1. Installing SANDRA

**How to Install the Standardized AND Reproducible Analyses (SANDRA) Library and Framework**

*Thomas Pronk*

[*tpronk@uva.nl*](mailto:tpronk@uva.nl)

*Version 1.5, 2017-06-29*

# Introduction

This is a guide on installing SANDRA (Standardized AND Reproducible Analyses) and setting up a SANDRA Analysis Framework (SAF). SANDRA is a library of R modules for processing data, with specialized features for decoding JASMIN and LOTUS datasets. An analysis framework is a standard way of organizing your analyses in *original* data, *interim* data, and analysis *scripts*.

# ****Part 1. Install Required Software****

## ****1a. Install R (and RStudio)****

Install one of the two applications listed below. I recommend to install RStudio for its ease of use.

1. R, which is a free and open source statistical programming language, with features that are similar to MATLAB. You can download it at <https://www.r-project.org/>)
2. A program for editing R-scripts. A plain text editor (such as Notepad/Kladblok) is a simple but sufficient solution. For a more powerful editor, consider installing **RStudio**: <https://www.rstudio.com>.

## ****1b. Only Mac Users: Install XQuartz****

If you use a Mac, you’ll also need to install XQuartz, which can be downloaded over here: <https://www.xquartz.org/>**. After installing XQuartz reboot your computer.**

**If you encounter the error message below, then see this page for a solution:** <http://tips.tutorialhorizon.com/2015/10/01/xcrun-error-invalid-active-developer-path-library-developer-commandline-tools-missing-xcrun/>

xcrun: error: invalid active developer path

(/Library/Developer/CommandLineTools), missing xcrun at:

/Library/Developer/CommandLineTools/usr/bin/xcrun

# ****Part**** 2. Run the SANDRA installation script

Create an *analysis folder* (also known as a directory or sub-directory) on your hard-drive, which will contain your analysis scripts and data. Next, **start RStudio (or R)** and copy-paste the installation script shown below into the **R console**. Be sure to paste the *whole* script (everything in the grey box). The most recent version of this script can be found here: <https://github.com/tpronk/sandra/blob/master/SAFs/template/Install%20SANDRA%20(from%20GitHub).R>

* Windows users can run the script as-is. A folder picker will appear that allows you to select the analysis folder you just created. *Note that you might need to minimize R/RStudio to make this folder picker visible.*
* Mac users need alter the installation script to specify the full path to their analysis folder, for example as follows:

installAnalysisFramework( “/studies/analysis” );

* If R asks you to “Select a CRAN Mirror”, you can pick any option in the list presented, though I recommend “0-Cloud”, which is the top-most entry in the list.

# Install & load package devtools

install.packages( "devtools" );

library( "devtools" )

# Install & load package SANDRA

install\_github( "tpronk/SANDRA/src", force = TRUE );

library( "sandra" );

safVersion = 1.00;

# Install SAF

installAnalysisFramework(

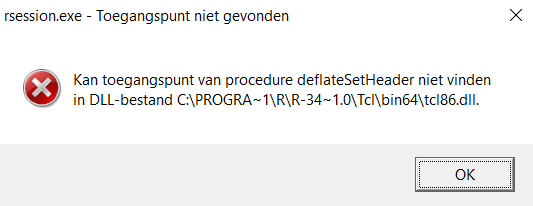
safVersion,

# Enter path to analysis folder here (without trailing slash)

# Only for Windows: leave empty to get a folder picker

);

If you encounter the error message below, it could help to restart RStudio and next only run the lines from the installation script from library( "sandra" ); onwards.



The installation script first installs the SANDRA R-package and next installs a SAF in the analysis folder that you specified. A SAF is a standard way of organizing your data. The installation scripts creates all scripts and directories required for the framework if they do not exist yet.

**Table 1.** Folder structure of a SAF

|  |  |
| --- | --- |
| **Folder or File** | **Explanation** |
| analysis/Load SANDRA.R | Loads SANDRA package and sets up your SAF for this analysis folder |
| analysis/interim | Your interim data (datasets produced by processing other datasets) |
| analysis/original | Your original unprocessed data |
| analysis/scripts | Your analysis scripts |

If everything went well, then the files and directories listed in Table 1 have been created. Finally, SANDRA confirms having successfully loaded the framework by showing the text seen below in the console.

sandra::FrameworkFileIO. Succesfully constructed FileIO

# ****Part**** 3. Reload your SAF

After installing, your analysis framework is installed and properly set up. You only need to install the SAF once, but you need to set it up each time you restart R to run an analysis. To setup your framework, run the script in ‘Load SANDRA.R’, which has just been created in your analysis folder.

# ****Acknowledgements****

My gratitude goes to many researchers that provided help. In particular, I would like to thank Joeri van Wijngaarden for co-developing tutorial materials, and Marilisa Boffo, Eline Lampers, Mae Nuijs, Ruby Smits, and Poppy Watson, for testing tutorials, finding bugs, and providing feedback on usability.