we do not demonstrate the archive feature. It helps to be a bit familiar with HMD and HFD file naming conventions. To get the population codes used in either database, run

```
getHFDcountries()
getHMDcountries()
```

To see the data products available for a given country, run

The HMD and HFD both require user registration, and users will have usernames and passwords for these databases.

### 4 Acknowledgements

Thanks to Joshua Goldstein and Carl Boe for supporting development of this R functionality, and to Vladimir Shkolnikov, Dmitri Jdanov, and Tomáš Sobotka for the invitation to present this material at the HFD side meeting to the 2015 PAA Annual Meeting. This work was supported by the National Institute On Aging of the U.S. National Institutes of Health (NIH) under Award Numbers R01-AG011552 and R01-AG040245. The content is solely the responsibility of the author and does not necessarily represent the official views of the NIH.

---

<sup>2</sup>The `README` is at the foot of the main repository page: <https://github.com/timriffe/TR1>

## References

- Human Fertility Database. Max Planck Institute for Demographic Research (Germany) and Vienna Institute of Demography (Austria). Available at [www.humanfertility.org](http://www.humanfertility.org).
- Human Mortality Database. University of California, Berkeley (USA) and Max Planck Institute for Demographic Research (Germany). Available at [www.mortality.org](http://www.mortality.org) or [www.humanmortality.de](http://www.humanmortality.de) (data downloaded on July 10, 2014).
- Jon Minton. Merging, exploring, and batch processing data from the human fertility database and human mortality database. Technical Report TR-2015-001, Max Planck Institute for Demographic Research (MPIDR), April 2015. URL <http://www.demogr.mpg.de/papers/technicalreports/tr-2015-001.pdf>.
- R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2012. URL <http://www.R-project.org/>. ISBN 3-900051-07-0.